T o o o o: B a a oa a o e o o e
ea oa e e

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 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.*, 197[illegible] 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

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

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 onnear

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 (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 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 (Tf 18.85uenc/F12 1 Tf 0.333 0 TD [(one)--372.4(60)5-or.4(60)5--372o13.3(60)5--46.3(60)5

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.