Tori of confusion: Binaural localization cues for sources within reach of a listener

Barbara G. Shinn-Cunningham[a)]
Departments of Cognitive and Neural Systems and Biomedical Engineering, Hearing Research Center, Boston University, Boston, Massachusetts 02215

Scott Santarelli and Norbert Kopco
Department of Cognitive and Neural Systems, Hearing Research Center, Boston University, Boston, Massachusetts 02215

(Received 19 August 1999; revised 5 November 1999; accepted 7 December 1999)

To a first-order approximation, binaural localization cues are ambiguous: many source locations give rise to nearly the same interaural differences. For sources more than a meter away, binaural localization cues are approximately equal for any source on a cone centered on the interaural axis (i.e., the well-known ''cone of confusion''). The current paper analyzes simple geometric approximations of a head to gain insight into localization performance for nearby sources. If the head is treated as a rigid, perfect sphere, interaural intensity differences (IIDs) can be broken down into two main components. One component depends on the head shadow and is constant along the cone of confusion (and covaries with the interaural time difference, or ITD). The other component depends only on the relative path lengths from the source to the two ears and is roughly constant for a sphere centered on the interaural axis. This second factor is large enough to be perceptible only when sources are within one or two meters of the listener. Results are not dramatically different if one assumes that the ears are separated by 160 deg along the surface of the sphere (rather than diametrically opposite one another). Thus for nearby sources, binaural information should allow listeners to locate sources within a volume around a circle centered on the interaural axis on a ''torus of confusion.'' The volume of the torus of confusion increases as the source approaches the median plane, degenerating to a volume around the median plane in the limit. © *2000 Acoustical Society of America.* [S0001-4966(00)04803-7]

PACS numbers: 43.66.Qp, 43.66.Pn [DWG]

INTRODUCTION

The most robust, static cues for determining sound source direction in anechoic space are differences between the signals reaching the left and right ears (i.e., the interaural intensity differences, or IIDs; and interaural time differences, or ITDs). Other cues, such as spectral content and overall sound intensity, depend on the ability of the listener to tease apart acoustic attributes that are due to source content from attributes that are due to source position.

A. Cones of confusion: Binaural cues for relatively distant sources

Interaural time differences result mainly from differences in the path length from the source location to the two ears. In the simplest approximation, iso-ITD locations form a hyperbolic surface of rotation symmetrical about the interaural axis (e.g., see von Hornbostel and Wertheimer, 1920; cited in Blauert, 1997, p. 179). For distances more than a meter from the head, these hyperbolic surfaces approximate cones centered on the interaural axis (i.e., the well-known ''cones of confusion''). A better approximation takes into account the effects of a spherical head on the path lengths to the ears (e.g., see Mills, 1972; Molino, 1973); however, even these iso-ITD contours depend only on the angle from source to interaural axis for distant sources. Empirical measurements of ITD as a function of source direction show that these approximations are quite accurate (e.g., see Mills, 1972).

For relatively distant sources, IIDs arise primarily because of acoustic interference of the head (e.g., see Mills, 1972). In particular, for frequencies whose wavelengths are small relative to the dimensions of the head, the ear farther from the source generally receives less energy than the nearer ear. For sources more than a meter away and assuming a simple spherical head model, the IID at a given frequency is roughly constant for all sources at the same angle from the interaural axis. Thus for distant sources, a spherical head model predicts that iso-IID contours fall on the same cones of confusion as iso-ITD contours (although the magnitude of the IID at a particular direction generally varies with frequency).

When human subjects localize sounds, they often make errors in which the perceived location falls near the same cone of confusion as the actual source, but is at the wrong location on the cone of confusion. Such errors can be explained by the fact that to a first-order approximation, binaural cues can only resolve source position to within a cone of confusion and other less robust cues must be used to disambiguate location on a particular cone of confusion. It is noteworthy that in a large percentage of cone-of-confusion er-

[a)]Author to whom correspondence should be addressed; electronic mail: shinn@cns.bu.edn

rors, the perceived source location is near the true location mirrored about a vertical plane passing through the interaural axis, resulting in ''front/back'' confusions (e.g. see Makous and Middlebrooks, 1990; Wenzel *et al.*, 1993; Wightman and Kistler, 1999).

One cue for resolving this confusion is the spectrum of the signal reaching the eardrum, which varies with source position due to the acoustic effects of the head, pinnae, and torso (e.g., see Shaw, 1997). However, despite the fact that the spectrum of the signal at the eardrum also depends on the spectrum of the source signal itself, a number of experiments support the idea that a major cue for resolving cone-of-confusion ambiguities is the spectral content of the signals reaching the eardrums (e.g., see Roffler and Butler, 1968; Butler and Planert, 1976; Butler and Humanski, 1992; Wenzel *et al.*, 1993; Gilkey and Anderson, 1995; Wightman and Kistler, 1997b; Hofman *et al.*, 1998; Kulkarni and Colburn, 1998).

ITD and IID cues are not perfectly constant for sources on the same cone of confusion because the ears are not diametrically opposed to one another, the head is not a perfect sphere, and the head and ears are not perfectly symmetric about the interaural axis (e.g., see Molino, 1973; Searle *et al.*, 1976; Searle *et al.*, 1976a; Middlebrooks *et al.*, 1989; Duda and Martens, 1998). Such asymmetries probably aid in localizing sound sources (e.g., see Searle *et al.*, 1976b). For instance, the pattern of IID across frequencies may help to resolve source location on a cone of confusion for distant sources (e.g., see Middlebrooks *et al.*, 1989; Duda, 1997; Wightman and Kistler, 1997b). However, such cues are not as robust or systematic as other binaural cues; they tend to be extremely complex functions of both frequency and source location (see discussions in Middlebrooks *et al.*, 1989; Wightman and Kistler, 1997a). Recent analysis suggests that torso reflections cause a peak in the IID between 2–5 kHz, with the frequency of the peak varying with angle around the interaural axis (Avendano *et al.*, 1999). This low-frequency IID peak is grossly front–back symmetric, a trait which may explain why front/back reversals are the most common cone-of-confusion errors.

It is clear that binaural cues (particularly IIDs) arising from distant sources actually differ to some degree for different locations on the same cone of confusion. However, the variations in binaural cues on any given cone of confusion tend to be smaller and/or less consistent than the variations across different cones of confusion. Thus analyzing binaural cues for a simplified, symmetrical head model can provide insight into sound localization behavior by describing how gross binaural cues vary with source location.

B. Interaural intensity differences as a cue for source distance

A number of researchers have previously pointed out that for a nearby point source, IIDs vary with source position differently than do ITDs. If one assumes that ITDs convey in which cone of confusion a source is contained (an angle that will be called the ''angle between the source cone and the interaural axis'' throughout this paper), then IIDs can be used to determine source distance (e.g., see Hartley and Fry, 1921; Firestone, 1930; Wightman and Firestone, 1930; Coleman, 1963; Hirsch, 1968; Molino, 1973; Brungart and Rabinowitz, 1996; Duda and Martens, 1998; Brungart and Durlach, 1999; Brungart and Rabinowitz, 1999). Due to limitations in available computational power, most early studies of the binaural differences arising from a point source near the listener were limited to a very restricted set of positions and frequencies (e.g., Hartley and Fry, 1921; Firestone, 1930; Wightman and Firestone, 1930; Hirsch, 1968; Molino, 1973). While these papers point out that IID cues can disambiguate source location once the ITD is known, they do not explicitly show how reliable IID spatial cues are or how these cues change with source location.

Recently, Duda and Martens (1998) and Brungart and Rabinowitz (1999) computed how the signals reaching a rigid spherical head vary with distance and direction of a point source. Both groups discuss the fact that IIDs provide distance information for nearby sources. Brungart and Rabinowitz (1999) further point out that the IID grows as the source moves lateral to the head, as well as increasing with frequency and decreasing with distance. In analyzing behavioral localization data for nearby sources, Brungart and Durlach (1999) show that the ability to judge source distance increases as source azimuth increases, consistent with the idea that the IID magnitude generally increases with source laterality. In a preliminary model of localization for nearby sources, Brungart (1998) assumes that IID decreases exponentially with source distance and that the exponent power increases with source azimuth. By taking into account the perceptual sensitivity to ITD and IID information, this model was able to predict an observed improvement in distance perception with increasing source azimuth for sources restricted to within 30 deg of the horizontal plane.

These studies demonstrate that IIDs provide unique information about source location for nearby sources, but that the amount of spatial information gained from the IID depends on the spatial position of the source. The current analysis shows quantitatively how IID depends on the angle between source cone and interaural axis and the distance of the source from the head. By calculating surfaces for which binaural cues are constant (and thereby specifying which source positions cannot be disambiguated on the basis of binaural cues), insight can be gained into sound source localization for nearby sources.

The analyses below examine how binaural cues vary with source location; however, they exclude any consideration of the acoustic effects of the shoulders, torso, or pinnae. Thus, while the current analyses may help to explain some aspects of sound localization for nearby sources, they cannot explain perceptual results that depend on these acoustic effects.

I. INTERAURAL DIFFERENCES FOR AN ACOUSTICALLY TRANSPARENT HEAD

For a spherical head, interaural differences will vary with frequency. However, if one treats the head as acoustically transparent, interaural differences depend only on source position relative to the two ears and are independent of frequency. This section considers this simplified case in order to gain insight into more realistic approximations (treated in later sections).

A. Interaural time differences

Ignoring the acoustic effects of the head itself, the ITD τ depends only on Δ, the difference in the path lengths to the two ears. Specifically,

$$\tau=\frac{\Delta}{c}, \tag{1}$$

where c is the speed of sound (343 m/s). By definition, an iso-ITD surface for point receivers in free space is the locus of positions at which Δ is constant. Assuming that two ears are located in a rectilinear coordinate system at $(r,0,0)$ and $(-r,0,0)$ (which puts the center of a head of radius r at the origin), this surface is given by

$$\frac{4}{\Delta^2}x^2-\frac{4}{4r^2-\Delta^2}(y^2+z^2)=1. \tag{2}$$

These iso-ITD surfaces form the ''traditional'' cones of confusion (e.g., see Blauert, 1997).

The just-noticeable difference (JND) for ITD is roughly 10–20 μs for a reference with 0 ITD and increases by roughly a factor of 2–3 for larger reference ITDs (e.g., see Durlach and Colburn, 1978). Of course, these JNDs measure the best performance that can be achieved in a simple discrimination experiment when there is no stimulus uncertainty (e.g., see Braida and Durlach, 1988). In fact, the ability to extract ITD information will generally be worse if the subject must attend to a large range of stimuli (e.g., see Koehnke and Durlach, 1989; Shinn-Cunningham *et al.*, 1998). In order to gain insight into the spatial information conveyed by the ITD, one can compute the iso-ITD surfaces that should lead to detectable changes in ITD as a function of spatial location. The left side of Fig. 1 shows ITD contours spaced at 50-μs increments for an arbitrary plane containing the interaural axis.

One can see from the symmetry of Eq. (2) that the iso-ITD contours are three-dimensional surfaces formed by rotating the depicted one-dimensional contours around the interaural axis. ITD information should allow listeners to determine a sound source location to within the volume delineated by an adjacent pair of iso-ITD surfaces. The gray area on the left of the graph shows the cross section of such a volume taken through an arbitrary plane containing the interaural axis for a source at the position labeled ''○.'' Of course, if the head were not acoustically transparent, the path from source to ear would not always be direct (i.e., sound would have to travel around the head) and the sound wave would be reflected and diffracted by the head. As a result, the acoustically transparent head analysis generally underestimates ITDs.

B. Interaural intensity differences

The IIDs that occur for sound sources very close to the listener help to disambiguate source positions that result in

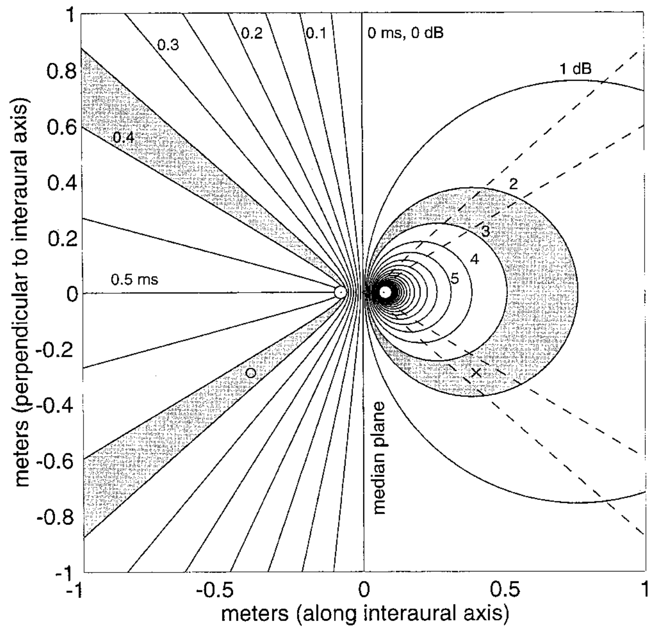

FIG. 1. Iso-ITD and iso-IID contours as a function of spatial location for an acoustically transparent head. The left side of the figure shows iso-ITD contours spaced every 50 μs. The right side of the figure shows iso-IID contours spaced every 1 dB. By symmetry, these contours are identical for any arbitrary plane containing the two ears (small circles). All spatial units are in meters. The abscissa is parallel to the interaural axis and the ordinate perpendicular to the interaural axis. The gray filled areas show the regions of space that are consistent with the ITD for a source at the position marked by ''○'' and the IID for a source at the position marked by ''×.'' The dashed line repeats the iso-ITD contours that delineate the ''cone of confusion'' for the source at ''×.'' The intersection of the area enclosed by these dashed lines and the filled area on the right are the only locations consistent with both ITD and IID cues for a source at ''×.''

nearly identical ITD values. Like ITDs, IIDs restrict source location to a two-dimensional surface. For sources near the head, iso-IID surfaces differ from iso-ITD surfaces and thus provide unique information about source position.

One major component of the IID for nearby sources arises due to differences between the path lengths to the two ears. The energy transmitted to a point in space from a uniformly radiating point source is inversely proportional to the square of the distance from the source. Assume that the ''left'' and ''right'' point receivers are located at distances of d_L and d_R, respectively, from a uniformly radiating point source. Then the square of the ratio d_L/d_R equals the ratio of the intensity reaching the right receiver over the intensity reaching the left receiver. The resulting IID will be constant for all positions at which the ratio $k=d_L/d_R$ is constant. The IID α (in dB) is given by

$$\alpha=20\log_{10}k. \tag{3}$$

Again assuming that two ears are located at $(r,0,0)$ and $(-r,0,0)$, the iso-IID surfaces are given by

$$(x-x_\alpha)^2+y^2+z^2=d_\alpha^2,$$

$$x_\alpha=-r\frac{(1+k^2)}{(1-k^2)}, \tag{4}$$

$$d_\alpha = \left| \frac{2rk}{1-k^2} \right|.$$

These iso-IID surfaces constitute perfect spheres of radius d_α, whose centers fall on the interaural axis at $(x_\alpha,0,0)$. Both the distance from the center of the sphere to the nearer ear and the radius of the sphere increase with decreasing IID magnitude (as k approaches one). As the IID approaches zero, the magnitudes of both d_α and x_α grow to infinity and the iso-IID sphere degenerates to the entire median plane. The iso-IID sphere degenerates to a point at the position of the nearer ear as the IID magnitude increases (k approaches zero or infinity).

The just-noticeable difference (JND) in IID is approximately 0.8 dB, independent of frequency and reference IID (e.g., see Mills, 1960; Hershkowitz and Durlach, 1969; Mills, 1972). The right side of Fig. 1 shows iso-IID contours at 1-dB separations for source positions on an arbitrary plane containing the interaural axis. As with iso-ITD curves, rotating these iso-IID curves around the interaural axis generates iso-IID surfaces in 3-space. In other words, gross IID information alone should allow subjects to determine source location to a volume of space whose bounding surfaces are iso-IID spheres separated by one JND. The shaded gray area on the right side of Fig. 1 shows the cross section of such a volume (through a plane containing the interaural axis) for a source at location ''×.''

C. Tori of confusion

The goal of this analysis is to estimate how well subjects can judge source position based only on robust, binaural cues. It is generally accepted that ITD and IID are separately computed in individual frequency channels. Both types of binaural information have a limited resolution; however, both are available to help determine source location. A listener should be able to determine source location to within the intersection of the volumes separately determined by IID and ITD information. In other words, based on binaural cues, a subject should be able to determine source location to within a volume whose four bounding surfaces are the two iso-IID spheres and two iso-ITD cones described above.

For the source at location ''×'' in Fig. 1, the listener should be able to judge the location of the source as somewhere within the gray area on the right half of the figure based on IID information alone. ITD information constrains the source to be between the dashed lines on the right side of the figure. In the horizontal plane, the intersection of these constraints forms two roughly square regions positioned symmetrically about the interaural axis. Rotating these areas around the interaural axis defines the locus of positions for which the binaural cues are consistent with those from a source at position ''×.''

The extent of the resulting volume of space varies dramatically with source position. For source positions that are near the head and on the interaural axis, IID cues alone will restrict the source position to a relatively small region of space. In contrast, both iso-IID and iso-ITD surfaces degenerate to the same surface for a source on the median plane so that gross binaural cues only determine source position to a broad swatch of space within about 5 deg of the median plane. For intermediate locations, iso-IID and iso-ITD surfaces are nearly perpendicular to one another and the combination of cues provides much more information than either cue taken alone. In these cases, the intersection of the IID and ITD volumes is a torus-shaped volume.[1] We call these volumes ''tori of confusion[2]'' (after the ''cones of confusion'') to reflect the additional constraints on source location derived from the IID for a nearby source. If the source is more than 2 meters away, the change in IID with source position is too gradual to provide spatial information (at least for an acoustically transparent head), and the source can only be localized to a volume around the correct cone of confusion.

This analysis demonstrates why one should not think of IID as providing distance information about nearby sources, *per se*. Instead, IID cues provide information about the location of the source in both distance and direction. In addition, distance perception (based on binaural information) does not depend on source azimuth, but rather on the angle between the source cone and the interaural axis. In the next section, these transparent-head iso-IID and iso-ITD surfaces are compared to those derived for a simple rigid, spherical head model.

II. INTERAURAL DIFFERENCES FOR A RIGID, SPHERICAL HEAD

Treating the head as rigid sphere, Rabinowitz and his colleagues (Rabinowitz *et al.*, 1993) derived how the pressure on the surface of a rigid sphere varies for a point source at an arbitrary location. This analysis has previously been applied to the problem of determining how sound pressure varies with source position for nearby sources (e.g., see Duda and Martens, 1998; Brungart and Rabinowitz, 1999). The current analysis focuses on how much spatial information binaural cues convey as a function of source position and how these results differ (and are similar to) the simple analysis given in the previous section.

The current computations also use the spherical head model presented by Rabinowitz *et al.* (1993). The point receivers (ears) are assumed to be at opposite ends of one diameter of a rigid sphere of radius 8 cm. From these assumptions, interaural differences depend only on source distance and cone-of-confusion angle (solutions will be symmetrical about the interaural axis). The complex pressure on the surface of the sphere was calculated for a uniformly radiating point source at all directions, distances between 12 cm and 10 m, and frequencies ranging from 20 to 20 000 Hz. The ratio of the RMS pressures on opposite sides of the sphere were then computed in order to determine interaural differences in intensity and time as a function of sound frequency and source position. These results were then smoothed in frequency using a 1/3-octave-wide kernel.

A. Interaural time differences

ITD was estimated from the phase of the ratio of the complex transfer functions for the right and left ears. The interaural phase difference was nearly linear (after unwrap-

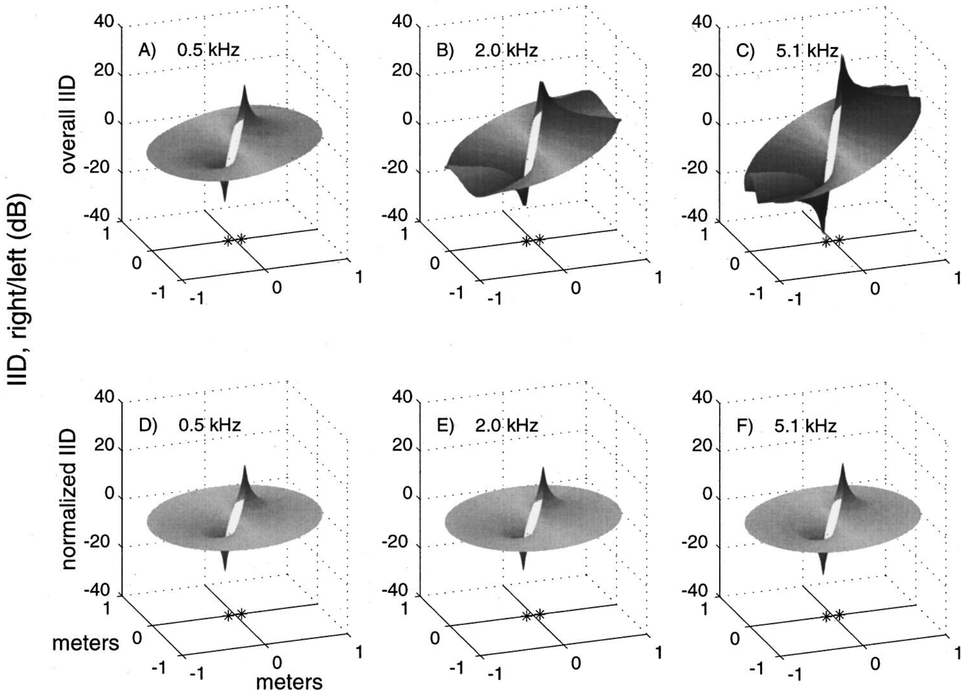

FIG. 2. IID as a function of spatial location in an arbitrary plane through the interaural axis using a spherical head model with diametrically opposed ears. The interaural intensity difference (right ear relative to left ear, in decibels) is plotted in the vertical dimension. The (x,y) position corresponds to the location of the source relative to a head centered at (0,0) of radius 8 cm. The locations of the ears are shown by asterisks in the x-y plane at the bottom of the figure. The median plane is located at $x=0$. The top row (panels A, B, and C) shows the overall IID for three different frequencies (500, 2000, and 5100 Hz). The bottom row (panels D, E, and F) shows the normalized IID (total IID less the IID that arises for a source infinitely far, i.e., 10 m, from the sphere) for the same three frequencies.

ping the phase) up to 6 kHz for all distances and angles. Since ITDs are generally assumed to be relatively unimportant above 2 kHz, a single ITD value was estimated for each source location by averaging the slope of the interaural phase versus frequency function for frequencies below 2 kHz.

At any particular location, the expected ITD using a spherical head model is larger than that predicted in the earlier analysis, as expected (i.e., the iso-ITD surfaces are more closely spaced). This difference in ITD magnitude arises primarily because the path length to the far car is longer for a rigid sphere compared to an ''invisible'' head, increasing the ITD for a given source location. Although the ITD increases slightly as the source approaches the ear along a particular cone of confusion, the resulting distance information is negligible for most source positions. As a result, the spatial information conveyed by the ITD for a rigid head follows roughly the same geometry as predicted by the analysis in the previous section; namely, the ITD determines source location to near a particular cone of confusion. In general, the volumes of the cones of confusion are smaller than for the case of an acoustically transparent head.

B. Interaural intensity differences

The top half (panels A, B, and C) of Fig. 2 plots the IID as a function of source position for an arbitrary plane through the interaural axis. In general, the magnitude of the IID increases as the source approaches either ear (consistent with the analysis for an acoustically transparent head). In addition, the IID becomes more complex and grows in magnitude as frequency increases. At low frequencies (500 Hz and below, see panel A), the IID function is shaped nearly identically to the IID function for an acoustically transparent head. However, even for low frequencies, the IID magnitude at a particular location is larger when the head is treated as a rigid sphere rather than acoustically transparent. As frequency increases, the IID begins to show a directional dependence, with the IID increasing as the source position moves away from the median plane (e.g., see panels B and C). The IID actually decreases for mid and high frequencies as the source position approaches the interaural axis. This results from the fact that the sound waves traveling around the head add in phase at the far ear as the source approaches the interaural axis, thereby increasing the intensity of the sound at the far ear and decreasing the IID (the so-called ''acoustic bright spot''). In the limit, for sources relatively far from the head, the iso-IID surfaces must approach the cones of confusion. For these distances, the IID at a given frequency depends only on the angle from the source to the interaural axis.

In order to gain further insight into the dependence of the IID on source position, the IID for a source very far from the head (i.e., for a source at 10 m) was calculated for all possible cone-of-confusion angles. For each source position, the IID that occurs for a source at 10 meters in that direction was subtracted from the overall IID to form ''normalized IIDs'' (i.e., the IID was normalized by subtracting the IID for a source from the same direction, but at an effectively infinite distance). The normalized IIDs are plotted in the bottom half of Fig. 2 (panels D, E, and F) for the same frequencies shown in the top half of the figure.

At low frequencies, IID is approximately zero for sources far from the head and the normalization has little effect (panels A and D). For higher frequencies, the IID for a distant source accounts for a large percentage of the overall IID. Once this distant-source IID is removed, the only positional dependence in the IID is nearly identical to the form predicted by the analysis for an acoustically transparent head, independent of frequency (compare panels B and C with E and F). In other words, while the magnitude of the normalized IID is slightly larger than the IID predicted for an acoustically transparent head, the shape is virtually identical.

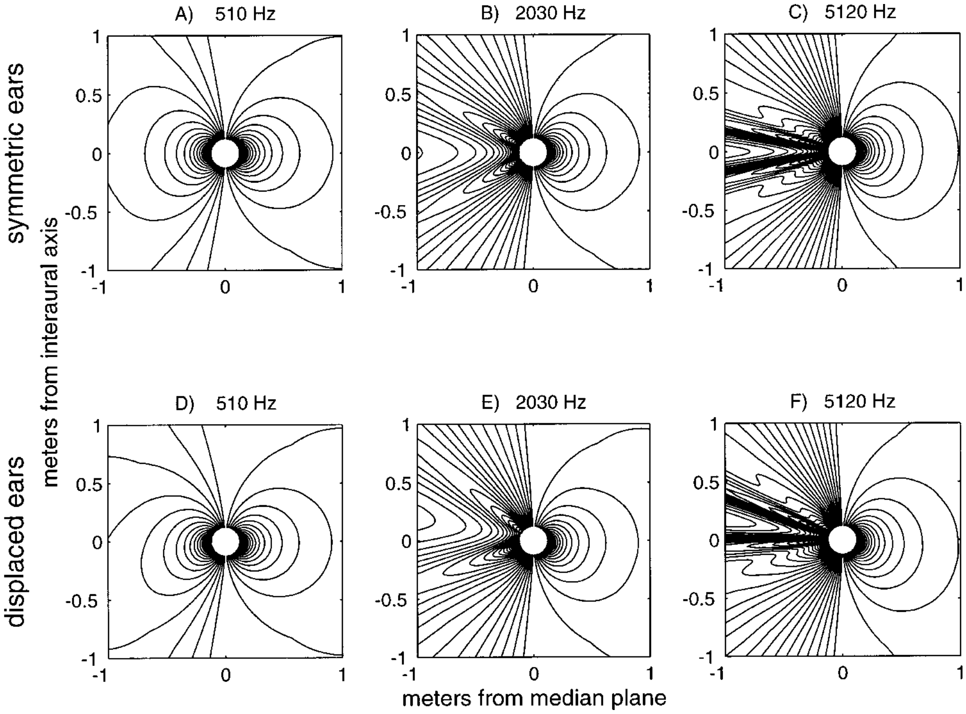

FIG. 3. Iso-IID contours as a function of spatial location in a plane through the interaural axis using a spherical head model. The spherical head is centered at (0,0) and the ears are located at ($\pm$0.08,0) m. The left half of each panel shows the overall iso-IID contours. The right half of each panel shows the iso-IID contours that remain after the IID that arises for a source infinitely far (i.e., 10 m) from the head is subtracted. Contours are shown for 1-dB increments, starting at 1 dB (contour nearest to the median plane). The top row shows the IID for three different frequencies (500, 2000, and 5100 Hz) assuming that the ears are diametrically opposed. In these panels (A, B, and C), the results are valid for an arbitrary plane containing the interaural axis. The bottom row shows the IID for the same three frequencies when the ears are displaced 10 deg behind the center of the head in the horizontal plane. In these panels (D, E, and F), the results are valid for the horizontal plane.

Iso-IID contours derived from the plots in Fig. 2 are shown in the top row (panels A, B, and C) of Fig. 3 (the bottom row is discussed in the next section). The left side of each panel shows the full iso-IID contours and the right side of each panel shows the corresponding normalized iso-IID contours. For low frequencies (e.g., see panel A), iso-IID contours are grossly similar to the iso-IID contours for an acoustically transparent head. At intermediate frequencies (e.g., panels B and C), the iso-IID contours become complex, varying both with the angle from the interaural axis and the relative distance from the source to the two ears. Once the head-shadow IID component (i.e., the component present for sources infinitely far from the head) is removed from the IID surfaces (right half of each panel), the remaining iso-IID contours depend mainly on the relative distance from the source to the two ears, like iso-IID contours for an acoustically transparent head. The main distinction between the two cases is that the IIDs for an acoustically transparent head are slightly smaller in magnitude than those that arise for a spherical head. As a result, the volumes of spatial uncertainty delineated by the iso-IID contours for the normalized IIDs are slightly smaller than from those predicted by an acoustically transparent head.

This analysis shows that the IIDs that arise for a rigid, spherical head model can be broken down into two components. The first component is frequency-dependent but distance-independent. As expected, the magnitude of this distant-source component increases dramatically with frequency. For frequencies below about 500 Hz, this factor is negligible; at high frequencies, this ''head shadow'' dominates the overall IID. The second component, the ''normalized IID,'' varies with the relative distance from the source to the two ears and conveys roughly the same spatial information predicted for an acoustically transparent head. While this relative-distance component increases with frequency, the frequency dependence is not pronounced.

C. Tori of confusion

ITD information for a rigid spherical head conveys information about the angle between the source cone and the interaural axis. Although the ITD increases slightly as the source approaches the head, these deviations from the perfect ''cones of confusion'' are small in perceptual terms (as pointed out previously; e.g., see Brungart, 1998; Brungart and Rabinowitz, 1999).

Spatial information conveyed by the IIDs varies dramatically with frequency. At low frequencies (below 500 Hz), the IID information is essentially the same as predicted for an acoustically transparent head. At high frequencies, the IID primarily conveys information about source direction, but also conveys some information about the relative distances from the source to the two ears. At intermediate frequencies, the IID varies with angle from the interaural axis and with relative distance to the two ears.

In other words, at both intermediate and high frequencies, part of the information conveyed in the IID covaries with the information conveyed in the ITD. Combining ITD and IID information will restrict the possible source position to a torus of confusion. However, if the source is broadband, combining spatial information in the IIDs in different frequency bands will restrict the source location to the same torus of confusion, since mid- and high-frequency IIDs contain spatial information similar to the information conveyed by ITD.

It must be pointed out that the spherical-head approximation becomes increasingly less accurate as frequency increases. In particular, above about 6 kHz, spectral notches and peaks that depend on the angle around the interaural axis (the angle around the torus of confusion) will begin to arise in the signals at the ears due to the acoustic effects of the pinnae. These notches and peaks will cause changes not only in the energy pattern at the individual ear drums, but will result in IIDs since the effects are different at the left and right ears. Also, for frequencies between 2 and 5 kHz, torso reflections may affect the IID (Algazi *et al.*, 1999). Thus the IID analysis for a rigid spherical head model is most useful for relatively low frequencies.

Overall, this analysis shows that when sources are within a meter of the listener, low- and mid-frequency IID information should allow listeners to localize a source to within a torus of confusion. This IID information is further refined by the ITD cues, which partially covary with IIDs at mid and high frequencies. Many of the observations made for the acoustically transparent head analysis continue to hold. For instance, the volume of a torus of confusion increases as the angle between the source cone and the interaural axis increases until it degenerates to the entire median plane. The toroidal volume decreases as the angle between the source cone and the interaural axis decreases and as the source moves closer to the nearer ear. For sources beyond 2 meters from the listener, IID changes so gradually with distance that it conveys no useful spatial information beyond that contained in the ITD and the torus of confusion degenerates into a cone of confusion.

III. INTERAURAL DIFFERENCES FOR DISPLACED EARS

Human ears are not diametrically opposed on the head. For instance, in their analysis of the range dependence of the HRTF, Duda and Martens (1998) assumed that the ears were located 10 deg behind the diameter parallel to the interaural axis. This asymmetry is relatively small, but complicates the geometry of iso-binaural surfaces. In particular, the rotational symmetry that is assumed in the above discussion no longer holds; instead of symmetrical tori of confusion, binaural cues will allow the listener to locate the source to within some skewed volume. In this section, we examine the effect of displacing the ears backward on the head.

The same rigid spherical head model was used to calculate the transfer function from a point source in space to point receivers on the surface of a rigid sphere (Rabinowitz *et al.*, 1993). However, the point receivers were assumed to be angularly displaced backward by 10 deg in the horizontal plane. Of course, for a head in which the ears are angularly displaced, different planes containing the x-axis yield slightly different iso-IID and iso-ITD contours (i.e., binaural differences are no longer constant on a cone of confusion). While full rotational symmetry no longer holds, binaural cues are mirror symmetric about the horizontal plane [i.e., the binaural differences that occur for a source at (x,y,z) are equal to those for a source at $(x,y,-z)$]. Such symmetry is consistent with up/down and down/up reversals that are occasionally reported in the literature (e.g., see Wenzel *et al.*, 1993). The only plane containing the x-axis and the two ears is the horizontal plane $(z=0)$; results in this plane show the greatest (front–back) asymmetry. As a result, the distortion of the iso-binaural contours from those seen in the analysis for the symmetrically placed ears will be greatest through the horizontal plane.

A. Interaural time differences

Iso-ITD contours that arise for a spherical head with the ears displaced are nearly indistinguishable from the iso-ITD contours for a head with diametrically opposed ears. The biggest difference is that the ITD is slightly smaller at positions near an azimuth of 90 deg when the ears are displaced. This is explained by the observation that if a source moves from 78 to 79 deg in the horizontal plane, it moves away from the left (far) ear and toward the right (near) ear, causing an increase in the ITD. However, a source at 80 deg azimuth is directly opposite the left ear, so as a source moves from azimuth 80 to 81 deg, it actually moves closer to the far ear (around the back of the head) as well as moving closer to the near ear. As a result, for sources located at azimuths between 80 and 100 deg, the difference in the path lengths to the two ears does not change significantly with azimuth and the magnitude of the ITD is reduced for locations to the side of the head. Thus the main effect of positioning the ears behind the center of the head is to decrease ITD magnitude.

B. Interaural intensity differences

The bottom row (panels D, E, and F) of Fig. 3 plots overall- and normalized-iso-IID contours for locations in the horizontal plane when the ears are displaced. It should be emphasized that nonhorizontal planes containing the x-axis would yield different iso-IID contours. In particular, since only the horizontal plane bisecting the head passes through the ears, the iso-IID contours for this plane show the greatest asymmetry (and reach the largest values).

Looking first at the iso-IID contours for the overall IID (left half of each panel), the IID function is no longer front–back symmetric (compare top and bottom rows in Fig. 3). There are two factors influencing the IID as a function of location in the horizontal plane. The first can be ascribed to differences in the ratio of the path lengths to the two ears. This factor is maximal for sources located 100 deg to the side of the median plane for sources in the horizontal plane containing the ears and increases as the source approaches the head. The second factor depends on the head shadow and varies with the direction between source and head as well as with frequency. The relative importance of this second factor increases with frequency.

The maxima of the normalized iso-IID functions (right half of each panel) are skewed toward an azimuthal angle of 100 deg (in line with the near ear). Nonetheless, this spatial dependence is similar to the spatial dependence exhibited for the symmetrical case. The IID information consists of a component that varies with azimuthal angle and a component that varies with the relative path length from source to the two ears.

C. Distorted tori of confusion

Because of the asymmetry for angularly displaced ears, the combination of IID/ITD information taken across fre-

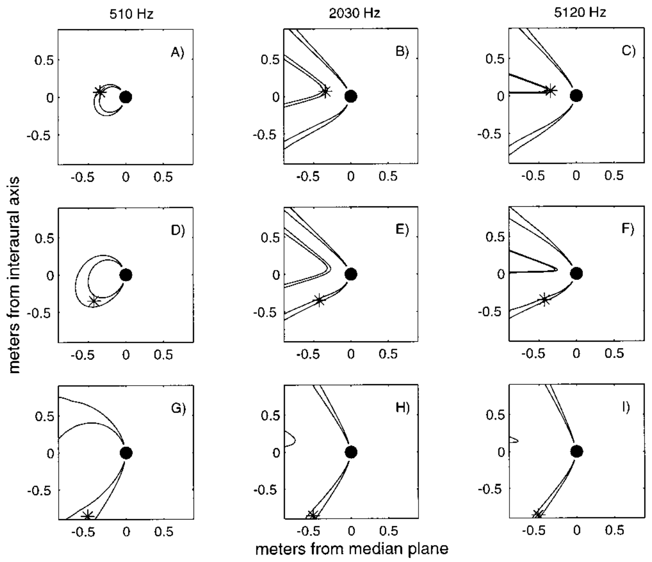

FIG. 4. Constraints on spatial location due to the IID for a spherical head model in which the ears are angularly displaced. Each row corresponds to one source location (plotted as an asterisk in each panel). The solid circle in the middle of the panel shows the head. Each column corresponds to a different frequency. In each plot, the source is constrained to fall within the area delineated by the drawn contours by the IID in the corresponding frequency range.

quency no longer forms a symmetrical torus of confusion. ITD information constrains the source to be on a distorted cone of confusion. In order to gain insight into how the IID would constrain locations for sources at various locations in the horizontal plane, Fig. 4 shows the locations in the horizontal plane that are consistent with the IID in different frequencies (columns) for a source in various locations (rows).

Figure 4 shows that the tori of confusion are skewed by the angular displacement of the ears, but many of the observations made for the symmetrical model still hold. In particular, IID information will constrain the source location to within some volume of space. The size of the volume decreases as the source nears an ear and increases as sources approach the median plane. In general, there are locations both in front of and behind the listener that could give rise to the observed IID cues at each frequency, consistent with front/back and back/front reversals. At lower frequencies, the source position is constrained in both distance and direction, similar to predictions from an acoustically transparent head analysis. As frequency increases, iso-IID contours vary more dramatically with source direction than source distance and provide information that is similar to the information in the ITD. For some source positions, the IID in moderate and high frequencies constrains the source to fall within one of two spatial bands (e.g., see panels C and D), roughly corresponding to two different cones of confusion (one of which contains the actual source location). This occurs because the IID is not monotonic with source azimuth (see Fig. 2) so that there can be two or more connected regions of space consistent with the IID in mid and high frequencies. Looking across these spatial constraints, broadband IID information (with or without ITD information) restricts the source location to within a volume of space that forms a distorted torus of confusion.

IV. SUMMARY AND DISCUSSION

These results show that when sources are close to one ear, IIDs vary dramatically with source position and frequency so that using only binaural cues, a broadband sound source can be located to somewhere on a ''torus of confusion.'' ITD cues determine source position to within a cone of confusion. IIDs vary with both distance and direction for sources, but are only significant for sources within one or two meters of the head and close to the interaural axis.

There are two main components of the IID: one that depends primarily on the ratio of the path lengths from the source to the two ears and one that depends on the direction from the source to the center of the head. The relative importance of these two factors depends on frequency, primarily because the magnitude of the second factor increases with frequency.

Of course, the acoustic signals at the ears of a real listener include many effects not considered here. In particular, the effects of the shoulders, torso, and pinnae are known to be acoustically significant and to affect localization judgments. However, the current results demonstrate how gross binaural cues vary with direction and distance. In addition, the current analysis is most accurate at low frequencies, precisely the frequencies at which IIDs do not occur for distant sources.

When one considers the problem of how to estimate source position from the acoustic cues available at the eardrums, the importance of the ''extra-large'' IIDs that can occur for sources very near the listener becomes clear. It has been reported that ITD cues dominate judgments of source direction for broadband sounds (Wightman and Kistler, 1992). In these experiments, IID and spectral shape cues derived from individually measured head-related transfer functions (HRTFs) were pitted against interaural phase information (derived from a different spatial location using the individualized HRTFs). In the study, the measured HRTFs were taken for relatively distant sources, where only the head shadow contributed to the IID. The results show that as long as the source signal contained low frequencies, judgments of source direction were dominated by the ITD cue. In contrast, many headphone experiments in which ITD and IID cues are pitted against one another (e.g., see the review in Durlach and Colburn, 1978, as well as Buell *et al.*, 1994; Buell and Trahiotis, 1997) show that a large, broadband IID favoring one ear biases judgments of source direction toward that ear, even when ITD cues indicate a different direction. From the current analysis, we see that a large, low-frequency IID only occurs for sources near the ear *and* close to the interaural axis. Thus in many dichotic headphone studies, the low-frequency IID in the imposed broadband IID restricts the possible source location to be a very small volume of space close to the ear. A parsimonious explanation for all of these results is that both ITD and IID information are used to determine source position, but that low-frequency ITD information is generally more reliable than head-shadow cues. Only very large, low-frequency IIDs (which can only occur when sources are very near one ear) are sufficiently reliable to overcome the dominance of low-frequency ITD information. In fact, if the auditory system treats ITD and IID information as separate, independent channels of spatial information, a maximum-likelihood estimation approach (e.g., see Duda, 1997), will predict dual source images (as are often reported in the literature) for ITD/IID pairings that are inconsistent.

It should also be pointed out that energy differences in the signals reaching the two ears are important for spatial unmasking in many real-world listening situations. Thus the pattern of spatial unmasking for sources close to the listener may be different than when sources are more than a meter away. If a source is near the interaural axis and close to the head, the ''better ear'' advantage will be more pronounced than when a source is at the same direction but distant from the head. However, the extra-large IIDs that arise in such a situation may actually decrease the binaural component of spatial unmasking (e.g., see Colburn and Durlach, 1965). The question of how these two factors play out under free-field listening conditions is relevant for understanding spatial unmasking in true ''cocktail party'' situations, where talkers may be fairly close to the listener.

Finally, it must be emphasized that above 6 kHz, large IIDs will arise due to the interaction of the pinnae with the sound waves. At intermediate frequencies, torso effects will alter the IID. However, the analysis presented here is most useful for predicting the IID for the very frequencies which are normally assumed to have zero IID for ''naturally occurring,'' free-field sources. This analysis demonstrates how the unique low-frequency IIDs for sources near the listener depend upon spatial position, and how these IIDs may allow listeners to determine where on the cone of confusion a source is located.

ACKNOWLEDGMENTS

This work was supported in part by AFOSR Grant No. F49620-98-1-0108. The authors gratefully acknowledge assistance from Jay Desloge and Douglas Brungart with code for the analysis of the spherical head model. Nat Durlach, Wesley Grantham, Fred Wightman, and an anonymous reviewer provided valuable feedback and criticisms of earlier drafts of this work. A more extensive description of this work is available as Technical Report CAS/CNS-TR-99-026 from the Boston University Department of Cognitive and Neural Systems.

[1]It should be pointed out that the cross-sectional area of the ''torus of confusion'' is not circular, although the word ''torus'' often conjures an image of a solid of rotation with a perfectly circular cross section. The more general definition of torus includes any solid created by rotation of a fixed cross-sectional area about an (in this case interaural) axis. Indeed, the cross-sectional shape of the torus of confusion varies with spatial location of the source. For sources very close to one ear, the cross-sectional area is roughly trapezoidal; for sources near the median plane, the cross section is much more asymmetrical and significantly larger.

[2]We have previously used the more colloquial term ''doughnuts of confusion'' to describe the volume of spatial uncertainty.

Algazi, V. R., Avendano, C., and Duda, R. O. (**1999**). ''Low-frequency ILD elevation cues,'' J. Acoust. Soc. Am. **106**(4), 2237.

Avendano, C., Algazi, V. R., and Duda, R. O. (**1999**). ''A head-and-torso model for low-frequency binaural elevation effects,'' in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY.

Blauert, J. (**1997**). *Spatial Hearing*, 2nd ed. (MIT Press, Cambridge, MA).

Braida, L. D., and Durlach, N. I. (**1988**). ''Peripheral and central factors in intensity perception,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 559–583.

Brungart, D. S. (**1998**). ''Preliminary model of auditory distance perception for nearby sources,'' in *Proceedings of the NATO AS1 on Auditory Computational Hearing*, Il Ciocco, Italy.

Brungart, D. S., and Durlach, N. I. (**1999**). ''Auditory localization of nearby sources II: Localization of a broadband source in the near field,'' J. Acoust. Soc. Am. **106**(4), 1956–1968.

Brungart, D. S., and Rabinowitz, W. M. (**1999**). ''Auditory localization of nearby sources I: Head-related transfer functions,'' J. Acoust. Soc. Am. **106**(3), 1465–1479.

Brungart, D. S., and Rabinowitz, W. R. (**1996**). ''Auditory localization in the near field,'' in *Proceedings of the Third International Conference on Auditory Display*, Palo Alto, CA.

Buell, T. N., and Trahiotis, C. (**1997**). ''Recent experiments concerning the relative potency and interaction of interaural cues,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. Anderson (Erlbaum, New York), pp. 139–150.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (**1994**). ''Lateralization of bands of noise as a function of combinations of interaural intensitive differences, interaural temporal differences, and bandwidth,'' J. Acoust. Soc. Am. **95**(3), 1482–1489.

Butler, R. A., and Humanski, R. A. (**1992**). ''Localization of sound in the vertical plane with and without high-frequency spectral cues,'' Percept. Psychophys. **51**, 182–186.

Butler, R. A., and Planert, N. (**1976**). ''The influence of stimulus bandwidth on localization of sound in space,'' Percept. Psychophys. **19**, 103–108.

Colburn, H. S., and Durlach, N. I. (**1965**). ''Time-intensity relations in binaural unmasking,'' J. Acoust. Soc. Am. **38**(1), 93–103.

Coleman, P. D. (**1963**). ''An analysis of cues to auditory depth perception,'' Psychol. Bull. **60**, 302–315.

Duda, R. O. (**1997**). ''Elevation dependence of the interaural transfer function,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. Anderson (Erlbaum, New York), pp. 49–76.

Duda, R. O., and Martens, W. L. (**1998**). ''Range dependence of the response of a spherical head model,'' J. Acoust. Soc. Am. **104**(5), 3048–3058.

Durlach, N. I., and Colburn, H. S. (**1978**). ''Binaural phenomena,'' in *Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 365–466.

Firestone, F. A. (**1930**). ''The phase difference and amplitude ratio at the ears due to a source of pure tone,'' J. Acoust. Soc. Am. **2**, 260–270.

Gilkey, R. H., and Anderson, T. R. (**1995**). ''The accuracy of absolute localization judgments for speech stimuli,'' J. Vestib. Res. **5**(6), 487–497.

Hartley, R., and Fry, T. (**1921**). ''The binaural location of pure tones,'' Phys. Rev. **18**, 431–442.

Hershkowitz, R. M., and Durlach, N. I. (**1969**). ''Interaural time and amplitude JND's for a 500 Hz tone,'' J. Acoust. Soc. Am. **46**, 1464–1467.

Hirsch, H. R. (**1968**). ''Perception of the range of a sound source of unknown strength,'' J. Acoust. Soc. Am. **43**(2), 373–374.

Hofman, P. M., Van Riswick, J. G. A., and Van Opstal, A. J. (**1998**). ''Relearning sound localization with new ears,'' Nat. Neurosci. **1**(5), 417–421.

Koehnke, J., and Durlach, N. J. (**1989**). ''Range effects in the identification of lateral position,'' J. Acoust. Soc. Am. **86**, 1176–1178.

Kulkarni, A., and Colburn, H. S. (**1998**). ''Role of spectral detail in sound-source localization,'' Nature (London) **396**, 747–749.

Makous, J. C., and Middlebrooks, J. C. (**1990**). ''Two-dimensional sound localization by human listeners,'' J. Acoust. Soc. Am. **87**(5), 2188–2200.

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

Mills, A. W. (**1960**). ''Lateralization of high-frequency tones,'' J. Acoust. Soc. Am. **32**, 132–134.

Mills, A. W. (**1972**). ''Auditory localization,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), pp. 303–348.

Molino, J. (**1973**). ''Perceiving the range of a sound source when the direction is known,'' J. Acoust. Soc. Am. **53**, 1301–1304.

Rabinowitz, W. R., Maxwell, J., Shao, Y., and Wei, M. (**1993**). ''Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere,'' Presence **2**(2), 125–129.

Roffler, S. K., and Butler, R. A. (**1968**). ''Factors that influence the localization of sound in the vertical plane,'' J. Acoust. Soc. Am. **43**, 1255–1259.

Searle, C. L., Braida, L. D., Cuddy, D. R., and Davis, M. F. (**1976a**). ''Binaural pinna disparity: Another auditory localization cue,'' J. Acoust. Soc. Am. **57**, 448–455.

Searle, C. L., Braida, L. D., Davis, M. F., and Colburn, H. S. (**1976b**). ''Model for auditory localization,'' J. Acoust. Soc. Am. **60**, 1164–1175.

Shaw, E. A. G. (**1997**). ''Acoustical features of the human external ear,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. Anderson (Erlbaum, New York), pp. 25–48.

Shinn-Cunningham, B. G., Durlach, N. J., and Held, R. M. (**1998**). ''Adapting to supernormal auditory localization cues I. Bias and resolution,'' J. Acoust. Soc. Am. **103** (6), 3656–3666.

von Hornbostel, E. M., and Wertheimer, M. (**1920**). ''Uber die Wahrnehmung der Schallrichtung [On the perception of the direction of sound],'' Sitzungsber. K. Preuss. Akad. Wiss. 388–396.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). ''Localization using nonindividualized head-related transfer functions,'' J. Acoust. Soc. Am. **94**, 111–123.

Wightman, E. R., and Firestone, F. A. (**1930**). ''Binaural localization of pure tones,'' J. Acoust. Soc. Am. **2**, 271–280.

Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661.

Wightman, F. L., and Kistler, D. J. (**1997a**). ''Factors affecting the relative salience of sound localization cues,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. Anderson (Erlbaum, New York), pp. 1–24.

Wightman, F. L., and Kistler, D. J. (**1997b**). ''Monaural sound localization revisited,'' J. Acoust. Soc. Am. **101**(2), 1050–1063.

Wightman, F. L., and Kistler, D. J. (**1999**). ''Resolution of front-back ambiguity in spatial hearing by listener and source movement,'' J. Acoust. Soc. Am. **105**(5), 2841–2853.