13 Binaural and Spatial Hearing

In this chapter, we explore several aspects of binaural hearing, that is, hearing with both ears instead of just one. We shall see that binaural hearing offers a number of advantages over monaural hearing, which have obvious implications for daily living. In particular, we will examine how we use hearing to determine the locations of sounds in space, which relies upon time and level differences between the ears and spectral variations caused by the pinna, head, and torso, all of which are affected by the direction of the sound source.

BINAURAL SUMMATION

Although Sivian and White (1933) did not find significant differences between the minimal audible field (MAF) for the better ear and binaural MAF, subsequent studies demonstrate that the intensity needed to reach threshold is lower when listening with two ears than with one. The essential finding is that, if one first corrects for any difference in monaural threshold between the two ears so that they are equal in terms of sensation level, then the binaural threshold will be approximately 3 dB better (lower) than the monaural thresholds (Causse and Chavasse, 1941, 1942a; Keys, 1947; Shaw et al., 1947). For example, to correct for the difference between monaural thresholds of 11 dB in the right ear and 16 dB in the left, the binaural stimulus would be presented 5 dB higher in the left ear. The resulting binaural threshold would be about 3 dB below these equated monaural thresholds. Hirsh (1948a) refers to this threshold advantage that occurs when listening with two ears as **binaural summation at threshold**. Similar binaural advantages have been demonstrated when the stimulus is white noise (Pollack, 1948) or speech (Keys, 1947; Shaw et al., 1947).

Loudness is also enhanced by binaural hearing. Based upon loudness level measurements, Fletcher and Munson (1933) concluded that a stimulus presented at a given SPL will sound twice as loud binaurally as monaurally. **Binaural summation of loudness** (Hirsh, 1948a) was shown as a function of sensation level by Causse and Chavasse (1942b), who performed loudness balances between binaurally and monaurally presented tones. At sensation levels close to threshold, they found that a binaural tone had to be about 3 dB lower in intensity than a monaural tone in order to produce the same sensation of loudness. This binaural advantage increased gradually with sensation level so that equal loudness was produced by a binaural tone 6 dB softer than the monaural stimulus at about 35 dB sensation level. This difference remained essentially constant at approximately 6 dB for higher sensation levels.

Perfect binaural summation means that a sound is twice as loud binaurally as it is monaurally. That loudness summation actually occurs at the two ears was questioned by Reynolds and Stevens (1960), who found that rate of binaural loudness growth had a slope of 0.6 compared to 0.54 for monaural loudness growth, and less than perfect binaural summation was found by Scharf and Fishken (1970). However, most findings suggest that a sound is twice as loud binaurally as it is monaurally (Fletcher and Munson, 1933; Hellman and Zwislocki, 1963; Marks, 1978, 1987). For example, Marks (1978) reported on the binaural summation of loudness for tones using magnitude estimation (and also loudness matches for corroboration). His findings are summarized in Fig. 13.1 The circles and squares show the loudness estimates for the left and right ears, respectively. The dotted lines show what the binaural estimates should be if summation is perfect. Notice that the actual binaural loudness estimates (shown by the triangles) fall almost exactly along the predicted functions. This indicates essentially perfect binaural summation at each frequency. Marks (1987) subsequently demonstrated complete binaural summation of loudness at 1000 Hz, as revealed by a 2:1 ratio of the loudness of a binaural tone to the monaural one. Recall from Chapter 11, in this context, that the calculated loudness of a binaural sound is taken to be twice that of a monaural sound (ANSI S3.4–2007).

DIFFERENTIAL SENSITIVITY

Various studies suggest that differential sensitivity for both intensity (Churcher et al., 1934; Harris, 1963; Rowland and Tobias, 1967; Jesteadt et al., 1977a, 1977b) and frequency (Shower and Biddulph, 1931; Pickler and Harris, 1955; Jesteadt et al., 1977a, 1977b) is better binaurally than when listening with only one ear. A problem, however, has been that the small differences detected between monaural and binaural difference limens (DLs) may have been the result of loudness summation. Pickler and Harris (1955) highlighted this problem. They found that the frequency DL was better binaurally than monaurally at low sensation levels. Recall from Chapter 9 that the effect of intensity upon differential sensitivity is greatest at low sensation levels, and that binaural hearing enhances sensitivity (or loudness) by roughly 3 to 6 dB. Thus, the smaller binaural DL may be due to summation rather than to some binaural mechanism for discrimination. To test this idea, Pickler and Harris adjusted the binaural signal level to account for the loudness advantage, and also tested DLs at a high level where differential sensitivity should not be affected by intensity. In both cases, the difference between monaural and binaural DLs disappeared. It was thus unclear whether the binaural DL is smaller than it is monaurally, or whether the difference just reflects a level difference.

This issue was essentially resolved in a study by Jesteadt et al. (1977a). They obtained intensity and frequency DLs at 70 dB SPL for 2500-, 10,000-, and 4000-Hz tones by using a two-interval forced-choice method. Their results are shown in Fig. 13.2 Note that binaural differential sensitivity is uniformly

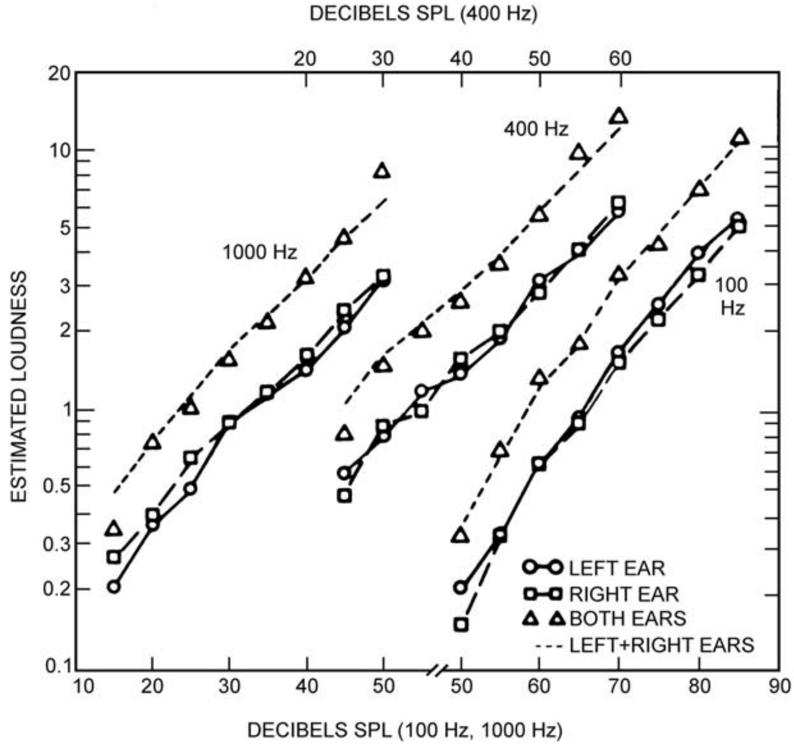

Figure 13.1 Loudness magnitude estimates for each ear and binaurally at 100, 400, and 1000 Hz. The dotted lines are predicted values for perfect summation. (See text.) *Source*: From Marks (1978), with permission of *J. Acoust. Soc. Am.*

better (the DL is smaller) than monaural, and that the difference is largely the same regardless of frequency. The ratio of the monaural to the binaural DL is on the order of 1.65 for intensity and 1.44 for frequency. The binaural–monaural differences obtained by Jesteadt et al. are not attributable to a loudness advantage for binaural hearing, because a difference of about 30 dB would have been needed to produce the observed binaural DL advantages (Shower and Biddulph, 1931; Jesteadt et al., 1977a, 1977b); and binaural summation is equivalent to only about 3 to 6 dB. Stellmack, Viemeister, and Byrne (2004) found that average intensity DLs were about 2 dB better binaurally (interaural intensity differences) than monaurally, which was significant for broadband noise but not for 4000-Hz tones.

BINAURAL FUSION AND BEATS

Even though the sounds of daily life reach the two ears somewhat differently in terms of time, intensity, and spectrum, we still perceive a single image. As Cherry (1961) pointed out, we perceive one world with two ears. More precisely, the similar but nonidentical signals reaching the two ears are fused into a single, coherent image (gestalt). This process is called **binaural fusion**.

Binaural fusion experiments require earphone listening because this allows us to precisely control the stimuli presented to the two ears, as well as how these signals are related. Generally, the experimenter is looking for a combination of stimuli that results in a *fused image lateralized to the center (midline) of the head*. The essential finding is that, although completely dissimilar signals are not fused, the auditory system does achieve binaural fusion as long as the signals presented to the two ears are similar in some way (Cherry, 1953; Cherry and Sayers, 1956; Sayers and Cherry, 1957; Broadbent, 1955; Leakey et al., 1958). The low frequencies, below roughly 1500 Hz, appear to be the most important. Thus, if each ear is presented with a 300-Hz tone at the same time, the subject will perceive a fused image in the center of his head.

A second example will demonstrate an important property of binaural fusion. If two different high-frequency tones are presented one to each ear, they will be heard as two separate signals. However, if a single low-frequency tone is superimposed upon both high frequencies so that they are caused to modulate at the frequency of the low tone, the listener will report a fused image (Leakey et al., 1958). This result shows that the auditory system uses the low-frequency envelopes of the complex signals (their macrostructures) for fusion even though the details of the signals (their microstructures) are different. Fusion of speech can be shown to occur, for example, when only the high-frequency components of the speech waveform are directed to one ear and only the lows are presented to the other (Broadbent, 1955). Even

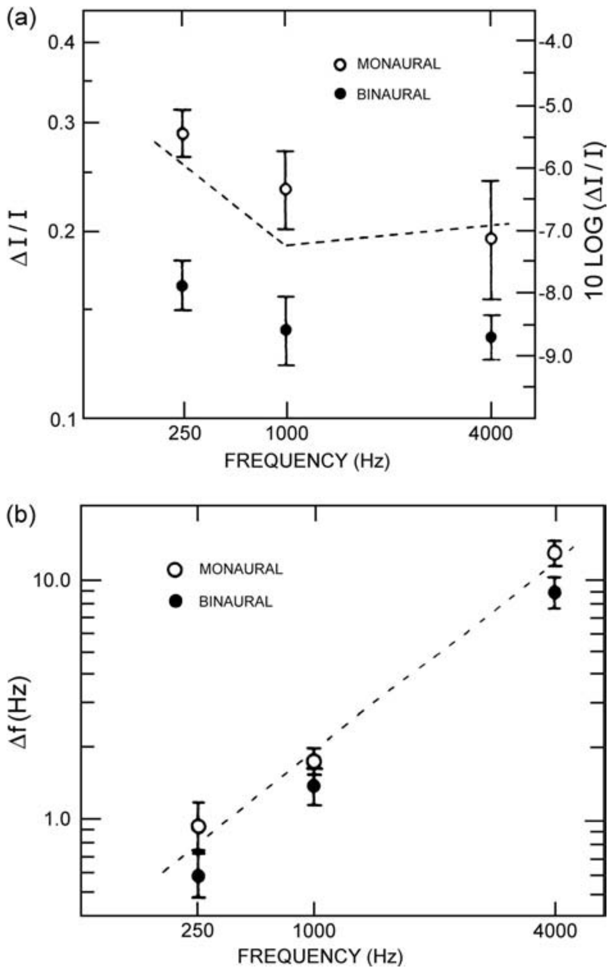

Figure 13.2 (a) Mean values of binaural and monaural ΔI//I. Dotted line shows predicted monaural values from Jesteadt et al. (1977a). (b) Mean binaural and monaural values of Δf. Dotted line shows predicted monaural DLs from Wier et al. (1977). *Source*: From Jesteadt et al. (1977b), with permission of *J. Acoust. Soc. Am.*

though neither ear alone receives enough of the speech signal for identification, the resulting fused image is readily identified.

The binaural fusion mechanism has been described in terms of a model by Cherry and Sayers (Cherry and Sayers, 1956; Sayers and Cherry, 1957) in which the central auditory nervous system carries out a running cross-correlation between the inputs to the two ears. In other words, the signals entering the ears are viewed as statistical events, and the fusion mechanism operates by looking for commonalities between the inputs coming from the two ears on an ongoing basis.

A very interesting phenomenon occurs when one tone is presented to the right ear and a second tone of slightly different frequency is presented to the left. The result is the perception of beats (see Chap. 12) in the fused image. Recall that beats occur when one combines two tones slightly different in frequency because phase differences between the tones result in alternating increases and decreases in amplitude. The intriguing aspect of **binaural beats** is that they occur even though the two signals are acoustically completely isolated from one another. Obviously, binaural beats must result from some interaction between the neural codings of the signals from the two ears taking place within the central nervous system. [Cells have been identified in the superior olive that are responsive to the envelope of binaural beats (e.g., Wernick and Starr, 1966). They are probably at least partially involved in subserving the perception of binaural beats.]

Binaural beats differ from monaural beats in several ways (Licklider et al., 1950; Tobias, 1963; Groen, 1964). Whereas monaural beats can be heard for interacting tones across the audible frequency range, binaural beats are associated with the lower frequencies, and the best responses are for tones between about 300 and 600 Hz. Binaural beats can still be heard even if the frequency difference between the ears is relatively wide, although the perception of the image changes with frequency separation (see below). In addition, binaural beats can be perceived even if there is a substantial difference in sound level between the ears. (Recall from Chap. 5 that phase locking to stimulus cycle occurs at the very lowest levels at which an auditory neuron responds.) There have also been reports that binaural beats can be detected if one of the tones is presented at a level below the behavioral threshold for that ear (Lehnhardt, 1961; Groen, 1964); however, subsequent experiments have failed to confirm these findings (Tobias, 1963; Gu et al., 1995).

Licklider et al. (1950) reported that perceptual differences occur as the frequency separation widens between the ears. When identical frequencies are presented to two ears, the listener hears a fused image. When the frequencies are 2–10 Hz apart, the subject reports loudness fluctuations, which give way to a perception of “roughness” when the frequency difference reaches about 20 Hz. As the frequency separation becomes wider and wider, the fused image appears first to split into two smooth tones, and these tones then migrate in perceived location to the respective ears.

AUDITORY SCENE ANALYSIS

Before proceeding to the question of directional hearing, let us consider the different but related issue of how we differentiate among the many sounds that surround us at any particular time. One very familiar experience of this kind is the aptly named **cocktail party effect**, which refers to our ability to follow what one person is saying when one or more other people are talking at the same time (Cherry, 1953). Of course, this is certainly not the only situation in which we separate one sound from the other. For example, we regularly separate the words of a song from the accompanying music, and are able to differentiate among the many sound sources in a busy work or home environment. This phenomenon has been described as **auditory scene analysis** (Bregman, 1990) or **sound source determination** (Yost, 1993, 1997; Yost and Sheft, 1993).

A variety of acoustical parameters have been considered as contributors to this process, including spectral profile, spectral separation, harmonicity, spatial separation, temporal onsets and offsets, temporal modulation, and temporal separation (Yost, 1993, 1997; Yost and Sheft, 1993). For example, the spatial separation of a signal (target) and a masker has been shown to reduce both energetic and informational masking (see Chap. 10; see, e.g., Arbogast et al., 2002; Hawley, Litovsky, and Culling, 2004; Gallum, Mason, and Kidd, 2005; Kidd, Mason, Brughera, and Hartmann, 2005; Wu, Wang, Chen, et al., 2005). This **spatial release from masking** (or **spatial unmasking**) occurs because (1) separating the target and the masker provides an acoustical advantage by increasing the target-to-masker ratio at one of the ears, and (2) the advantages provided by binaural processing, which are discussed in this chapter.

Auditory scene analysis may also be addressed in the context of gestalt psychology principles for the grouping of objects in the visual field [1] (Bregman, 1990; Bregman and Ahad, 1996). We will briefly consider a few basic aspects of this approach as a foundation for further study.

Fundamentally, auditory scene analysis involves grouping the sounds impinging on the listener's ears into *perceptual units* called **streams** based on certain criteria or *grouping principles*. For example, the grouping factors of **proximity** and **similarity** pertain to how close or far apart sounds are in terms of their physical parameters. In other words, sounds tend to be grouped together when they are close and/or similar with respect to parameters such as frequency, spectral shape, timing, harmonicity (harmonics of a common fundamental frequency), intensity, and direction or spatial origin. On the other hand, sounds that are far apart or dissimilar in terms of these parameters tend not to be grouped together, but are perceived as separate streams. This is illustrated in Fig. 13.3 which represents the perception of alternating higher- and lower-frequency tones. In the upper frame, the two tones are relatively close in frequency, and they are perceptually grouped into a *single stream* of alternating pitches (*ABABABAB*) that is heard to be coming from the same sound source. However, when the two tones are far apart in frequency, they are heard as *two separate streams* of interrupted tones (*A*. . . *A*. . . *A*. . . *A* and *B*. . . *B*. . . *B*. . . *B*) coming from different sound sources, as in the lower frame. The former case illustrates **stream fusion** (or **stream integration**), and the latter case is **stream segregation**.

Other gestalt grouping principles are also involved in auditory streaming including, among others, common fate and good continuation. **Common fate** is the tendency for stimuli that change together over time to be grouped perceptually, implying a common origin. This would apply to sounds that have similar onsets, offsets, and variations in frequency, harmonicity, or level

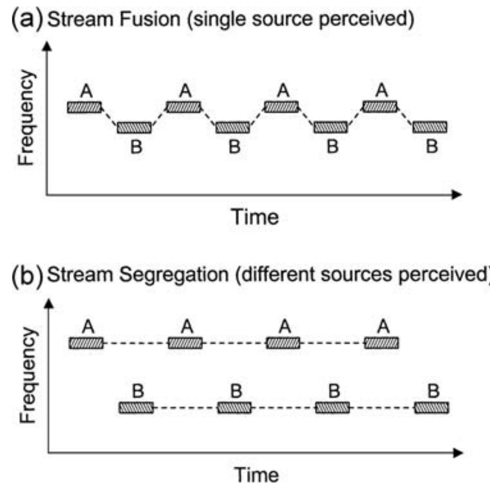

Figure 13.3 Both panels show a sequence of alternating higher-(*A*) and lower-(*B*) frequency tones. (a) When the two frequencies are close enough, they are heard as a single stream of alternating pitches (*ABABABAB*), which is coming from one sound source. (b) When the two frequencies are sufficiently different, they are heard as two separate streams of interrupted tones (*A*. . . *A*. . . *A*. . . *A* and *B*. . . *B*. . . *B*. . . *B*), each coming from a different sound source.

over time. **Good continuation** applies to smooth changes in the physical parameters of a sound so that abrupt changes or discontinuities imply a change in the source.

Auditory streaming involves both primitive and schema-based processes (Bregman, 1990). **Primitive processes** are innate, automatic, and unconscious, and may be viewed as operating on the acoustical aspects of the signal in a bottom-up fashion. Thus, the streaming of alternating high- and low-frequency tones illustrated in Fig. 13.3 is an example of a primitive process. On the other hand, **schema-based processes** operate in a top-down fashion, involving learned information and cognitive effort to "hear out" a signal (e.g., listening for a familiar tune being played within the din of sounds in a university cafeteria). Increasing the speed of a schema-based streaming task results in poorer performance, but this is not the case for tasks that involve primitive processes. Primitive processes are symmetrical; for example, there is no difference between listening for just the higher tones or just the lower tones when presented with a sequence containing both, as in Fig. 13.3 However, schema-based processes are *asymmetrical*. Thus, it is easier to extract familiar names than the other sounds with which they are intermixed in the cacophony of a busy cafeteria.

DIRECTIONAL HEARING

Localization

How do we determine the direction of a sound source? Intuitively, we expect that some sort of comparison between the two ears must be involved. We are usually concerned with binaural

[1] See Palmer (1999) for an extensive review of visual perception.

listening in a sound field (stereophony), but we sometimes use earphones for personal entertainment or to precisely control experimental conditions, as described above. Interestingly, stereophonic and headphone listening can result in different perceptions of space. Sounds presented from loudspeakers are perceived to be coming from *outside* the head (*externalized*) from a source that can be **localized** in the environment. On the other hand, sounds presented from earphones are generally perceived to be *inside* the head (*internalized*), coming from an apparent source that is **lateralized** along a plane between the two ears. This difference between *extracranial* **localization** and *intracranial* **lateralization** is easily experienced by comparing the way music from the same compact disc album sounds through loudspeakers versus earphones. In general, identical sounds impinging at the same time upon the two ears are localized directly in front of (or behind) the listener or, through earphones, from an apparent source lateralized in the center of the head. However, one should be aware that externalization can be a matter of degree, as opposed to being an all-or-none experience (see Blauert, 1997; Hartmann and Wittenberg, 1996). For a review of this topic, see Durlach et al. (1992).

Horizontal directions are expressed as angles of **azimuth** around the head, illustrated in Fig. 13.4a Sounds coming from straight ahead have an azimuth 0° and those coming from directly behind have an azimuth of 180°. Other azimuths are usually given as the number of degrees right (+) or left (−) of center. For example, a loudspeaker that is off center toward the right by 45° has an azimuth of 45° right or +45°, and a loudspeaker located 45° off center toward the left has an azimuth of 45° left or −45°. Azimuths are sometimes expressed in terms of the total number of degrees going around the head toward the right, in which case 45° right would be 45°, and 45° left would be 315°. Vertical directions are expressed as angles of **elevation** (usually along the medial plane from front to back), as illustrated in Fig. 13.4b. In this case, 0° elevation means straight ahead, an elevation of 90° is directly above the head, and 180° is directly behind the head.

The traditional **duplex theory** explains localization on the basis of *time* differences between the ears at lower frequencies and *level* differences between the ears at higher frequencies (Lord Rayleigh, 1907).[2] Consider the arrangement in Fig. 13.5a. The signal from the speaker, which is off to the right, must follow a longer path to the far (left) ear than to the near (right) ear. As Fig. 13.5b. shows, low frequencies have wavelengths that are longer than the path around the head so that they "bend around" the head to the far ear (diffraction). Thus, **interaural time differences (ITDs)** are expected to provide localization cues for the lower frequencies, where the wavelength of the tone is larger than the distance the signal must travel from the near (right) ear to the far (left) ear. In contrast, higher frequencies have wavelengths smaller than the head so that they are

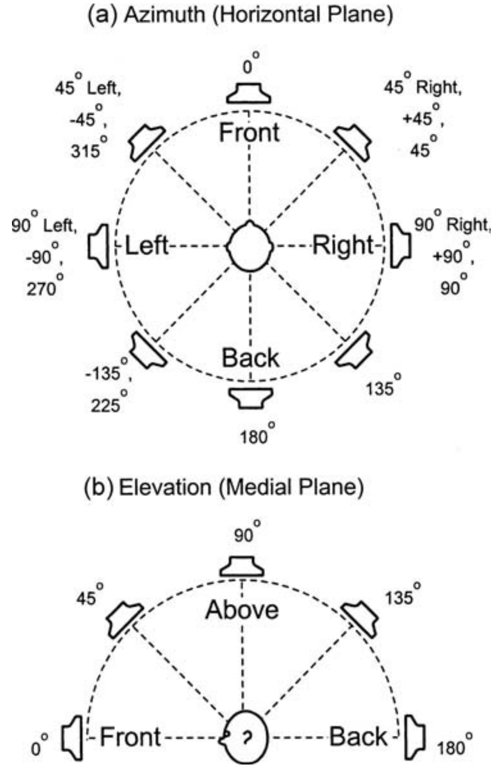

Figure 13.4 (a) Angles of azimuth horizontally around the head. Various ways of expressing azimuth angles are illustrated. (b) Angles of elevation vertically around the head in the medial plane.

"blocked" in the path to the far (left) ear (Fig. 13.5c). This **head shadow** causes a reduction in the intensity of the signal at the far ear, producing sound level differences between the ears. Thus, **interaural level differences (ILDs)** or **interaural intensity differences (IIDs)** are expected to provide localization cues for the higher frequencies. Our thresholds for interear differences are as small as approximately 10 μs for ITDs (Klumpp and Eady, 1956) and about 1 dB for ILDs (Mills, 1960; Blauert, 1997).

The traditional approach to interaural differences involves modeling the head as a solid sphere around which the ear-to-ear distance approximates 22 to 23 cm (Woodworth, 1938). This results in a time delay of roughly 660 μs for the sound to get from the near ear to the far ear, which in turn corresponds to a frequency of 1500 Hz. Thus, the greatest time delay occurs when a sound source is directly to one side or the other (90° azimuth), for which the ITD would be 660 μs, denoted by the peak of the curve in Fig. 13.6 Below 1500 Hz, the wavelength is greater than the distance around the head, and the phase difference at the two ears provides an unambiguous localization cue. However, the phase discrepancy becomes ambiguous (except for the first wavelength) as the frequency increases to 1500 Hz, where its wavelength approximates the distance around the head, resulting in localization errors. At higher frequencies, the wavelength

[2] Lord Rayleigh was John William Strutt (1842–1919).

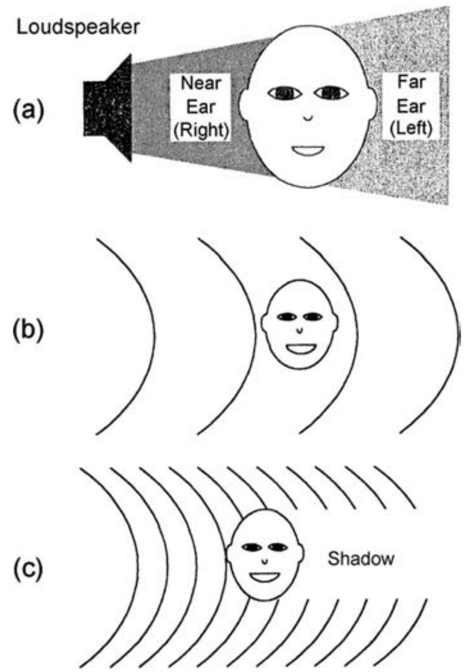

Figure 13.5 (a) Relationship between a loudspeaker and the two ears. (b) Low frequencies bend around the head due to their large wavelengths. (c) High frequencies have wavelengths smaller than head diameter so that an acoustic shadow results at the far ear.

is shorter than the size of the head so that the resulting head shadow produces ILDs. The ILDs thus provide localization cues for the higher frequencies.

Feddersen et al. (1957) measured the interaural time and level differences for human heads as functions of angle around the head (azimuth) and frequency. Figure 13.6 shows that their ITD measurements were in good agreement with the Woodworth model. Notice that there was no difference between the ears when the signals (clicks) come from directly in front or behind (0° and 180°), because the ears are equidistant from the sound source in both cases. Interaural time differences developed as the loudspeaker moved around the head, bringing it closer to one ear than the other. The ITD increased to a maximum of about 660 μs when the loudspeaker was directly in front of one ear (90° azimuth), where the distance (and therefore the time delay) between the ears is greatest. Feddersen et al. also found that ILDs depend on both frequency and azimuth. As expected, ILDs were negligible at 200 Hz and increased with frequency to as much as about 20 dB at 6000 Hz. The ILDs were 0 dB directly in front and behind (at 0° and 180° azimuth), where the sound source is equidistant between ears and increased to as much as about 20 dB (depending on frequency), as the loudspeaker moved closer to side or the other, reaching a maximum where the loudspeaker directly in front of one ear (at 90°).

Woodworth's sphere model actually falls short of providing an accurate description of the interaural cues associated with the human head. This is not surprising because the configuration

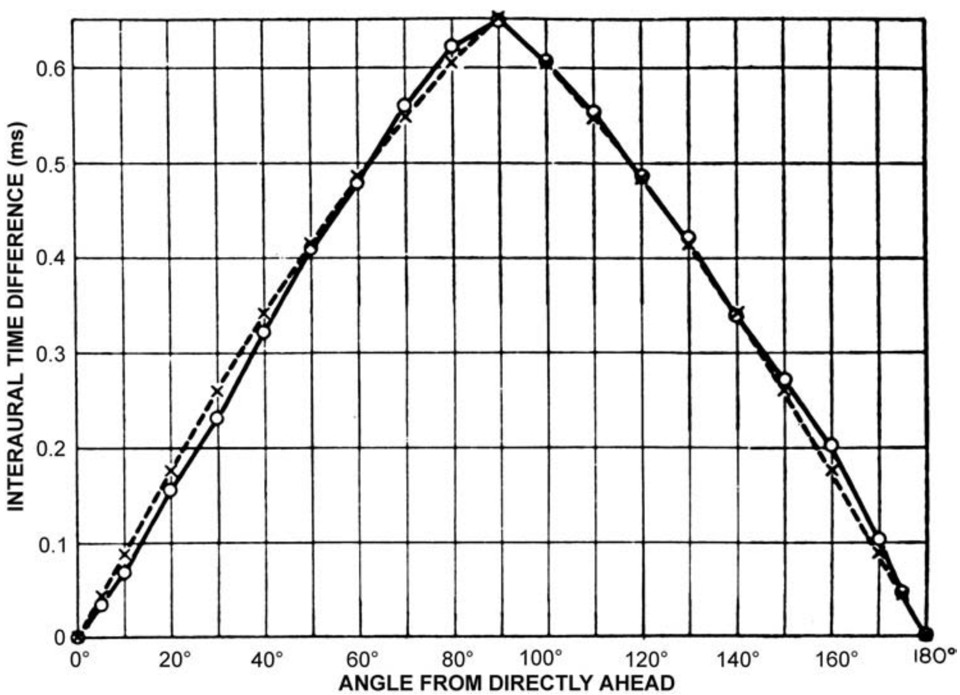

Figure 13.6 Interaural time differences for different loudspeaker azimuths. Crosses are calculated differences based on Woodworth's (1938) solid sphere model. Circles are measurements by Feddersen et al. (1957). *Source*: From Feddersen et al. (1957), permission of *J. Acoust. Soc. Am.*

of the head differs considerably from that of a sphere. Kuhn (1977) found that the ITDs for the higher frequencies came close to the 660 μs value predicted by the sphere model, but they do not change with frequency below 500 Hz or above 3000 Hz. He also found that the ITDs below about 500 Hz were about 800 to 820 μs instead of 660 μs. Middlebrooks (1999a) found that maximum interaural time delays among adults ranged from 657 to 792 μs. Taken together with other findings (e.g., Nordlund, 1962; Abbagnaro et al., 1975; Roth et al., 1980; Bronkhorst and Plomp, 1988), the implication is that the sphere model does not explain the nature of low-frequency ITDs, but applies principally to the higher frequencies and to the leading edges of clicks and click-like signals, for which the major directional cues are not the phase-derived ITDs. Even directional hearing for high-frequency complex sounds, which is principally associated with ILDs, is also affected by interaural time differences. Moreover, the duplex model alone falls short of completely explaining directional hearing because it does not adequately explain front-back distinctions, directionality above and below the horizontal plane, and monaural localization.

In addition to the effects of the ear canal resonance (Chap. 3), the spectrum of the sound arriving at the eardrum is affected by the pinna, head, and torso. In particular, **spectral cues** (also known as **pinna cues** and **monaural spectral cues**) at high frequencies introduced by the pinnae are important for the perception of elevation, front-back distinctions, and monaural localization (Blauert, 1969/70, 1997; Hebrank and Wright, 1974; Weinrich, 1982; Musicant and Butler, 1984; Middlebrooks and Green, 1991; Middlebrooks, 1992, 1997; Shaw, 1997; Wightman and Kistler, 1997). These pinna cues also contribute to the extracranialization of sound sources (Plenge, 1974; Blauert, 1997). Low-frequency cues (below about 3000 Hz) associated with head diffraction and torso reflections appear to be involved in vertical localization (e.g., Gardner, 1973; Algazi et al., 2001).

The manner in which the amplitude spectrum of a sound is modified going from a source in the environment to the eardrum is shown by the **head-related transfer function (HRTF)**. Families of HRTFs are shown for sources located *horizontally* at many azimuths around the head in Fig. 13.7 (Shaw, 1974),[3] and *vertically* at many elevations around the medial plane of the head in Fig. 13.8

The curves in these figures make it clear that the spectrum of the sound that finally arrives at the eardrum depends on the direction from which it came (see, e.g., Shaw, 1997). This is most easily seen for horizontal directions in Fig. 13.7 by comparing the shapes of HRTFs for sounds coming from the front (top panel), side (middle panel), and back (bottom panel) of the head. Also notice that the sound level reaching the eardrum gets weaker as the source moves from the same side of the head around to the opposite side of the head. Consider a sound source located 45° to the right. In this case, the sound reaching the right (*near*) eardrum is affected by the HRTF labeled 45°, and the sound arriving at the left (*far*) eardrum is modified by the HRTF labeled −45°. (The 45° and −45° HRTFs are found in the upper panel in Fig. 13.7, or refer back to Fig. 13.2 for an uncluttered view.) It is easy to realize how these differences translate into ILDs. The same 0° curve applies to both ears if the source is directly ahead, and the same 180° curve applies to both ears if the sound is coming from directly behind (180°), in which case the ILDs are 0 dB.

Changes in elevation cause changes in the high-frequency aspects of the HRTF due to directionally dependent filtering by the pinna (Hebrank and Wright, 1974; Shaw, 1997). This is seen in Fig. 13.8 as a "pinna notch" that shifts between about 5000 and 11,000 Hz as the sound source moves up around the head. Scanning upward from the bottom of the figure, notice that there is a notch at about 6000 Hz when the sound source is *below* the head. The notch gets *higher* in frequency as the sound source moves *upward* toward the *front* and then continues up toward the top of the head (*above*), where the notch is essentially absent. Continuing around, the notch then gets *lower* in frequency as the source moves *downward* toward the *back* of the head and then *below* again.

Classical studies of localization were conducted by Stevens and Newman (1936) and Sandel, Teas, Feddersen, and Jeffress (1955). Stevens and Newman (1936) sat their subjects in a chair elevated about 12 f above the roof of the Harvard Biological Laboratories building to minimize the possible confounding effects of echoes and reverberation. The sound source was a loudspeaker mounted on a boom arm that extended 12 f from the listener. The loudspeaker was rotated around the subject in 15° steps from 0° to 180°, and the task was to listen for the signal and report its apparent direction. It is important to point out that their subjects regularly made *front-back confusions*. For example, sounds presented from 30° right of center *in front* of the subject (30° off center from 0° azimuth) were confused with sounds presented from 30° right of center *behind* (30° off center from 180° azimuth). These front-back reversals were treated as equivalent, correct responses, and the location of the sound source was judged relative to 0° or 180°, whichever was closer. With this in mind, Stevens and Newman's findings are shown as a function of frequency in Fig. 13.9 Localizations were most accurate below 1000 Hz and above 4000 Hz, with the greatest errors between about 2000 and 4000 Hz.

Sandel et al. (1955) asked subjects to localize sounds in an anechoic (echo-free) room. Loudspeakers placed at 0° and at 40° right and left were used to generate "phantom" sound sources at various azimuths, depending upon the phases of the signals. The subject indicated the perceived location of the tone source with an "acoustic pointer," which was a speaker that rotated on a boom around his or her head. A noise from this speaker alternated with the test tones, and the subject's task was to place the noise loudspeaker (pointer) at the apparent location

[3] This data may be found in tabular form in Shaw and Vaillancourt (1985).

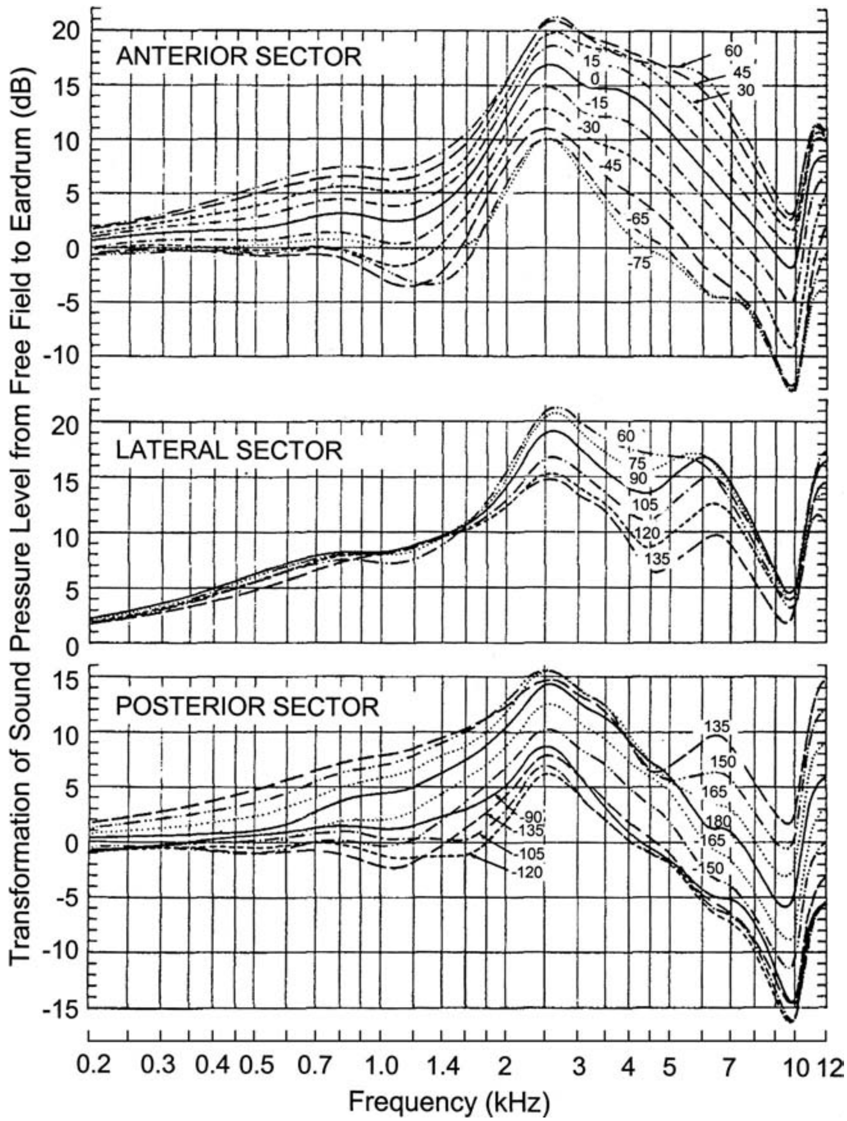

Figure 13.7 Horizontal head-related transfer functions for sound sources located at many angles of azimuth (θ) around the head based on data from 12 studies. *Source*: From Shaw (1974) with permission of *J. Acoust. Soc. Am.*

of the test tone source. They found that ITDs accounted for the localization of tones below about 1500 Hz, and that these were used to localize the high frequencies. Many random errors at 1500 Hz suggested that interaural cues are ambiguous around this frequency. These results were essentially consistent with those of Stevens and Newman.

The sizable localization errors and front/back confusions observed in these early studies imply that the directional cues provided by tonal signals are limited and ambiguous. It is interesting to note in this context that Stevens and Newman found better localization results for noises than for tones, which they attributed to quality (spectral) differences and ILDs for the high-frequency energy in these noises. In contrast, accuracy is substantially improved and front/back errors are reduced when localizing *broad-band* signals (e.g., Butler and Planert, 1976; Oldfield and Parker, 1984; Butler, 1986; Makous and Middlebrooks, 1990). Broad-brand stimuli provide the listener with multiple cues across frequencies, including both interaural differences and spectral shape information (Wightman and Kistler, 1993, 1997). Recall, here, that the spectral shape cues due to pinna effects are found in the higher frequencies. It is interesting to note, for example, that Musicant and Butler (1984) found that front/back distinctions were best when their stimuli included high frequencies (above 4000 Hz), and that

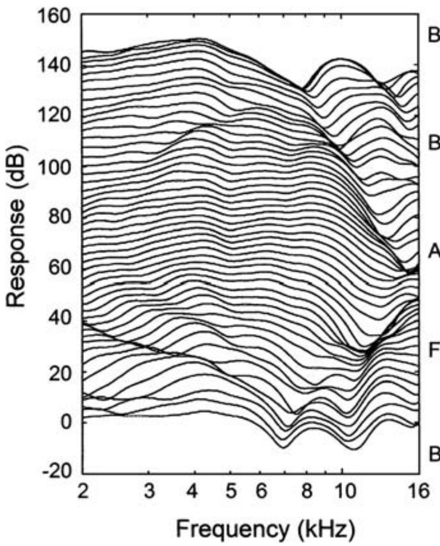

Figure 13.8 Vertical head-related transfer functions for sound sources located at many angles of elevation along the medial plane around the head of a representative subject. *Source*: Figure provided courtesy of Dr. Richard O. Duda, used with permission.

performance dropped considerably when the various depressions of the pinnae were occluded.

We learn more about the contributions of the higher frequencies from a study by Middlebrooks (1992). In this study, subjects were asked to localize high-frequency narrow-band noises presented from loudspeaker locations corresponding to 360 combinations of azimuth and elevation. The subject's task was to point the subject's nose at the perceived sound source location, and this response was in turn monitored by using special instrumentation to determine head orientation. The results demonstrated that *azimuth* was accurately localized and related to ILDs, while *front/back* and *elevation* localizations were related to spectral cues.

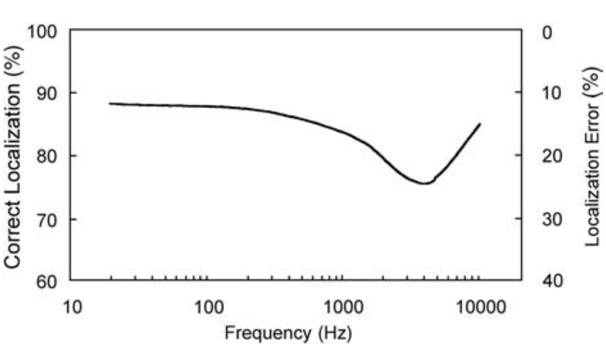

Figure 13.9 Accuracy of localization in percent (left axis) and localization error in percent (right axis) as a function of frequency, based on data reported by Stevens and Newman (1936).

Head Movements

There has been some controversy about whether the dynamic cues provided by head movements can improve localization accuracy (see, e.g., Middlebrooks and Green, 1991; Wightman and Kistler, 1993). Classical papers by Wallach (1939, 1940) presented findings and convincing arguments supporting the relevance of head movements as a means of reducing localization ambiguities, but some studies found that any benefits provided by head movements were either small or not significant (Pollack and Rose, 1967; Thurlow and Runge, 1967; Fisher and Freedman, 1968). However, questions about the impact of head movements on localization appear to have been resolved by contemporary experiments using real and virtual sound methods (e.g., Bronkhorst, 1995; Perrett and Noble, 1997a, 1997b; Wightman and Kistler, 1999; Macpherson and Middlebrooks, 2002). Although there is considerable intersubject variability, localization is improved by head movements, especially in terms of reducing front/back and vertical confusions.

Lateralization

Lateralization experiments have helped clarify and expand upon what we know about directional cues, because the use of earphones allows us to precisely control and manipulate the signals presented to the ears. Many studies have examined the effects of ITD and ILD cues upon lateralization (e.g., Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Mills, 1960; Yost et al., 1971; Yost, 1974; Grantham, 1984; Yost and Dye, 1988). While exact procedures vary, the general approach is to present two stimuli to the subject that differ with respect to interaural time (phase) or level, and to determine whether this interaural disparity results in a perceptible change in lateralization. The overall findings essentially agree with the localization data. That is, ITDs are most important up to about 1500 Hz, and ILDs take over as the primary lateralization cue for higher frequencies.

Yost (1974) performed a particularly interesting lateralization experiment addressing the discrimination of interaural time (actually phase) differences. He presented subjects with two stimuli. The first included a particular interaural time (actually phase) difference, θ. This difference, of course, resulted in a lateralization toward one side of the head analogous to the azimuth position. The second stimulus was the same except that the phase difference between the ears was larger by a slight amount $\Delta\theta$. Thus, it was $\theta + \Delta\theta$. The subjects had to detect the value of $\Delta\theta$ by discriminating between the two stimuli (θ vs. $\theta + \Delta\theta$). For any value of θ, the smaller the value of $\Delta\theta$ needed for a change in apparent lateralization, the better the discrimination of interaural phase. We might think of $\Delta\theta$ as analogous to the smallest perceptible change in azimuth (similar to the minimal audible angle discussed later in this chapter). The results are summarized in Fig. 13.10 Note that $\Delta\theta$ is smallest (best) when θ is 0° or 360°. These values of θ are midline lateralizations because 0° and 360° correspond to a zero phase disparity between the ears. Thus, the most acute interaural phase discriminations are made at the midline. On the other hand,

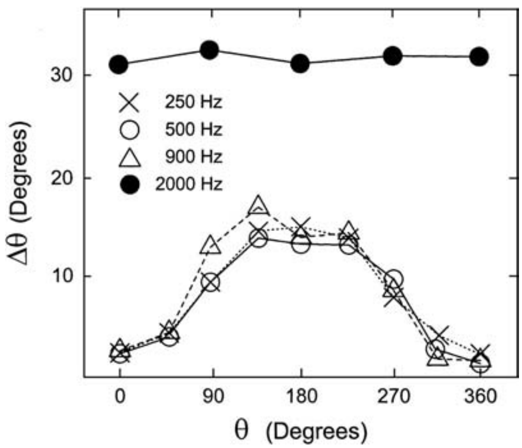

Figure 13.10 Changes in interaural phase (Δθ) required in order to detect a difference in lateralization from a standard (reference) phase difference (θ). *Source*: Adapted from Yost (1974), with permission of *J. Acoust. Soc. Am.*

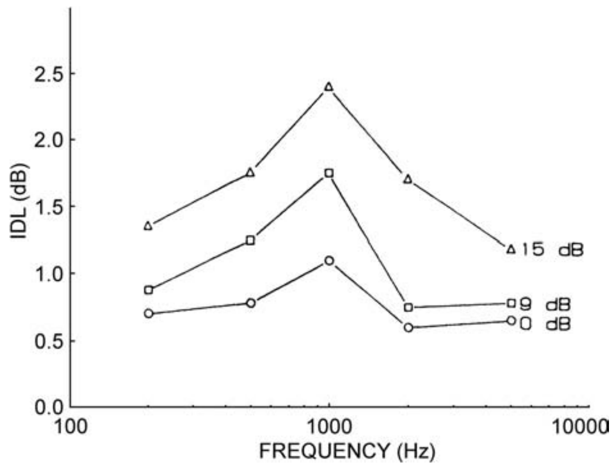

Figure 13.11 Thresholds for interaural differences in level (IDL) in decibels as a function of frequency from 200 to 5000 Hz. (Note that the scale of the ordinate is from 0 to 2.5 dB.) These findings were obtained when the standard stimulus itself was lateralized at midline (circles, marked 0 dB), halfway between midline and the left ear (squares, marked 15 dB), and at the left ear (triangles, marked 15 dB). The 0-, 9-, and 15-dB values are the interaural level differences needed to place the standard image at these locations. *Source*: From Yost and Dye (1988), with permission of *J. Acoust. Soc. Am.*

interaural phase discrimination was poorest (Δθ as largest) when θ was 180°; that is, when the signals were lateralized directly to one side.

Figure 13.10 also shows that Δθ was essentially the same for the frequencies up to 900 Hz. In contrast, interaural phase discrimination was substantially poorer at 2000 Hz, where it was constant at about 30°. Interaural phase had no effect at all for 4000 Hz (not shown on the graph). Thus, interaural phase was shown to be an important cue at low frequencies but unimportant for highs, in a manner consistent with the localization data.

Lateralization experiments have revealed that discrimination for ILDs is constant as a function of frequency within a range of about 2 dB, but with a curious increase in the size of the ILD at 1000 Hz (Mills, 1960; Grantham, 1984; Yost and Dye, 1988). These findings are illustrated in Fig. 13.11 The excellent interaural level discrimination at the high frequencies is, of course, expected from the previous discussion. Three sets of findings are shown in Fig. 13.11 depending upon whether the standard stimulus against which the discrimination was made was itself perceived to be (lateralized) at the midline, halfway between midline and the left ear, or at the left ear. These three lateralizations of the standards signal were achieved by using ILDs of 0, 9, and 15 dB, respectively (Yost, 1981). A comparison of these three curves makes it clear that ILD discrimination is most acute at midline, is least sensitive when the standard is lateralized off to one side, and is at some intermediate value when the discrimination is made between the midline and at the left ear.

Interaural time differences come into play whenever the signal contains low frequencies. For example, Yost et al. (1971) found that the lateralization of clicks is impaired by removing their low-frequency components, but not by eliminating the highs; others found that ITDs could result in lateralization differences for high-frequency noise bursts and clicks (Klumpp and Eady, 1956; Hafter and DeMaio, 1975).

Lateralization has been used to study the relative salience of ILDs and ITDs by establishing the trade-off between the time and intensity cues. This was done by asking the subject to adjust the ILD (or ITD) until a midline image was perceived. Harris (1960) found trading ratios of 25 μs/dB for clicks with energy below 1500 Hz and 60 μs/dB for clicks with energy above 1500 Hz. These trading ratios imply that ITDs have greater salience for low frequencies (because a smaller ITD is needed to center the image) and ILDs are more salient for the highs (because a larger ITD is needed). However, these studies were marked by inconsistent results between subjects. For example, Moushegian and Jeffress (1959) found that the trading ratio for a 500-Hz tone was about 20 to 25 μs/dB for two subjects but only about 2.5 μs/dB for another subject.

Subjects in time-intensity trading studies often heard two lateralized images instead of one (Whitworth and Jeffress, 1961; Hafter and Jeffress, 1968; Jeffress and McFadden, 1971; Hafter and Carrier, 1972). One image was called the *time image* because it depended on ITDs (especially <1500 Hz) but was essentially unaffected by ILDs. The other was an *intensity image* and was responsive to both ILDs and ITDs at all frequencies. Typical trading ratios (for clicks) were on the order of 2 to 35 μs/dB for the time image and 85 to 150 μs/dB for the intensity image (Hafter and Jeffress, 1968). It may be that the intersubject differences reported by Harris (1960) and by Moushegian and Jeffress (1959) were due to responses to the time image by some subjects and to the intensity image by others.

Although high-frequency lateralization is typically associated with ILDs, it can be influenced by ITDs, such as when the high-frequency signal is amplitude-modulated at a low rate (e.g., Henning, 1974; Neutzel and Hafter, 1976, 1981). Let us borrow some examples from Henning (1974) to illustrate this. Consider three kinds of signals (each lasting 250 ms) presented binaurally to subjects with interaural time differences: (1) a low frequency (300 Hz), (2) a high frequency (3600 Hz), and (3) a high frequency (3900 Hz) that was sinusoidally amplitude-modulated (SAM) at 300 Hz. Listeners could lateralize the 300-Hz tone based on ITDs but could not do this for the 3600-Hz tone, but the 3900-Hz SAM signal also could be lateralized on the basis of ITDs (as well, in fact, as for the 300-Hz tone).

An interesting effect called **binaural interference** [4] occurs when a listener's experience of binaural phenomena for high frequencies is affected by the simultaneous presence of low frequencies (e.g., McFadden and Pasanen, 1976**;** Trahiotis and Bernstein, 1990; Buell and Hafter, 1991; Stellmack and Dye, 1993; Bernstein and Trahiotis, 1995; Heller and Trahiotis, 1995; Hill and Darwin, 1996; Best, Gallun, Carlile, and Shinn-Cunningham, 2007). This disruption is usually reported as poorer sensitivity for ITDs or lateralization changes at a high frequencies (signal) when the low frequency (interferer) are also present compared to when the high frequency signal is presented alone. Binaural interference appears to occur when the signal and interferer are grouped into a single perceptual unit, such as when they turn on and off simultaneously; however, it is reduced or eliminated when cues are provided that allow them to be segregated, such as when the signal and interferer frequencies turn on and off at different times.

Virtual Auditory Space Localization

The ability to take advantage of earphone testing to study directional hearing has been dramatically enhanced with the use of **virtual auditory space (VAS)** techniques. This approach uses an individual's own HRTFs for both ears (based on many sound source directions) to produce test signals that simulate naturally occurring free-field cues when they are presented to the subject through earphones (Wightman and Kistler, 1989a). These virtual stimuli have been found to accurately represent acoustical cues for localization (Wightman and Kistler, 1989a), and result in spatial position judgments similar to those found with real free-field signals, although front/back and vertical errors are more common (e.g., Wightman and Kistler, 1989b; Bronkhorst, 1995).

Unlike the *intracranially lateralized* images produced by the earlier earphone methods, VAS techniques appear to produce perceptions that are actually *localized extracranially* (e.g., Macpherson and Middlebrooks, 2002). Subjects in these virtual localization studies perceived a single apparent source as opposed to the split images often reported in the earlier lateralization studies (Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002). Wightman and Kistler (1992) compared the relative strengths of ITDs, ILDs, and spectral cues in establishing the perceived location of a sound. This was done by presenting broad-band noise signals containing conflicting localization cues. For example, the ITD would indicate that the sound source was at one location but the ILD and spectral cues would indicate that it was at a different location. In spite of the conflicting cues, their subjects perceived a single sound source at a location determined by the ITD. However, the dominance of the ITD cue was lost when the low frequencies were removed from the noise (by filtering), in which case the perceived location was determined by the ILD and spectral cues. More recently, Macpherson and Middlebrooks (2002) found that ITDs had greater salience than ILDs when localizing low-pass sounds, and that ILDs had greater salience than ITDs for high-pass sounds.

Considerable individual differences exist in the dimensions of the skull, pinna, etc., which are related to differences in HRTFs (Middlebrooks, 1999a). Testing with VAS techniques has made it possible to determine whether these individual differences affect directional hearing. To do this involves comparing an individual's localization performance for virtual sources based on his or her *own* ears versus virtual sources based on *another person's* ears (Wenzel et al., 1993; Møller et al., 1996; Middlebrooks, 1999b). Overall, it has been found that localization becomes less accurate when "listening through somebody else's ears" compared to "listening through your own ears," principally involving front/back confusions and elevation errors. It is interesting to note that this kind of manipulation of VAS sound sources has been found to produces changes in the spatial responses of the primary auditory cortex in the ferret (Mrsic-Flogel et al., 2001).

Minimum Audible Angle

Another aspect of directional hearing involves determining the smallest difference in location between two sound sources that results in a different perceived location. Since the two sound sources are viewed relative to the head, this is the same as asking what is the smallest angle (or difference in azimuth) that a listener can discriminate. Mills (1958, 1963, 1972) studied this phenomenon in depth and called it the **minimal audible angle (MAA)**. Specifically, he tested the MAA as a function of frequency when the sound sources were located in front of the subject (0°), and when they were 30°, 45°, 60°, and 75° off to the side. The logistics of the basic task are illustrated in Fig. 13.12 where we see that the listener must distinguish the difference between two points in space. Notice that the figure actually shows two different conditions, one in which the MAA is being determined when both sound sources are directly in front of the listener (at 0° azimuth) and another one in which the two sound sources are off to one side (at 45° azimuth).

[4] This term should not be confused with *clinical* binaural interference, in which some hearing-impaired patients have the atypical experience of poorer performance binaurally than monaurally.

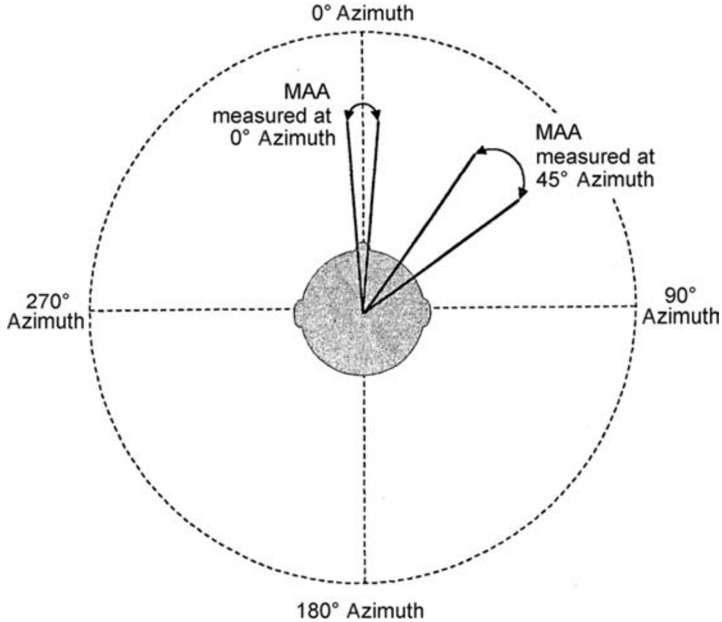

Figure 13.12 The minimal audible angle (MAA) is the smallest angular difference that can be perceived between two sound sources. This illustration shows the arrangement when the MAA is being measured under two different conditions. In one case, the MAA is being measured for loudspeakers directly in front of the listener (at 0° azimuth); in the other case, the loudspeakers are located at 45° azimuth.

In his classical studies, Mills found that the MAA was smallest (best) for the frequencies below about 1500 Hz and above approximately 2000 Hz, and was largest (poorest) between these frequencies. This result reflects the ambiguity of localization cues in the vicinity of 1500 Hz, thus confirming the previously mentioned findings. Mills also found that MAAs were most acute (approximately 1–2°) when the sound sources were directly in front of the head, and that they increased dramatically to very high values when the sources were at the side of the head. This result occurs because small changes in location in front of the head result in large interaural differences (especially ITDs). However, when the sources are off to one side of the head (facing one ear), the interaural differences remain largely the same in spite of relatively large changes in angle between the loudspeakers. We might thus conceive of a **cone of confusion** (Fig. 13.13) to one side of the head, within which the interaural differences do not vary when the sound sources change location (Mills, 1972). This image demonstrates the importance of head movements in localization, since these movements keep changing the position of the cone of confusion—the zone of ambiguity—thereby minimizing its detrimental effect.

The MAA described by Mills involved the discrimination of two stimuli presented sequentially. Perrott (1984) expanded upon the concept of the MAA using stimuli that were presented at the same time, describing the **concurrent minimum audible angle (CMAA)**. As for the MAA, the CMAA is also most acute for sounds presented directly in front of the subject and least sensitive when the stimuli are presented off to one side. However, the CMAA is also affected by spectral differences between the two sounds whose locations are to be discriminated. Subjects were unable to distinguish a difference in the angular locations of the stimuli when their frequencies differed by only 15 Hz. When the signals differed in frequency by 43 Hz, the size of the CMAA increased from 4.5° when the two signals were presented from 0° azimuth to an angle of about 45° when they were

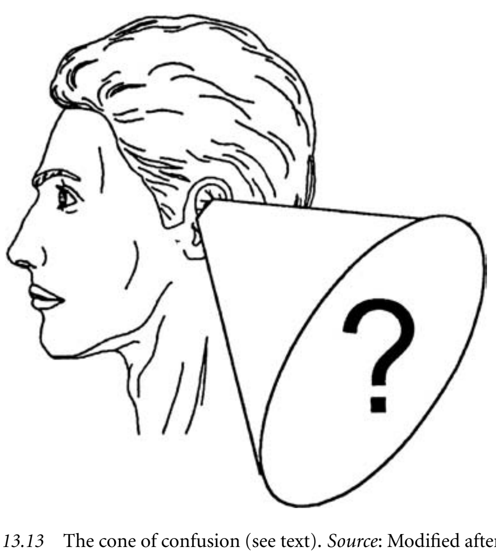

Figure 13.13 The cone of confusion (see text). *Source*: Modified after Mills (1972).

presented from 67° off to the left. On the other hand, signals 101 Hz apart had CMAAs of about 10° when the two sources were anywhere from straight ahead (0° azimuth) up to 55° off center. The CMAA then increased dramatically to about 30° when these signals were presented from 67° to the left. (An intermediate relationship between CMAA and azimuth was obtained for signals differing in frequency by 72 Hz.) As Perrott has pointed out, the CMAA involves the issue of sound source identification ("what") as well as that of localization ("where").

Minimum Audible Movement Angle

The MAA and CMAA describe the smallest perceptible difference in location between two *stationary* sound sources, A and B. We may also ask a similar question with regard to the *motion* of a sound source. How far (in degrees) must B travel away from A in order for us to know that B is moving? This smallest change in location that is needed for us to detect motion is called the **minimum audible movement angle (MAMA)** and has been the topic of considerable study (Harris and Sergeant, 1971; Perrott and Musicant, 1977; Grantham, 1986; Perrott and Tucker, 1988; Perrott et al., 1989; Saberi and Perrott, 1990a; Chandler and Grantham, 1992). Taken together, the evidence suggests that minimum audible movement angles depend on the same parameters that apply for stationary sources, plus the effect of how fast the moving sound source is moving. Using broad-band noise sources presented from directly in front of the head (0° azimuth), Chandler and Grantham (1992) found that the mean MAMA was 5.2° when the moving loudspeaker had a velocity of 10° per second (°/s), and 5.7° at 20°/s. However, the mean MAMA increased to 8.2° at 45°/s, 12° at 90°/s, and 17.3° at 180°/s. For noises presented from speakers at 60° azimuth, mean MAMAs increased from 8.3° at 10°/s to 28.4° at 180°/s. Notice that binaural sensitivity for motion becomes less sensitive for faster velocities. Saberi and Perrott (1990a) found that the minimum audible movement angle was most acute for velocities between 1.8°/s and 11°/s, and that it increased for slower velocities as well as for faster ones. Within this range, they were able to measure mean MAMAs as small as about 2° (approaching the size of the stationary MAA).

Distance

How do we judge distance with hearing? The answer to this question is not clear, although there has been renewed interest in this area. This section will outline several of the factors that have been considered as potential cues for perceiving the distance from a sound source. The interested student will find informative reviews in Coleman (1963), Blauert (1997), and Zahorik (1996). We will begin with sound level and the ratio of direct-to-reverberant energy. It appears that both can be salient distance cues, and that there is flexibility in their relative prominence in making distance judgments, depending on the nature of the sound (Zahorik, 2002a).

Sound level in a free field drops by 6 dB for each doubling of distance from the source (Chap. 1). Thus, sound level provides the listener with a distance cue. However, more than the expected decrease of 6 dB is required to perceive a doubling of distance so that apparent distance underestimates actual distance (Bekesy, 1938; Cochran, Throop, and Simpson, 1968; Gardner, 1969; Blauert, 1997; Petersen, 1990; Begault, 1991). For example, Blauert (1997) demonstrated that a doubling of perceived distance requires sound level reduction of about 20 dB instead of the expected value of 6 dB. Perceptual underestimates of actual distances appear to be ubiquitous findings in distance perception studies (Zahorik, 2002a).

Sounds reaching a listener in a real room involve both the *direct sound* from the source and *reverberation*, composed of multiple *reflections* from the walls, ceiling, floors, and various objects within the room. Under these conditions, the direct energy decreases with distance from the source (due to the inverse square law), but the reverberant energy remains pretty much uniform. As a result, the *ratio of direct-to-reverberant energy* changes with distance, enabling it to provide a distance perception cue for the listener (Blauert, 1997). The salience of the direct-to-reverberant energy ratio is fairly well established, and it is commonly found that distance performance is better in reverberant settings than in anechoic (echo-free) situations (Mershon and King, 1975; Mershon and Bowers 1979; Mershon et al., 1989; Wagenaars, 1990; Nielsen, 1993; Bronkhorst and Houtgast, 1999; Zahorik, 2002a, 2002b). Moreover, Zahorik (2002a) found that the direct-to-reverberant ratio had greater perceptual weight as a distance cue than sound level for noise signals. However, it appears to provide a coarsely grained distance cue because the threshold for discriminating direct-to-reverberant ratios is about 5 to 6 dB, roughly corresponding to a 2.5-fold change in distance (Zahorik, 2002b).

Spectral shape is another potential acoustical cue for distance perception when dealing with relatively long distances (Coleman, 1968; Butler et al., 1980; Little, Mershon, Cox, 1992; Blauert, 1997). The spectrum of a sound changes with distance from the sound source due to absorption as sound travels through the air, which causes the high frequencies to be attenuated with distance a great deal more than the lows.

Brungart (1999) showed that *binaural cues*, specifically *ILDs* for relatively *low* frequencies (<3000 Hz), contribute to distance perception when the sound source is less than 1 m from the head. This might seem odd because ILDs are generally associated with the high frequencies. However, low-frequency ILDs become significant when the source is close to the head and increase with proximity (Brungart and Rabinowitz, 1999; Duda and Martens, 1998).

There is at least some evidence which suggests that *familiarity* or *experience* with the sound source has some influence on distance perception. For example, experience with speech may provide the listener with a frame of reference that would be helpful in making distance judgments based on sound level, and it has been found that sound level is a salient cue for judging distance for speech (Gardner, 1969; Brungart and Scott, 2001). In addition, the accuracy of distance judgments is better for speech played forward compared to backward (McGregor et al., 1985). Also suggesting that familiarity improves auditory distance perception, Coleman (1962) found that the accuracy of distance judgments for unfamiliar sounds (1-s noise bursts) improved as experience with the stimulus accumulated over the course of repeated test trials. (The accuracy of azimuth localizations did not change with repeated trials.) Interestingly, Zahorik (2002a) found that sound level has a greater perceptual weight than the direct-to-reverberant ratio when making distance judgments for speech signals (a familiar sound), whereas the opposite is true for noise bursts (an unfamiliar sound). On the other hand, Nielsen (1991) did not find differences in judged distances for speech, noise, and two kinds of musical signals.

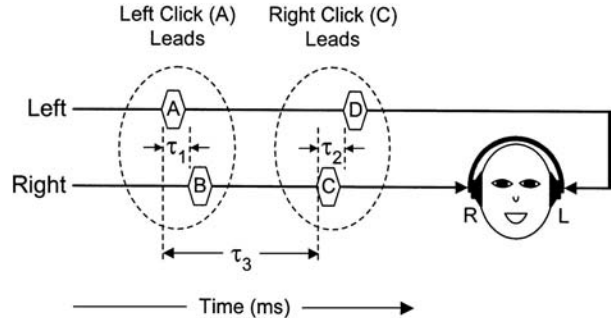

Figure 13.14 Arrangement of clicks presented through earphones to demonstrate the precedence effect, as used by Wallach, Newman, and Rosenzweig (1949).

PRECEDENCE EFFECT

Consider two apparently unrelated situations. The first involves listening to a radio news broadcast through both speakers of a home sound system. (We are not using a stereo music CD because we want identical signals from both speakers.) Sitting equidistant from the speakers causes us to perceive a phantom sound source between them. However, sitting close to one speaker (so that the signal reaches our ears sooner from that direction than from the other speaker) gives us the impression that all of the sound is coming from the closer speaker. This occurs even though the other speaker is still on.

The second situation involves listening to someone talking in a hard-walled room. In this case, the sounds reaching our ears include the direct sound from the talker's lips plus reflections of these sounds from the walls. Because the reflected sounds take an indirect route (via the walls), they reach our ears later than the direct sound and also from different directions. Yet, we hear a single sound coming from the direction of the earlier-arriving direct sound (although the reflections will "color" the quality of what we hear).

These situations illustrate a phenomenon known as the **precedence effect**, **Haas effect**, or the **first wavefront principle** (Gardner, 1968; Blauert, 1997). In general terms, when a sound coming from one direction is very quickly followed by a second sound (the echo) from another direction, then the perceived sound will be dominated by the earlier-arriving signal. In other words, we could say that **echo suppression** has occurred.

The precedence effect may be described in terms of how listeners perceive a sequence of four clicks presented through earphones, as in the classic experiment by Wallach, Newman, and Rosenzweig (1949). Figure 13.14 shows that the first click (A) went to the left ear followed by click B to the right ear after a very short delay ($\tau 1$). If presented alone, this pair of clicks was heard as a *fused image coming from the left*. Click C went to the right ear followed after another very short delay ($\tau 2$) by click D to the left ear, and by itself this pair was heard as a *fused image from the right*. The composite four-click sequence was heard as a *fused image from the left*; that is, its perception was dominated by the left-leading onset rather than the right leading at its offset. Thus, the first-arriving signal determined the perceived location of the fused sound. Wallach et al. found that this precedence effect occurred for intervals between the two click pairs ($\tau 3$) up to 40 ms. However, longer durations of $\tau 3$ caused the listener to hear two separate signals, one at each ear.

Haas' (1949, 1951) classic demonstration of the precedence effect involved presenting speech from two loudspeakers, with a delay in the onset of the signal from one speaker compared to the other. These delays are analogous to interval $\tau 3$ in the Wallach et al. study. A *fused image* coming from the *leading* loudspeaker was heard for delays up to 35 ms. Longer delays caused listeners to detect the presence of the second (delayed) sound, although the signal was still localized toward the leading side. Delays longer than about 50 ms caused listeners to hear one sound from the leading speaker and a distinct echo coming from the delayed speaker. It is interesting to note that the precedence effect can is robust with regard to some modifications of the temporal, spectral, and interear characteristics of the stimuli (Dixon and Colburn, 2006).

Various studies have expanded on the classical descriptions of the precedence effect, revealing that the general phenomenon encompasses several identifiable aspects (for reviews, see Blauert, 1997; Litovsky et al., 1999). **Fusion** is the perception of the leading and trailing signals as a single, unified image, and occurs for delays up to about 5 to 10 ms for clicks (e.g., Ebata et al., 1968; Freyman et al., 1991; Yang and Grantham, 1997; Litovsky and Shinn-Cunningham, 2001) and roughly 40 to 50 ms for speech (e.g., Haas, 1951; Lochner and Burger, 1958). Longer delays cause two separate images to be heard. The delay at which the perception splits into two images is called the **echo threshold**.

The perceived location of the fused image is affected by the size of the delay between the two signals. **Summing localization** occurs for delays shorter than 1 ms, in which case the perceived location of the fused image is affected by both the leading and lagging clicks (Blauert, 1997), as illustrated in Fig. 13.15a. **Localization dominance** occurs when the location of the fused image

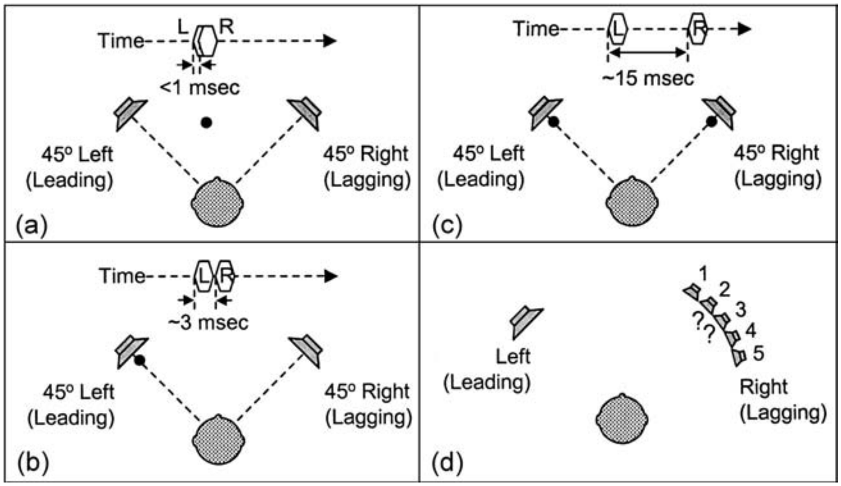

Figure 13.15 Aspects of the precedence effect shown in terms of signals presented from two loudspeakers, located 45° right and left of the listener. The left signal leads and the right signal(s) lags. In frames a to c, the filled circles indicate the locations of the perceived images, and the delay between the onsets of the two signals is shown along the top of each frame. In frame d, the task is to discriminate between the signals from the array of speakers on the (lagging) right side.

is determined by the leading signal (Fig. 13.15b). This occurs when the delay between the first and second clicks is between about 1 and 5 ms. Localization dominance breaks down as the delay lengthens, with the image splitting in two at the echo threshold, beyond which two images are heard, each coming from its own side (Fig. 13.15c).

Another aspect of the precedence effect is illustrated in Fig. 13.15d. Here, the listener is asked to discriminate between the azimuth locations of clicks on the lagging side (i.e., between speakers 1, 2, 3, etc.). Discrimination is poorer when the leading signal is present compared to what it is without the leading signal (e.g., Perrott et al., 1989; Freyman et al., 1991; Litovsky and Macmillan, 1994). An analogous effect occurs with earphones, where the presence of the leading signal affects the difference limens for ITDs or ILDs for the delayed signal (e.g., Zurek, 1980; Saberi and Perrott, 1990b; Shinn-Cunningham et al., 1993; Tollin and Henning, 1998). This effect occurs for delays up to about 5 ms and is aptly called **discrimination suppression**.

As suggested by the examples at the beginning of this section, the precedence effect suppresses the effects of reflections which would otherwise interfere with our perception of the direct sound, including speech. For example, Lochner and Burger (1964) found that speech discrimination was unaffected by reflections arriving up to 30 ms after the direct sound, although later-arriving reflections resulted in reduced intelligibility. Reflections that arrive beyond the time when the precedence effect is operative result in distortions and masking of the speech signal (e.g., Bolt and MacDonald, 1949; Kurtovic, 1975; Nabelek, 1976; Nabelek and Robinette, 1978; Gelfand and Silman, 1979; Nabelek and Dagenais, 1986; see Chap. 14).

Several interesting phenomena are encountered when the precedence effect is viewed in the context of the listener's immediate experience with the listening situation. Suppose we measure a listener's echo threshold several times: immediately after hearing leading/lagging click trains of various lengths (e.g., 3, 6, 9, 12, and 15 click pairs), and also without any preceding click pairs. [The upper left part of Fig. 13.16 (labeled a) shows an example of a train of 12 leading/lagging click pairs presented from right and left loudspeakers.] We would find that the echo threshold increases, as the preceding click train gets longer (reaching a maximum at 12 pairs). Thus, there is a *build up of echo suppression* that is dependent upon the listener's immediate experience with the listening task (Clifton and Freyman, 1989; Freyman et al., 1991; Clifton et al., 1994). What's more, this build up of echo suppression appears to be *asymmetrical*, being greater when the right click leads than when the left click is leads (Clifton and Freyman, 1989; Grantham, 1996).

A *breakdown of echo suppression* called the **Clifton effect** occurs when the leading and lagging clicks switch directions (Clifton, 1987; Clifton and Freyman, 1989). The basic demonstration of the Clifton effect involves presenting the listener with a train of click pairs coming from right and left loudspeakers (Fig. 13.16). The leading click in each pair is from the left and the lagging click is from the right, so that each click pair is heard as a fused image coming from the left side (labeled a in the figure). After 12 pairs, the order of the clicks is *switched* so that now the right click leads and the left one lags. At this point, we would expect to hear a fused image from the now-leading right side, but this does not happen. Instead, the *precedence effect breaks down* and the listener hears *both* clicks, each coming from its own side (b in the figure). Then, after several right/left click pairs are presented, the listener again hears a fused image from

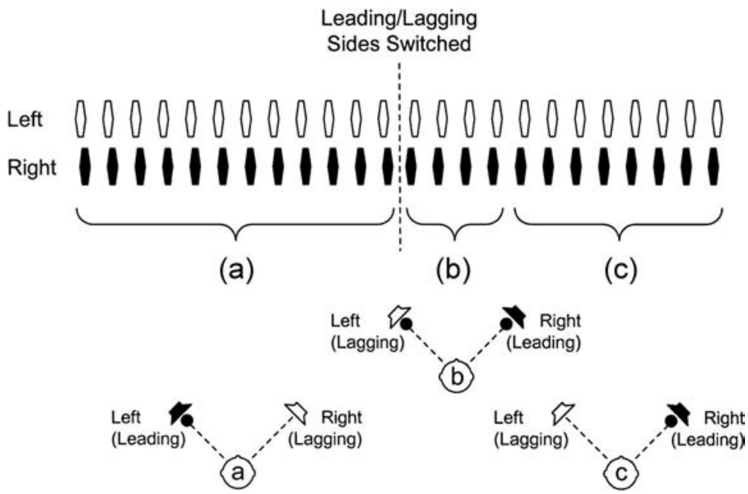

Figure 13.16 The Clifton effect. (a) Left-leading click pairs are heard as a fused image coming from the left (represented by the filled circle in the lower drawing). (b) Switching to right-leading click pairs causes the precedence effect to break down, and both clicks are heard coming from their respective sides (represented by the two filled circles). (c) The precedence effect is then re-established, with a fused signal heard from the now-leading right side (represented by the filled circle).

the now-leading right side, which indicates that the precedence effect has been re-established (c in the figure).

A fascinating illusion called the **Franssen effect** (Franssen, 1960, 1962) is illustrated in Fig. 13.17 Here, a low-frequency tone (e.g., 500 Hz) is presented from two loudspeakers. The tone from the left speaker has an abrupt onset and then fades away over the course of, say, 100 ms. The tone from the right speaker builds up, while the left tone is fading away and may stay on for a few seconds. The amplitude envelopes of the two tones over time are represented in Fig. 13.17a. As expected, the listener initially localizes a fused image coming from the left speaker, attributable to the abrupt onset (Fig. 13.17b). What is odd, however, is that the listener continues localizing the tone to the *left* even after the *left* signal is *off* and the *only* signal is coming from the *right* speaker (Fig. 13.17c). This illusion can actually persist for quite some time (Berkley, 1983, 1987). Interestingly, Hartmann and Rakerd (1989) showed that the Franssen effect fails to occur when the environment is anechoic (echo-free), and explained the illusion based on the **plausibility hypothesis**: The Franssen effect occurs in typical rooms, where reflections cause the ITDs for the steady-state tone from the right speaker (as in Fig. 13.17) to become so atypical that they are implausible as localization cues. The listener discounts these implausible cues and attributes the tone's direction to the unambiguous cue provided by the abrupt onset of the tone burst from the left speaker. The illusion does not occur in anechoic rooms because the ITDs are not distorted by reflections, and therefore provide the listener with plausible localization cues.

The build-up effect (including its right-left asymmetry), Clifton effect, and Franssen illusion reveal that the precedence

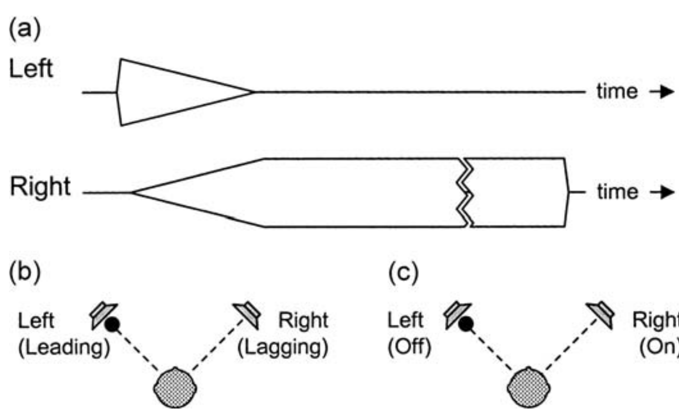

Figure 13.17 The Franssen effect (illusion). (a) Amplitude envelopes over time of the tones coming from the left and right loudspeakers. (b) Fused (filled circle) image initially localized to the leading left side. (c) Image (filled circle) continues being localized to the left side even after the left signal is off.

effect is influenced by, for example, experiences with and expectations about the listening task and environment, and the plausibility of the cues. Thus, higher-level central processes are involved in the precedence effect (e.g., Hafter et al., 1988; Hartmann and Rakerd, 1989; Freyman et al., 1991; Clifton et al., 1994; Grantham, 1996; Clifton and Freyman, 1997; Hartmann, 1997; Litovsky et al., 1999). There is also considerable physiological evidence of central involvement in the precedence effect (e.g., Litovsky, 1998; Litovsky et al., 1999; Litovsky and Delgutte, 2002). For example, in their comprehensive review, Litovsky et al. (1999) used physiological data from several studies in the literature to compare the time frames over which evidence of echo suppression occurred at various levels in the nervous system of cats. Suppression increased going from lower to higher levels of the auditory pathway from the auditory nerve to the auditory cortex.

MASKING LEVEL DIFFERENCES

The term **masking level difference (MLD)** may be a bit confusing at first glance. Obviously it refers to some sort of difference in masking. Consider a typical masking experiment (Chap. 10) in which a signal S is barely masked by a noise N. This can be done in one ear **(monotically),** as in Fig. 13.18a, or by presenting an *identical* signal and noise to both ears **(diotically),** as in Fig. 13.18b. (Identical stimuli are obtained by simply directing the output of the same signal and noise sources to both earphones in phase.) For brevity and clarity, we will adopt a shorthand to show the relationships among the stimuli and ears. The letter **m** will denote a monotic stimulus and **o** will refer to a diotic stimulus. Thus, **SmNm** indicates that the signal and noise are presented to one ear, and **SoNo** means that the same signal and the same noise are simultaneously presented to both ears. Either of these conditions can be used as our starting point.

Suppose that we add an identical noise to the unstimulated ear of Fig. 13.18a, so that the signal is still monotic but the noise is now diotic **(SmNo),** as in Fig. 13.18c. Oddly enough, the previously masked signal now becomes audible again! Starting this time from the masked situation in Fig. 13.18b (SoNo), we can make the signal audible again by reversing the phase of (inverting) the noise between the ears (Fig. 13.18d) or by reversing the phase of the signal between the ears (Fig. 13.18e). The phase reversal is indicated by the Greek letter π, since the stimuli are now 180° (or one radian, π) out of phase between the ears. (The phase reversal is accomplished by simply reversing the positive and negative poles at one of the earphones.) These new conditions are thus called **SoNπ** and **SπNo,** respectively. Note that the binaural advantage occurs only when the stimuli are in some way *different* at the two ears (dichotic). These fascinating observations were first reported in 1948 by Hirsh (1948b) for tonal signals and by Licklider (1948) for speech.

We may now define the MLD as the difference (advantage) in masked threshold between dichotically presented stimuli and signals that are presented monotically (or diotically). It is not surprising to find that the MLD is also referred to as **binaural unmasking**, **binaural release from masking**, or the **binaural masking level difference (BMLD)**. We shall express the magnitude of the MLD as the difference in decibels between a particular dichotic arrangement and either the SmNm or SoNo conditions. Other MLD conditions are discussed below.

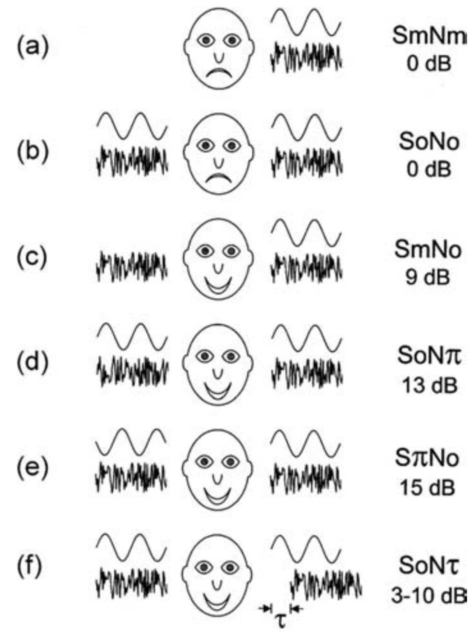

Figure 13.18 Masking level differences (MLDs) for various conditions (see text).

The size of the MLD varies from as large as about 15 dB for the SπNo condition (Green and Henning, 1969) to as little as 0 dB, depending upon a variety of parameters. Typical MLD magnitudes associated with various dichotic arrangements are shown in Fig. 13.18 The MLD becomes larger as the spectrum level of the masking noise is increased, especially when the noise is presented to both ears (No) at the same level (Hirsh, 1948a, 1948b; Blodgett et al., 1962; Dolan and Robinson, 1967; Dolan, 1968; McFadden, 1968).

The largest MLDs are obtained when either the signal (SπNo) or the noise (SoNπ) is opposite in phase at the two ears. The large MLDs obtained from these antiphasic conditions have been known since it was first described (Hirsh, 1948b) and have been repeatedly confirmed (e.g., Jeffress et al., 1952; Colburn and Durlach, 1965). Recall from Chapter 5 that the firing patterns of auditory nerve fibers are phase-locked to the stimulus, particularly at low frequencies. Thus, the large MLDs associated with antiphasic conditions may be related to this phase-locking in the neural coding of the stimuli (Green and Henning, 1969). Furthermore, since the degree of phase-locking is greatest at low frequencies, and decreases as frequency becomes higher,

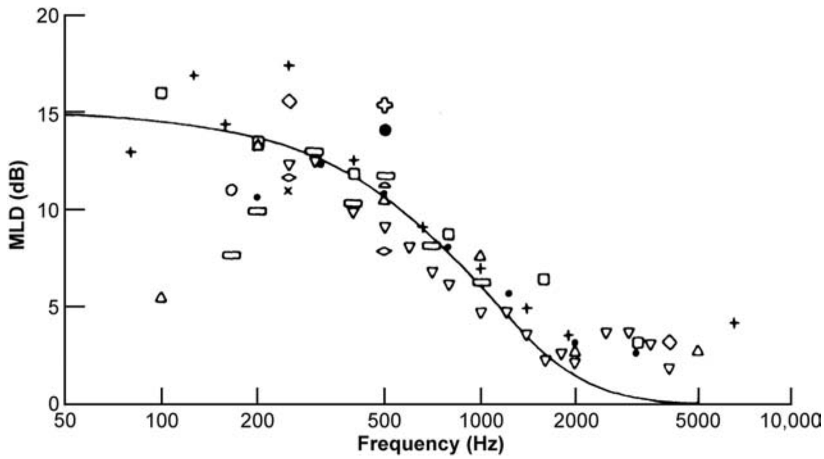

Figure 13.19 Magnitude of the MLD (SπNo–SoNo) as a function of frequency for many studies. *Source*: Adapted from Durlach, Binaural signal detection: equalization and cancellation theory, in *Foundations of Modern Auditory Theory 2* (J.V. Tobias, ed.), © 1972 by Academic Press.

we would expect the size of the MLD to be related to stimulus frequency as well.

Figure 13.19 shows the relationship between MLD size and stimulus frequency from several studies, as summarized by Durlach (1972). As expected, the MLD is largest for low frequencies—about 15 dB for 250 Hz—and decreases for higher frequencies until a constant of about 3 dB is maintained by about 1500 to 2000 Hz. (Look at the individual data points in Fig. 13.19 rather than at the smooth line, which is a statistical approximation of the actual results.) Note that the MLD (at least for SπNo) does not fall to zero above 1500 Hz, and recall in this context that statistically significant phase-locking is maintained as high as 5000 Hz (Chap. 5).

There is very good agreement about the size of the MLD for the frequencies above 250 Hz. At lower frequencies, there is a great deal of variation in the MLD sizes reported by different studies (e.g., Hirsh, 1948b; Webster, 1951; Durlach, 1963; Rabiner et al., 1966; Dolan, 1968). This is shown in Fig. 13.19 by the substantial spread among the data points for the lower frequencies. Much of this variation in the lower frequencies may be explained on the basis of differences in noise level. In particular, Dolan (1968) showed that the MLDs at 150 and 300 Hz increase with the spectrum level of the masker, attaining a value of approximately 15 dB when the noise spectrum level is 50 dB or more. Thus, the MLD becomes rather stable once moderate levels of presentation are reached.

We have been assuming that the noises at the two ears (SoNo) are derived from the same noise source, insuring that the waveforms are exactly the same at both ears. Another way to indicate identical waveforms is to say that the noises are perfectly correlated. Had we used two separate noise generators, then the noises would no longer be perfectly correlated. We would then say that the noises are uncorrelated (Nu). Robinson and Jeffress (1963) added noises from the same (correlated) and different (uncorrelated) generators to study how noise correlation affects the size of the MLD. They found that the MLD resulting from uncorrelated noises is on the order of 3 to 4 dB, and that the MLD becomes larger as the degree of correlation decreases. The MLD resulting from uncorrelated noise may contribute to our ability to overcome the effects of reverberation. This relation was demonstrated by Koenig et al. (1977), who found that room reverberation decorrelates the noise reaching the two ears, resulting in an MLD of about 3 dB.

As only a certain critical bandwidth contributes to the masking of a tone (Chap. 10), it is not surprising that the MLD also depends upon the critical band (Sondhi and Guttman, 1966; Mulligan et al., 1967) around a tone. In fact, the MLD actually increases as the noise band narrows (Metz et al., 1967; Wightman, 1971). As it turns out, a very narrow band of noise looks much like a sinusoid that is being slowly modulated in frequency and amplitude. If we present such a narrow band noise to both ears and delay the wavefront at one ear relative to the other, then the degree to which the noises are correlated will change periodically as a function of the interaural time delay.

With this in mind, consider the arrangement in Fig. 13.18f. The noises presented to the two ears are from the same generator, but the noise is delayed at one ear relative to the other (Nτ). The interaural time delay decorrelates the noises in a manner dependent upon the time delay. This situation (**SoNτ**) results in MLDs, which are maximal when the time delay corresponds to half-periods of the signal and minimal when the time delays correspond to the period of the signal (Rabiner et al., 1966; Langford and Jeffress, 1964). Figure 13.20 shows example at several frequencies. The effect is clearest at 500 Hz. The period of 500 Hz is 2 ms, and the half-period is thus 1 ms. As the figure shows, the MLDs are largest at multiples of the half-period (in the vicinity of 1 ms and 3 ms for 500 Hz) and are smallest at multiples of the full period (about 2 ms and 4 ms for 500 Hz). Also notice that successive peaks tend to become smaller and smaller.

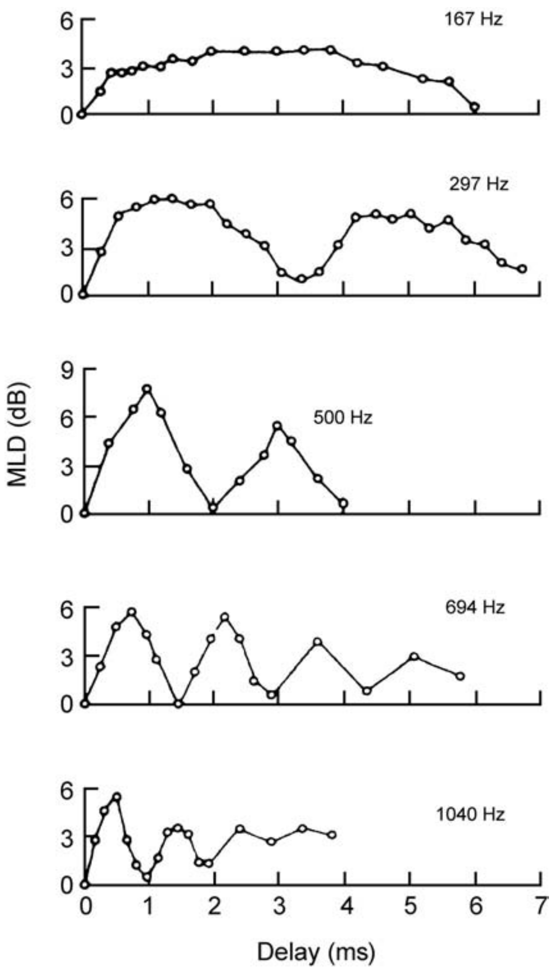

Figure 13.20 The MLD as a function of interaural time delay. *Source*: Adapted from Rabiner et al. (1966), with permission of *J. Acoust. Soc. Am.*

Licklider (1948) reported MLDs for speech about the same time that the phenomenon was described for tones (Hirsh, 1948a, 1948b). Interestingly enough, the unmasking of speech is associated with the MLDs for pure tones within the spectral range critical for speech perception (Carhart et al., 1967, 1968; Levitt and Rabiner, 1967a). This was shown quite clearly in a study by Levitt and Rabiner (1967a) that used monosyllabic words as the signal and white noise as the masker. The subjects were asked to indicate whether the test words were detectable in the presence of the noise while different parts of the speech spectrum were reversed in phase between the ears. They found MLDs (SπNo) on the order of 13 dB when the frequencies below 500 Hz were reversed in phase, indicating that the MLD for speech detection is primarily determined by the lower frequencies in the speech spectrum.

The MLD for speech detection obviously occurs at a minimal level of intelligibility. That is, a signal whose intelligibility is zero (see Chap. 14) may still be barely detectable. Increasing the presentation level of the words would result in higher intelligibility, that is, a larger proportion of the words would be correctly repeated. The level at which half of the words are correctly repeated may be called the 50% intelligibility level; the 100% intelligibility level is the test level at which all of the words are repeated correctly. The speech MLD is quite large at near-detection levels. However, MLDs for speech are smallest at higher presentation levels where overall intelligibility is good (Schubert and Schultz, 1962; Green and Yost, 1975; Carhart et al., 1967, 1968; Levitt and Rabiner, 1967a).

Levitt and Rabiner (1967a, 1967b) suggested the term **binaural intelligibility level difference** (**BILD** or **ILD**)[5] to indicate the amount of unmasking for speech intelligibility. The BILD is the difference between the levels at which a particular percentage of the test words are correctly repeated for a dichotic condition and for SoNo. Levitt and Rabiner (1967a) found that the BILD was on the order of only 3 to 6 dB for an intelligibility level of 50%. Thus, the release from masking increases from about 3 to 13 dB as the intelligibility level goes down toward bare detection. At the lowest intelligibility level, where one can detect but not repeat the words, the BILD and MLD for speech are synonymous. In this light, we may consider the MLD as the limiting case of the BILD (Levitt and Rabiner, 1967b).

Although the speech MLD depends on frequencies below 500 Hz, Levitt and Rabiner (1967a) found that the entire speech spectrum makes a significant contribution to the BILD. In a later paper they presented a numerical method for predicting the BILD (Levitt and Rabiner, 1967b). The procedure assumes that the SπNo condition reduces the noise level in a manner that depends upon frequency and signal-to-noise ratio. That is, the lower frequencies are given greater relative importance at low signal-to-noise ratios where overall intelligibility is poor. This technique makes predictions that are in close agreement with the empirical data (Licklider, 1948; Schubert and Schultz, 1962; Levitt and Rabiner, 1967a).

Some work has also been done on MLDs for differential sensitivity and loudness (Townsend and Goldstein, 1972; Henning, 1973). We shall not go into detail in these areas, except to point out that MLDs for loudness and discrimination of intensity occur mainly at near-detection levels and become insignificant well above threshold. This situation, of course, is analogous to what we have seen for BILDs as opposed to MLDs for speech signals. However, Marks (1987) reported that binaural loudness summation for a 1000 Hz tone under MLD-like masking conditions was substantially greater than a doubling of loudness over the monaural condition.

Although a variety of models has been proposed to explain the mechanism of MLDs, we will take a brief look at two of the

[5] We will use BILD to avoid confusion because ILD also means interaural level difference.

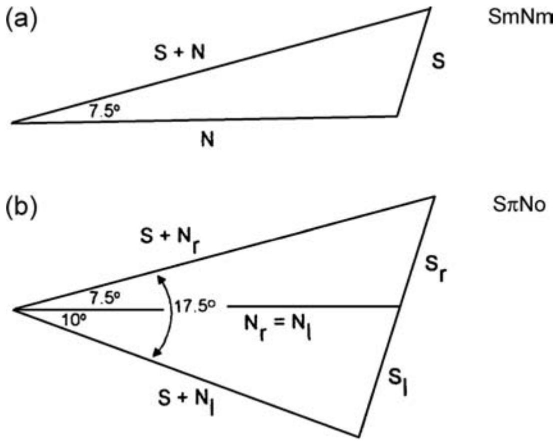

Figure 13.21 Vector diagrams for the (a) SmNm and (b) SπNo conditions (see text). *Source*: Adapted from Jeffress, Binaural signal detection: Vector theory, in *Foundations of Modern Auditory Theory, vol. 2* (J.V. Tobias, ed.), © 1972 by Academic Press.

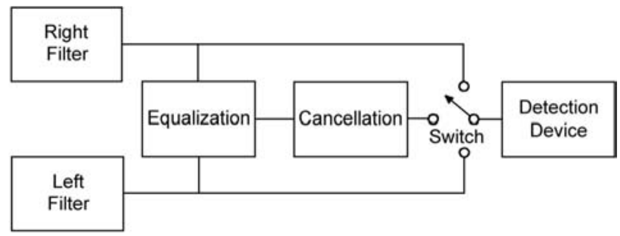

Figure 13.22 A simplified block diagram of Durlach's equalization–cancellation (EC) model.

best-known approaches, the Webster–Jeffress lateralization theory (Webster, 1951; Jeffress, 1972) and Durlach's equalization–cancellation model (Durlach, 1963, 1972; Culling and Summerfield, 1995; Breebaart, van de Par, and Kohlrausch, 2001).

The Webster–Jeffress **lateralization theory** attributes MLDs to interaural phase and level differences. Recall that only a certain critical band contributes to the masking of a tone, and that this limited-bandwidth concept also applies to MLDs. Basically, the lateralization model compares the test tone to the narrow band of frequencies in the noise that contributes to its masking. Changing phase between the ears (SπNo) results in a time-of-arrival difference at a central mechanism, which provides the detection cue. The vector diagrams in Fig. 13.21 show how this system might operate for SπNo versus SmNm. Figure 13.21a shows the interaction between the noise (N) and signal (S) amplitudes as vectors 7.5° apart. The resulting S + N amplitude is too small for detection. Figure 13.21b shows the situation when the signal is reversed in phase at the two ears (SπNo). Now, the right and left noise amplitudes (Nr and Nl) are equal (Nr = Nl), but the signal vectors for the two ears (Sr and Sl) point in opposite directions due to the reversed phase. Thus, the signal in the right ear leads the noise in phase by 7.5°, while in the left ear it lags in phase by 10° so that the phase difference between the resulting S = N at the two ears is 17.5°. In other words, S + N at the right ear is now 17.5° ahead of that in the left. If the signal is a 500-Hz tone, this lag corresponds to a time-of-arrival advantage for the right S + N of 97 μs. In addition, the lengths of the S + N vectors indicate an amplitude advantage that is also available as a detection cue, since the right S + N is now substantially longer (more intense) than the left. That is, the SπNo condition causes the combined S + N from the right ear to reach a central detection mechanism sooner and with greater amplitude. This causes a lateralization of the result (to the right in this example), which the model uses as the basis for detection of the signal.

Durlach's **equalization–cancellation (EC) model** is shown schematically in Fig. 13.22. The stimuli pass through critical-band filters at the two ears, and then follow both monaural and binaural routes to a detection device, which decides whether the signal is present. The detection device switches among the three possible channels (two monaural and one binaural), and will use the channel with the best signal-to-noise ratio as the basis for a response. The monaural channels go straight to the detection mechanism. The binaural channel, however, includes two special stages.

In the first stage, the inputs from the two ears are adjusted to be equal in amplitude (the equalization step). Then the inputs from the two ears are subtracted one from the other in the cancellation step.[6] Of course, if the signal and noise were identical in both ears (SoNo), then the entire binaural signal would be canceled. In this case, the detection device would choose among the monaural inputs so that no MLD resulted. However, for the SπNo condition, the subtraction cancels the in-phase noise and actually enhances the out-of-phase signal. Thus, if the EC model works perfectly, the signal-to-noise ratio will be improved infinitely. In reality, the mechanism operates less than perfectly so that cancellation is not complete. This imperfection is due to atypical stimuli that necessitate unusual types of equalization, or to random jitter in the process, which causes the cancellation mechanism to receive inputs that are imperfectly equalized.

REFERENCES

Abbagnaro, LA, Bauer, BB, Torick, EL. 1975. Measurements of diffraction and interaural delay of a progressive sound wave caused by the human head-II. *J Acoust Soc Am* 58, 693–700.

Algazi, VR, Avendano, C, Duda, RO. 2001. Elevation localization and head-related transfer function analysis at low frequencies. *J Acoust Soc Am* 109, 1110–1122.

American National Standards Institute, ANSI S3.4–2007.2007. American National Standard Procedure for the Computation of Loudness of Steady Sounds. New York, NY: ANSI.

[6] Recall the rules for algebraic subtraction $(+1) - (+1) = 0$, as for the in-phase noise; $(+1) - (-1) = +2$, as for the out-of-phase signal.

Arbogast, TL, Mason, CR, Kidd, G. 2002. The effect of spatial separation on informational and energetic masking of speech. *J Acoust Soc Am* 112, 2086–2098.

Begault, DR. 1991. Preferred sound intensity increase for sensation of half distance. *Percept Motor Skill* 72, 1019–1029.

Bekesy, G. 1938. Über die Entstehung der Entfernungsempfindung beim Hören. Akustiche Zeitschrif 3, 21–31.

Berkley, DA. 1983. Room acoustics and listening. *J Acoust Soc Am* 44, S17.

Berkley, DA. 1987. Hearing in rooms. In: WA Yost, G Gourevitch (eds.), *Directional Hearing*. New York, NY: Springer, 249–260.

Bernstein, LR, Trahiotis, C. 1995. Binaural interference effects measured with masking-level difference and with ITD- and IID-discrimination paradigms. *J Acoust Soc Am* 98, 155–163.

Best, V, Gallun, FJ, Carlile, S, Shinn-Cunningham, BG. 2007. Binaural interference and auditory grouping. *J Acoust Soc Am* 121, 1070–1076.

Blauert, J. 1969/70. Sound localization in the median plane. *Acustica* 22, 205–213.

Blauert, J. 1997. *Special Hearing: The Psychophysics of Human Sound Localization.* Revised Edition. Cambridge, MA: MIT Press.

Blodgett, HC, Jeffress, LA, Whitworth, RH. 1962. Effect of noise at one ear on the masking threshold for tones at the other. *J Acoust Soc Am* 34, 979–981.

Bolt, RH, MacDonald, AD. 1949. Theory of speech masking by reverberation. *J Acoust Soc Am* 21, 577–580.

Breebaart, J, van de Par, S, Kohlrausch, A. 2001. Binaural processing model based on contralateral inhibition. I. Model structure. *J Acoust Soc Am* 110, 1074–1088.

Bregman, AS. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: Bradford Books, MIT Press.

Bregman, AS, Ahad, PA. 1996. *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound [Compact Disc]*. Cambridge, MA: MIT Press.

Broadbent, DE. 1955. A note on binaural fusion. *Q J Exp Psychol* 7, 46–47.

Bronkhorst, AW. 1995. Localization of real and virtual sources. *J Acoust Soc Am* 98, 2542–2553.

Bronkhorst, AW, Houtgast, T. 1999. Auditory distance perception in rooms. *Nature* 397, 517–520.

Bronkhorst, AW, Plomp, R. 1988. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J Acoust Soc Am* 83, 1508–1516.

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, Scott, KR. 2001. The effects of production and presentation level on the auditory distance perception of speech. *J Acoust Soc Am* 106, 1465–1479.

Buell, TN, Hafter, ER. 1991. Combination of binaural information across frequency bands. *J Acoust Soc Am* 90, 1894–1900.

Butler, RA. 1986. The bandwidth effect on monaural and binaural localization. *Hear Res* 21, 67–73.

Butler, RA, Levy, ET, Neff, WD. 1980. Apparent distance of sounds recorded in echoic and anechoic chambers. *J Exp Psychol Human Percept Perf* 6, 745–750.

Butler, RA, Planert, N. 1976. The influence of stimulus bandwidth on localization of sound in space. *Percept Psychophys* 19, 103–108.

Carhart, R, Tillman, T, Dallos, P. 1968. Unmasking for pure tones and spondees: Interaural phase and time disparities. *J Speech Hear Res* 11, 722–734.

Carhart, R, Tillman, T, Johnson, K. 1967. Release of masking for speech through interaural time delay. *J Acoust Soc Am* 42, 124–138.

Causse, R, Chavasse, P. 1941. Recherches sur les seuil de l'audition binauriculaire compare au seuil monauriculaire en fonction de la frequence. *Comp R Soc Biol* 135, 1272–1275.

Causse, R, Chavasse, P. 1942a. Difference entre le seuil de l'audition binauriculaire et le seuil monauriculaire de la frequence. *Comp R Soc Biol* 136, 301.

Causse, R, Chavasse, P. 1942b. Difference entre l'ecoute binauriculaire et monauriculaire por la perception des intensities supraliminaires. *Comp R Soc Biol* 139, 405.

Chandler, DW, Grantham, DW. 1992. Minimum audible movement angle in the horizontal plane as a function of frequency and bandwidth, source azimuth, and velocity. *J Acoust Soc Am* 91, 1625–1636.

Cherry, C. 1961. Two ears-but one world. In: WA Rosenblith (ed.), *Sensory Communication.* Cambridge, MA: MIT Press, 99–117.

Cherry, EC. 1953. Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25, 975–979.

Cherry, EC, Sayers, BMcA. 1956. "Human cross-correlator"—A technique for measuring certain parameters of speech perception. *J Acoust Soc Am* 28, 889–896.

Churcher, BG, King, AJ, Davies, H. 1934. The minimal perceptible change of intensity of a pure tone. *Phil Mag* 18, 927–939.

Clifton, RK. 1987. Breakdown of echo suppression in the precedence effect. *J Acoust Soc Am* 82, 1834–1835.

Clifton, RK, Freyman, RL. 1989. Effect of click rate and delay on breakdown of the precedence effect. *Percept Psychophys* 462, 139–145.

Clifton, RK, Freyman, RL. 1997. The precedence effect: Beyond echo suppression. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments.* Hillsdale, NJ: Erlbaum, 233–255.

Clifton, RK, Freyman, RL, Litovsky, RY, McCall, D. 1994. Listeners' expectations about echoes can raise or lower echo threshold. *J Acoust Soc Am* 95, 1525–1533.

Cochran, P, Throop, J, Simpson, WE. 1968. Estimation of distance of a source of sound. *Am J Psychol* 81, 198–206.

Colburn, HS, Durlach, NI. 1965. Time-intensity relations in binaural unmasking. *J Acoust Soc Am* 38, 93–103.
Coleman, PD. 1962. Failure to localize the source distance of an unfamiliar sound. *J Acoust Soc Am* 34, 345–346.
Coleman, PD. 1963. Analysis of cues to auditory depth perception in free space. *Psychol Bull* 60, 302–315.
Coleman, PD. 1968. Dual role of frequency spectrum in determination of auditory distance. *J Acoust Soc Am* 44, 631–632.
Culling, JF, Summerfield, Q. 1995. Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *J Acoust Soc Am* 98, 785–797.
Dixon, RM, Colburn, HS. 2006. The influence of spectral, temporal and interaural stimulus variations on the precedence effect. *J Acoust Soc Am* 119, 2947–2964.
Dolan, TR. 1968. Effects of masker spectrum level on masking-level differences at low signal frequencies. *J Acoust Soc Am* 44, 1507–1512.
Dolan, TR, Robinson, DE. 1967. An explanation of masking-level differences that result from interaural intensive disparities of noise. *J Acoust Soc Am* 42, 977–981.
Duda, RO, Martens, WL. 1998. Range dependence of the response of a spherical head model. *J Acoust Soc Am* 104, 3048–3058.
Durlach, NI. 1963. Equalization and cancellation theory of binaural masking level differences. *J Acoust Soc Am* 35, 1206–1218.
Durlach, NI. 1972. Binaural signal detection: Equalization and cancellation theory. In: JV Tobias (ed.), *Foundations of Modern Auditory Theory*, Vol. 2. New York, NY: Academic Press, 369–462.
Durlach, NI, Colburn, HS. 1978. Binaural phenomena. In: EC Carterette, MP Friedman (eds.), *Handbook of Perception, Vol. I -Hearing*. New York, NY: Academic Press, 365–466.
Durlach, NI, Rigopulos, A, Pang, XD, Woods, WS, Kulkarni, A, Colburn, HS, Wenzel, EM. 1992. On the externalization of auditory images. *Presence* 1, 251–257.
Ebata, M, Sone, T, Nimura, T. 1968. On the perception of direction of echo. *J Acoust Soc Am* 44, 542–547.
Feddersen, WE, Sandel, TT, Teas, DC, Jeffress, LA. 1957. Localization of high-frequency tones. *J Acoust Soc Am* 29, 988–991.
Fisher, HG, Freedman, SJ. 1968. The role of the pinna in auditory localization. *J Aud Res* 8, 15–26.
Fletcher, H, Munson, W. 1933. Loudness: Its definition, measurement and calculation. *J Acoust Soc Am* 5, 82–108.
Franssen, NV. 1960. Some considerations on the mechanism of directional hearing. Ph.D. Dissertation. Delft, The Netherlands: Technische Hogeschool.
Franssen, NV. 1962. *Stereophony*. Eindhoven, The Netherlands: Phillips Technical Library. [English translation, 1964.]
Freyman, RL, Clifton, RK, Litovsky, RY. 1991. Dynamic processes in the precedence effect. *J Acoust Soc Am* 90, 874–884.
Gallum, FJ, Mason, CR, Kidd, G. 2005. Binaural release from informational masking in a speech identification task. *J Acoust Soc Am* 118, 1605–1613.
Gardner, MB. 1968. Historical background of the Haas and/or precedence effect. *J Acoust Soc Am* 43, 1243–1248.
Gardner, MB. 1969. Distance estimation of 0 degrees or apparent 0 degree-oriented speech signals in anechoic space. *J Acoust Soc Am* 45, 47–53.
Gardner, MB. 1973. Some monaural and binaural facets of median plane localization. *J Acoust Soc Am* 54, 1489–1495.
Gelfand, SA, Silman, S. 1979. Effects of small room reverberation upon the recognition of some consonant features. *J Acoust Soc Am* 66, 22–29.
Grantham, DW. 1984. Interaural intensity discrimination: Insensitivity at 1000 Hz. *J Acoust Soc Am* 75, 1190–1194.
Grantham, DW. 1986. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *J Acoust Soc Am* 79, 1939–1949.
Grantham, DW. 1996. Left-right asymmetry in the buildup of echo suppression in normal-hearing adults. *J Acoust Soc Am* 99, 1118–1123.
Green, DM. 1966. Interaural phase effects in the masking of signals of different durations. *J Acoust Soc Am* 39, 720–724.
Green, DM, Henning, GR. 1969. Audition. *Ann Rev Psychol* 20, 105–128.
Green, DM, Yost, WA. 1975. Binaural analysis. In: WD Keidel, WD Neff (eds.), *Handbook of Sensory Physiology, Vol. V/2: Auditory System*. New York, NY: Springer-Verlag, 461–480.
Groen, JJ. 1964. Super- and subliminate binaural beats. *Acta Otol* 57, 224–230.
Gu, X, Wright, BA, Green, DM. 1995. Failure to hear binaural beats below threshold. *J Acoust Soc Am* 97, 701–703.
Haas, H. 1949. The influence of a single echo on the audibility of speech. *Library Com* 363. Garston, Watford, UK: Dept. Sci. Indust. Rest.
Haas, H. 1951. Über den Einfluss eines Einfachechos aud die Hörsamkeit von Sprache. *Acustica* 1, 49–58. [English translation: *J Audiol Eng Soc* 20, 146–159 (1972).]
Hafter, ER, DeMaio, J. 1975. Difference thresholds for interaural delay. *J Acoust Soc Am* 57, 181–187.
Hafter, ER, Bourbon, WT, Blocker, AS, Tucker, A. 1969. Direct comparison between lateralization and detection under antiphasic masking. *J Acoust Soc Am* 46, 1452–1457.
Hafter, ER, Buell, TN, Richards, V. 1988. Onset-coding in lateralization: Its form, site, and function. In: GM Edelman, WE Gall, WM Cowan (eds.), *Auditory Function: Neurobiological Bases of Hearing*. New York, NY: Wiley, 647–676.
Hafter, ER, Carrier, SC. 1970. Masking-level differences obtained with a pulsed tonal masker. *J Acoust Soc Am* 47, 1041–1048.
Hafter, ER, Carrier, SC. 1972. Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely. *J Acoust Soc Am* 51, 1852–1862.
Hafter, ER, Jeffress, LA. 1968. Two-image lateralization of tones and clicks. *J Acoust Soc Am* 44, 563–569.
Harris, GG. 1960. Binaural interactions of impulsive stimuli and pure tones. *J Acoust Soc Am* 32, 685–692.

Harris, JD. 1963. Loudness discrimination. *J Speech Hear Dis Monogr Suppl* 11.

Harris, JD, Sergeant, RL. 1971. Monaural/binaural minimum audible angle for a moving sound source. *J Speech Hear Res* 14, 618–629.

Hartmann, WM. 1997. Listening in a room and the precedence effect. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. Hillsdale, NJ: Erlbaum, 191–210.

Hartmann, WM, Rakerd, B. 1989. Localization of sound in rooms. IV: The Franssen effect. *J Acoust Soc Am* 86, 1366–1373.

Hartmann, WM, Wittenberg, A. 1996. On the externalization of sound images. *J Acoust Soc Am* 99, 3678–3688.

Hawley, ML, Litovsky, RY, Culling, JF. 2004. The benefits of binaural hearing in a cocktail party: Effect of location and type of interferer. *J Acoust Soc Am* 115, 833–843.

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.

Heller, LM, Trahiotis, C. 1995. Extents of laterality and binaural interference effects. *J Acoust Soc Am* 99, 3632–3637.

Hellman, RP, Zwislocki, J. 1963. Monaural loudness function at 1000 cps, and interaural summation. *J Acoust Soc Am* 35, 856–865.

Henning, GB. 1973. Effect of interaural phase on frequency and amplitude discrimination. *J Acoust Soc Am* 54, 1160–1178.

Henning, GB. 1974. Delectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* 55, 84–90.

Hill, NI, Darwin, CJ. 1996. Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning. *J Acoust Soc Am* 100, 2352–2364.

Hirsh, IJ. 1948a. Binaural summation: A century of investigation. *Psychol Bull* 45, 193–206.

Hirsh, IJ. 1948b. The influence of interaural phase on interaural summation and inhibition. *J Acoust Soc Am* 20, 536–544.

Hirsh, IJ, Burgeat, M. 1958. Binaural effects in remote masking. *J Acoust Soc Am* 30, 827–832.

Jeffress, LA. 1972. Binaural signal detection: Vector theory. In: JV Tobias (ed.), *Foundations of Modern Auditory Theory*, Vol. 2. New York, NY: Academic Press, 349–368.

Jeffress, LA, Blodgett, HC, Deatheredge, BH. 1952. The masking of tones by white noise as a function of interaural phases of both components. *J Acoust Soc Am* 24, 523–527.

Jeffress, LA, McFadden, D. 1971. Differences of interaural phase and level in detection and lateralization. *J Acoust Soc Am* 49, 1169–1179.

Jesteadt, W, Wier, CC, Green, DM. 1977a. Comparison of monaural and binaural discrimination of intensity and frequency. *J Acoust Soc Am* 61, 1599–1603.

Jesteadt, W, Wier, CC, Green, DM. 1977b. Intensity discrimination as a function of frequency and sensation level. *J Acoust Soc Am* 61, 169–177.

Keys, J. 1947. Binaural versus monaural hearing. *J Acoust Soc Am* 19, 629–631.

Kidd, GK, Mason, CR, Brughera, A, Hartmann, WM. 2005. The role of reverberation in the release from masking due to spatial separation for speech intelligibility. *Acta Acustica with Acustica* 91, 526–536.

Klumpp, RG, Eady, HR. 1956. Some measurements of interaural time difference thresholds. *J Acoust Soc Am* 28, 859–860.

Koenig, AH, Allen, JB, Berkley, DA, Curtis, TH. 1977. Determination of masking-level differences in a reverberant environment. *J Acoust Soc Am* 61, 1374–1376.

Kuhn, GF. 1977. Model for the interaural time differences in azimuthal plane. *J Acoust Soc Am* 62, 157–167.

Kurtovic, H. 1975. The influence of reflected sound upon speech intelligibility. *Acustica* 33, 32–39.

Langford, TL, Jeffress, LA. 1964. Effect of noise cross-correlation on binaural signal detection. *J Acoust Soc Am* 36, 1455–1458.

Leakey, DM, Sayers, BMcA, Cherry, EC. 1958. Binaural fusion of low and high frequency sounds. *J Acoust Soc Am* 30, 222–223.

Lehnhardt, E. 1961. Die akustische Korrelation. *Arch Ohren Nasen Kehlkopfheilk* 178, 493–497.

Levitt, H, Rabiner, LR. 1967a. Binaural release from masking for speech and grain in intelligibility. *J Acoust Soc Am* 42, 601–608.

Levitt, H, Rabiner, LR. 1967b. Predicting binaural gain in intelligibility and release from masking for speech. *J Acoust Soc Am* 42, 820–829.

Licklider, JCR. 1948. The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am* 20, 150–159.

Licklider, JCR, Webster, JC, Hedlun, JM. 1950. On the frequency limits of binaural beats. *J Acoust Soc Am* 22, 468–473.

Litovsky, RY. 1998. Physiological studies on the precedence effect in the inferior colliculus of the kitten. *J Acoust Soc Am* 103, 3139–3152.

Litovsky, RY, Colburn, HS, Yost, WA, Guzman, SJ. 1999. The precedence effect. *J Acoust Soc Am* 106, 1633–1654.

Litovsky, RY, Delgutte, B. 2002. Neural correlates of the precedence effect in the inferior colliculus: Effect of localization cues. *J Neurophysiol* 87, 976–994.

Litovsky, RY, Macmillan, NA. 1994. Sound localization precision under conditions of the precedence effect: Effects of azimuth and standard stimuli. *J Acoust Soc Am* 96, 752–758.

Litovsky, RY, Shinn-Cunningham, BG. 2001. Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression. *J Acoust Soc Am* 109, 346–358.

Little, AD, Mershon, DH, Cox, PH. 1992. Spectral content as a cue to perceived auditory distance. *Perception* 21, 405–416.

Lochner, JPA, Burger, JF. 1958. The subjective masking of short time delayed echoes, their primary sounds, and their contribution to the intelligibility of speech. *Acustica* 8, 1–10.

Lochner, JPA, Burger, JF. 1964. The influence of reflections on auditorium acoustics. *J Sound Vib* 1, 426–454.
Lord Rayleigh (Strutt, JW). 1907. Our perception of sound duration. *Phil Mag* 13, 214–232.
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.
Makous, JC, Middlebrooks, JC. 1990. Two-dimensional sound localization by human listeners. *J Acoust Soc Am* 87, 2188–2200.
Marks, LE. 1978. Binaural summation of loudness of pure tones. *J Acoust Soc Am* 64, 107–113.
Marks, LE. 1987. Binaural versus monaural loudness: Supersummation of tone partially masked by noise. *J Acoust Soc Am* 81, 122–128.
McFadden, D. 1968. Masking-level differences determined with and without interaural disparities in masking intensity. *J Acoust Soc Am* 44, 212–223.
McFadden, D, Pasanen, EG. 1976. Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am* 59, 634–639.
McGregor, P, Horn, AG, Todd, MA. 1985. Are familiar sounds ranged more accurately? *Percept Motor Skills* 61, 1082.
Mershon, DH, Ballenger, WL, Little, AD, McMurtry, PL, Buchanan, JL. 1989. Effects of room reflectance and background noise on perceived auditory distance. *Perception* 18, 403–416.
Mershon, DH, Bowers, JN. 1979. Absolute and relative cues for the auditory perception of egocentric distance. *Perception* 8, 311–322.
Mershon, DH, King, LE. 1975. Intensity and reverberation as factors in the auditory perception of echocentric distance. *Percept Psychophys* 18, 409–415.
Metz, PJ, von Bismark, G, Durlach, NI. 1967. I. Further results on binaural unmasking and the EC model. II. Noise bandwidth and interaural phase. *J Acoust Soc Am* 43, 1085–1091.
Middlebrooks, JC. 1992. Narrow-band sound localization related to external ear acoustics. *J Acoust Soc Am* 92, 2607–2624.
Middlebrooks, JC. 1997. Spectral shape cues for sound localization. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. Hillsdale, NJ: Erlbaum, 77–97.
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, Green, DM. 1991. Sound localization by human listeners. *Ann Rev Psychol* 42, 135–159.
Mills, AW. 1958. On the minimal audible angle. *J Acoust Soc Am* 30, 237–246.
Mills, AW. 1960. Lateralization of high-frequency tones. *J Acoust Soc Am* 32, 132–134.
Mills, AW. 1963. Auditory perception of spatial relations. *Proc Int Cong Tech Blind*,Vol. 2. New York, NY: American Foundation for the Blind.
Mills, AW. 1972. Auditory localization. In: JV Tobias (ed.), *Foundations of Modern Auditory Theory*, Vol. 2. New York, NY: Academic Press, 301–348.
Møller, H, Sorensen, MF, Jensen, CB, Hammershoi, D. 1996. Binaural technique: Do we need individual recordings? *J Audio Eng Soc* 44, 451–469.
Moushegian, G, Jeffress, LA. 1959. Role of interaural time and intensity differences in the lateralization of low-frequency tones. *J Acoust Soc Am* 31, 1441–1445.
Mrsic-Flogel, TD, King, AJ, Jenison, RL, Schnupp JWH. 2001. Listening through different ears alters special response fields in ferret primary auditory cortex. *J Neurophysiol* 86, 1043–1046.
Mulligan, BE, Mulligan, MJ, Stonecypher, JF. 1967. Critical band in binaural detection. *J Acoust Soc Am* 41, 7–12.
Musicant, A, Butler, R. 1984. The influence of pinnae-based spectral cues on sound localization. *J Acoust Soc Am* 75, 1195–1200.
Nabelek, AK. 1976. Reverberation effects for normal and hearing-impaired listeners. In: SK Hirsh, DH Eldridge, IJ Hirsh, SR Silverman (eds.), *Hearing and Davis: Essays Honoring Hallowell Davis*. St. Louis, MO: Washington University Press, 333–341.
Nabelek, AK, Dagenais, PA. 1986. Vowel errors in noise and in reverberation by hearing impaired listeners. *J Acoust Soc Am* 80, 741–748.
Nabelek, AK, Robinette, L. 1978. Influence of the precedence effect on word identification by normal hearing and hearing-impaired subjects. *J Acoust Soc Am* 63, 187–194.
Neutzel, JM, Hafter, ER. 1976. Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration. *J Acoust Soc Am* 60, 1339–1346.
Neutzel, JM, Hafter, ER. 1981. Discrimination of interaural delays in complex waveforms: Spectral effects. *J Acoust Soc Am* 69, 1112–1118.
Nielsen, SH. 1991. Distance perception in hearing. Unpublished Ph.D. Dissertation. Aalborg, Denmark: Aalborg University.
Nielsen, SH. 1993. Auditory distance perception in different rooms. *J Audio Eng Soc* 41, 755–770.
Nordlund, B. 1962. Physical factors in angular localization. *J Acoust Soc Am* 54, 75–93.
Oldfield, SR, Parker, PA. 1984. Acuity of sound localisation: A topography of auditory space. I. Normal hearing conditions. *Perception* 13, 581–600.
Palmer, SE. 1999. *Vision Science: Photon to Phenomenology*. Cambridge, MA: MIT Press.
Perrett, S, Noble, W. 1997a. The contribution of head motion cues to localization of low-pass noise. *Percept Psychophys* 59, 1018–1026.

Perrett, S, Noble, W. 1997b. The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am* 102, 2325–2332.

Perrot, DR, Marlborough, K. 1989. Minimum audible movement angle: Marking the end points of a path traveled by a moving sound source. *J Acoust Soc Am* 85, 1773–1775.

Perrott, DR. 1984. Concurrent minimum audible angle: A reexamination of the concept of auditory special acuity. *J Acoust Soc Am* 75, 1201–1206.

Perrott, DR, Marlborough, K, Merrill, P, Strybel, TZ. 1989. Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate. *J Acoust Soc Am* 85, 282–288.

Perrott, DR, Musicant, A. 1977. Minimum audible movement angle: Binaural localization of moving sources. *J Acoust Soc Am* 62, 1463–1466.

Perrott, DR, Tucker, J. 1988. Minimum audible movement angle as a function of signal frequency and the velocity of the source. *J Acoust Soc Am* 83, 1522–1527.

Petersen, J. 1990. Estimation of loudness and apparent distance of pure tones in a free field. *Acustica* 70, 61–65.

Pickler, AG, Harris, JD. 1955. Channels of reception in pitch discrimination. *J Acoust Soc Am* 27, 124–131.

Plenge, G. 1974. On the difference between localization and lateralization. *J Acoust Soc Am* 56, 944–951.

Pollack, I. 1948. Monaural and binaural threshold sensitivity for tones and white noise. *J Acoust Soc Am* 20, 52–58.

Pollack, I, Rose, M. 1967. Effects of head movements on the localization of sounds in the equatorial plane. *Percept Psychophys* 2, 591–596.

Rabiner, LR, Lawrence, CL, Durlach, NI. 1966. Further results on binaural unmasking and the EC model. *J Acoust Soc Am* 40, 62–70.

Reynolds, GS, Stevens, SS. 1960. Binaural summation of loudness. *J Acoust Soc Am* 32, 1337–1344.

Robinson, DE, Jeffress, LA. 1963. Effect of varying the interaural noise correlation on the delectability of tonal signals. *J Acoust Soc Am* 35, 1947–1952.

Roth, RL, Kochhar, RK, Hind, JE. 1980. Interaural time differences: Implications regarding the neurophysiology of sound localization. *J Acoust Soc Am* 68, 1643–1651.

Rowland, RC, Tobias, JV. 1967. Interaural intensity difference limens. *J Speech Hear Res* 10, 745–756.

Saberi, K, Perrott, DR. 1990a. Minimum audible movement angles as a function of sound source trajectory. *J Acoust Soc Am* 88, 2639–2644.

Saberi, K, Perrott, DR. 1990b. Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. *J Acoust Soc Am* 87, 1732–1737.

Sandel, TT, Teas, DC, Feddersen, WE, Jeffress, LA. 1955. Localization of sound from single and paired sources. *J Acoust Soc Am* 27, 842–852.

Sayers, BMcA, Cherry, EC. 1957. Mechanism of binaural fusion in the hearing of speech. *J Acoust Soc Am* 29, 973–987.

Scharf, B, Fishken, D. 1970. Binaural summation of loudness: Reconsidered. *J Exp Psychol* 86, 374–379.

Schenkel, KD. 1964. Über die Abhängigkeit der Mithörschwellen von der interauralen Phasenlage des Testchalls. *Acustica* 14, 337–346.

Schubert, ED, Schultz, MC. 1962. Some aspects of binaural signal selection. *J Acoust Soc Am* 34, 844–849.

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.

Shaw, EAG. 1997. Acoustical features of the external ear. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. Hillsdale, NJ: Erlbaum, 25–47.

Shaw, EAG, Vaillancourt, MM. 1985. Transformation of sound–pressure level from the free field to the eardrum presented in numerical form. *J Acoust Soc Am* 78, 1120–1123.

Shaw, WA, Newman, EB, Hirsh, IJ. 1947. The difference between monaural and binaural thresholds. *J Exp Psycho* 37, 229–242.

Shinn-Cunningham, BG, Zurek, PM, Durlach, NI. 1993. Adjustment and discrimination measurements of the precedence effect. *J Acoust Soc Am* 93, 2923–2932.

Shower, EG, Biddulph, R. 1931. Differential pitch sensitivity of the ear. *J Acoust Soc Am* 3, 275–287.

Sivian, LJ, White, SD. 1933. On minimal audible fields. *J Acoust Soc Am* 4, 288–321.

Sondhi, MM, Guttman, N. 1966. Width of the spectrum effective in the binaural release of masking. *J Acoust Soc Am* 40, 600–606.

Stellmack, MA, Dye, RH. 1993. The combination of interaural information across frequencies: The effects of number and spacing of components, onset asynchrony, and harmonicity. *J Acoust Soc Am* 93, 2933–2947.

Stellmack, MA, Viemeister, NF, Byrne, AJ. 2004. Monaural and interaural intensity discrimination: Level effects and the " binaural advantage." *J Acoust Soc Am* 116, 1149–1159.

Stevens, SS, Newman, EB. 1936. The localization of actual sources of sound. *Am J Psychol* 48, 297–306.

Thurlow, WR, Runge, PS. 1967. Effect of induced head movements on localization of direction of sounds. *J Acoust Soc Am* 42, 480–488.

Tobias, JV. 1963. Application of a "relative" procedure to the binaural-beat problem. *J Acoust Soc Am* 35, 1442–1447.

Tollin, DJ, Henning, GB. 1998. Some aspects of the lateralization of echoed sound in man. I: Classical interaural delay-based precedence. *J Acoust Soc Am* 104, 3030–3038.

Townsend, TH, Goldstein, DP. 1972. Suprathreshold binaural unmasking. *J Acoust Soc Am* 51, 621–624.

Trahiotis, C, Bernstein, LR. 1990. Detectability of interaural delays over select spectral regions: Effects of flanking noise. *J Acoust Soc Am* 87, 810–813.

Wagenaars, WM. 1990. Localization of sound in a room with reflecting walls. *J Audio Eng Soc* 38, 99–110.
Wallach, H. 1939. On sound localization. *J Acoust Soc Am* 10, 270–274.
Wallach, H. 1940. The role of head movements and vestibular and visual cues in sound localization. *J Exp Psychol* 27, 339–368.
Wallach, H, Newman, EB, Rosenzweig, MR. 1949. The precedence effect in sound localization. *Am J Psychol* 62, 315–336.
Webster, FA. 1951. The influence of interaural phase on masked thresholds. I. The role of interaural time-duration. *J Acoust Soc Am* 23, 452–462.
Weinrich, S. 1982. The problem of front-back localization in binaural hearing. *Scand Audiol Suppl* 15, 135–145.
Wenzel, EM, Arruda, M, Kistler, DJ, Wightman, FL. 1993. Localization using nonindividualized head-related transfer functions. *J Acoust Soc Am* 94, 111–123.
Wernick, JS, Starr, A. 1966. Electrophysiological correlates of binaural beats in superior-olivary complex of cat. *J Acoust Soc Am* 40, 1276.
Whitworth, RH, Jeffress, LA. 1961. Time vs. intensity on the localization of tones. *J Acoust Soc Am* 33, 925–929.
Wier, CC, Jesteadt, W, Green, DM. 1977. Frequency discrimination as a function of frequency and sensation level. *J Acoust Soc Am* 61, 178–184.
Wightman, FL. 1971. Detection of binaural tones as a function of masker bandwidth. *J Acoust Soc Am* 50, 623–636.
Wightman, FL, Kistler, DJ. 1989a. Headphone simulation of free-field listening. I: Stimulus synthesis. *J Acoust Soc Am* 85, 858–867.
Wightman, FL, Kistler, DJ. 1989b. Headphone simulation of free-field listening. II: Psychophysical validation. *J Acoust Soc Am* 85, 868–878.
Wightman, FL, Kistler, DJ. 1992. The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91, 1648–1661.
Wightman, FL, Kistler, DJ. 1993. Sound localization. In: WA Yost, AN Popper, RR Fay (eds.), *Human Psychoacoustics*. New York, NY: Springer-Verlag, 155–192.
Wightman, FL, Kistler, DJ. 1997. Factors affecting the relative salience of sound localization cues. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. Hillsdale, NJ: Erlbaum, 1–23.
Wightman, FL, Kistler, DJ. 1999. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am* 105, 2841–2853.
Woodworth, RS. 1938. *Experimental Psychology*. New York, NY: Holt, Rhinehart, and Winston.
Wu, X, Wang, C, Chen, J, Qu, H, Li, W, Wu, Y, Schneider, BA., Li, L. 2005. The effect of perceived spatial separation on informational masking of Chinese speech. *Hear Res* 199, 1–10.
Yang, X, Grantham, DW. 1997. Echo suppression and discrimination suppression aspects of the precedence effect. *Percept Psychophys* 59, 1108–1117.
Yost, WA. 1974. Discriminations of interaural phase differences. *J Acoust Soc Am* 55, 1299–1303.
Yost, WA. 1981. Lateral position of sinusoids presented with interaural intensive and temporal differences. *J Acoust Soc Am* 70, 397–409.
Yost, WA. 1993 Overview: Psychoacoustics. In: WA Yost, AN Popper, RR Fay (eds.), *Human Psychophysics*. New York, NY: Springer-Verlag, 1–12.
Yost, WA. 1997. The cocktail party problem: Forty years later. In: R Gilkey, T Anderson (eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*. Hillsdale, NJ: Erlbaum, 329–347.
Yost, WA, Dye, RH Jr. 1988. Discrimination of interaural differences of level as a function of frequency. *J Acoust Soc Am* 83, 1846–1851.
Yost, WA, Sheft, S. 1993. Auditory perception. In: WA Yost, AN Popper, RR Fay (eds.), *Human Psychophysics*. New York, NY: Springer-Verlag, 193–236.
Yost, WA, Wightman, FL, Green, DM. 1971. Lateralization of filtered clicks. *J Acoust Soc Am* 50, 1526–1531.
Zahorik, P. 1996. Auditory distance perception: A literature review. Wisconsin: University of Wisconsin-Madison. http://www.waisman.wisc.edu/∼zahorik/papers/dist.html.
Zahorik, P. 2002a. Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111, 1832–1846.
Zahorik, P. 2002b. Direct-to-reverberant energy ratio sensitivity. *J Acoust Soc Am* 112, 2110–2117.
Zurek, PM. 1980. The precedence effect and its possible role in the avoidance of interaural ambiguities. *J Acoust Soc Am* 67, 952–964.
Zwislocki, J, Feldman, RS. 1956. Just noticeable differences in dichotic phase. *J Acoust Soc Am* 28, 860–864.