

SEPTEMBER 01 1999

Auditory localization of nearby sources. Head-related transfer functions

Douglas S. Brungart; William M. Rabinowitz



J Acoust Soc Am 106, 1465–1479 (1999)

<https://doi.org/10.1121/1.427180>

Selectable Content List

Effect of temperature on the acoustic response and stability of size-isolated protein-shelled ultrasound contrast agents and SonoVue

Intra- and inter-speaker variation in eight Russian fricatives

Age-related reduction of amplitude modulation frequency selectivity

Low frequency ambient noise dynamics and trends in the Indian Ocean, Cape Leeuwin, Australia

Estimating cochlear impulse responses using frequency sweeps



View
Online



Export
Citation

CrossMark

Related Content

Auditory localization of nearby sources. II. Localization of a broadband source

J Acoust Soc Am (October 1999)

Auditory localization of nearby sources. III. Stimulus effects

J Acoust Soc Am (December 1999)

Auditory localization of a nearby point source

J Acoust Soc Am (October 1996)



Advance your science and career
as a member of the

ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE



Auditory localization of nearby sources. Head-related transfer functions

Douglas S. Brungart^{a)} and William M. Rabinowitz^{b)}

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 29 June 1998; revised 29 March 1999; accepted 26 May 1999)

Although researchers have long recognized the unique properties of the head-related transfer function (HRTF) for nearby sources (within 1 m of the listener's head), virtually all of the HRTF measurements described in the literature have focused on source locations 1 m or farther from the listener. In this study, HRTFs for sources at distances from 0.12 to 1 m were calculated using a rigid-sphere model of the head and measured using a Knowles Electronic Manikin for Acoustic Research (KEMAR) and an acoustic point source. Both the calculations and the measurements indicate that the interaural level difference (ILD) increases substantially for lateral sources as distance decreases below 1 m, even at low frequencies where the ILD is small for distant sources. In contrast, the interaural time delay (ITD) is roughly independent of distance even when the source is close. The KEMAR measurements indicate that the direction of the source relative to the outer ear plays an important role in determining the high-frequency response of the HRTF in the horizontal plane. However, the elevation-dependent characteristics of the HRTFs are not strongly dependent on distance, and the contribution of the pinna to the HRTF is independent of distance beyond a few centimeters from the ear. Overall, the results suggest that binaural cues play an important role in auditory distance perception for nearby sources. © 1999 Acoustical Society of America.

[S0001-4966(99)04209-5]

PACS numbers: 43.66.Qp [DWG]

INTRODUCTION

Over the past 100 years, considerable efforts have been made to characterize the relationship between the location of a sound source in space and the sound pressure generated by that source at the eardrums of a human listener. Traditionally, this relationship has been represented by the head-related transfer function, or HRTF. Many researchers have measured the HRTF for relatively distant sound sources, located 1 m or more from the listener. (Møller *et al.*, 1995; Gardner and Martin, 1995; Pralong and Carlile, 1994; Carlile and Pralong, 1994; Middlebrooks *et al.*, 1989; Wightman and Kistler, 1989; Mehrgardt and Mellert, 1977; Shaw, 1974). However, despite the recognition by early researchers that the HRTF varies substantially with distance at distances less than 1 m, very little is known about the characteristics of the HRTF for nearby sources.

Nearly all previous work examining the HRTF for close sources has been based on a rigid-sphere model of the head. Stewart (1911a, 1911b) derived mathematical expressions for the interaural level difference (ILD) and interaural time delay (ITD) on a rigid-sphere head with ears located at diametrically opposite points on its surface, and found that the ILD increases substantially as a nearby source approaches the head, while the ITD remains roughly independent of distance even for close sources. Hartley and Frey (1921) expanded Stewart's calculations, and noted that a combination of distance-invariant ITDs and distance-varying ILDs could

provide binaural distance cues for nearby sources. More recently, interaural time and intensity differences have been calculated with an updated sphere model (Brungart and Rabinowitz, 1996; Brungart *et al.*, 1996), and validated with acoustic measurements of the HRTF on a bowling-ball head (Duda and Martens, 1998). A common feature of all of the sphere model predictions is that the HRTF changes substantially with distance at distances less than 1 m, but is virtually independent of distance for sources beyond 1 m. Therefore, it is useful to introduce terminology to differentiate between these near and far regions. In our discussion of auditory localization near the head, we use the term "proximal region" to describe the region within 1 m of the listener's head, and the term "distal region" to describe the region beyond 1 m.¹

The rigid-sphere model of the head can provide useful insights into the general behavior of the HRTF for close sources, but it does not fully represent the complexities of the actual HRTF. In order to obtain a clearer picture of the distance dependence of the HRTF, an approximation that accounts for the irregular shape of the head and outer ear as well as diffraction by the neck and torso is needed. Measurements on an anthropomorphic manikin can provide this type of approximation, without the difficulties involved in measuring the HRTF on human subjects. With the exception of some ITD and ILD measurements on a wax dummy head in an early study by Firestone (1930), no comprehensive measurements of the HRTF on an anthropomorphic head (human or manikin) are available in the literature. This paper describes an extensive set of HRTF measurements made with a KEMAR manikin at distances ranging from 0.12 to 1 m. The results are compared to the predictions of the sphere model,

^{a)}Currently at Human Effectiveness Directorate, Air Force Research Laboratory, WPAFB, OH. Electronic mail: douglas.brungart@he.wpafb.af.mil

^{b)}Currently at Bose Corporation, Framingham, MA.

and are discussed in relation to their implications for the localization of nearby sources.

I. METHODS

A. HRTF calculations for a rigid-sphere model of the head

Useful approximations of the HRTF for a nearby source, and the corresponding interaural time and intensity differences, can be obtained from mathematical descriptions of the acoustic properties of a rigid sphere. Our model was adapted from the work of Rabinowitz and colleagues (1993), who examined the frequency scalability of the HRTF for an enlarged head. The present variation of the model maintains a fixed head size and varies distance, rather than varying head size at a fixed distance (Brungart and Rabinowitz, 1996). A more detailed review of the application of the model, as well as a comparison of the model's predictions to acoustic measurements made on the surface of a bowling ball, are available in Duda and Martens (1998).

The model approximates the head as a rigid sphere, of radius a , with "ears" located at diametrically opposed points on the surface of the head. The sound source is a point velocity source radiating spherical acoustic waves, and is located at distance r from the center of the head and at angle α from the interaural axis. The expression for the complex sound pressure at the ear, denoted by $\mathbf{P}_s$, is given by

$$\mathbf{P}_s(r, a, \alpha, f) = \frac{c \rho_0 u_0}{2 \pi a^2} \sum_{m=0}^{\infty} \left(m + \frac{1}{2} \right) L_m(\cos \alpha) \times \frac{H_m(2 \pi f r / c)}{H'_m(2 \pi f a / c)} e^{-j \pi / 2}, \quad (1)$$

where f is the sound frequency, c is the speed of sound, ρ_0 is the density of air, u_0 is the volume velocity of the (infinitesimal) source, L_m is the Legendre polynomial function, and H_m is the spherical Hankel function. To obtain the HRTFs from the model, the pressure at the ear must be divided by the reference free-field pressure at the center of the head $\mathbf{P}_{ff}$, which is simply the pressure of a point source of strength u_0 at distance r :

$$\mathbf{P}_{ff}(r, f) = \frac{2 \rho_0 f u_0}{r} e^{j(-2 \pi f r / c + \pi / 2)}. \quad (2)$$

A remaining complication is the transformation of the angle α between the location of the source and the location of the ear to the more standard spherical coordinates used to define HRTFs. Throughout this paper, the coordinate system has its origin at the midpoint of the interaural axis. Azimuth (θ) will be defined as 0 degrees directly in front of the head, 90 degrees directly to the left, and -90 degrees directly to the right. Elevation (ϕ) will be 0 degrees in the horizontal plane, 90 degrees directly above, and -90 degrees directly below. In this coordinate system, $\alpha = \arccos(\sin(\theta) \cos(\phi))$ and the transfer function at the left ear is

$$\mathbf{H}_L(r, a, \theta, \phi, f) = \frac{\mathbf{P}_s(r, a, \arccos(\sin(\theta) \cos(\phi)), f)}{\mathbf{P}_{ff}(r, f)}. \quad (3)$$

B. Procedures for HRTF measurements

1. Facilities

The HRTF measurements were made in a corner of the anechoic chamber of the Armstrong Laboratory at Wright-Patterson Air Force Base.² All of the HRTF measurements were made with a KEMAR manikin, which was fitted bilaterally with Zwislocki couplers and 1.2-cm pressure microphones (Briel & Kjaer 4134) at the eardrum locations. The microphone outputs were connected to a Briel & Kjaer 5935 dual microphone power supply, and then passed through a patch panel into the control room. The KEMAR was mounted on a metal stand equipped with optically encoded stepper motors that allow programmable control of the manikin's azimuth and elevation within a fraction of a degree.

In order to avoid the complications involved in near-field measurements with a conventional cone loudspeaker (due to their size, directionality, and complex near-field characteristics), a special sound source approximating an acoustic point source was developed for these measurements. An electrodynamic horn driver (Electro-Voice DH1506) was connected to a 3-m-long section of Tygon tubing with internal diameter of 1.2 cm and walls 1.5 mm thick. For convenience, the end of this tube was routed through a PVC pipe sleeve, 2.5 cm in diameter and 64 cm in length, with the end of the Tygon tube projecting 2 cm from the end of the pipe and with foam material sealing the space between the tube and the interior of the sleeve. The sleeve was used to clamp the point source to a tripod stand that was used to position the source during the measurements. The relatively small diameter of the tube (1.2 cm) allows its opening to act approximately as an acoustic point source in that it is relatively nondirectional even at high frequencies. At 15 kHz, for example, the 3-dB beamwidth of the source was found to be 120 degrees. Furthermore, the small size allows a precise placement of the source and eliminates the potential problem of secondary reflections off the source. The source's frequency response is irregular (Fig. 1); however, its shape is consistent and easily removed from the transfer function measurements, and its effective frequency range is from 200 Hz to 15 kHz.

Reference measurements were made with a 1.2-cm (0.5 in.) free-field microphone (Briel & Kjaer 4165) located at the position of the center of the manikin head with the manikin removed. Measurements were made at 0.125, 0.25, 0.50, and 1.00 m before and after the HRTF measurements. Changes in the frequency response of the source over the course of the measurements were found to be negligible (within ± 1.5 dB), and the signal measured at the reference microphone was essentially independent of distance except for the inverse relation of overall amplitude to distance.

The measurements were controlled by a computer located in a small room adjacent to the chamber. The two rooms were connected by a patch panel which passed the left and right microphone signals, the sound source signal, and the motor controller signal. The computer was also connected via GPIB bus to a dynamic signal analyzer (Hewlett Packard HP35665A) which was used both to generate the source signal and to measure the transfer functions. The mi-

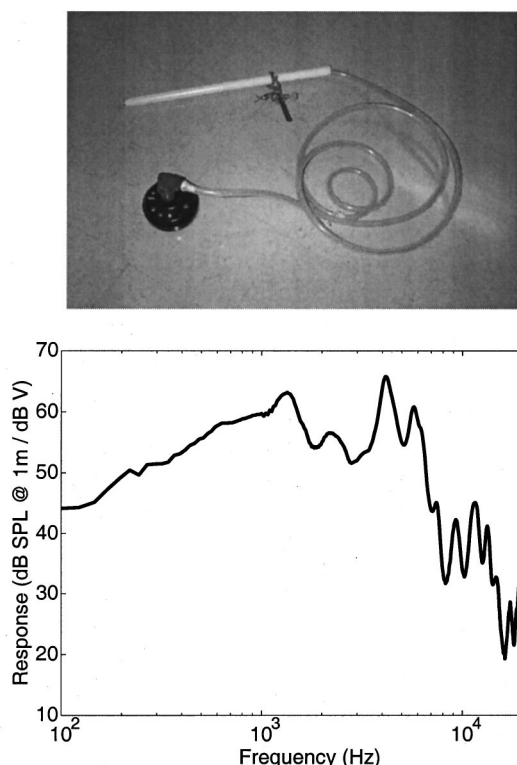


FIG. 1. Acoustic point source. The top panel illustrates the construction of the point source, which essentially consists of an electrodynamic horn driver connected to a 3-m-long flexible tube with a small amount of acoustic foam inserted into the opening to act as a terminating impedance and cut down on standing waves in the tube. The end of the tube is encased in a rigid PVC sleeve to allow easy placement of the tube opening. The sleeve in this picture is curved to allow easy manual placement of the source, and equipped with an electromagnetic sensor to record the location of the tip. In the HRTF measurements described in this paper, only the 64-cm section of the sleeve was used. The bottom panel shows the frequency response of the point source measured with a free-field microphone at a distance of 1 m. The measurement was made in an anechoic chamber using a periodic chirp stimulus and a 1024-point FFT. One-third octave smoothing has been applied to the response. Note the drop-off in the response at frequencies above 6 kHz.

crophone signals were connected to the analyzer inputs. The source signal was amplified (Crown D-75) before being passed through the patch panel to the sound driver.

2. Measurement procedure

The analyzer was operated in transfer-function mode, which measures the ratio of the signal spectra at its two inputs. The analyzer's source output was connected to one input, and one of the ear microphones was connected to the second input. A periodic chirp source signal was used to maximize the signal-to-noise ratio. The periodic chirp is a periodic sine-wave sweep with its period equal to the length of the time record of the fast Fourier transform (FFT). Because of its periodicity, no windowing is required for FFT calculations with the chirp stimulus. At each source position, 64 FFT measurements were combined using root-mean-square (rms) averaging.

All measurements were made at two frequency ranges: from 100 Hz to 12.9 kHz and from 7.8 to 20.6 kHz. Each measurement consisted of a 400-point FFT with 32-Hz resolution. The first 240 points of the low-frequency FFT (from

100 to 7748 Hz) and the first 360 points of the high-frequency FFT (from 7780 to 19 268 Hz) were concatenated into a single 600-point transfer function with 32-Hz resolution from 100 Hz to 19.2 kHz. The two-part measurement was used for two reasons. It allowed higher resolution over the frequency range than a single measurement and, more importantly, it allowed independent ranging of the analyzer's input gain at low and high frequencies. The transfer function of the point source drops off considerably above 7 kHz (see Fig. 1). By dividing the measurement into two parts, the analyzer could be set to optimize the signal-to-noise ratio in both frequency bands.

After each measurement, the amplitude and phase of the transfer function at each frequency value were saved into separate files. The magnitude of the HRTF at each location was found by dividing the amplitude spectrum at that location by the free-field reference measurement corresponding to the source distance. The ITDs were calculated from the difference between the phase responses at the left and right ears.

3. Calibration

Prior to each set of measurements, the KEMAR manikin was adjusted to place the center of the interaural axis directly over the axis of rotation of the stand. When the source is close to the head, correct placement is particularly important because small deviations between the center of the head and the center of rotation will cause changes in the distance to the source as the manikin is rotated in azimuth and severely corrupt the measurements. The centering of KEMAR was accomplished automatically via a series of acoustic measurements with the source placed in front of the manikin at 0-degree elevation. The manikin was first centered in azimuth by rotating the head until the magnitude of the ITD was reduced below $2 \mu\text{s}$. Then the roll of the manikin was verified by rotating KEMAR 180 degrees in azimuth and again verifying that the ITD remained near zero. With a nearby source, any left or right tilt of the manikin will prevent the source from falling on the median plane at both 0 and 180 degrees in azimuth. Finally, the manikin head was centered in elevation by verifying that the time delay from the source to the left ear at 0 degrees in azimuth was equivalent (within $5 \mu\text{s}$) to the delay from the source to the left ear at 180 degrees in azimuth. If the manikin were tilted forward or backward, the distances (and delays) for these two measurements would differ. After the elevation was adjusted, the centering measurements were repeated until the manikin was acceptably centered in azimuth, roll, and elevation.

II. THE HRTF FOR PROXIMAL-REGION SOURCES IN THE HORIZONTAL PLANE

A. Sphere model predictions

The magnitudes of the HRTFs predicted by the sphere-based mathematical model of the head are shown in Fig. 2. The principal characteristics of these transfer functions can be summarized by four observations:

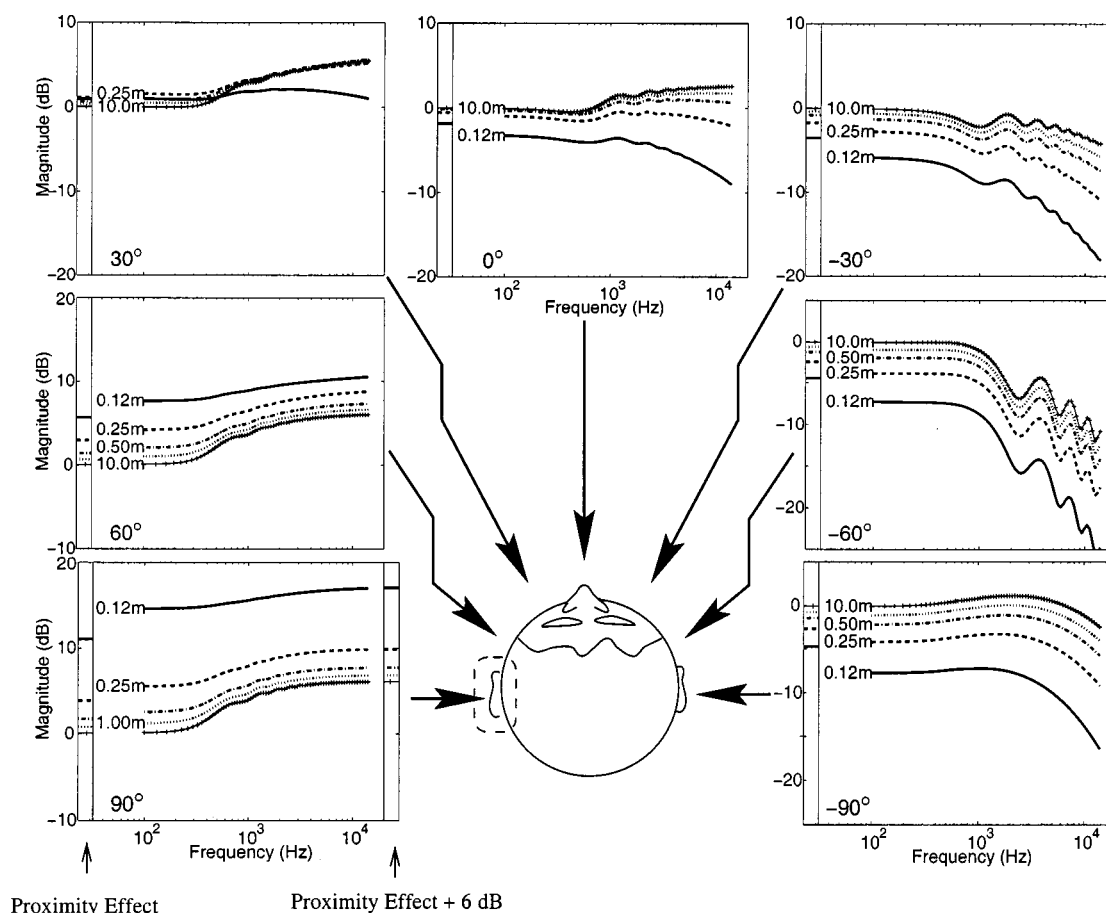


FIG. 2. The HRTF for horizontal-plane sources from 0.125 to 10 m when the head is modeled as a rigid sphere 18 cm in diameter. The HRTFs were calculated by dividing the pressure at the left ear by the free-field pressure at the center of the head [Eq. (3)]. Results are shown for source locations at 30-degree intervals in azimuth in the front hemisphere only, as the sphere model is symmetric across the frontal plane. Note that each panel represents a 30-dB range of responses, but that the values change to accommodate the data. The bars at the left side of each graph show the source proximity effect, which is the gain of the HRTF ignoring diffraction by the head. In the 90-degree panel, the bar at the right side of the panel shows the source proximity effect plus the 6-dB high-frequency pressure doubling effect, and illustrates that the combination of these two effects fully explains the high-frequency asymptotes of the HRTFs at this location. See text for details.

- (1) The HRTF increases as the source rotates toward the ear, and decreases as the source rotates away from the ear.
- (2) The HRTF increases with frequency when there is a direct path between the sound source and the ear, and decreases with frequency when the ear lies in the acoustic shadow of the head.
- (3) The HRTF decreases with distance when there is a direct path from the source to the ear, and increases with distance when the ear is shadowed by the head.
- (4) The HRTF changes rapidly as distance decreases below 0.5 m, but slowly (< 1 dB) as distance increases from 1 to 10 m.

1. Analysis of sphere model results

Several of the characteristics of the sphere-model HRTFs can be explained intuitively with relatively simple acoustic concepts. There are also some low-frequency diffraction effects in the HRTFs which defy intuitive explanation. These are explained in more detail below.

a. Head shadowing. Head (or sphere) shadowing is the attenuation in the HRTF that occurs when the head obscures the direct path from the sound source to the ear. The magnitude of the shadow is related to the size of the head relative

to the wavelength of the sound, so the attenuation at the shadowed ear increases with frequency. As a result, the HRTFs for contralateral sources resemble low-pass filters (see the HRTFs for -30 , -60 , and -90 degrees in Fig. 2).

As a result of the convex shape of the spherical head, the shadowed region of the sphere's surface will always extend into the hemisphere closest to the source when the source distance is finite. A crude illustration of this phenomenon is provided in Fig. 3, which shows a sound source located 20 cm from the center of a spherical head 20 cm in diameter. Note that the direct path from the source to the surface of the sphere is blocked at all points to the right of point *T*, where the tangent line *ST* intersects with the surface of the sphere. Thus the direct path *SA* from the source to point *A* on the sphere is blocked by the surface of the sphere, and point *A* is shadowed by the head. In this example, the shadowed zone extends over two-thirds of the surface of the sphere (the entire region where the arc from point *C* closest to the source is greater than 60 degrees). The size of this shadowed zone increases as the source moves closer to the head, and decreases as the source moves farther away.

As a consequence of this type of shadowing, it is possible for both the ipsilateral and contralateral ears to fall

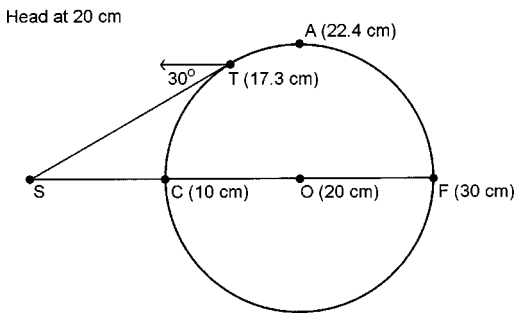


FIG. 3. Head shadowing and the source-proximity effect for a nearby source. In this figure, sound propagates from an infinitesimal point source (S) to the surface of a spherical head with a diameter of 20 cm, centered at O . Note that the head shadow, which is defined as the region where the direct path from the source to the ear is occluded by the head, includes the entire contralateral hemisphere and extends into the ipsilateral hemisphere where the direct path from the source is tangent to the surface of the sphere (point T). Thus point A is in the head-shadowed region in this example. Also, note that point A is more than twice as far away from the source as the closest point (C) and that (F) is three times as far away as the closest point (C). Thus, even if the effects of diffraction by the sphere are ignored, there would be a substantial difference in sound level between points C , A , and F due to source proximity effects.

inside the shadowed zone of the sphere. In the case of the spherical head used in the model (with a radius of 9 cm), the ipsilateral ear lies in the head shadow for a sound source at 30 degrees at all distances less than 18 cm. Thus, at 30 degrees in Fig. 2, the ear lies in the acoustic shadow of the head at a source distance of 0.12 m, and is unshadowed at all greater source distances. As a result, the high-frequency component of the sound reaching the ear is substantially attenuated when the source is at 12 cm, and the magnitude of the 12-cm transfer function switches from greatest to smallest at this point. The amount of attenuation due to the head shadow also increases as the ear moves deeper inside the shadowed region. Since the size of the shadowed region increases as source distance decreases, this results in increased high-frequency attenuation for nearby sources at the shadowed ear. At all points to the right of 0 degrees, the source is shadowed at all distances, and the ordering of the transfer functions is reversed (i.e., the magnitude of the transfer function decreases with distance in the unshadowed zone (60 and 90 degrees), and increases with distance in the shadowed zone (0, -30, and -60 degrees).

b. Source proximity. The inverse relationship between pressure and distance for a spherically radiating sound wave can substantially influence the HRTF when the source is near the head. This *source-proximity* effect can be viewed as the portion of the HRTF which is not a result of diffraction by the head, i.e., as the HRTF for a pressure sensor suspended in free space at the location of the ear. Since the HRTF is defined as the ratio of the pressure at the ear to the pressure at the center of the head, the source-proximity effect is simply the ratio of the distance from the source to the center of the head to the distance from the source to the ear. An example of the source-proximity effect is provided in Fig. 3. In this figure, which assumes that a head with diameter 20 cm is located 20 cm from the source, the intensity of the sound reaching its surface would be 4.8 dB stronger at the point closest to the source (C in the top panel of Fig. 3) than at the

point where the direct path from the source is tangent to the sphere (T). If the direct sound could pass through the head unimpeded, the sound reaching the point farthest from the source (F) would be attenuated 9.5 dB relative to the sound reaching the point closest to the source.

The magnitude of the source-proximity effect is shown along the left side of each panel in Fig. 2. At 10.0 m, the effect is negligible (near 0 dB) for all source directions. The effect increases as source distance decreases, and at 0.12 m the source-proximity effect produces more than 10 dB of gain at 90 degrees and almost 5 dB of attenuation at -90 degrees. The proximity effect is frequency independent. In particular, it helps interpret the large changes in the HRTF at low frequencies that are evident only for nearby sources. However, the source-proximity effect cannot explain all of the low-frequency behavior of the HRTF. In general, the ordering of the low-frequency HRTFs by distance is consistent with the source-proximity effect at each azimuth location, but the magnitude of the low-frequency gain or attenuation is greater than that predicted by source proximity.

c. High-frequency pressure doubling. The magnitude of the HRTF at the ipsilateral ear generally increases with frequency due to high-frequency reflections off the surface of the sphere. When the source is located at 90 degrees, the sound wave impinging on the head is perpendicular to the surface of the sphere at the location of the ear, and at sufficiently high frequencies the sound wave is specularly reflected off the surface of the sphere back in the direction of the source. The direct and reflected sound waves combine in phase to produce a 6-dB pressure gain at the location of the ear. This 6-dB increase in high-frequency gain is evident in the 10 m HRTF at 90 degrees in Fig. 2. In fact, for all of the distances, the high-frequency magnitude of the HRTF at 90 degrees is essentially equivalent to the combination of the source-proximity effect and the high-frequency pressure doubling effect, as shown at the right side of the panel in Fig. 2. As the source rotates away from 90 degrees, the sound waves from the source are no longer perpendicular to the surface of the sphere at the ear and only a portion of the sound wave is reflected at the ear canal opening, resulting in a high-frequency gain slightly less than 6 dB at 60 and 30 degrees. Note that high-frequency pressure doubling does not occur in the contralateral HRTFs.

d. Acoustic bright spot. When the ear is located directly opposite the source (-90 degrees in Fig. 2), all possible sound paths from the source around the sphere to the ear are cylindrically symmetric and, consequently, all components of the diffracted sound wave combine in phase at the ear. This in-phase combination results in a local maximum in the HRTF at all frequencies, a phenomenon known as the acoustic "bright spot." As a result the high-frequency responses of the HRTFs at -90 degrees are substantially greater than the corresponding components of the HRTFs at -60 degrees at each source distance. The effects of constructive and destructive interference on the contralateral hemisphere of the head are seen as ripples of the high-frequency HRTF response.

e. Other diffraction effects. The low-frequency response of the sphere-model HRTFs cannot be explained fully with

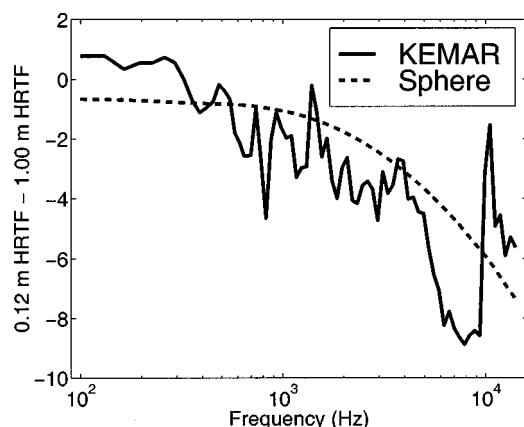


FIG. 4. The difference between the mean HRTFs at 0.12 and 1.0 m. The sphere model and KEMAR HRTFs at 0.12 and 1.0 m were averaged (in dB units) across all locations in the horizontal plane. The results illustrate that the HRTF, on average, decreases in magnitude more quickly at high frequencies than at low frequencies as the source approaches the head. This effect is more pronounced in the KEMAR measurements than in the sphere model. The decrease in spectral content at high frequencies as distance decreases could potentially serve as a spectral distance cue in the proximal region.

the simple concepts described above. As noted before, the source-proximity effect only partially accounts for the low-frequency HRTF responses. The remaining differences result from diffraction by the sphere. Note that the diffraction effects tend to increase the low-frequency gain at the ipsilateral ear and decrease the gain at the contralateral ear. The magnitude of these diffraction effects increases as distance decreases.

2. Low-pass filtering of proximal-region sources

The combination of the low-frequency diffraction effects in the ipsilateral hemisphere and the head-shadowing effects in the contralateral hemisphere produces a consistent change in the shape of the HRTF with source distance. The high-frequency response of the HRTF is generally lower relative to the low-frequency response of the HRTF when the source is close than when the source is more distant. This effect occurs at all source locations, in both the ipsilateral and contralateral hemispheres. The high-frequency response is generally 4–6 dB lower relative to the low-frequency response when the source is at 0.12 m than when the source is at 1 m (Fig. 4). This relationship implies that a sound source at a fixed angle relative to the head will appear to be low-pass filtered as the source approaches the head. Although this effect is modest, it could be used as a monaural distance cue in the proximal region, and it is consistent with previous observations that sound sources appear to “darken” in timbre as they approach the head (Begault, 1987; von Békésy, 1960).

3. Contour plots of sphere-model HRTFs

Contour plots provide a different perspective of the sphere-model HRTFs (Fig. 5, left panels). These plots are particularly useful for viewing the rippled nature of the acoustic bright spot in the contralateral HRTFs. At 2500 Hz, there is a single peak around -90° (location B). As

frequency increases, this peak decreases in width and additional peaks form on either side until, at high frequencies, the central peak is very sharp and is surrounded by multiple ridges on either side (location A). The increases in the high-frequency response of the ipsilateral HRTFs due to pressure doubling are also apparent in the contour plots (location C).

B. KEMAR measurements

The HRTFs measured with the KEMAR manikin are shown in Fig. 6. Numerous distal-region HRTFs, both from manikins and from human listeners, are available in the literature, and the directional characteristics of the distal-region HRTFs are well documented. This discussion will focus on a comparison between the HRTFs calculated with the sphere model and those measured with KEMAR, and on the distance-dependent characteristics of the KEMAR HRTFs. It is important to note that these HRTFs measured with the KEMAR manikin may differ from those measured on any individual subject. However, the general pattern of the distant-dependent changes in the HRTFs with the KEMAR manikin should be representative of those in the HRTFs of human listeners. Although it is unlikely that the high-frequency characteristics of the KEMAR HRTFs will match those of any particular human listener, the relative changes in those HRTFs (such as systematic shifts in the azimuth or frequency of notches in the HRTF with distance) provide some insights about the general distant-dependent behavior of the HRTF for nearby sources.

1. Comparison of KEMAR measurements with sphere model

The sphere model best fits the KEMAR measurements at low frequencies (below 1 kHz). For comparison, the magnitude of the sphere-model HRTF at 100 Hz is shown beside the KEMAR measurements in Fig. 6. The fit of the model to the measurements is best at the contralateral ear and worst near the boundary between the shadowed and unshadowed zones (30° and 150° degrees). There are also some discrepancies between the model and measurements at 0.12 m near 90° degrees.

As frequency increases, the KEMAR HRTFs begin to diverge from those of the sphere model. At 2.9 kHz, the quarter-wavelength resonance of the ear canal causes a large peak in the KEMAR HRTFs at all directions and distances. At higher frequencies, the KEMAR transfer functions exhibit a complex series of direction-dependent peaks and notches which arise from the geometry of the pinnae and are not reflected in the sphere model. Five major characteristics of the sphere HRTFs are preserved in the KEMAR HRTFs.

- (1) The magnitude of the HRTFs generally increases with frequency when there is a direct path from the source to the ear. In part, this increase is probably a result of the reflections that cause the 6-dB gain in the sphere model. The pinna, which is shaped somewhat like a cone, also provides some gain at high frequencies. Note that the overall high-frequency gain at the ipsilateral ear is greater in the KEMAR measurements than in the sphere model.

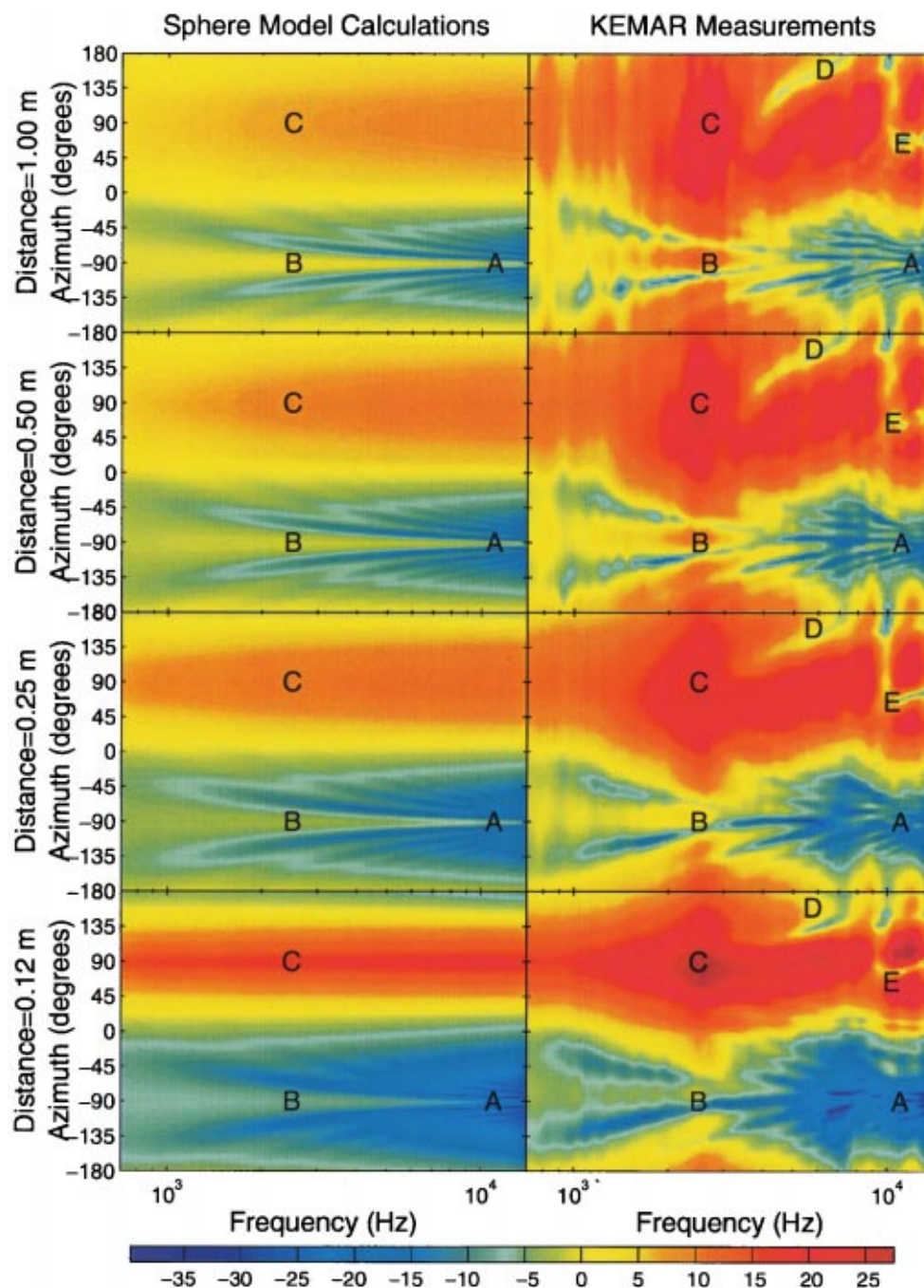


FIG. 5. Contour plots of the HRTFs predicted by the sphere model (left column) and measured with KEMAR (right column). Azimuth is shown at 3-degree intervals, and frequency is shown at 1/12-oct intervals from 500 Hz to 14 kHz. The magnitude of the transfer function at each point (in dB) is represented by color, as shown by the legend across the bottom of the figure. Five reference points (labeled A–E) are present on the contour plots (see text for details).

- (2) The high-frequency responses of the HRTFs are attenuated when the ear is in the acoustic shadow of the head.
- (3) The overall gain of the HRTFs increases as distance decreases when a direct path exists between the source and the ear, and the overall attenuation of the HRTFs increases as distance decreases when the ear is shadowed by the head. Note that, as in the sphere model, the ear is first shadowed by the head at 30 and 150 degrees when the source is at 0.12 m, and that the ordering of the HRTFs at high frequencies reverses at these locations.
- (4) Overall, the magnitude of the HRTF increases more rapidly at low frequencies than at high frequencies as the

source approaches the head (Fig. 4). Thus, the sound reaching the eardrums is effectively low-pass filtered as the source approaches the head. This effect is more dramatic in the KEMAR HRTFs than in the sphere model, and may serve as a monaural distance cue in the proximal region.

- (5) Although its structure is more complicated, the acoustic bright spot seen in the sphere model is also evident in the KEMAR measurements. This is best seen in the contour plots of the KEMAR measurements (Fig. 5). The peak at intermediate frequencies occurs slightly to the left of -90 degrees in the KEMAR measurements due to the

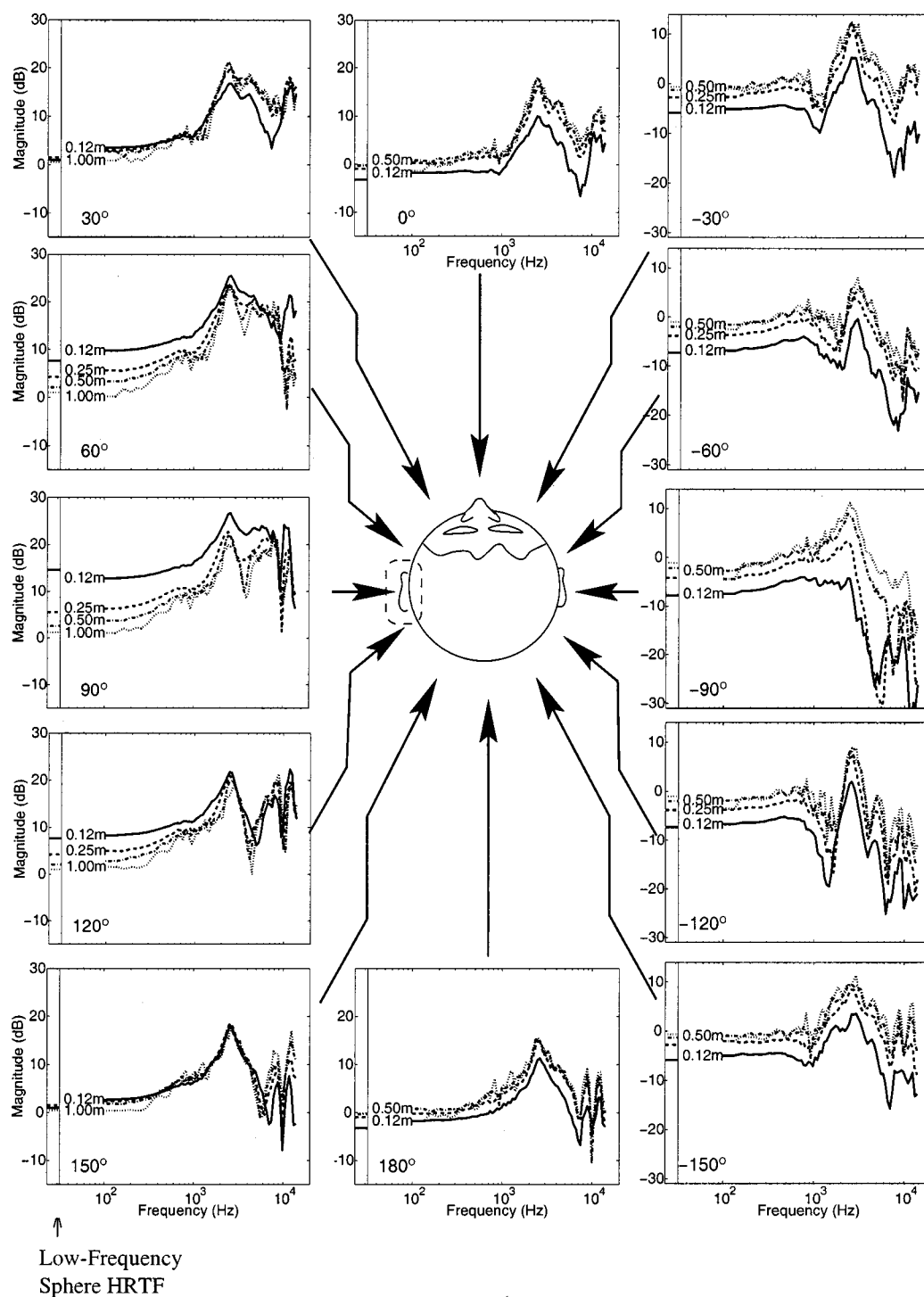


FIG. 6. The HRTF for horizontal-plane sources from 0.125 to 1.0 m measured with the KEMAR manikin. The HRTFs were calculated by dividing the pressure at the left ear by the free-field pressure at the center of the head. The bars at the left side of each panel show the low-frequency asymptote of the HRTF calculated with the sphere model (as shown in Fig. 2).

asymmetries of the manikin head (location B). At higher frequencies, the periodic interference pattern around the bright spot is seen in the KEMAR measurements, but it is more erratic than in the sphere model. Note that one notch from the interference pattern appears to extend into the ipsilateral hemisphere at 6.5 kHz (location D).

2. The parallax effect in the high-frequency HRTF

When a sound source is distant, it is located in roughly the same direction relative to the head and to the left and

right ears. However, when the source is close to the head, the angle from the source to the ear can differ substantially from the angle of the source relative to the head (Fig. 7). This discrepancy between the angles of the source relative to the head and ear (or relative to the left and right ears) is known as acoustic parallax (Suzuki *et al.*, 1998). The parallax effect changes systematically with the distance of the sound source; the closer the source comes to the head, the more its angle relative to the ipsilateral ear is pushed away from the interaural axis. As a result, the high-frequency features of the

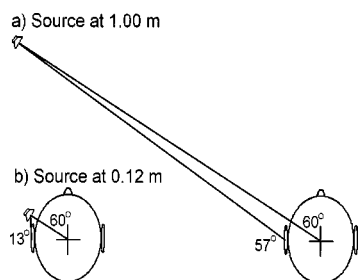


FIG. 7. The parallax effect for nearby sound sources. When a sound source is relatively distant, the direction of the source relative to the ear is approximately the same as the direction of the source relative to the center of the head. For example, with a head width of 18 cm, a source at 60 degrees azimuth in the horizontal plane (relative to the head) is located at 57 degrees relative to the left ear. However, when a sound source is near the head, the angles from the ear to the source and the head to the source can differ substantially. A source 0.12 m from the head at 60 degrees azimuth relative to the head is located at 13 degrees relative to the ear. To the extent that the high-frequency features of the HRTF depend on the direction of the source relative to the pinna, this parallax effect can cause a lateral shift in the HRTF. See text for details.

ipsilateral HRTF associated with the directional properties of the outer ear tend to occur at increasingly lateral azimuth locations as the source approaches the head.

A direct illustration of this effect is provided by the high-frequency HRTFs in Fig. 8. In this figure, two prominent features of the HRTF have been labeled: a notch in the HRTF for sources in the rear hemisphere that increases in frequency from 3.5 kHz to 7 kHz as the source moves behind the head (point D), and a dip in the HRTF for frequencies from 9 to 13 kHz at a narrow range of azimuths around 65 degrees (point E). These points are also marked on the color contour plots in Fig. 5. At 1.00 m, the reference points D and E are centered directly over the HRTF features. As distance decreases, however, both features are shifted laterally (toward 90 degrees) in azimuth. As distance decreases from 1.00 to 0.12 m, the notch at point D is shifted systematically from 160 to 140 degrees, while the dip at point E is shifted from 65 degrees to approximately 80 degrees.

Since the parallax effect is based on the directional frequency response of the pinna, it is important to note that this frequency response varies much less with distance than does the frequency response of the head. This is a direct result of the much smaller dimensions of the pinna—since the pinna is only about 1/5 as large as the head, the “near field” of the pinna should be much smaller. The contribution of the pinna to the proximal-region HRTF has been measured by comparing the KEMAR HRTF with the pinna in place to the KEMAR HRTF with the pinna removed (Brungart, 1998). The measurements verify that the frequency response of the ear does not vary substantially with distance until the source is within 4 cm of the ear (≈ 15 cm from the center of the head).

III. INTERAURAL DIFFERENCES FOR PROXIMAL-REGION SOURCES IN THE HORIZONTAL PLANE

A. Interaural level differences

Interaural level differences were calculated at 0.125, 0.25, 0.50, and 1.0 m with the KEMAR HRTFs, and at each

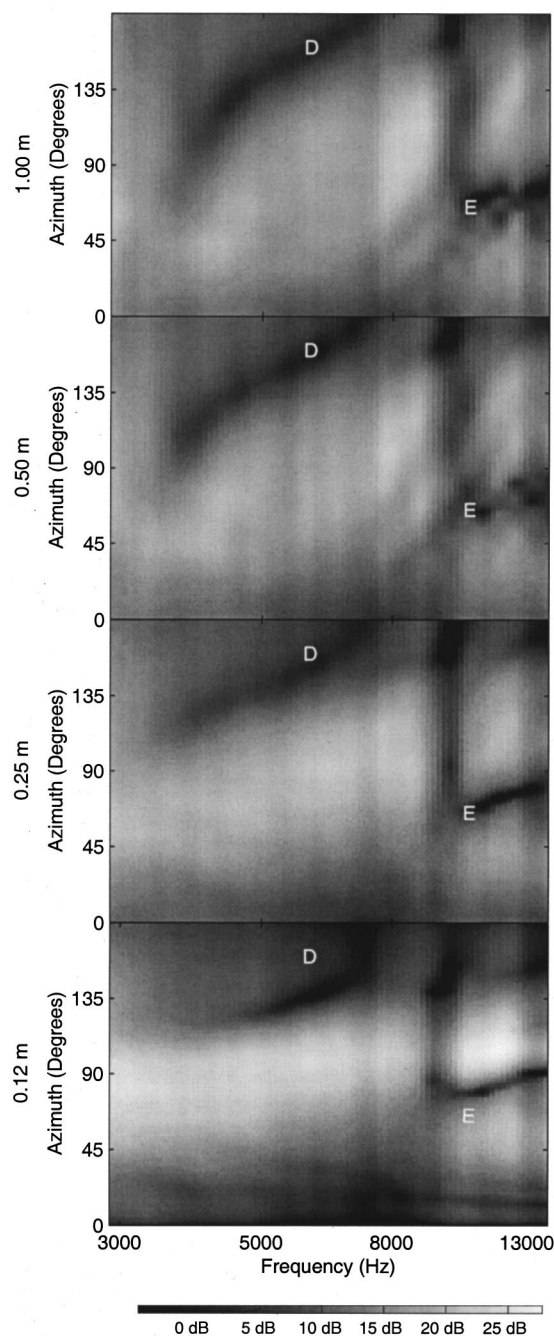


FIG. 8. Illustration of the parallax effect in the high-frequency HRTF. These contour plots show the high-frequency ipsilateral HRTF as a function of frequency and azimuth. The magnitude in (dB) is represented by the gray-scale shading, as indicated in the scale at the bottom of the figure. Two reference points are indicated on each of the panels: D (at 5800 Hz and 160 degrees) and E (at 10 300 Hz and 65 degrees). Note that these points are also indicated in the color panels of Fig. 5. As distance decreases, the features of the HRTF are shifted away from the fixed reference points and toward 90 degrees.

of these distances plus 10.0 m with the sphere-model HRTFs. The data are shown in polar form in Fig. 9. The important characteristics of the ILDs can be summarized as follows:

- (1) The ILD is always near 0 dB in the median plane and generally increases as the source moves lateral to the head. This result follows directly from the directional

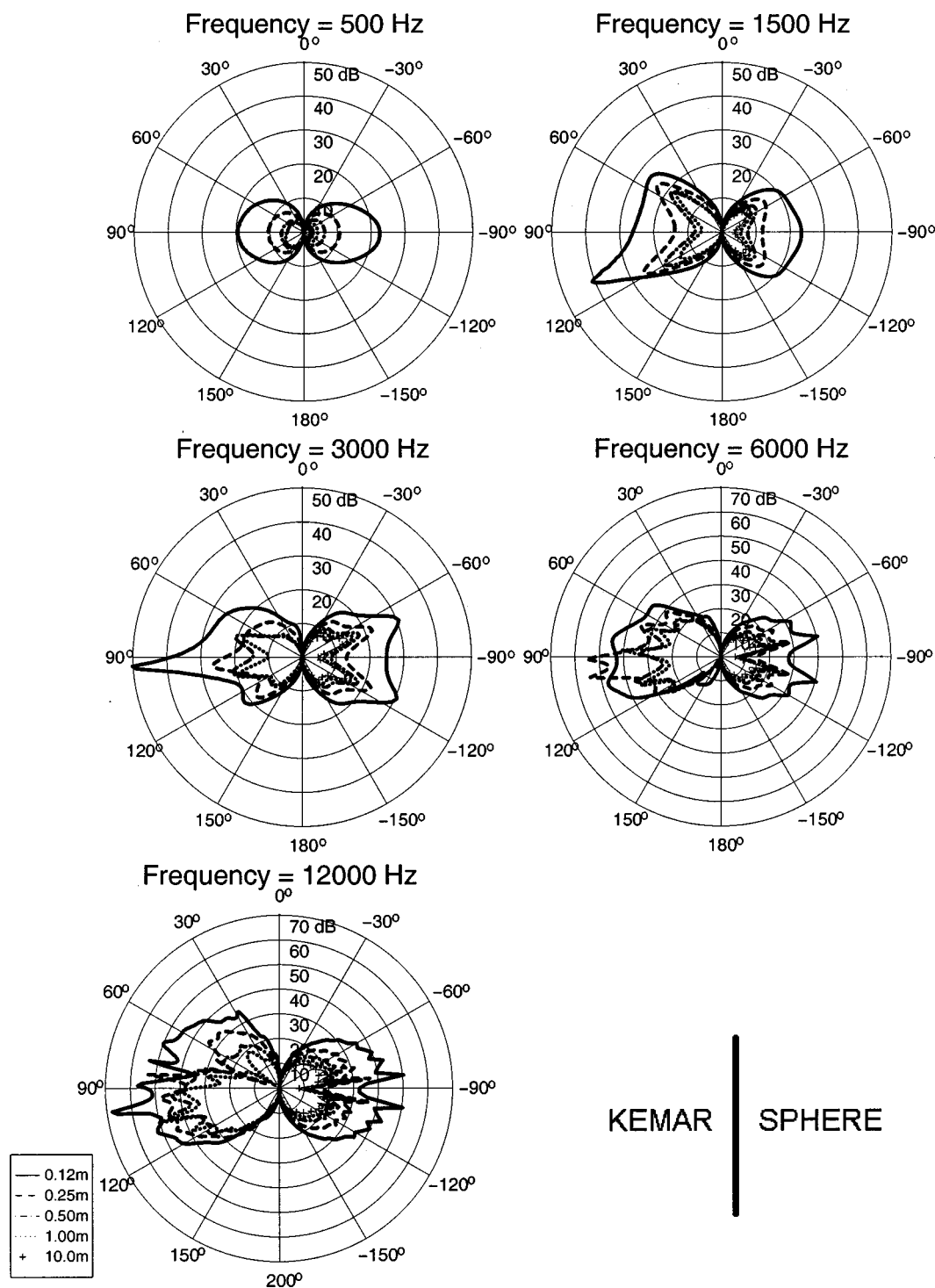


FIG. 9. Interaural intensity differences. Source location in azimuth is represented by angle, and the magnitude of the IID (in dB) is represented by the radius at each angle. Results are provided at five frequencies, ranging from 500 Hz to 12 kHz. The left side of each polar plot shows the IID from the KEMAR measurements, while the right side shows the IID calculated from the sphere model. Note that the radial scale in the plots at 6000 Hz and 12 000 Hz is larger than in the lower-frequency plots.

dependence of the HRTF, which increases in magnitude as the ear rotates toward the source and decreases in magnitude as the ear rotates away from the source.

- (2) The ILD generally increases as frequency increases. This behavior results from the tendency of the HRTF to increase with frequency when there is a direct path from the source to the ear and to decrease with frequency when the ear is shadowed by the head. Both effects contribute to the enlarged ILD at high frequencies.

- (3) The ILD increases as distance decreases, and dramatically so as the source distance drops below 0.5 m. This distance dependence occurs because the magnitude of the HRTF increases as distance decreases at the ipsilateral ear, and decreases as distance decreases at the contralateral ear. This increase in ILD may be most salient perceptually at low frequencies. For HRTFs measured at distances greater than 1 m, low-frequency ILDs are very small. At 0.12 m, however, they can be on the order of

20 dB and may provide a substantial cue for distance. High-frequency ILDs also increase as distance decreases, but they can also be quite large when the source is distant, and listeners may be less sensitive to a change in ILD from 40 to 60 dB (at high frequencies) than to one from 2 to 20 dB (at low frequencies). Low-frequency ILDs are a unique feature of the proximal-region HRTFs which clearly differentiate them from distal-region HRTFs.

- (4) The acoustic bright spot directly opposite the location of the source causes a local minimum in the ILD near ± 90 degrees at intermediate frequencies (1500 and 3000 Hz). In the sphere plots, this minimum also occurs at higher frequencies, but in the KEMAR HRTFs the irregular shape of the head causes the bright spot to break into an erratic series of peaks and nulls which influence the ILD around 90 degrees. At 500 Hz, the bright spot does not influence the ILD.
- (5) As was the case with the HRTFs, the sphere model most accurately reflects the behavior of the ILDs measured with KEMAR at low frequencies. The ILD at 500 Hz is a smooth function both for the sphere model and for the KEMAR measurements. At higher frequencies, asymmetries of the KEMAR head cause its ILD to deviate from the sphere model. At these frequencies, the KEMAR ILDs are generally larger than the ILDs predicted by the sphere model. This discrepancy is primarily the result of the directional properties of the pinna, which provides substantial mid- to high-frequency gain at the ipsilateral ear.

B. Interaural time delays

The interaural time delay for a source in the horizontal plane was calculated at 0.125, 0.25, 0.50, 1, and 10 m with the sphere model and measured at 0.12, 0.25, 0.50, and 1 m with the KEMAR manikin (Fig. 10). These time delays were calculated directly from the average slope of the unwrapped phase from 100 to 6500 Hz of the interaural transfer function (the left HRTF divided by the right HRTF). The time delays from the sphere model are necessarily symmetric across both the median and frontal planes. At 10 m, the time delay peaks at approximately $700 \mu\text{s}$ at 90 degrees. As distance decreases, the magnitude of the time delay increases slightly. At 90 and -90 degrees, where the time delay increase is largest, it changes by about $100 \mu\text{s}$ as the distance decreases from 10 to 0.125 m. The majority of this magnitude increase occurs as the source moves from 0.25 to 0.125 m and the remainder from 0.50 to 0.25 m; there is virtually no dependence between the time delay and distance beyond 0.50 m.

The KEMAR ITDs from the measurements are generally similar to those from the sphere model, although there is some asymmetry between the front and rear hemispheres. Also, the ITDs with KEMAR exhibit a broader peak when the source is near the interaural axis. As with the time delays predicted by the sphere model, the KEMAR time delays increase by approximately $100 \mu\text{s}$ as distance decreases from 1 to 0.12 m for a lateral source.

There is, however, one important difference between the KEMAR ITDs and those predicted by the sphere model. In

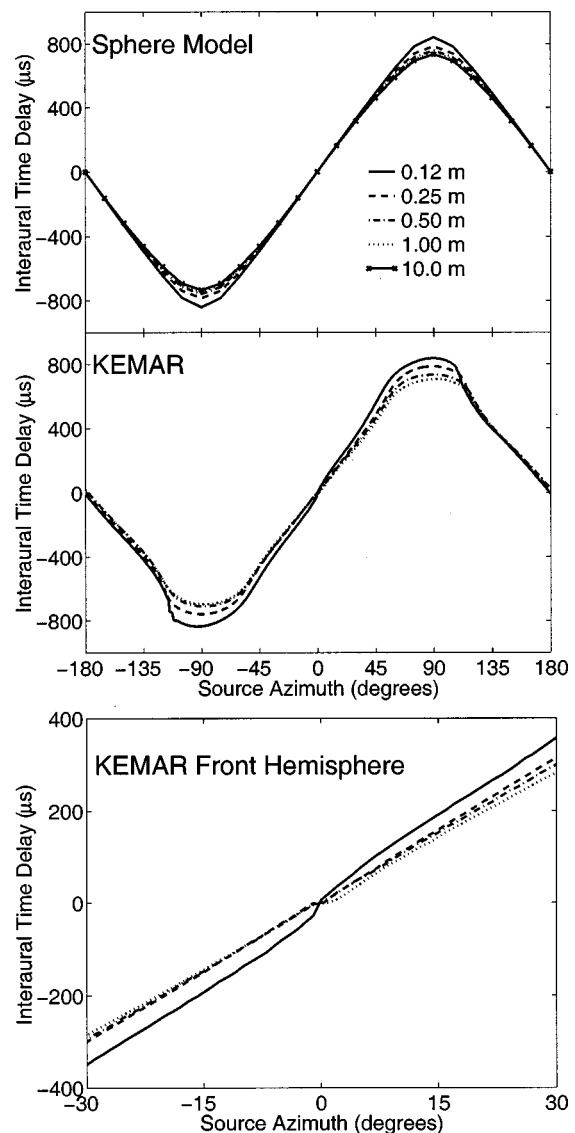


FIG. 10. The ITDs for the sphere model of the head (upper panel) and KEMAR (middle and lower panels). The delay was determined from the best linear fit of the unwrapped phase difference between the left and right ears (see text for details). The sphere model results assume a head diameter of 18 cm. The bottom panel is an enlarged view of the middle panel around 0 degrees. Positive delays indicate a lag at the right ear, and negative delays indicate a lag at the left ear.

the manikin measurements, the magnitudes of the ITDs increase substantially in the front hemisphere when the distance of the source decreases from 25 to 12 cm. The majority of this increase occurs within a few degrees of the median plane, resulting in a substantial increase in the slope of the ITD at 12 cm near 0 degrees azimuth (bottom panel of Fig. 10). The slope of the ITD as a function of azimuth increases from 11 to $17 \mu\text{s}/\text{degree}$ as a sound source directly in front of the listener moves from 25 to 12 cm. At locations more than 10 degrees from the median plane, the slope of the ITD as a function of azimuth is approximately the same at 12 and 25 cm, and the magnitude of the ITD at 12 cm remains approximately $50 \mu\text{s}$ greater than the ITD at 25 cm over a large portion of the front hemisphere. The increase in ITD for a very close source directly in front of the face is most likely caused by the interference of the nose with the propa-

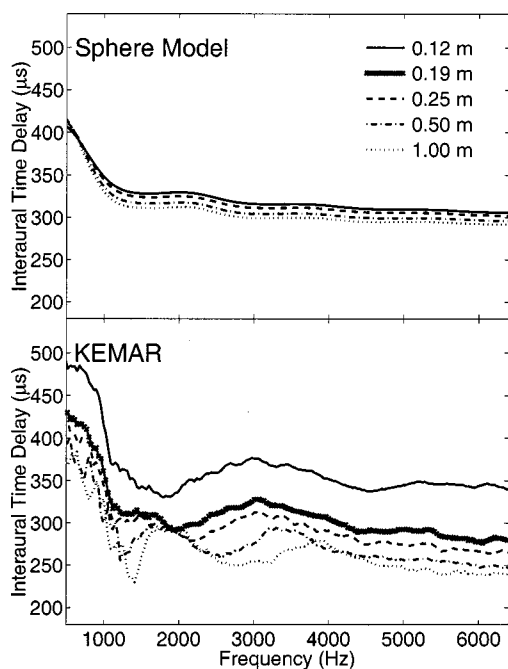


FIG. 11. The ITD (phase delay) for a source at 30 degrees azimuth calculated with the sphere model (upper panel) and measured with KEMAR (lower panel), as a function of frequency. Note that the plot at 19 cm is shown only for the KEMAR measurements.

gation of the sound around the head, since at 12 cm the sound source is located within 1 cm of the tip of the nose when the source is at 0 degrees. Since listeners are most sensitive to changes in ITD when the ITD is near zero (Hershkowitz and Durlach, 1969), it is likely that the increase in the rate of change of the ITD as a function of azimuth at locations directly in front of the face also results in increased sensitivity to changes in the azimuth of a sound source in that region.

The frequency-averaged values of the ITD shown in Fig. 10 provide a reasonable overall measure of the changes in the ITD with distance. However, the ITD is not constant across all frequencies, so it is useful to examine the frequency dependence of the ITD as a function of distance. The frequency dependence of the ITD for a nearby source at 30 degrees azimuth is shown in Fig. 11. These time delay values were calculated directly from the unwrapped interaural phase spectrum at each frequency. Both the sphere model calculations and the KEMAR measurements indicate that the time delay increases substantially at low frequencies (below 1500 Hz), a phenomenon reported by Kuhn (1977). In the sphere model, the time delay for a source at 1 m is $\approx 290 \mu\text{s}$ at high frequencies ($> 14 \text{ kHz}$) and $\approx 430 \mu\text{s}$ at low frequencies. Kuhn's model of ITD, based on Woodworth's geometrical model of the ITD, would predict an ITD of $278 \mu\text{s}$ at high frequencies and $420 \mu\text{s}$ at low frequencies for a distant sound source. Thus the sphere-model calculations are basically in agreement with the predictions of Kuhn's model. Note that the sphere-model calculations indicate very little change in the time delay with the distance of a sound source at 30 degrees azimuth.

The ITD measurements with the KEMAR manikin for a source at 30 degrees (bottom panel of Fig. 11) indicate a

much larger distance dependence in the ITD measurements. In particular, they indicate an increase in ITD of approximately $70 \mu\text{s}$ as the distance of the source decreases from 25 to 12 cm. In this plot, an additional measurement at 19 cm has been added to illustrate how quickly the time delay increases as the source gets close to the face: more than half of the total increase in ITD from 1 m to 12 cm occurs between 19 and 12 cm. Note that the frequency dependence of the ITD is largely independent of the distance of the source. The ITDs increase systematically at low frequencies at each source distance. The ITD measured with KEMAR at 1 m is similar to the ITD for a clothed KEMAR manikin, as reported by Kuhn (Kuhn, 1977). Both the sphere model and the KEMAR measurements indicate that the dependence of the interaural time delay on distance is substantially weaker than the dependence of the interaural level difference on distance. At the closest distances, the magnitude of the ILD increases dramatically, while the ITD generally increases by no more than 10%–12% (with the exception of very close sources directly in front of the face). The reason for this discrepancy is simple: the ITD depends (roughly) on the arithmetic difference between the source distance to the ipsilateral ear and the distance to the contralateral ear. The ILD (ignoring head-diffraction effects) depends on the ratio of these distances. As the source approaches the head, the ratio of distances to the ipsilateral and contralateral ears increases much faster than the arithmetic difference between the distances, so the ILD increases more dramatically than the ITD. As discussed later, this disparity between the distance dependence of ILD and ITD is even greater when perceptual factors are considered.

IV. ELEVATION DEPENDENCE OF THE PROXIMAL-REGION HRTF

The high-frequency characteristics of the HRTF at the ipsilateral ear change substantially with elevation. This is illustrated in Fig. 12, which shows the high-frequency HRTF at three elevation locations and three source distances. The pattern of peaks and notches in the transfer functions is substantially different at each elevation location. At -30 degrees elevation, for example, there is a notch at approximately 7 kHz in the HRTF that stretches across the entire ipsilateral hemisphere which is not found at either of the other two elevations. Similar patterns have been reported in previous HRTF studies (Carlile and Pralong, 1994; Shaw, 1974) and they will not be discussed in detail here. It is apparent, however, that these patterns do vary with elevation, and that they could provide a salient cue for evaluating the elevation of a sound source.

The HRTF patterns appear relatively independent of source distance. Specifically, in Fig. 12, the major features of the HRTFs are considerably more consistent across the three distances when elevation is fixed (within each row of the figure) than across the three elevations when distance is fixed (within each column). If this result generalizes to all elevations, it would imply that elevation cues are roughly independent of distance and that the same mechanisms that mediate elevation perception in the distal region may also be used in the proximal region.

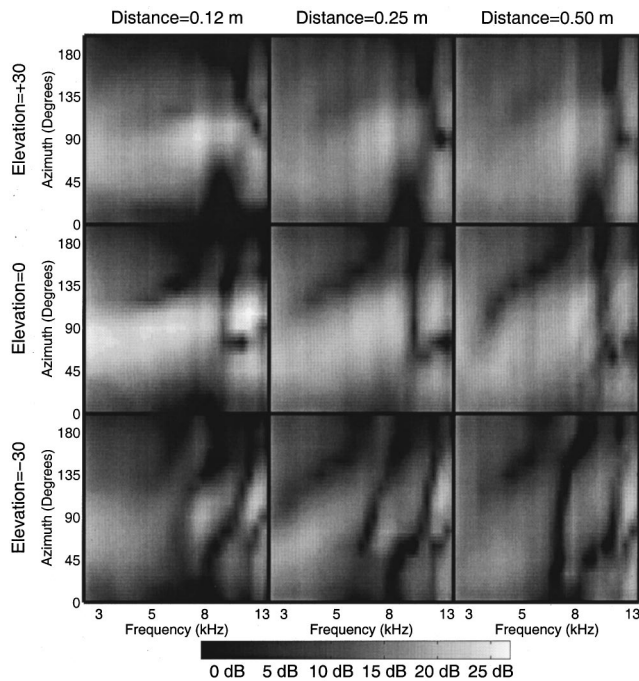


FIG. 12. Surface plots of the left-ear HRTFs measured with the KEMAR manikin at +30, 0, and -30 degrees in elevation. As in Fig. 8, the darkness of the plots indicates the magnitude of the HRTF at each point: brighter areas indicate greater gain in the HRTF. The plots are limited to high frequencies at the ipsilateral ear, where the most salient elevation cues occur, and are shown with 1/12-oct resolution in frequency and 15 degrees resolution in azimuth (interpolation has been used to smooth the plots).

V. PERCEPTUAL IMPLICATIONS OF THE PROXIMAL-REGION HRTFS

The distance-dependent attributes of the proximal-region HRTFs and consequent changes in interaural differences may have important perceptual implications. The ILDs increase rapidly as distance decreases, especially at distances less than 0.5 m, while the ITDs increase only slightly. Perceptual issues may further increase this disparity. Hershkowitz and Durlach (1969) found that listeners could discriminate changes in ILD on the order of 0.8 dB over a wide range of baseline ILDs. Therefore, the distance-dependent changes in ILD from 0.125 to 1 m (Fig. 13) span a range of up to 15 jnd's at 500 Hz and 30 or more jnd's at 3 kHz. In contrast, the jnd for ITD was approximately 15 μ s at ITDs below 400 μ s and increased rapidly for ITDs greater than 400 μ s (Hershkowitz and Durlach, 1969). The changes in ITD in the proximal region span a few jnd's, at most. Thus, subjects can be expected to perceive large changes in ILD as the distance of a nearby source changes while the perceived ITD should remain relatively constant.

The combination of perceptually invariant ITDs and strongly distance-dependent ILDs suggests a possible strategy for determining the distance of a nearby source in the horizontal plane based on binaural information from the HRTFs. The ITD information, which is relatively independent of distance, could be used to identify the azimuthal direction of the sound source. With source azimuth determined, ILD could then be used to estimate the distance of a sound source, provided the source is outside the median plane (where the ILD is near zero at all distances). The de-

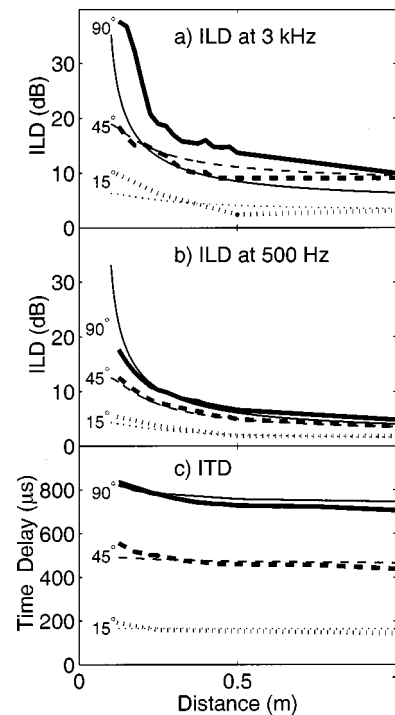


FIG. 13. Effects of source distance on interaural differences for the sphere model (thin line) and KEMAR (thick line). Results for (a) IID at 3 kHz, (b) IID at 500 Hz, and (c) ITD for three azimuths (15, 45, and 30 degrees) at 0-degree elevation.

pendence of ILDs on azimuth and distance predicts the following characteristics in proximal-region distance estimation performance:

- (1) Distance accuracy would be best for sources to the side of the listener and worst for sources directly in front, behind, or anywhere in the median plane.
- (2) At a fixed azimuth, the jnd in percentage distance would increase as distance increases. In Fig. 13, the slope of the curve relating ILD to distance increases substantially as the source approaches the head. Thus the percent decrease in distance necessary to produce a fixed increase in ILD decreases as distance decreases. If the jnd in distance is defined as the percent decrease in distance necessary to produce a single jnd in interaural intensity, the distance jnd will decrease with distance.
- (3) Provided a source is sufficiently broadband to allow the listener to perceive both ITD and ILD information, distance perception would not require information about the spectral shape or intensity of the source. The key advantage to binaural depth perception, in contrast to depth perception based on intensity, spectral cues, or reverberation, is that the binaural information is derived from the difference signal between the left and right ears and does not depend on the characteristics of the source (except its frequency range).

Note that this model is similar in concept to one suggested by Hirsch (1968), further explored by Greene (1968), and expanded by Molino (1970). Hirsch's model demonstrated the possibility of determining the distance of a sound source based on the relationship between the ILD and ITD. His model, which ignored diffraction by the head and as-

sumed that the ears were detectors in free space, predicted that distance could be calculated directly from the ratio of ITD to ILD. Molino's expanded model, based on a spherical head, required the azimuth location of the source to be known *a priori*. Greene and Molino used threshold data for ITD and ILD to calculate the predicted accuracy of distance perception using this model and, predictably, found that distance perception in the distal region would be very poor. Molino noted that predicted accuracy would be much greater in the proximal region, due to the dramatic increase in ILDs in that region. The present data indicate not only that the changes in ILD in the proximal region are large, and should be easily perceptible, but also that the relative invariance of the ITD in the proximal region may allow listeners to determine azimuth directly from the ITD, without any other external knowledge of source direction. Outside of the horizontal plane, the situation becomes more complex, but it is still possible to determine the azimuth, elevation, and distance of a sound source from the HRTFs. Recall that the KEMAR measurements at 30 and -30 degrees indicate that the high-frequency, elevation-dependent features in the HRTF, which are believed to be important elevation cues, are roughly independent of distance. If these high-frequency cues could be used to determine the elevation of the source, and compensate for the elevation-dependent changes in the ILD and ITD, the azimuth and distance of the source could still be accurately determined from interaural cues. This model would imply greater distance accuracy in the horizontal plane than in the median plane, since ILDs decrease as a source moves directly above or below a listener.

Another possible distance cue indicated by the HRTF measurements is a modest increase in the relative low-frequency gain as distance decreases. As a result, the signal reaching the eardrums from a nearby source is effectively low-pass filtered as the source approaches the head (Fig. 4). This low-pass filtering may explain the "darkening" of a very near sound source reported by von Békésy (1960). Begault (1987) also noted the "darkening" in the timbre of very near sound sources, and suggested that the tendency to perceive larger increases in loudness at low frequencies than at higher frequencies from an equivalent increase in sound pressure level (the so-called Fletcher–Munson curve) might provide an explanation. Since the pressure level increases at all frequencies as the source approaches the ear, the Fletcher–Munson curve suggests that the perceived increase in loudness would be greater at low frequencies. The "darkening" of proximal-region stimuli reported in the literature is probably a combination of the boost in low-frequency gain due to proximal-region acoustic effects shown in the HRTFs and the nonuniform perception of increasing loudness across frequency.

The implications of the proximal-region properties of the HRTF in directional localization are less obvious than those in distance localization. There is evidence that low-frequency time delays (Wightman and Kistler, 1992) dominate the perception of azimuth when they are available. Since time delays vary only slightly as distance decreases, it is not likely that azimuth perception will be significantly different in the proximal region than in the distal region

when the spectrum of the source extends into low frequencies. However, the increase in the slope of the ITD as a function of azimuth for nearby sources directly in front of the nose might increase the sensitivity to changes in the azimuth of the source in that region. Note that, when time delay information is not available from low-frequency time delays or high-frequency envelope delays (Middlebrooks *et al.*, 1989), it is possible that localization ability may degrade substantially. Without time delay information, there is no obvious mechanism for determining the relative contributions of distance and direction to the ILD. In other words, there is no way to determine whether a certain ILD is the result of a distant sound source near the interaural axis or a nearby sound source near the median plane. The consequences of this confusion on localization performance are unclear. The increase in ILDs at close distances will, however, decrease the jnd in azimuth to the extent that the jnd is limited by the change in ILD. The change in ILD with head orientation could also provide a strong dynamic distance cue when exploratory head motions are allowed. In addition, the lateral shift in the high-frequency pinna cues for nearby sources caused by the acoustic parallax effect may introduce a lateral bias if the auditory system is using these cues in the perception of azimuth.

The high-frequency pinna cues that are believed to determine elevation were found to be roughly independent of distance over the limited range of elevations measured. These cues are believed to dominate the perception of elevation, so it is unlikely that the localization of elevation is strongly dependent on distance in the proximal region.

VI. CONCLUSIONS

The major conclusions of this study can be summarized as follows:

- (1) The dominant distance-dependent feature of both the sphere model and KEMAR HRTFs in the proximal region is an increase in the interaural level difference with decreasing source distance across all frequencies. When a source is near the interaural axis, the change in ILD can span 20–30 dB (25 or more jnd's) as distance moves from 1 to 0.12 m, providing a potentially strong binaural distance cue in the proximal region. It is particularly interesting that large ILDs occur at low frequencies exclusively in the proximal region. In the distal region, low-frequency ILDs are small at all source directions.
- (2) In contrast to the ILD, the ITD is roughly independent of distance in the proximal region.
- (3) Both the sphere model and the KEMAR measurements indicate that the average low-frequency gain of the HRTF increases more rapidly than the average high-frequency gain as the source approaches the head. Thus, the sound reaching the ears is effectively low-pass filtered as the source approaches the head. This filtering may serve as a spectral distance cue in the proximal region.
- (4) There is a substantial parallax effect in the ipsilateral HRTF due to the discrepancy between the orientation of the source relative to the ear and the orientation of the

source relative to the head. This parallax effect causes a lateral shift of 20° or more in the high-frequency features of the HRTF as the source approaches the head.

- (5) For a fixed azimuth, the high-frequency characteristics of the HRTFs vary systematically with elevation but are relatively invariant to changes in distance. As a result, elevation localization may be similar in the proximal and distal regions.

In summary, it is clear that the distance-dependent attributes of the HRTF in the proximal region provide possible binaural distance cues which are unavailable for more distant sources. There are also substantial differences in the azimuth-dependent characteristics of the HRTF, particularly in the ILD, when the source is close the head, but their effects are probably mitigated by the relative insensitivity of the perceptually dominant interaural time delay to source distance. The high-frequency elevation-dependent characteristics of the HRTF also appear to be relatively independent of distance. Later papers in this series will discuss experiments measuring localization accuracy in the proximal region and whether listeners are able to make use of the available distance cues in that region.

ACKNOWLEDGMENTS

The authors would like to thank Nat Durlach, Steve Colburn, and William Peake for their assistance in the preparation of this manuscript. The work was also partially supported by AFOSR Grant Nos. F49620-96-1-0202 and F49620-98-1-0108.

¹In the past, we have used the terms “near field” and “far field” to describe these somewhat arbitrary regions. However, because the terms “near field” and “far field” have very specific meanings in physical acoustics (and have been used loosely to describe a number of other phenomena in other areas of acoustics), we have decided to eliminate any ambiguities about our terminology by introducing the terms “proximal region” and “distal region” and precisely defining them in the context of auditory localization.

²This chamber (10×10×10 m³) currently contains the Auditory Localization Facility (ALF). The ALF is a large, wire-frame geodesic sphere used in localization experiments at the Armstrong Laboratory (McKinley *et al.*, 1994). Because of the ALF, HRTF measurements were made in a corner of the chamber. Acoustic measurements with a free-field microphone verified that the presence of the ALF did not substantially affect the anechoic conditions for the measurements in this study.

Begault, D. R. (1987). “Control of Auditory Distance,” Ph.D. thesis, University of California, San Diego.
 Brungart, D. (1998). “Near-Field Auditory Localization.” Ph.D. thesis, Massachusetts Institute of Technology.
 Brungart, D., and Rabinowitz, W. (1996). “Auditory localization in the near-field,” in Proceedings of the Third International Conference on Auditory Display, Santa Fe Institute.

Brungart, D. S., Rabinowitz, W. M., and Durlach, N. (1996). “Auditory localization of a nearby point source,” *J. Acoust. Soc. Am.* **100**, 2593.
 Carlile, S., and Pralong, D. (1994). “The location dependent nature of perceptually salient features of the head-related transfer function,” *J. Acoust. Soc. Am.* **95**, 3445–3459.
 Duda, R., and Martens, W. (1998). “Range dependence of the response of a spherical head model,” *J. Acoust. Soc. Am.* **104**, 3048–3058.
 Firestone, F. (1930). “The phase difference and amplitude ratio at the ears due to a source of pure tone,” *J. Acoust. Soc. Am.* **2**, 260–270.
 Gardner, W., and Martin, K. (1995). “HRTF measurements of a KEMAR,” *J. Acoust. Soc. Am.* **97**, 3907–3908.
 Greene, D. (1968). “Comments on ‘Perception of the range of a sound source of unknown strength,’” *J. Acoust. Soc. Am.* **44**, 634.
 Hartley, R., and Fry, T. (1921). “The binaural location of pure tones,” *Phys. Rev.* **18**, 431–442.
 Hershkowitz, R., and Durlach, N. (1969). “Interaural time and amplitude JNDs for a 500-Hz tone,” *J. Acoust. Soc. Am.* **46**, 1464–1467.
 Hirsch, H. (1968). “Perception of the range of a sound source of unknown strength,” *J. Acoust. Soc. Am.* **43**, 373–374.
 Kuhn, G. (1977). “Model for the interaural time differences in the azimuthal plane,” *J. Acoust. Soc. Am.* **62**, 157–167.
 McKinley, R., Ericson, M., and D’Angelo, W. (1994). “Three dimensional audio displays: development, applications, and performance,” *Aviat., Space Environ. Med.* **65**, A31–A38.
 Mehrgardt, S., and Mellert, V. (1977). “Transformation characteristics of the external human ear,” *J. Acoust. Soc. Am.* **61**, 1567–1576.
 Middlebrooks, J., Makous, J. C., and Green, D. M. (1989). “Directional sensitivity of sound-pressure levels in the human ear canal,” *J. Acoust. Soc. Am.* **86**, 89–107.
 Molino, J. (1970). “Perceiving the range of a sound source when the direction is known,” *J. Acoust. Soc. Am.* **53**, 1301–1304.
 Møller, H., Sørensen, M., Hammershøi, D., and Jensen, C. B. (1995). “Head-related transfer functions of human subjects,” *J. Audio Engineering Society*, **43**, 300–320.
 Pralong, D., and Carlile, S. (1994). “Measuring the human head related transfer functions: A novel method for the construction and calibration of a miniature ‘in ear’ recording system,” *J. Acoust. Soc. Am.* **95**, 3435–3444.
 Rabinowitz, W., 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**, 125–129.
 Shaw, E. (1974). “Transformation of sound pressure level from the free field to the eardrum in the horizontal plane,” *J. Acoust. Soc. Am.* **56**, 1848–1861.
 Stewart, G. (1911a). “The acoustic shadow of a rigid sphere with certain applications in architectural acoustics and audition,” *Phys. Rev.* **33**, 467–479.
 Stewart, G. (1911b). “Phase relations in the acoustic shadow of a rigid sphere; phase difference at the ears,” *Phys. Rev.* **34**, 252–258.
 Suzuki, Y., Kim, H.-Y., Takane S., and Sone, T. (1998). “A modeling of distance perception based on auditory parallax model,” in Proceedings of the 16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America.
 von Békésy, G. (1960). *Experiments in Hearing* (Wiley, New York).
 Wightman, F., and Kistler, D. (1989). “Headphone simulation of free-field listening. I: Stimulus synthesis,” *J. Acoust. Soc. Am.* **85**, 858–877.
 Wightman, F., and Kistler, D. (1992). “The dominant role of low-frequency interaural time differences in sound localization,” *J. Acoust. Soc. Am.* **91**, 1648–1660.