

Chapter 6

Sound localization

JOHN C. MIDDLEBROOKS*

*Departments of Otolaryngology, Neurobiology and Behavior, Cognitive Sciences, and Biomedical Engineering,
University of California at Irvine, Irvine, CA, USA*

INTRODUCTION

Spatial hearing permits a listener to identify the locations of sound sources and it aids in detection and recognition of sounds in the presence of other competing sounds. Of these functions, sound localization *per se* probably is the best understood, and that will be the main topic of this chapter. In most natural listening conditions, normal human listeners can localize sounds within a few degrees of accuracy in both horizontal and vertical dimensions (Stevens and Newman, 1936; Makous and Middlebrooks, 1990; Carlile et al., 1997; Yost et al., 2013). Figure 6.1 shows the performance of a listener in localizing brief broadband sounds.

Localization is a special challenge for the auditory system. In the visual and somatosensory systems, stimulus location is mapped directly on the sensory surfaces, i.e., the retina and the skin, respectively. In the auditory system, it is frequency that is mapped on the cochlear sensory surface, not location. Sound location must be inferred from spatial acoustic cues that arise from the physical interactions of sounds with the head and external ears. Those cues must be analyzed within the central auditory system and, in some way, integrated to create a neural representation of auditory space. We will see that this is not possible in every stimulus condition: examples are failures of horizontal localization of tones with frequencies around 3 kHz and failures of vertical or front/back localization of any sound that is restricted in bandwidth.

In this chapter, we will consider the acoustic cues on which listeners rely for source localization, the brainstem pathways in which those cues are analyzed, and the representations of sound locations in the midbrain and cerebral cortex. We treat the horizontal (lateral) dimension separately from vertical and front/back

dimensions because it turns out that localization in those dimensions relies on different acoustic cues and, necessarily, different brainstem pathways. We will consider the somewhat special problems of localization in distance and perception of sound motion. We will address some of the mechanisms that permit a listener to defeat spatial ambiguity potentially caused by the sound reflections in typical acoustic environments. Finally, we will consider present-day understanding of the representation of sound locations within the auditory cortex.

SOME TERMS AND TECHNIQUES

We begin by briefly defining some terms and describing some experimental techniques that have been used in spatial hearing research. The most direct way of testing sound localization seemingly would be to present sounds at varying locations remote from the listener and to have the listener make some judgment of those locations. Experimental conditions in which the sound source is at a distance from the listener are known as “open-field” conditions. Most often, investigators will strive for the particular open-field condition known as a “free field” in which the sound field is free of obstacles and in which there are no reflections from room surfaces. Although free-field conditions are possible in natural environments (e.g., a grassy or snowy field), such experiments most often are conducted in special anechoic chambers in which the walls, floor, and ceiling are treated to absorb sound. Generally, free-field conditions also imply that sound sources are presented in the “far field,” more than about 1 m from the center of the listener’s head. “Near-field” conditions, discussed below in the context of distance perception, introduce changes in localization cues that require special attention. Study of sensitivity to

*Correspondence to: John C. Middlebrooks, Ph.D., 116 Medical Science Building E, University of California at Irvine, Irvine, CA 92697-5310, USA. Tel: +1-949-824-0119, E-mail: j.midd@uci.edu

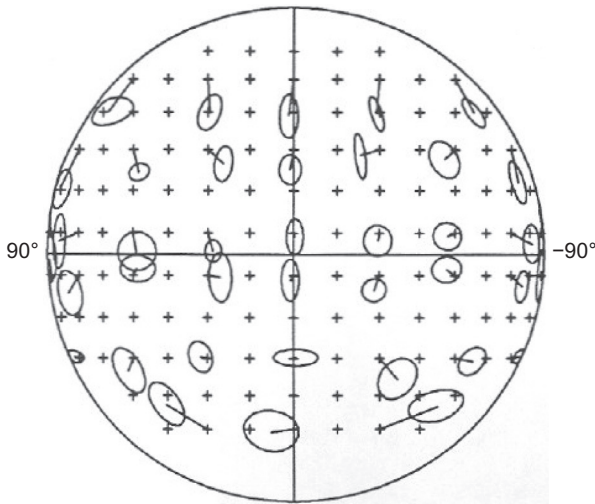


Fig. 6.1. Sound localization by a human listener. The coordinate system is drawn as if looking toward the listener, with 90° at the listener's right side. Stimuli were 150-ms broadband noise bursts presented from the locations indicated with plus signs, 1.2 m from the center of the listener's head. Listeners reported the locations by orienting to face the source while the head orientation was monitored electronically. Source locations were spaced in 10° intervals, although for clarity, responses are shown only in 20° intervals. The line segments connect each source location to the mean of five response locations. Horizontal and vertical axes of the ellipses indicate ± 1 SD of responses in horizontal and vertical dimensions. (Reproduced from Makous and Middlebrooks, 1990.)

particular acoustical cues to sound location might be conducted in “closed-field” conditions, meaning that sounds are presented through headphones. Typically in such experiments, a listener will attempt to distinguish a “dichotic” presentation, in which there is a difference in sound level or timing between the two ears, from a “diotic” presentation, in which sounds are identical at the two ears.

Finally, we consider “virtual auditory space” conditions, in which an investigator will attempt to present through headphones sounds containing all the natural localization cues, thereby eliciting a sensation of a realistic sound source at a well-defined location at some distance from the listener (Wightman and Kistler, 1989a, b). The procedure requires detailed measurements of filter characteristics of the listener's ears, typically derived by recording sounds from the two ear canals while a broadband probe sound is presented from systematically varying locations; the filters are referred to as “head-related transfer functions” (HRTFs). Arbitrary sounds then are processed with the HRTFs from the two ears and are presented through headphones, with or without some experimental manipulation. As we will see in a later section, head and ear acoustics vary among listeners. For that reason, optimal virtual audio is achieved with use of HRTFs

recorded from a listener's own ears, although use of “generic” HRTFs is acceptable in many situations. Indeed, many studies have employed HRTFs recorded from an acoustic manikin, such as the Knowles electronics manikin for auditory research (KEMAR). Virtual audio techniques have become refined and more widely accessible in the past decade and have provided many useful insights.

HORIZONTAL LOCALIZATION

Duplex theory

“Horizontal localization” primarily refers to the localization of a sound source to the left or right relative to the midline plane. Discrimination of front from back might be thought of as a component of horizontal localization, but the mechanisms of front/back localization are more like those of vertical localization and will be considered in that section. The foundation of our understanding of horizontal localization can be traced to the work of the English physicist and baron, Lord Rayleigh (Strutt, 1907). He worked largely at the end of the 19th century, often with Lady Rayleigh as the experimental subject, often using tuning forks as the sound sources. Rayleigh observed that left-versus-right locations of tuning forks could be discriminated across a broad range of frequencies. He appreciated that, at relatively high frequencies, the head is larger than the wavelength of the sound and casts an acoustic shadow. For that reason, the sound level at the ear nearer to a sound source would be greater than that at the farther ear, thereby providing a cue to the sound-source location. He reported: “When the pitch is pretty high, there is no doubt that this [intensity theory] explanation is adequate.” That explanation was more doubtful for sounds of lower frequency, at which the sound wavelength can be several times longer than the diameter of the human head. In that condition, the sound wave “bends” around the head and the difference in sound level at the two ears is negligible. After a series of ingenious experiments using pairs of slightly mistuned tuning forks or sounds directed independently to the ears with “India-rubber” tubes, Rayleigh determined that “the only alternative to the intensity theory is to suppose that the judgment is founded upon the difference of *phases* at the two ears.” That is, the difference in path lengths from a source to the two ears introduces differential delays, resulting in an interaural phase difference. The notion of a listener being sensitive to interaural phase differences was somewhat difficult to accept, because it implies that the auditory system is sensitive to interaural time differences (ITDs) of less than a tenth of a millisecond near the threshold of location discrimination. Nevertheless, such sensitivity has been confirmed with modern psychophysical (i.e., quantitative

perceptual) procedures and has been demonstrated at the level of brainstem neurons in animal physiology experiments.

Rayleigh's view of horizontal sound localization has come to be known as the "duplex" theory: that sound localization at low frequencies relies on differences in phase at the two ears, or ITDs, and that high-frequency localization relies on interaural differences in sound level (ILDs). Early free-field studies of localization of tonal stimuli (Stevens and Newman, 1936; Sandel et al., 1955; Mills, 1958) confirmed that sound location and location discrimination were accurate for sound frequencies below ~ 1.5 kHz, consistent with listeners' use of ITDs, and for sound frequencies above ~ 4 kHz, in the range at which sound wavelengths shorter than the diameter of the head result in substantial ILDs. There was an intermediate range of frequencies, roughly 2–4 kHz, at which neither ITD nor ILD provided a reliable cue and at which localization errors were greatest. Fortunately, this range of sounds that cannot be localized well has little impact on everyday sound localization because one is rarely asked to localize sounds limited to the 2–4-kHz frequency band.

The basic elements of the duplex theory have been quantified with modern psychophysical measures, and their anatomic and physiologic correlates are largely understood. Rayleigh's work was conducted using more or less pure-tone stimuli. The demonstrations that ITDs and ILDs are used for localization of low- and high-frequency sounds, respectively, were extended to broader-band sounds by Wightman and Kistler (1992), who used virtual auditory-space technology to deliver sounds in which ITDs and ILDs signaled opposing directions. Those authors demonstrated that, when sounds had broad frequency spectra containing low frequencies (below ~ 2 kHz), listeners' localization judgments were dominated by ITD cues irrespective of conflicting ILDs. When low frequencies were eliminated by high-pass filtering, the ITD dominance was lost. Macpherson and Middlebrooks (2002) extended those observations by showing that localization of high-pass (4–16 kHz) sounds is dominated by ILD cues, even in the presence of conflicting ITD cues.

Interaural time-difference cues

ITDs result from differences in the travel time from a sound source to the ears closer to and farther from the source. Woodworth (1938) modeled the ears and head as two points on a rigid sphere. Based on the geometry, he calculated ITD as:

$$\text{ITD} = a/c [\theta_{\text{rad}} + \sin(\theta_{\text{rad}})]$$

with a equal to the radius of the sphere, c the speed of

sound (~ 344 m/s at 20°C), and θ_{rad} equal to the angle in radians of the sound source relative to the midline; $\theta_{\text{rad}} = \theta_{\text{deg}} \times \pi/180$, where θ_{deg} is the angle in degrees. Woodworth used $a = 8.75$ cm as the effective radius of the sphere, whereas measures based on the best fits to actual ITD measurements in adult humans range from 8.75 to 10.57 cm (Kuhn, 1977; Middlebrooks, 1999a). The Woodworth model fits the data well for steady-state tones at frequencies ≥ 2 kHz. Kuhn (1977) showed that the effective radius of the head increases with frequency decreasing below ~ 2 kHz, as the wavelength of sound approaches the dimensions of the head, such that the effective head radius is larger by a factor of 1.5 at 0.5 kHz. Similarly, maximum values of ITD increase with frequency decreasing from 2 to 0.5 kHz.

The thresholds for detection of ITDs in pure-tone stimuli by well-practiced human listeners are around 10–20 μs (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Brughera et al., 2013). That is remarkable acuity when one considers that it is accomplished with neural machinery having time constants of several hundreds of microseconds or longer. Thresholds are minimal (i.e., sensitivity is greatest) for frequencies from 0.7 to 1.0 kHz, they increase slowly for frequencies < 0.7 kHz, and they increase dramatically above 1.0 kHz such that listeners are essentially insensitive to ITDs at frequencies 1.4 kHz and higher (Brughera et al., 2013). These thresholds refer to ITDs in the cycle-by-cycle fine structure of the waveforms of low-frequency sounds having periods on a time scale of ~ 0.7 –3 ms; it is fine-structure ITDs that dominate localization judgments. Listeners also are sensitive to ITDs in the envelopes of sounds, where "envelope" refers to variation in the amplitude of complex sounds on a time scale longer than ~ 10 ms. Thresholds for envelope ITDs tend to be 2–10 times larger (i.e., less sensitive) than for low-frequency fine structure (Henning, 1974; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1981; Blauert, 1982; Bernstein and Trahiotis, 1994, 2002). Envelope ITDs are rather ineffective in eliciting a sensation of a sound displaced away from the midline (Trahiotis and Bernstein, 1986), and listeners give relatively little weight to envelope ITDs when judging locations of high-frequency sounds (Macpherson and Middlebrooks, 2002).

Our understanding of the brainstem mechanisms of ITD sensitivity is shaped by the influential model of a "place" mechanism for ITD representation proposed by Jeffress (1948). In the Jeffress model (Fig. 6.2) pathways originate at the two ears, and signals travel with varying delays to converge on an array of "tertiary" bin-aural comparator neurons. The delay lines would compensate for differences in times of arrival of sounds at the two ears such that, for a given ITD, signals from

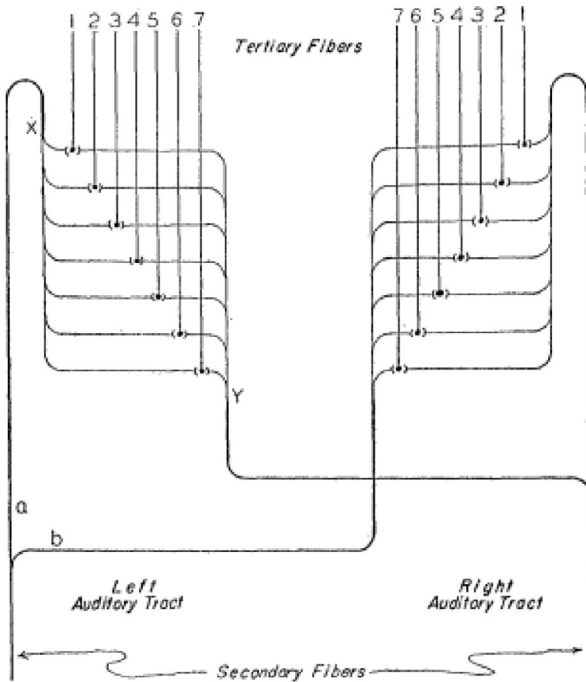


Fig. 6.2. A hypothetical mechanism for detection of interaural time differences. (Reproduced from [Jeffress, 1948](#).)

the two ears would arrive simultaneously at certain tertiary fibers, eliciting an action potential. In contrast, mistimed signals traveling on different delay lines would not elicit responses. The Jeffress model supposed that there would be comparator neurons for every realistic ITD. The bank of delay lines and comparators presumably would be replicated for each frequency band.

The Jeffress ITD-versus-place model appears to correspond very well to the functional organization of the avian nucleus laminaris ([Carr and Konishi, 1990](#); [Overholt et al., 1992](#); [Koppl and Carr, 2008](#)), which is a structure homologous to the medial superior olive (MSO) in mammals. In mammals, ITD-sensitive neurons much like those implied by Jeffress's hypothetic tertiary neurons have been identified in the MSO ([Goldberg and Brown, 1969](#); [Yin and Chan, 1990](#)). The pathways to the MSO neurons begin with auditory nerve fibers firing action potentials phase-locked to the fine structure of low-frequency sounds. The auditory nerve fibers synapse on the spherical bushy cells of the anterior ventral cochlear nucleus (AVCN) with high-temporal-fidelity end-bulbs of Held ([Ryugo and Seishiro, 1991](#)). Each principal neuron in the MSO receives excitatory projections from both the ipsilateral AVCN and, via the trapezoid body, the contralateral AVCN. The MSO neurons are activated by synchronized inputs from the ipsi- and contralateral inputs. The MSO also receives inhibition input, from the medial and lateral nuclei of the trapezoid body (MNTB and LNTB). The role of that inhibition in ITD

sensitivity presently is a topic of debate ([Pecka et al., 2008](#); [Roberts et al., 2013](#); [van der Hiejden et al., 2013](#)).

The MSO principal cells exhibit robust sensitivity to ITD, and for many years there was a general acceptance that something like the Jeffress model lived in the mammalian MSO. Recent studies, however, have cast some doubt on that notion. Specifically, the Jeffress model predicts that there must be neurons tuned to a range of ITDs at each frequency. In contrast, at least in small-headed mammals like gerbils, preferred ITDs of MSO neurons tend to vary with neurons' characteristic frequencies (CFs), and low-CF neurons tend to prefer interaural delays longer than the animal ever would experience in nature ([Brand et al., 2002](#)). These neurons tuned to unrealistic ITDs seemingly would be worthless in a strict place model of ITD because they would never receive their optimal stimulus. Conversely, there would be a dearth of neurons tuned to realistic ITDs. Nevertheless, a neuron that responds maximally to large, even unrealistic, ITDs will necessarily have a spike-rate-versus-ITD function that varies most steeply across ITDs near zero, corresponding to the range of sound-source locations at which spatial hearing is most acute. A model is gaining favor in which locations of low-frequency sounds are coded by the relative spike rates of neurons tuned to large ipsi- and contralateral-leading ITDs ([McAlpine et al., 2001](#)). Whether such a rate difference model applies to large-headed animals like humans presently is controversial ([Joris and Yin, 2007](#)). Relevant to that controversy is a recent study that attempted to model human detection of ITDs, tested at narrowly spaced pure-tone frequencies ([Brughera et al., 2013](#)). Across the frequency range in which ITD thresholds could be measured, 0.25 to near 1.4 kHz, the data were fit best by a hybrid model consisting of a Jeffress-like place model at the higher frequencies and a rate difference model at lower frequencies.

Interaural level-difference cues

The head casts an acoustic shadow for sounds having wavelengths shorter than the dimensions of the head, such that the sound level at the ear closer to the sound source is greater than that at the far, shadowed, ear. The resulting ILD potentially is a cue to the sound-source location. Conversely, sounds at frequencies lower than ~ 1 kHz have wavelengths longer than the diameter of the head, and ILDs at those frequencies are negligible. The magnitudes and spatial dependence of ILDs can be measured by recording from the ear canals with miniature microphones while presenting probe sounds in the free field ([Shaw, 1974](#); [Mehrgardt and Mellert, 1977](#)). Given a sound source around 90° from the midline, ILDs are around 20 dB at 4 kHz, increasing to ~ 35 dB at

10 kHz (Middlebrooks et al., 1989). At any given frequency, ILDs are roughly proportional to the sine of the angle of the sound source relative to the midline. The threshold for detection of an ILD in a high-frequency tone is ~ 1 dB (Mills, 1960; Hafter et al., 1977).

The lowest-level brainstem nucleus containing neurons sensitive to ILD is the lateral superior olive (LSO) (Boudreau and Tsuchitani, 1968). Principal cells in the LSO receive excitatory inputs from spherical bushy cells in the ipsilateral AVCN. The contralateral pathway is inhibitory, involving an excitatory projection from globular bushy cells in the contralateral AVCN, crossing the midline in the trapezoid body, and terminating in the MNTB. Neurons in the MNTB send inhibitory glycinergic projections to the LSO. The ipsilateral excitation and contralateral inhibition that converge on single LSO neurons result in sigmoid spike-rate-versus-ILD in the LSO: activity is low for ITDs favoring the contralateral ear, increases most steeply for ILDs around 0 dB, and is high for ipsilaterally favoring ILDs.

Monaural conditions

Loss of hearing in one ear disrupts interaural difference cues for horizontal localization. If one ear of a normal-hearing listener is plugged, the listener will tend to localize all sounds to the side of the open ear. It is thought that most individuals who have acquired hearing loss in one ear after some stage of maturity are unable to localize sounds. A study of monaural localization included five individuals who had congenital deafness in one ear and normal hearing in the other (Slattery and Middlebrooks, 1994). Two of those unilaterally deaf patients localized all sounds to the side of the hearing ear, resembling normal-hearing listeners wearing an earplug. The other three unilaterally deaf individuals localized reasonably well. They could localize sources on both sides of head with errors within a factor of 2 of those of normal-hearing controls and with accuracy far better than that of normal-hearing controls wearing an earplug. The interpretation is that the successful unilaterally deaf localizers had learned to make use of spectral-shape cues for horizontal localization in a manner analogous to the use of such cues by normal-hearing listeners for vertical and front/back localization, as described in the next section.

VERTICAL AND FRONT/BACK LOCALIZATION

Interaural differences in time and level are effective cues for the angle of displacement of a sound source relative to the midline plane. Knowledge of that lateral displacement, however, narrows the range of possible sound sources only to a “cone of confusion” centered

on the interaural axis (Woodworth, 1938) – a given ITD or ILD could be produced by a source at any vertical or front/back location on such a cone. If asked to localize a pure tone, a listener can respond fairly accurately if the possible sources are restricted to the front half of the horizontal plane. If the source location can vary throughout the horizontal plane, however, the listener will make front/back confusions. If the source also is free to vary in elevation, pure-tone localization is pretty hopeless.

Vertical and front/back localization can improve dramatically if the sound source has a relatively broad, flat spectrum, like that of many natural sounds; the important spectral range is ~ 4 kHz and higher. Direction-dependent filtering of such a spectrum by the head and pinnae (i.e., the external flaps of the outer ears) provides spectral-shape cues that are the principal cues for vertical and front/back source location. Specifically, the various convolutions of the pinnae introduce resonances and antiresonances that result in spectral peaks and notches in the spectrum of sound reaching the tympanic membranes. Those spectral features vary in frequency according to the angle of incidence of sound relative to the pinna. Batteau (1967) first considered the filtering action of the pinna in the time domain, but most other authors have emphasized the effects of the pinna on spectral amplitude. The importance of pinna acoustics for localization has been demonstrated by filling or otherwise disrupting the cavities of the pinna, which results in prominent disruptions of vertical localization (Musicant and Butler, 1984; Humanski and Butler, 1988). Spectral-shape cues often are referred to as “monaural spectral cues” – monaural because the necessary spectral analysis can occur at each ear individually. Spectral cues are available to both ears of a normal-hearing listener, but when a sound source is located to one side or the other, the greatest weight in localization judgments is given to spectral cues from the pinna on the same side as the source (Macpherson and Sabin, 2007).

Successful use of spectral-shape cues for localization requires that listeners be familiar with the filter properties of their own ears and that the sound source is relatively broad and flat in spectrum. Listeners appear to hold to the broad, flat spectrum assumption even when that assumption obviously is invalid. When asked to localize 1/6- or 1/3-octave band-passed sounds, listeners report erroneous vertical and front/back locations that vary with center frequency of the stimulus, irrespective of the actual source locations (Blauert, 1983; Butler, 1987; Middlebrooks, 1992). In such narrowband vertical-localization experiments, comparison of the patterns of localization judgments with the spectral transformations by listeners’ pinnae (quantified by their HRTFs) suggests that listeners tended to misinterpret the narrowband sounds as broadband sounds that had

been filtered by their pinnae to introduce a spectral peak, even though it was clear to the listener that the source spectrum was narrowband.

A quantitative model predicted vertical localization judgments successfully by assuming that listeners correlate the spectra of sounds in their ear canals with internal libraries of spectral templates, corresponding to their HRTFs for various sound directions, and report the elevation corresponding to the best-fitting template (Middlebrooks, 1992). That model, and several subsequent template-matching schemes (Hofman and Van Opstal, 1998; Chung et al., 2000; Langendijk and Bronkhorst, 2002; Bremen et al., 2010), did not offer any specific physiologic mechanism for template matching. Visual inspection of HRTFs reveals prominent spectral notches that vary in frequency according to sound direction, and a number of studies have seized on notches as the key vertical localization cue (Hebrank and Wright, 1974; Bloom, 1977; Musicant et al., 1990), but one is not aware of a conclusive demonstration that spectral notches signal vertical sound location exclusive of other spectral features.

The sizes and detailed shapes of heads and ears vary among listeners and, not surprisingly, their HRTFs also vary. A perceptual consequence of that interlistener variability can be demonstrated by recording HRTFs from one individual and processing sounds through those HRTFs to present to a different listener. Such an ear swap can result in prominent vertical localization errors, especially as an increase in the incidence of front/back confusions (Wenzel et al., 1993; Middlebrooks, 1999b). The differences among ears is not haphazard. Basic features of HRTFs are conserved among listeners but tend to scale in frequency roughly in proportion to listeners' physical sizes (Middlebrooks, 1999a). Hofman and colleagues (1998) have demonstrated that vertical localization can adapt to physical changes in the acoustics of a person's pinnae over the course of ~6 weeks. All these results are consistent with the view that the aspects of head and ear shape that are relevant for vertical localization are similar among most individuals, and that corresponding spectral templates are fine-tuned based on listeners' experience.

It is difficult to point to a single brainstem nucleus that could accomplish the broadband spectral template matching that is needed for vertical localization. Nevertheless, there is good evidence that type IV cells in the DCN are sensitive to spectral notches like those found in HRTFs (Young et al., 1992; Imig et al., 2000). Those cells are excited by broadband sounds, and that excitatory response is inhibited by the introduction of a spectral notch close to the cell's best frequency. Neurons in the inferior colliculus that appear to receive a direct projection from the contralateral DCN show excitatory

responses that are selective to the center frequencies of notches in broadband sounds. These observations demonstrate that neurons in the DCN can code at least one element of the spectral cues for vertical localization. A necessary role of the DCN in vertical localization behavior has been shown by demonstrations that lesions of dorsal acoustic stria, which carries the output of the DCN, can result in deficits in unconditioned (Sutherland et al., 1998) or conditioned (May, 2000) orienting responses to elevated sounds.

The spectral features of HRTFs that vary as a function of vertical source location also vary systematically with horizontal location (Shaw, 1974). For that reason, one might expect spectral-shape cues to contribute to horizontal localization. Indeed, as mentioned in a previous section, there is some evidence of use of spectral cues for horizontal localization by some unilaterally deaf individuals (Slattery and Middlebrooks, 1994). In listeners with normal binaural hearing, however, the interaural difference cues (i.e., ITD and ILD) appear to dominate horizontal judgments, with little or no influence of spectral-shape cues (Macpherson and Middlebrooks, 2002). Interaural difference cues are less vulnerable to peculiarities in the source spectrum than are spectral-shape cues. It seems as if, when confronted with conflicting cues, the auditory system favors the cues that are the most reliable.

DISTANCE LOCALIZATION

Sound localization in the third dimension, distance, is not nearly as accurate as that in the horizontal and vertical dimensions. In optimal conditions (e.g., distance <2 m, broadband noise bursts, reflective room acoustics, target located to one side), reported distances can correlate closely with actual distances; correlation coefficients were >0.8 (Kopco and Shinn-Cunningham, 2011). In studies using unfavorable conditions (e.g., far field, recorded sounds presented at randomized levels, anechoic environment, target in front), distance judgments show random-chance performance levels (Coleman, 1962; Gardner, 1969; Holt and Thurlow, 1969). Listeners tend to underestimate distances that are greater than ~1.5 m and tend to overestimate distances that are <1.5 m (Zahorik, 2002; Zahorik et al., 2005). Distance judgments generally are more accurate for lateral sounds than for sounds in the midline plane, both for far and for nearby sources (Holt and Thurlow, 1969; Brungart et al., 1999; Zahorik, 2002; Kopco and Shinn-Cunningham, 2011). Distance judgments are influenced by familiarity of the sound source, source sound pressure level, source spectrum, source radiation pattern, azimuth of the source relative to the listener, acoustics of the surroundings, and near-versus-far range of

distances. Visual cues also influence sound distance judgments, but they are beyond the scope of this review.

The principal acoustic cues for distance perception are intensity (i.e., sound level arriving at the listener's ears), direct-to-reverberant (D/R) energy ratio, and ILD. The relative importance of these cues varies widely across conditions. The intensity cue arises from the physical attenuation of a sound with distance. Given a point sound source in anechoic conditions, sound intensity arriving at the listener will decay by 6 dB with every doubling of the distance; the rate of decay is lower in reflective surroundings or if the source is directional. High and low intensities arriving at the listener signal near and far source distances, respectively. Of course, use of the intensity cue is confounded by uncertainty about source levels. Familiarity with the source can improve use of the intensity cue. For instance, a quiet word from a nearby speaker or a distant shout might arrive at the listener with equal intensities, but the listener could recognize the vocal effort and attribute the low intensity of the shout to attenuation across a greater distance (Gardner, 1969). Listeners give particular importance to intensity cues for the distances of sources in the midline plane, where D/R cues are weak and ILD cues are negligible (Zahorik, 2002).

D/R energy ratios are probably the most reliable distance cues in reverberant conditions as, for instance, in most office- to classroom-sized rooms. In such a room, a listener will receive sound directly from the source and also indirectly from reflections from multiple room surfaces. As discussed, the direct energy tends to decay with increasing distance, whereas the reverberant energy reaching the listener is largely independent of the source-to-listener distance. For those reasons, the ratio of D/R energy tends to decay with increasing source distance. The D/R ratio at a particular ear tends to be maximum for sound sources straight to the side, declining for midline source locations, and declining further for contralateral locations (Zahorik, 2002; Kopco and Shinn-Cunningham, 2011). In a study that simulated the spatial cues available in a small classroom, Kopco and Shinn-Cunningham (2011) demonstrated that distance judgments for simulated sources in the range of 15–170 cm correlated with D/R cues more than with any other potential cue.

ILDs are essentially independent of distances >1 m, but ILDs can increase markedly as lateral sources approach closer than ~ 1 m (Brungart and Rabinowitz, 1999); ILDs are negligible for midline sources regardless of distance. Near-field ILDs are present even at low frequencies at which they are absent in the far field. Use of ILD for distance judgments potentially could be confounded with uncertainty about the lateral location of the source. That confound can be resolved, however,

by identification of lateral location on the basis of ITD cues; ITD cues give a reliable signal for lateral location, largely independent of distance (Brungart and Rabinowitz, 1999). Brungart (1999) has argued that ILDs are the dominant cues for judgments of the distances of nearby sources, and that is almost certainly true for anechoic conditions. Results from reverberant conditions, however, suggest that D/R cues dominate distance judgments when room reflections are available (Kopco and Shinn-Cunningham, 2011).

No brainstem structure has been identified that is sensitive specifically to distance cues aside from the LSO, which was discussed above as the likely site of ILD analysis.

MOTION PERCEPTION

One's everyday experience affirms the perceptual salience of a moving sound source. A sound source that moves in distance relative to a listener produces both Doppler shifts in frequency and modulation in loudness. We will focus here, however, on the problem of motion in azimuth or elevation, i.e., change in direction. It is a matter of some controversy whether such directional motion perception is a product of motion-specific brain mechanisms, or whether it is simply another aspect of localization. There certainly is no evidence in hearing for the degree of specialization for motion detection that is well known in the visual system, in which neurons are specific for the direction and velocity of visual motion (Hubel and Wiesel, 1962) and in which a motion-specific cortical area has been characterized (i.e., area MT of primates) (Mikami et al., 1986). In psychophysical studies, humans can discriminate velocities of visual targets differing by $<10\%$, whereas discrimination of auditory target velocities require 20–70% differences (Carlile and Best, 2002).

A number of early studies represented auditory motion sensitivity by the minimum audible movement angle (MAMA), which was the smallest detectable change in location of a moving stimulus. The MAMA may be compared with the minimum audible angle (MAA), which is the smallest detectable difference in location of two static sounds presented successively. In the motion studies considered in an early review (Middlebrooks and Green, 1991), MAMAs all were substantially larger than MAAs, indicating that motion did not improve (and generally impaired) detection of a change in location. Based on such data, a number of investigators have concluded that auditory motion detection does not reflect sensitivity to specific dynamic cues such as changing ITD or ILD. Instead it might be that the auditory system simply takes two or more “snapshots” of source locations at successive times and interprets

changes in location as evidence for motion (Grantham, 1986). At the time of that 1991 review, we concluded that “As yet there is no compelling evidence for motion-sensitivity systems in the auditory system.” Several more recent observations tip the balance somewhat in favor of motion-sensitive systems, although it is not yet clear that the sum of evidence reaches the threshold for “compelling.” The newer observations to be considered include those of motion aftereffects, neurons sensitive to dynamic ITDs and ILDs, and cortical areas in humans activated by moving sounds.

“Motion aftereffects” refer to phenomena in which prolonged exposure to a moving sound influences perception of the motion of a following probe sound; this is analogous to the “waterfall illusion” in vision. Grantham (1986) found that exposure to a moving noise source could result in a bias of the judgment of the direction of movement of a probe. The auditory aftereffect was considerably less robust than is seen in vision in that the magnitude of the auditory aftereffect was smaller and its duration was only a few seconds; the visual-motion aftereffect can last for a minute or more. Grantham speculated that the auditory aftereffect that he observed was due to some combination of response bias (a tendency to report probe motion contrary to that of the adaptor) and a sensory effect consisting of a loss of sensitivity to probe velocity. Grantham (1992) later demonstrated that exposure to a moving adaptor could decrease motion sensitivity as indicated by enlarged MAMAs. The notion of generalized suppression of motion sensitivity by a moving adaptor is supported by a study of auditory evoked potentials (Magezi et al., 2013). The auditory event potentials elicited by a stationary probe stimulus differed depending on the characteristics of a preceding auditory adaptor. There was no significant difference in the responses conditioned by leftward compared to rightward adaptor motion, suggesting a lack of directionality in a putative motion system. There was, however, a significant difference between effects of a bidirectional-moving adaptor compared to a stationary adaptor, consistent with the notion of generalized suppression of motion sensitivity.

Studies of ITD sensitivity neurons in the inferior colliculus have demonstrated that the response to a particular ITD can vary according to the context of a varying ITD (Spitzer and Semple, 1991, 1998). That is, the best ITD of a neuron can be influenced by the range of ITDs that is traversed in a sweep of varying ITD. Similarly, a study of ILD sensitivity in the inferior colliculus (Sanes et al., 1998) demonstrated that the ILD sensitivity of neurons depended on their history of prior stimulation by a dynamic stimulus. One might regard these demonstrations of sensitivity to dynamic ITDs and ILDs as a substrate for motion perception. Alternatively, as pointed

out by Spitzer and Semple (1998), these phenomena might be “non-adaptive by-products” of neural mechanisms that might account for the psychophysical observations of degraded spatial acuity for moving sounds (i.e., that MAMAs are wider than MAAs).

Several human cortical imaging studies using positron emission tomography or functional magnetic resonance imaging have identified cortical areas that show enhanced activity during presentation of sounds that simulate sound-source motion compared to control conditions of static locations (Griffiths et al., 1994, 1998; Baumgart et al., 1999; Lewis et al., 2000; Parvani et al., 2002; Warren et al., 2002; Krumbholz et al., 2005). Involved areas generally are located in the right hemisphere, including the insula, planum temporale, and posterior parietal cortex. Although one could readily accept those results as evidence of motion-specific brain areas, one must also entertain the alternative interpretation that those cortical areas simply are sensitive to static sound-source location and that they respond more strongly to the broad range of locations provided by the moving source than to a single fixed source. Indeed, studies that have attempted to control for the number of stimulus locations have failed to show motion selectivity (Smith et al., 2004, 2007). Similarly, there are multiple interpretations possible in a case study in which a right-sided lesion involving the insula, the planum temporale, and the posterior parietal cortex resulted in a pronounced deficit in sensitivity to simulated moving sounds (Griffiths et al., 1996). Although the loss of sensitivity to dynamic sounds was clear, that patient also exhibited a deficit in “fixed lateralization,” including elevated thresholds for detection of static ITD and ILD. In support of the argument that responses to changes in location can appear as motion sensitivity, Getzmann and Lewald (2012) recorded very similar “motion-onset responses” (i.e., auditory evoked potentials recorded from the human scalp, elicited by sound movement) in response to simulated-motion sounds that moved continuously, moved abruptly from one sound hemifield to the other, or moved randomly among a range of static locations.

In summary, listeners are sensitive to motion of sound sources, and their judgments of the movement of a probe sound can be influenced by a recent history of motion stimulation. Similarly, the sensitivity of neurons to spatial cues can be modulated by prior stimulation. Clearly, there are brain regions that respond well to moving sounds, but it is difficult to distinguish responses to continuous motion from responses to multiple static stimuli. Although the presence of motion-sensitivity systems in the auditory system is an attractive notion, evidence for such systems still is less than compelling.

LOCALIZATION IN REVERBERANT SPACES: THE PRECEDENCE EFFECT

Most of our listening occurs in reverberant environments in which sounds reflect from walls, floor, ceiling, and other structures. That means that when a sound is emitted, the listener first is exposed to the sound traveling on a direct path from the source, followed by a few early reflections, followed by the reverberation due to the interaction of all the reflections. We already have considered how the ratio of D/R energy can aid estimation of source distance, but one might expect the reverberations to confound estimation of source direction. Our everyday experience, as well as a large body of research, indicates that the auditory system is quite effective in limiting localization judgments to the sound that arrives first, via the direct path, and disregarding the directions of sound reflections. The group of auditory phenomena encompassing the factors that facilitate localization in the presence of reflections is known as the “precedence effect”; one of several good reviews is by [Litovsky and colleagues \(1999\)](#).

The precedence effect has most often been studied using brief sounds presented from two free-field locations or, under headphones, from two locations simulated using pairs of ITDs. A range of perceptions is reported depending on the delay between the leading and lagging sound. When the delay is less than ~ 5 ms, the two sounds tend to fuse and the listener reports hearing only a single sound. At 0 ms delay, the listener tends to localize the sound midway between the two sources. As the delay increases from 0 to ~ 1 ms, the perceived source location gradually drifts toward that of the leading source. This is the range of “summing localization,” which is exploited for the purpose of stereophonic sound reproduction. At further increasing delays, “localization dominance” occurs in which the perceived location corresponds to that of the leading source. At a delay > 5 ms, known as the “echo threshold,” the listener begins to report hearing two sounds. Even though two sounds are heard in this condition, the localization judgment is still dominated by the leading source and discrimination of multiple lagging-source locations is impaired (this is “discrimination suppression”). Only after delays of > 10 ms is a listener able to judge accurately the location of the lagging-sound source. The various delays given here are minimum values, appropriate for brief sources. The respective delays generally all are longer under conditions of longer-duration sounds.

The physiologic mechanisms of the precedence effect involve some combination of the binaural neurons of the superior olivary complex plus suppressive mechanisms at midbrain and cortical levels. In anesthetized cats, correlates of fusion are present at the level of the inferior

colliculus in the sense that single neurons respond reliably to the first of two sounds and respond to the lagging sound only after some delay ([Yin, 1994](#)). The delay at which the neural response to the lagging sound recovers to half that of the leading sound varies widely across neurons but has a median of 20 ms, not much longer than that observed in human psychophysics. In unanesthetized cats, locations of sound sources can be estimated from the spike patterns of auditory cortical neurons. In the primary auditory cortex, neurons reliably indicated the locations of leading sounds. Responses of those neurons to lagging sounds were suppressed, and accurate estimates of the locations of lagging sounds could be obtained from only a small subset of neurons ([Mickey and Middlebrooks, 2005](#)).

The notion that the precedence effect requires considerable central integration is supported by studies of its development in humans ([Muir et al., 1989](#); [Litovsky et al., 1999](#)). Infants can distinguish between two sound sources within hours of birth, but localization dominance appears first at 4–5 months of age, maturing fully by about 5 years.

CENTRAL REPRESENTATION OF SOUND-SOURCE LOCATIONS

We have considered the various cues for sound-source locations and described brainstem pathways that process those cues. Now we address the question of how and where in the auditory system information from the spatial cues might be integrated to form neural representations of the locations of sources in auditory space. Existing evidence suggests that there are at least two very different modes of central spatial representation: a topographic representation in the superior colliculus (SC) and distributed representations in the auditory cortex.

Spatial topography in the superior colliculus

The SC is a sensorimotor integrative structure that directs orienting movements of the head and eyes toward sources of visual, tactile, and auditory stimuli. Studies conducted under anesthetized conditions in a number of non-primate species have demonstrated topographic maps of auditory space in the intermediate and deep layers of the SC: cat ([Gordon, 1973](#); [Middlebrooks and Knudsen, 1984](#)); guinea pig ([Palmer and King, 1982](#)); ferret ([King and Hutchings, 1987](#)); rat ([Gaese and Johnen, 2000](#)); a map also is seen in a homologous structure, the optic tectum, in barn owls ([Knudsen, 1982](#)). Here, “topographic” means that neurons in the SC respond maximally to sounds presented within a more-or-less restricted region of space (often called the “best area”) and that best areas vary systematically in space according to locations of neurons within the SC. Most SC

neurons that respond to sounds also respond to visual stimuli. The auditory map lies in approximate spatial register with visual maps in the SC. This is only “approximate” because auditory spatial sensitivity extends further lateral than do visual responses, consistent with the 360° spatial sensitivity of hearing compared with the more restricted visual fields. Auditory and visual spatial tuning for near-midline locations tends to show fairly close mutual alignment whereas SC neurons with more eccentric receptive fields tend to show auditory best areas shifted systematically lateral to the centers of visual receptive fields.

The static auditory space maps that have been described in the SC in non-primate species arguably reflect the use of general anesthesia in studying those species. Under those conditions, neurons show sensitivity to source locations relative to the positions of head and external ears, i.e., in head-centered coordinates. In non-human primates, study of the SC generally has been conducted in awake conditions in which the animal is required to make eye movements to visual and/or auditory targets (Jay and Sparks, 1987). Under such conditions, the spatial sensitivity is interpreted more easily as a motor error signal reflecting the gaze shift needed to acquire the sensory target, and the auditory topography appears integrated within a map of motor error. Auditory spatial signals appear to transform within the SC from head-centered to eye-centered coordinates (Lee and Groh, 2012).

Distributed spatial representation in the auditory cortex

Normal auditory spatial perception and behavior requires intact auditory cortex. Experimental lesions or reversible inactivation of the auditory cortex in laboratory animals (Thompson and Cortez, 1983; Jenkins and Merzenich, 1984; Kavanagh and Kelly, 1987; Malhotra et al., 2004) and clinical lesions of auditory cortex in humans (Sanchez-Longo and Forster, 1958; Klingon and Bontecou, 1966; Zatorre and Penhune, 2001) result in deficits in the ability to report the location of a sound source, either by walking to the source (by animals) or by some other means of report by humans (see Chapter 32).

Given the apparent role of auditory cortex in sound localization, it may come as a surprise that no topographic representation of auditory space has been identified in the cortex, despite the efforts of several research groups to find such a map (Middlebrooks and Pettigrew, 1981; Imig et al., 1990; Rajan et al., 1990). In the primary auditory cortex (A1), a large proportion of neurons responds with more than half of neurons’ maximum responses to sounds from all tested locations; such “omnidirectional” neurons constitute around half the

sample in some species and some experimental conditions (Harrington et al., 2008; Werner-Reiss and Groh, 2008; Zhou and Wang, 2012). Among the neurons that exhibit sharper spatial sensitivity, many have “contralateral hemifield” spatial tuning, meaning that they respond strongly to sound sources located throughout the contralateral half of space, with their responses cutting off near the midline. A smaller proportion of units responds most strongly to near-midline or ipsilateral sources. In area A1 of the cat, study in awake and in awake-behaving conditions (Mickey and Middlebrooks, 2003; Lee and Middlebrooks, 2013) reveals spatial tuning that is somewhat more restricted and biased toward the frontal midline than is seen in anesthetized conditions (Harrington et al., 2008). Nevertheless, it would be difficult to argue that spatial sensitivity in awake conditions is qualitatively different from that observed with anesthesia. Contrary to the requirement for a topographic representation, there has been no indication of a systematic progression of best area as a function of cortical area in either awake or anesthetized conditions in any species. Instead, the cortical distribution of spatial tuning tends to be modular in that neurons within a restricted cortical region tend to show one spatial preference while neighboring regions can show quite different preferences (Middlebrooks and Pettigrew, 1981; Rajan et al., 1990; Middlebrooks and Bremen, 2013). There is some indication of a systematic progression of spatial receptive field cutoffs in area A1 in the pallid bat (Razak, 2011), but no such organization has been seen in any of the other species that have been examined.

Examination of the temporal response patterns of auditory cortical neurons has suggested a form of spatial representation that is qualitatively different from the topographic place codes that are familiar in visual and somatosensory systems. In many cases, response patterns vary systematically with sound-source location such that the locations of stimulus sound sources can be identified with some accuracy for source locations throughout 360° of space (Middlebrooks et al., 1994, 1998). Figure 6.3 shows an example of one neuron in the second auditory area (A2) of an anesthetized cat. The dot raster plot in Figure 6.3A shows how the temporal response pattern varied with sound-source location, and Figure 6.3B shows estimates of source location based on computer recognition of those response patterns. The property of single neurons signaling source locations throughout a broad range of locations has been referred to as “panoramic” location coding (Middlebrooks et al., 1994). That individual neurons can signal locations panoramically and that such neurons are distributed throughout the auditory cortex indicate that the representation of any particular source location is distributed throughout the auditory cortex.

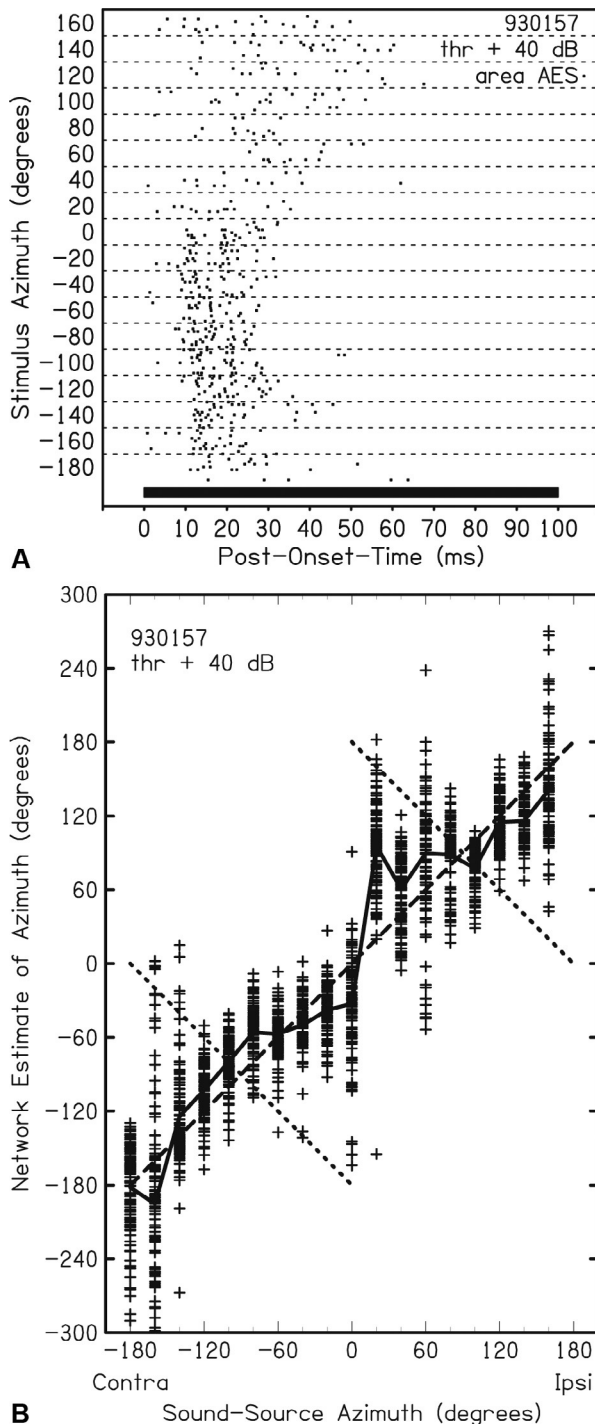


Fig. 6.3. Spatial sensitivity of a cortical neuron in area A2 of an anesthetized cat. **(A)** In this dot-raster plot, each row of dots represents the pattern of action potentials elicited by a 100-ms broadband noise burst presented from the indicated location in azimuth. Responses to eight such bursts are shown for each azimuth. **(B)** Estimates of source locations based on spike patterns from the neuron in **(A)**. Each symbol represents one estimate based on a bootstrapped response pattern based on responses to eight stimulus presentation. The dashed line

Cortical neurons transmit information about sound-source locations both with their spike counts and with the timing of spikes. Several studies have demonstrated that the accuracy of location estimation based on cortical spike patterns is degraded when temporal information is degraded or eliminated (Middlebrooks et al., 1994; Furukawa and Middlebrooks, 2002; Mickey and Middlebrooks, 2003). Indeed, in some conditions first-spike latencies transmit more location-related information than do spike counts (Furukawa and Middlebrooks, 2002). The accuracy of panoramic coding by cortical spike patterns improves when information is combined across multiple neurons (Furukawa et al., 2000; Miller and Recanzone, 2009; Lee and Middlebrooks, 2013). In one study, information from ensembles of 16 of the most informative neurons signaled source locations throughout space with accuracy comparable to the performance of cats in a psychophysical localization task (Furukawa et al., 2000).

The sensitivity of cortical neurons for locations of sound sources varies considerably depending on anesthetic and behavioral conditions and on the presence of other sounds. As noted above, the spatial sensitivity of neurons in cat area A1 tends to be sharper and the distribution of best areas shifted closer to the frontal midline in awake compared to anesthetized conditions (Mickey and Middlebrooks, 2003; Lee and Middlebrooks, 2013). In awake cats, spatial sensitivity of neurons in A1 sharpens on individual trials on which the animal performs a listening task compared to trials on which it is idle, and many neurons show additional sharpening when the task requires localization compared to a non-spatial task (Lee and Middlebrooks, 2011). A recent study of cortical responses to temporally interleaved sequences of sounds from spatially separated sources demonstrated that the selectivity of neurons for the location of a target source sharpens dramatically in the presence of a competing sound from another location (Middlebrooks and Bremen, 2013). These observations all support the picture of a dynamic, highly distributed cortical representation of auditory space rather than a rigid point-to-point topography.

Despite their panoramic spatial sensitivity, in most cases the spatial acuity of cortical neurons is greatest for sound sources located near the frontal midline (Stecker et al., 2005). This accords with human psychophysical results showing highest acuity (Mills, 1958) and

indicates loci of perfect performance, and the solid curve connects the mean values of estimates for each source location. The estimates involved recognition of spike patterns using an artificial neural network. AES, anterior ectosylvian sulcus. (Reproduced from Middlebrooks et al., 1998.)

greatest localization accuracy (Stevens and Newman, 1936; Makous and Middlebrooks, 1990) for near-midline targets. Several investigators have considered the possibility that sound locations are represented by the opponent activity of a small number (i.e., possibly two or three) of neural channels, each tuned to a rather broad region of space; channels tuned to the right or left half of space might be one example. Opponent-channel models draw support from human psychophysical studies of gap detection, spatial release from masking, and selective adaptation in humans (Boehnke and Phillips, 1999; Phillips et al., 2003; Dingle et al., 2012, 2013), from cortical single-neuron recordings in animals (Stecker et al., 2005), and from far-field magnetic or electric recordings in humans (Salminen et al., 2009; Magezi and Krumbholz, 2010; Briley et al., 2012). Models that assume summed activity across large populations of similarly tuned neurons capture the generally contralateral bias of neural spatial tuning in each hemisphere and can account well for results of low-resolution scalp recordings. Nevertheless, the variety of spatial sensitivity that can be recorded at the level of single cortical neurons and that is not captured by limited-channel models suggests that two- or three-channel models should be regarded as only first-order approximations of the details of cortical spatial processing.

Cortical areas specialized for sound localization

In the previous section, we emphasized the primary auditory cortex. Here, we consider additional auditory cortex areas that do or do not show specialization for localization, we consider the issue of laterality of spatial representation in humans and non-humans, and we consider hypothetical “what” and “where” processing streams in the human and non-human primate cortex.

Single-neuron recordings in cats have demonstrated at least some spatial sensitivity in each of six distinct cortical areas in which such sensitivity was evaluated (Middlebrooks et al., 1998; Harrington et al., 2008). Nevertheless, the sharpness of spatial tuning and the sound-level invariance of spatial tuning varies considerably among areas, with the sharpest spatial tuning seen in the posterior auditory field (PAF) and the dorsal zone (DZ), intermediate sharpness of tuning in A1, A2, and the anterior ectosylvian area, and the weakest tuning seen in the anterior auditory field (AAF). That ranking of the sharpness of spatial tuning parallels the results of behavioral experiments involving reversible inactivation of restricted cortical areas. Inactivation of PAF or DZ disrupted trained cats’ ability to walk to the location of a sound source (Malhotra et al., 2004). In a related study, inactivation of AAF disrupted performance of

a task requiring discrimination of temporal patterns while sparing localization performance (Lomber and Malhotra, 2008). In the same cats, inactivation of PAF disrupted localization but spared the temporal discrimination. An interpretation is that intact PAF and DZ are necessary for sound localization, whereas the spatial sensitivity in AAF might support some other aspect of spatial hearing, such as spatial release from masking (Middlebrooks et al., 2002).

Studies in cats and monkeys suggest specialization of cortical areas for localization in particular regions of space. In awake cats (Lee and Middlebrooks, 2013), best areas of neurons in area DZ show a bias toward representation of frontal space, with a large majority of spatially tuned neurons showing best areas centered within 45° of the frontal midline. In contrast, area PAF in awake cats shows a more uniform representation of space, with about equal numbers of spatially tuned neurons showing best areas centered in the frontal or contralateral quadrant of space. Analysis of signaling of sound-source location by responses of ensembles of 16 neurons shows most accurate localization of near-midline sources by DZ neurons and of more peripheral sources by PAF neurons (Lee and Middlebrooks, 2013). In the awake macaque (Woods et al., 2006), distributions of best azimuths in most of the tested cortical areas show large proportions of neurons showing far-peripheral best areas. In contrast, the largest proportion of neurons showing best areas around the frontal midline was in the core area (R). Miller and Recanzone (2009) tested identification of sound-source locations based on responses of ensembles of cortical neurons in the macaque data set (Woods et al., 2006). Across all source locations, localization was most accurate based on responses of neurons in the caudolateral area (CL). Examination of the figures, however, shows that accurate localization by CL neurons was seen only for far-contralateral targets. Localization of sources around the frontal midline was most accurate based on responses of neurons in R; note that the frontal midline is where psychophysical spatial acuity is highest. It is difficult to draw homologies between particular cortical areas in cats and macaques, but both species show analogous areas that apparently are specialized for frontal (areas DZ and R) and more peripheral (PAF and CL) localization.

Rauschecker and Tian (2000) proposed a hypothesis that has provided a very influential framework for study of cortical spatial processing in human and non-human primates. The hypothesis draws an analogy with the visual system, in which ventrally and dorsally directed hierarchical “streams” are specialized for recognizing, respectively, “what” an object is or “where” it is (Ungerleider and Mishkin, 1982). In the auditory system, the “where” pathway was proposed to originate in the

caudal part of the superior temporal gyrus (encompassing the planum temporale in humans and auditory areas CL and caudomedial (CM) in monkeys) and to project to the dorsolateral prefrontal cortex, both directly and by way of the parietal cortex (Romanski et al., 1999; Rauschecker and Tian, 2000). Those parietal and prefrontal areas also participate in visual spatial processing. In contrast, the “what” pathway was proposed to originate from anterior lateral belt areas and to project toward the temporal pole. Human functional imaging studies support the “where” part of the hypothesis in that auditory spatial tasks tend to activate parietal and prefrontal areas that also are activated during visual spatial tasks (Rauschecker and Scott, 2009; Recanzone and Cohen, 2010). Functional imaging offers less consistent support for the hypothetic special role of the caudal auditory cortex as the origin of the spatial-processing stream. Several studies that used simulated moving sounds or sounds that varied in their spatial distribution failed to demonstrate specific activation of posterior auditory areas (Griffiths et al., 1994, 1998; Zatorre et al., 2002). Other studies have demonstrated specific activation of the planum temporale, but only when the stimuli incorporated perceived movement or complex spectral changes (Baumgart et al., 1999; Warren et al., 2002; Krumbholz et al., 2005; Ahveninen et al., 2006).

Physiologic studies of the auditory cortex in non-human primates do not consistently support the notion that caudal areas are specialized for spatial processing. The prediction would be that caudal belt areas would show sharper spatial tuning or a higher percentage of sharply tuned neurons than do R or A1 core areas or middle or R belt areas. In marmoset monkeys, overall distributions of best areas and modulation of responses by source location were “markedly similar” between area A1 and the CL and CM fields (Zhou and Wang, 2012). One study in macaques demonstrated that neurons in CL tended to have narrower spatial tuning for monkey calls than did neurons in anterolateral or middle lateral (Tian et al., 2001). Another study in macaques is often cited in support of the what/where hypothesis (Recanzone et al., 2000). That study, however, found no significant difference in the proportion of spatially sensitive neurons in cortical areas A1 and CM; no R belt areas were studied. Another macaque study (Woods et al., 2006) found that a middle belt area consistently had a higher proportion of spatially sensitive neurons than did CL and that at moderate-to-high sound levels area A1 had a higher proportion of spatially sensitive neurons than did either CL or CM; again, no R belt area was sampled. As noted above, computational analysis of those data demonstrated that neurons in area CL signaled far-contralateral locations accurately but that signaling of near-midline regions by CL neurons was

poor compared to that by neurons in area R (Miller and Recanzone, 2009). Inasmuch as psychophysical localization is most accurate for near-midline targets, it is difficult to take those physiologic results as evidence that CL is the principal auditory cortical area driving sound localization.

The cortical representation of auditory space appears to be strictly contralateral in carnivores and non-human primates. Restricted auditory cortical lesions or inactivation disrupt the trained ability to walk to the source of a sound contralateral to the lesion or inactivation (Thompson and Cortez, 1983; Jenkins and Merzenich, 1984; Kavanagh and Kelly, 1987; Malhotra et al., 2004). The contralaterality of auditory spatial representation in humans is less clear. Early clinical studies suggested that contralesional deficits were most common (Sanchez-Longo and Forster, 1958; Klingon and Bontecou, 1966). In conflict with a notion of contralateral dominance, however, functional imaging studies tend to show enhanced activity in right parietal areas during auditory spatial tasks, regardless of the laterality of the auditory stimulus (Griffiths et al., 1998; Bushara et al., 1999; Zatorre et al., 1999). A study of patients who underwent partial temporal-lobe resections for relief of epilepsy demonstrated a substantial right-sided dominance of auditory spatial representation (Zatorre and Penhune, 2001). In that study, most patients having right-sided resections demonstrated symmetric bilateral sound localization deficits, whereas most patients having left-sided resections showed no localization deficit. Substantial deficits were observed even in cases in which the resection was entirely rostral to Heschl’s gyrus, involving hypothetic “what” pathways but sparing the primary auditory cortex and the putative “where” area on the planum temporale.

BEYOND LOCALIZATION

Sound localization can aid a cat in catching a mouse and can help a person orient to a talker of interest in a crowded party. Beyond just localization, however, normal-hearing listeners can exploit spatial hearing mechanisms to detect and recognize sounds of interest in the presence of other competing sounds. Conversely, people with mild-to-moderate hearing loss report that among their greatest disabilities are the inability to segregate multiple talkers and to understand speech in noisy environments (Gatehouse and Nobel, 2004). Spatial hearing has long been recognized as contributing to so-called “auditory scene analysis” (Cherry, 1953; Bregman, 1990), and there has been growing interest in spatial release from masking, which is the improvement of sound detection and recognition when the signal is separated in space from maskers (Edmonds and

Culling, 2005; Ihlefeld and Shinn-Cunningham, 2008; Marrone et al., 2008). Spatial release from masking shares many properties with localization – the dominance of low-frequency ITD cues (Kidd et al., 2010; Middlebrooks and Onsan, 2012; Bremen and Middlebrooks, 2013) is one important example. Understanding of relevant spatial cues and low-level mechanisms gained from ongoing research will lead to enhanced processing schemes for hearing aids and cochlear implants that will benefit both localization and spatial release. There is evidence, however, that localization and spatial release from masking might employ discrete cortical structures. Specifically, cortical lesions can disrupt localization while sparing spatial release from masking or (less commonly) can disrupt spatial release from masking while sparing localization (Thiran and Clarke, 2003; Duffour-Nikolov et al., 2012). One is hopeful that enhanced non-invasive imaging and encephalographic techniques in humans and invasive cortical recording techniques in behaving animals will yield new understanding of these relevant cortical mechanisms.

REFERENCES

- Ahveninen J, Jääskeläinen IP, Raij T et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proc Natl Acad Sci U S A* 103: 14608–14613.
- Batteau DW (1967). The role of the pinna in human localization. *Proc R Soc Lond B Biol Sci* 168: 158–180.
- Baumgart F, Gaschler-Markefski B, Woldorff MG et al. (1999). A movement sensitive area in auditory cortex. *Nature* 400: 724–726.
- Bernstein LR, Trahiotis C (1994). Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise. *J Acoust Soc Am* 95: 3561–3567.
- Bernstein LR, Trahiotis C (2002). Enhancing sensitivity to interaural delays at high frequencies by using “transposed stimuli”. *J Acoust Soc Am* 112: 1026–1036.
- Blauert J (1982). Binaural localization: Multiple images and application in room- and electroacoustics. In: RW Gatehouse (Ed.), *Localization of Sound: Theory and Application*. Amphora, Groton, CT, pp. 65–84.
- Blauert J (1983). *Spatial Hearing*. MIT Press, Cambridge, Mass.
- Bloom PJ (1977). Creating source elevation illusions by spectral manipulation. *J Audio Eng Soc* 25: 560–565.
- Boehnke SE, Phillips DP (1999). Azimuthal tuning of human perceptual channels for sound location. *J Acoust Soc Am* 106: 1948–1955.
- Boudreau JC, Tsuchitani C (1968). Binaural interaction in the cat superior olive S segment. *J Neurophysiol* 31: 442–454.
- Brand A, Behrend O, Marquardt T et al. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417: 543–547.
- Bregman AS (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, MA.
- Bremen P, Middlebrooks JC (2013). Weighting of spatial and spectro-temporal cues for auditory scene analysis by human listeners. *PLoS One* 8: e59815.
- Bremen P, van Wanrooij MM, van Opstal AJ (2010). Pinna cues determine orienting response modes to synchronous sounds in elevation. *J Neurosci* 30: 194–204.
- Briley PM, Kitterick PT, Summerfield AQ (2012). Evidence for opponent process analysis of sound source location in humans. *J Assoc Res Otolaryngol* 14: 83–101.
- Brughera A, Dunai L, Hartmann WM (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *J Acoust Soc Am* 133: 2839–2855.
- Brungart DS (1999). Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am* 106: 3589–3602.
- Brungart DS, Rabinowitz WM (1999). Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am* 106: 1465–1479.
- Brungart DS, Durlach NI, Rabinowitz WM (1999). Auditory localization in nearby sources. II. Localization of a broad-band source. *J Acoust Soc Am* 106: 1956–1968.
- Bushara KO, Weeks RA, Ishii K et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat Neurosci* 2: 759–766.
- Butler RA (1987). An analysis of the monaural displacement of sound in space. *Percept Psychophys* 41: 1–7.
- Carlile S, Best V (2002). Discrimination of sound source velocity in human listeners. *J Acoust Soc Am* 111: 1026–1035.
- Carlile S, Leong P, Hyams S (1997). The nature and distribution of errors in sound localization by human listeners. *Hear Res* 114: 179–196.
- Carr CE, Konishi M (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* 10: 3227–3246.
- Cherry CE (1953). Some experiments on the recognition of speech, with one and two ears. *J Acoust Soc Am* 25: 975–979.
- Chung W, Carlile S, Leong P (2000). A performance adequate computational model for auditory localization. *J Acoust Soc Am* 107: 432–445.
- Coleman PD (1962). Failure to localize the source distance of an unfamiliar sound. *J Acoust Soc Am* 34: 345–346.
- Dingle R, Hall S, Phillips D (2012). The three-channel model of sound localization mechanisms: interaural level differences. *J Acoust Soc Am* 131: 4023–4029.
- Dingle RN, Hall SE, Phillips DP (2013). The three-channel model of sound localization mechanisms: Interaural time differences. *J Acoust Soc Am* 133: 417–424.
- Duffour-Nikolov C, Tardif E, Maeder P et al. (2012). Auditory spatial deficits following hemispheric lesions: Dissociation of explicit and implicit processing. *Neuropsychol Rehabil* 22: 674–696.
- Edmonds BA, Culling JF (2005). The spatial unmasking of speech: evidence for within-channel processing of interaural time delay. *J Acoust Soc Am* 117: 3069–3078.
- Furukawa S, Middlebrooks JC (2002). Cortical representation of auditory space: Information bearing features of spike patterns. *J Neurophysiol* 87: 1749–1762.
- Furukawa S, Xu L, Middlebrooks JC (2000). Coding of sound-source location by ensembles of cortical neurons. *J Neurosci* 20: 1216–1228.

- Gaese BH, Johnen A (2000). Coding for auditory space in the superior colliculus of the rat. *Eur J Neurosci* 12: 1739–1752.
- Gardner MB (1969). Distance estimation of 0 deg. or apparent 0 deg.-oriented speech signals in anechoic space. *J Acoust Soc Am* 45: 47–53.
- Gatehouse S, Nobel W (2004). The speech, spatial, and qualities of hearing scale (SSQ). *Int J Audiol* 43 (85): 99.
- Getzmann S, Lewald J (2012). Cortical processing of change in sound location: Smooth motion versus discontinuous displacement. *Brain Res* 1466: 119–127.
- Goldberg JM, Brown PB (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *J Neurophysiol* 32: 613–636.
- Gordon B (1973). Receptive fields in deep layers of cat superior colliculus. *J Neurophysiol* 36: 157–178.
- Grantham DW (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *J Acoust Soc Am* 79: 1939–1949.
- Grantham D (1992). Adaptation to auditory motion in the horizontal plane: Effect of prior exposure to motion on motion detectability. *Percept Psychophys* 52: 144–150.
- Griffiths TD, Bench CJ, Frackowiak RSJ (1994). Human cortical areas selectively activated by apparent sound movement. *Curr Biol* 4: 892–895.
- Griffiths TD, Rees A, Witton C et al. (1996). Evidence for a sound movement area in the human cerebral cortex. *Nature* 383: 425–427.
- Griffiths T, Rees G, Rees A et al. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nat Neurosci* 1: 74–78.
- Haft ER, Dye RH, Nuetzel JM et al. (1977). Difference thresholds for interaural intensity. *J Acoust Soc Am* 61: 829–834.
- Harrington IA, Stecker GC, Macpherson EA et al. (2008). Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hear Res* 240: 22–41.
- Hebrank J, Wright D (1974). Spectral cues used in the localization of sound sources on the median plane. *J Acoust Soc Am* 56: 1829–1834.
- Henning GB (1974). Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* 55: 84–90.
- Hofman PM, Van Opstal JA (1998). Spectro-temporal factors in two-dimensional human sound localization. *J Acoust Soc Am* 103: 2634–2648.
- Hofman PM, Van Riswick JGA, Van Opstal JA (1998). Relearning sound localization with new ears. *Nat Neurosci* 1: 417–421.
- Holt RE, Thurlow Wr (1969). Subject orientation and judgment of distance of a sound source. *J Acoust Soc Am* 46: 1584–1585.
- Hubel DH, Wiesel TN (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160: 106–154.
- Humanski RA, Butler RA (1988). The contribution of the near and far ear toward localization of sound in the sagittal plane. *J Acoust Soc Am* 83: 2300–2310.
- Ihlefeld A, Shinn-Cunningham B (2008). Spatial release from energetic and informational masking in a selective speech identification task. *J Acoust Soc Am* 123: 4369–4379.
- Imig TJ, Irons WA, Samson FR (1990). Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J Neurophysiol* 63: 1448–1466.
- Imig TJ, Bibikov NG, Poirier P et al. (2000). Directionality derived from pinna-cue spectral notches in cat dorsal cochlear nucleus. *J Neurophysiol* 83: 907–925.
- Jay MF, Sparks DL (1987). Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J Neurophysiol* 57: 35–55.
- Jeffress LA (1948). A place theory of sound localization. *J Comp Physiol Psychol* 41: 35–39.
- Jenkins WM, Merzenich MM (1984). Role of cat primary auditory cortex for sound-localization behavior. *J Neurophysiol* 52: 819–847.
- Joris PH, Yin TCT (2007). A matter of time: internal delays in binaural processing. *Trends Neurosci* 30: 70–78.
- Kavanagh GL, Kelly JB (1987). Contribution of auditory cortex to sound localization by the ferret (*Mustela putorius*). *J Neurophysiol* 57: 1746–1766.
- Kidd Jr G, Mason CR, Best V et al. (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *J Acoust Soc Am* 128: 1965–1978.
- King AJ, Hutchings ME (1987). Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: A map of auditory space. *J Neurophysiol* 57: 596–624.
- Klingon GH, Bontecou DC (1966). Localization in auditory space. *Neurol* 16: 879–886.
- Klumpp RG, Eady HR (1956). Some measurements of interaural time difference thresholds. *J Acoust Soc Am* 28: 859–860.
- Knudsen EI (1982). Auditory and visual maps of space in the optic tectum of the owl. *J Neurosci* 2: 1177–1194.
- Kopco N, Shinn-Cunningham B (2011). Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am* 130: 1530–1541.
- Koppl C, Carr CE (2008). Maps of interaural time difference in the chicken's brainstem nucleus laminaris. *Biol Cybern* 98: 541–559.
- Krumbholz K, Schönwiesner M, Rübsamen R et al. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *Eur J Neurosci* 21: 230–238.
- Kuhn GF (1977). Model for the interaural time differences in the azimuthal plane. *J Acoust Soc Am* 62: 157–167.
- Langendijk EHA, Bronkhorst AW (2002). Contribution of spectral cues to human sound localization. *J Acoust Soc Am* 112: 1583–1596.
- Lee J, Groh JM (2012). Auditory signals evolve from hybrid-to eye-centered coordinates in the primate superior colliculus. *J Neurophysiol* 108: 227–242.
- Lee C-C, Middlebrooks J (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nat Neurosci* 14: 108–114.

- Lee C-C, Middlebrooks J (2013). Specialization for sound localization in fields A1, DZ, and PAF of cat auditory cortex. *J Assoc Res Otolaryngol* 14: 61–82.
- Lewis JW, Meauchamp MS, DeYoe EA (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex* 10: 873–888.
- Litovsky RY, Colburn HS, Yost WA et al. (1999). The precedence effect. *J Acoust Soc Am* 106: 1633–1654.
- Lomber S, Malhotra S (2008). Double dissociation of ‘what’ and ‘where’ processing in auditory cortex. *Nat Neurosci* 11: 609–616.
- Macpherson EA, Middlebrooks JC (2002). Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am* 111: 2219–2236.
- Macpherson EA, Sabin AT (2007). Binaural weighting of monaural cues for sound localization. *J Acoust Soc Am* 121: 3677–3688.
- Magezi DA, Krumbholz K (2010). Evidence for opponent-channel coding of interaural time differences in human auditory cortex. *J Neurophysiol* 104: 1997–2007.
- Magezi D, Buetler K, Chouiter L et al. (2013). Electrical neuroimaging during auditory motion aftereffects reveals that auditory motion processing is motion sensitive but not direction selective. *J Neurophysiol* 109: 321–331.
- Makous JC, Middlebrooks JC (1990). Two-dimensional sound localization by human listeners. *J Acoust Soc Am* 87: 2188–2200.
- Malhotra S, Hall AJ, Lomber SG (2004). Cortical control of sound localization in the cat: Unilateral cooling deactivation of 19 cerebral areas. *J Neurophysiol* 92: 1625–1643.
- Marrone N, Mason CR, Kidd Jr G (2008). Tuning in the spatial dimension: Evidence from a masked speech identification task. *J Acoust Soc Am* 124: 1146–1158.
- May BJ (2000). Role of the dorsal cochlear nucleus in the sound localization behavior of cats. *Hear Res* 148: 74–87.
- McAlpine D, Jiang D, Palmer AR (2001). A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4: 396–401.
- McFadden D, Pasanen EG (1976). Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am* 59: 634–639.
- Mehrgardt S, Mellert V (1977). Transformation characteristics of the external human ear. *J Acoust Soc Am* 61: 1567–1576.
- Mickey BJ, Middlebrooks JC (2003). Representation of auditory space by cortical neurons in awake cats. *J Neurosci* 23: 8649–8663.
- Mickey BJ, Middlebrooks JC (2005). Sensitivity of auditory cortical neurons to the locations of leading and lagging sounds. *J Neurophysiol* 94: 979–989.
- Middlebrooks JC (1992). Narrow-band sound localization related to external ear acoustics. *J Acoust Soc Am* 92: 2607–2624.
- Middlebrooks JC (1999a). Individual differences in external-ear transfer functions reduced by scaling in frequency. *J Acoust Soc Am* 106: 1480–1492.
- Middlebrooks JC (1999b). Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *J Acoust Soc Am* 106: 1493–1510.
- Middlebrooks JC, Bremen P (2013). Spatial stream segregation by auditory cortical neurons. *J Neurosci* 33: 10986–11001.
- Middlebrooks JC, Green DM (1991). Sound localization by human listeners. *Annu Rev Psychol* 42: 135–159.
- Middlebrooks JC, Knudsen EI (1984). A neural code for auditory space in the cat’s superior colliculus. *J Neurosci* 4: 2621–2634.
- Middlebrooks JC, Onsan ZA (2012). Stream segregation with high spatial acuity. *J Acoust Soc Am* 132: 3896–3911.
- Middlebrooks JC, Pettigrew JD (1981). Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J Neurosci* 1: 107–120.
- 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.
- Middlebrooks JC, Clock AE, Xu L et al. (1994). A panoramic code for sound location by cortical neurons. *Science* 264: 842–844.
- Middlebrooks JC, Xu L, Eddins AC et al. (1998). Codes for sound-source location in nontopographic auditory cortex. *J Neurophysiol* 80: 863–881.
- Middlebrooks JC, Xu L, Furukawa S et al. (2002). Cortical neurons that localize sounds. *Neuroscientist* 8: 73–83.
- Mikami A, Newsome W, Wurtz RH (1986). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *J Neurophysiol* 55: 1308–1327.
- Miller LM, Recanzone GH (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proc Natl Acad Sci U S A* 106: 5931–5935.
- 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.
- Muir D, Clifton R, Clarkson M (1989). The development of a human auditory localization response: A U-shaped function. *Can J Psychol* 43: 199–216.
- Musican AD, Butler RA (1984). The influence of pinnae-based spectral cues on sound localization. *J Acoust Soc Am* 75: 1195–1200.
- Musican 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.
- Nuetzel JM, Hafter ER (1981). Discrimination of interaural delays in complex waveforms: Spectral effects. *J Acoust Soc Am* 69: 1112–1118.
- Overholt EM, Rubel EW, Hyson RL (1992). A circuit for coding interaural time differences in the chick brainstem. *J Neurosci* 12: 1698–1708.
- Palmer AR, King AJ (1982). The representation of auditory space in the mammalian superior colliculus. *Nature* 299: 248–249.
- Parvani F, Macaluso E, Warren J et al. (2002). A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Curr Biol* 12: 1584–1590.

- Pecka M, Brand A, Behrend O et al. (2008). Interaural time difference processing in the mammalian medial superior olive: The role of glycinergic inhibition. *J Neurosci* 28: 6914–6925.
- Phillips DP, Vigneault-MacLean BK, Boehnke SE et al. (2003). Acoustic hemifields in the spatial release from masking of speech by noise. *J Am Acad Audiol* 14: 518–524.
- Rajan R, Aitkin LM, Irvine DRF (1990). Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J Neurophysiol* 64: 888–902.
- Rauschecker JP, Scott SK (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12: 718–724.
- Rauschecker JP, Tian B (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci U S A* 97: 11800–11806.
- Razak KA (2011). Systematic representation of sound locations in the primary auditory cortex. *J Neurosci* 31: 13848–13859.
- Recanzone GH, Cohen YE (2010). Serial and parallel processing in the primate auditory cortex revisited. *Behav Brain Res* 206: 1–7.
- Recanzone GH, Guard DC, Phan ML et al. (2000). Correlation between the activity of single auditory cortical neurons and sound-localization behaviour in the macaque monkey. *J Neurophysiol* 83: 2723–2739.
- Roberts MT, Seeman SC, Golding NL (2013). A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry. *Neuron* 78: 923–935.
- Romanski LM, Tian B, Fritz J et al. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 2: 1131–1136.
- Ryugo DK, Seishiro S (1991). Synaptic connections of the auditory nerve in cats: relationship between endbulbs of held and spherical bushy cells. *J Comp Neur* 305: 35–48.
- Salminen N, May P, Alku P et al. (2009). A population rate code of auditory space in the human cortex. *PLoS One* 4: e7600.
- Sanchez-Longo LP, Forster FM (1958). Clinical significance of impairment of sound localization. *Neurol* 8: 119–125.
- Sandel TT, Teas DC, Feddersen WE et al. (1955). Localization of sound from single and paired sources. *J Acoust Soc Am* 27: 842–852.
- Sanes DH, Malone BJ, Semple MN (1998). Role of synaptic inhibition in processing of dynamic binaural level stimuli. *J Neurosci* 18: 794–803.
- Shaw EAG (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J Acoust Soc Am* 56: 1848–1861.
- Slattery III WH, Middlebrooks JC (1994). Monaural sound localization: Acute versus chronic impairment. *Hear Res* 75: 38–46.
- Smith KR, Okada K, Saberi K et al. (2004). Human cortical auditory motion areas are not motion selective. *Neuroreport* 15: 1523–1526.
- Smith KR, Saberi K, Hickok G (2007). An event-related fMRI study of auditory motion perception: no evidence for a specialized cortical system. *Brain Res* 1150: 94–99.
- Spitzer MW, Semple N (1991). Interaural phase coding in auditory midbrain: Influence of dynamic stimulus features. *Reports* 254: 721–724.
- Spitzer M, Semple M (1998). Transformation of binaural response properties in the ascending auditory pathway: Influence of time-varying interaural phase disparity. *J Neurophysiol* 80: 3062–3076.
- Stecker GC, Harrington IA, Middlebrooks JC (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS Biol* 3: 520–528.
- Stevens SS, Newman EB (1936). The localization of actual sources of sound. *Am J Psychol* 48: 297–306.
- Strutt JW (1907). On our perception of sound direction. *Philos Mag* 13: 214–232.
- 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.
- Thiran AB, Clarke S (2003). Preserved use of spatial cues for sound segregation in a case of spatial deafness. *Neuropsychologia* 41: 1254–1261.
- Thompson GC, Cortez AM (1983). The inability of squirrel monkeys to localize sound after unilateral ablation of auditory cortex. *Behav Brain Res* 8: 211–216.
- Tian B, Reser D, Durham A et al. (2001). Functional specialization in rhesus monkey auditory cortex. *Science* 292: 290–293.
- Trahiotis C, Bernstein LR (1986). Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth. *J Acoust Soc Am* 79: 1950–1957.
- Ungerleider LG, Mishkin M (1982). Two cortical visual systems. In: DJ Ingle, MA Goodale, RJWOe Mansfields (Eds.), *Analysis of Visual Behaviour*, MIT Press, Cambridge, pp. 549–586.
- van der Hiejden M, Lorteije JAM, Plaushka A et al. (2013). Directional hearing by linear summation of binaural inputs at the medial superior olive. *Neuron* 78: 936–948.
- Warren JD, Zielinski BA, Green GGR et al. (2002). Perception of sound-source motion by the human brain. *Neuron* 34: 139–148.
- Wenzel EM, Arruda M, Kistler DJ et al. (1993). Localization using nonindividualized head-related transfer functions. *J Acoust Soc Am* 94: 111–123.
- Werner-Reiss U, Groh J (2008). A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. *J Neurosci* 28: 3747–3758.
- 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.
- Woods TM, Lopez SE, Long JH et al. (2006). Effects of stimulus azimuth and intensity on the single-neuron activity in

- the auditory cortex of the alert macaque monkey. *J Neurophysiol* 96: 3323–3337.
- Woodworth RS (1938). *Experimental Psychology*. Holt, Rinehart, and Winston, New York.
- Yin TCT (1994). Physiological correlates of the precedence effect and summing localization in the interior colliculus of the cat. *J Neurosci* 14: 5170–5186.
- Yin TCT, Chan JCK (1990). Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* 64: 465–488.
- Yost WA, Loisel L, Dorman M et al. (2013). Sound source localization of filtered noises by listeners with normal hearing: A statistical analysis. *J Acoust Soc Am* 133: 2876–2882.
- Young ED, Spirou GA, Rice JJ et al. (1992). Neural organization and responses to complex stimuli in the dorsal cochlear nucleus. *Philos Trans R Soc Lond B Biol Sci* 336: 407–413.
- Zahorik P (2002). Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111: 1832–1846.
- Zahorik P, Brungart DS, Bronkhorst AW (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica* 91: 409–420.
- Zatorre RJ, Penhune VB (2001). Spatial localization after excision of human auditory cortex. *J Neurosci* 21: 6321–6328.
- Zatorre RJ, Mondor T, Evans AC (1999). Functional activation of right parietal and frontal cortex during auditory attention of space and frequency. *Neuroimage* 10: 544–554.
- Zatorre RJ, Bouffard PA, Belin P (2002). Where is ‘where’ in the human auditory cortex? *Nat Neurosci* 5: 905–909.
- Zhou Y, Wang X (2012). Level dependence of spatial processing in the primate auditory cortex. *J Neurophysiol* 108: 810–826.
- Zwislocki J, Feldman RS (1956). Just noticeable differences in dichotic phase. *J Acoust Soc Am* 28: 860–864.