How We Localize Sound

William M. Hartmann

Citation: Physics Today **52**, 11, 24 (1999); doi: 10.1063/1.882727
View online: https://doi.org/10.1063/1.882727
View Table of Contents: https://physicstoday.scitation.org/toc/pto/52/11
Published by the American Institute of Physics

How We Localize Sound

Relying on a variety of cues, including intensity, timing, and spectrum, our brains recreate a three-dimensional image of the acoustic landscape from the sounds we hear.

William M. Hartmann

For as long as we humans have lived on Earth, we have been able to use our ears to localize the sources of sounds. Our ability to localize warns us of danger and helps us sort out individual sounds from the usual cacophony of our acoustical world. Characterizing this ability in humans and other animals makes an intriguing physical, physiological, and psychological study (see figure 1).

John William Strutt (Lord Rayleigh) understood at least part of the localization process more than 120 years ago.[1] He observed that if a sound source is to the right of the listener's forward direction, then the left ear is in the shadow cast by the listener's head. Therefore, the signal in the right ear should be more intense than the signal in the left one, and this difference is likely to be an important clue that the sound source is located on the right.

Interaural level difference

The standard comparison between intensities in the left and right ears is known as the interaural level difference (ILD). In the spirit of the spherical cow, a physicist can estimate the size of the effect by calculating the acoustical intensity at opposite poles on the surface of a sphere, given an incident plane wave, and then taking the ratio. The level difference is that ratio expressed in decibels.

As shown in figure 2, the ILD is a strong function of frequency over much of the audible spectrum (canonically quoted as 20–20 000 Hz). That is because sound waves are effectively diffracted when their wavelength is longer than the diameter of the head. At a frequency of 500 Hz, the wavelength of sound is 69 cm—four times the diameter of the average human head. The ILD is therefore small for frequencies below 500 Hz, as long as the source is more than a meter away. But the scattering by the head increases rapidly with increasing frequency, and at 4000 Hz the head casts a significant shadow.

Ultimately, the use of an ILD, small or large, depends on the sensitivity of the central nervous system to such differences. In evolutionary terms, it would make sense if the sensitivity of the central nervous system would somehow reflect the ILD values that are actually physically present. In fact, that does not appear to be the case. Psychoacoustical experiments find that the central nervous system is about equally sensitive at all frequencies. The smallest detectable change in ILD is approximately 0.5 dB, no matter what the frequency.[2] Therefore the ILD is a potential localization cue at any frequency where it is physically greater than a decibel. It is as though Mother Nature knew in advance that her offspring would walk around the planet listening to portable music through headphones.

BILL HARTMANN *is a professor of physics at Michigan State University in East Lansing, Michigan (hartmann@pa.msu.edu; http://www.pa.msu.edu/acoustics). He is the author of the textbook* Signals, Sound, and Sensation *(AIP Press, 1997).*

The spherical-head model is obviously a simplification. Human heads include a variety of secondary scatterers that can be expected to lead to structure in the higher-frequency dependence of the ILD. Conceivably, this structure can serve as an additional cue for sound localization. As it turns out, that is exactly what happens, but that is another story for later in this article.

In the long-wavelength limit, the spherical-head model correctly predicts that the ILD should become uselessly small. If sounds are localized on the basis of ILD alone, it should be very difficult to localize a sound with a frequency content that is entirely below 500 Hz. It therefore came as a considerable surprise to Rayleigh to discover that he could easily localize a steady-state low-frequency pure tone such as 256 or 128 Hz. Because he knew that localization could not be based on ILD, he finally concluded in 1907 that the ear must be able to detect the difference in waveform phases between the two ears.[3]

Interaural time difference

For a pure tone like Rayleigh used, a difference in phases is equivalent to a difference in arrival times of waveform features (such as peaks and positive-going zero crossings) at the two ears. A phase difference $\Delta\phi$ corresponds to an interaural time difference (ITD) of $\Delta t = \Delta\phi/(2\pi f)$ for a tone with frequency f. In the long-wavelength limit, the formula for diffraction by a sphere[4] gives the interaural time difference Δt as a function of the azimuthal (left–right) angle θ:

$$\Delta t = \frac{3a}{c} \sin\theta, \qquad (1)$$

where a is the radius of the head (approximately 8.75 cm) and c is the speed of sound (34 400 cm/s). Therefore, $3a/c = 763\ \mu$s.

Psychoacoustical experiments show that human listeners can localize a 500 Hz sine tone with considerable accuracy. Near the forward direction (θ near zero), listeners are sensitive to differences $\Delta\theta$ as small as 1–2°. The idea that this sensitivity is obtained from an ITD initially seems rather outrageous. A 1° difference in azimuth corresponds to an ITD of only 13 μs. It hardly seems possible that a neural system, with synaptic delays on the order of a millisecond, could successfully encode such small time differences. However, the auditory system, unaware of such mathematical niceties, goes ahead and does it anyway. This ability can be proved in headphone experiments, in which the ITD can be presented independently of the ILD. The key to the brain's success in this case is parallel processing. The binaural system apparently beats the unfavorable timing dilemma by transmitting timing information through many neurons. Estimates of the number of neurons required, based on statistical decision theory, have ranged from 6 to 40 for each one-third-octave frequency band.

There remains the logical problem of just how the

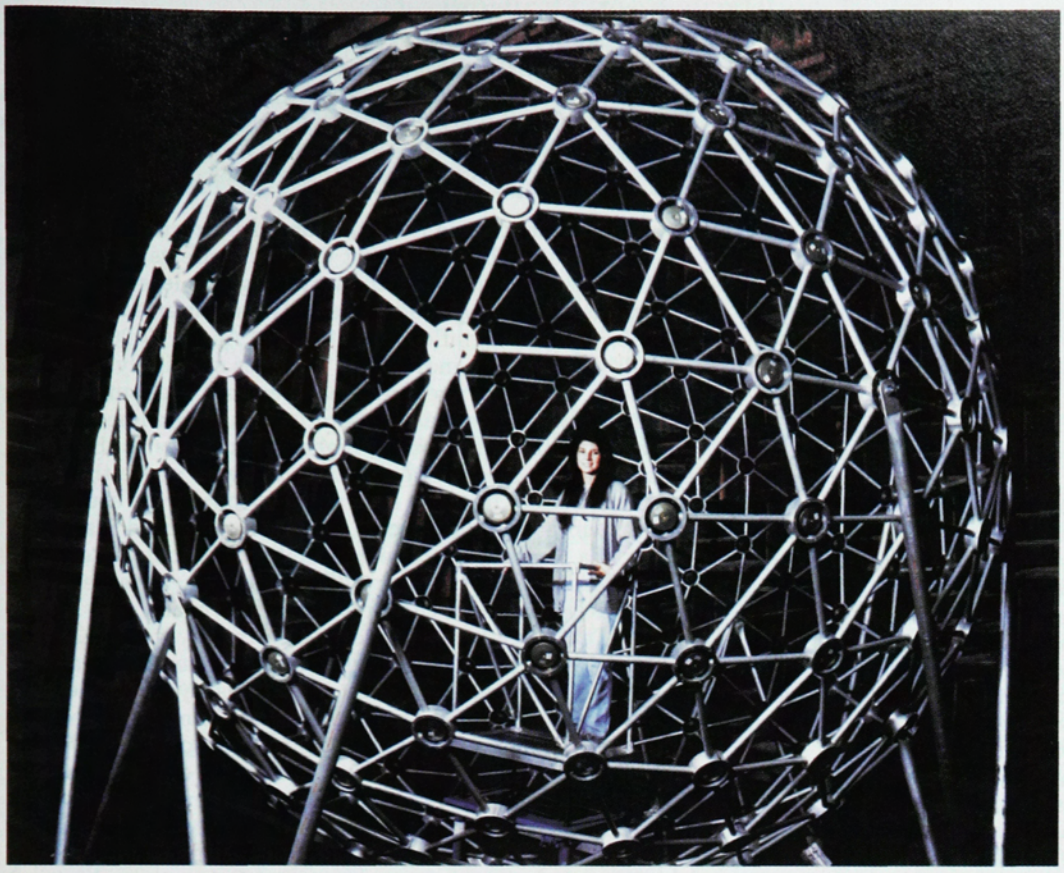


FIGURE 1. THE SOUND LOCALIZATION FACILITY at Wright Patterson Air Force Base in Dayton, Ohio, is a geodesic sphere, nearly 5 m in diameter, housing an array of 277 loudspeakers. Each speaker has a dedicated power amplifier, and the switching logic allows the simultaneous use of as many as 15 sources. The array is enclosed in a 6 m cubical anechoic room: Foam wedges 1.2 m long on the walls of the room make the room strongly absorbing for wavelengths longer than 5 m, or frequencies above 70 Hz. Listeners in localization experiments indicate perceived source directions by placing an electromagnetic stylus on a small globe. (Courtesy of Mark Ericson and Richard McKinley.)

auditory system manages to use ITDs. There is now good evidence that the superior olive—a processing center, or "nucleus," in the midbrain—is able to perform a cross-correlation operation on the signals in the two ears, as described in the box on page 27.

The headphone experiments with an ITD give the listener a peculiar experience. The position of the image is located to the left or right as expected, depending on the sign of the ITD, but the image seems to be *within* the listener's head—it is not perceived to be in the real external world. Such an image is said to be "lateralized" and not localized. Although the lateralized headphone sensation is quite different from the sensation of a localized source, experiments show that lateralization is intimately connected to localization.

Using headphones, one can measure the smallest detectable change in ITD as a function of the ITD itself. These ITD data can be used with equation 1 to predict the smallest detectable change in azimuth $\Delta\theta$ for a real source as a function of θ. When the actual localization experiment is done with a real source, the results agree with the predictions, as is to be expected if the brain relies on ITDs to make decisions about source location.

Like any phase-sensitive system, the binaural phase detector that makes possible the use of ITDs suffers from phase ambiguity when the wavelength is comparable to the distance between the two measurements. This problem is illustrated in figure 3. The equivalent temporal viewpoint is that, to avoid ambiguity, a half period of the wave must be longer than the delay between the ears. When the delay is exactly half a period, the signals at the two ears are exactly out of phase and the ambiguity is complete. For shorter periods, between twice the delay and the delay itself, the ITD leads to an apparent source location that is on the opposite side of the head compared to the true location. It would be better to have no ITD sensitivity at all than to have a process that gives such misleading answers. In fact, the binaural system solves this problem in what appears to be the best possible way: The binaural system rapidly loses sensitivity to any ITD at all as the frequency of the wave increases from 1000 to 1500 Hz—exactly the range in which the interaural phase difference becomes ambiguous.

One might imagine that the network of delay lines and coincidence detectors described in the box vanishes at frequencies greater than about 1500 Hz. Such a model would be consistent with the results of pure-tone experiments, but it would be wrong. In fact, the binaural system can successfully register an ITD that occurs at a high frequency such as 4000 Hz, if the signal is modulated. The modulation, in turn, must have a rate that is less than about 1000 Hz. Therefore, the failure of the binaural timing system to process sine tones above 1500 Hz cannot be thought of as a failure of the binaural neurons tuned to high frequency. Instead, the failure is best described in the temporal domain, as an inability to track rapid variations.

To summarize the matter of binaural differences, the physiology of the binaural system is sensitive to amplitude cues from ILDs at any frequency, but for incident plane waves, ILD cues exist physically only for frequencies above about 500 Hz. They become large and reliable for frequencies above 3000 Hz, making ILD cues most effective at high frequencies. In contrast, the binaural physiology is capable of using phase information from ITD cues only at low frequencies, below about 1500 Hz. For a sine tone of intermediate frequency, such as 2000 Hz, neither cue works well. As a result, human localization ability tends to be poor for signals in this frequency region.

The inadequacy of binaural difference cues

The binaural time and level differences are powerful cues for the localization of a source, but they have important limitations. Again, in the spherical-head approximation,

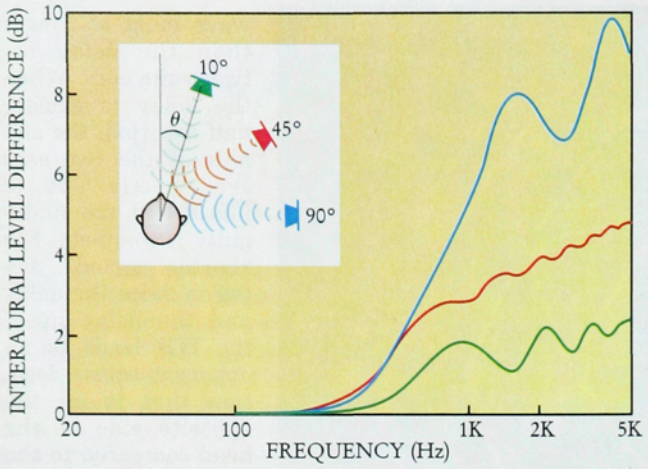


FIGURE 2. INTERAURAL LEVEL DIFFERENCES, calculated for a source in the azimuthal plane defined by the two ears and the nose. The source radiates frequency f and is located at an azimuth θ of 10° (green curve), 45° (red), or 90° (blue) with respect to the listener's forward direction. The calculations assume that the ears are at opposite poles of a rigid sphere.

the inadequacy of interaural differences is evident because, for a source of sound moving in the midsagittal plane (the perpendicular bisector of a line drawn through both ears), the signals to left and right ears—and therefore binaural differences—are the same. As a result, the listener with the hypothetical spherical head cannot distinguish between sources in back, in front, or overhead. Because of a fine sensitivity to binaural differences, this listener can detect displacements of only a degree side-to-side, but cannot tell back from front! This kind of localization difficulty does not correspond to our usual experience.

There is another problem with this binaural difference model: If a tone or broadband noise is heard through headphones with an ITD, an ILD, or both, the listener has the impression of laterality—coming from the left or right—as expected, but, as previously mentioned, the sound image appears to be within the head, and it may also be diffuse and fuzzy instead of compact. This sensation, too, is unlike our experience of the real world, in which sounds are perceived to be externalized. The resolution of front–back confusion and the externalization of sound images turn on another sound localization cue, the anatomical transfer function.

The anatomical transfer function

Sound waves that come from different directions in space are differently scattered by the listener's outer ears, head, shoulders, and upper torso. The scattering leads to an acoustical filtering of the signals appearing at left and right ears. The filtering can be described by a complex response function—the anatomical transfer function (ATF), also known as the head-related transfer function (HRTF). Because of the ATF, waves that come from behind tend to be boosted in the 1000 Hz frequency region, where-

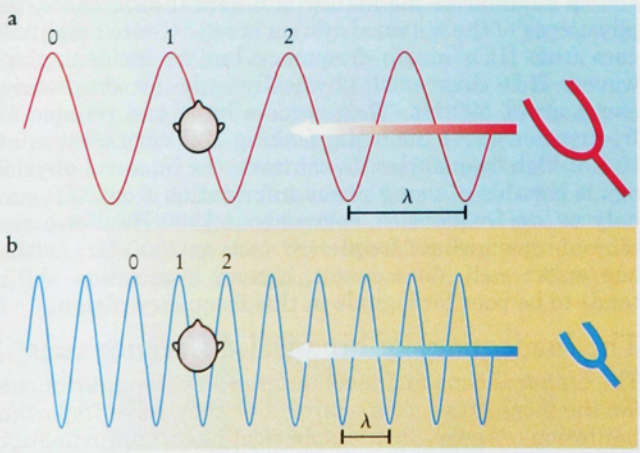


FIGURE 3. INTERAURAL TIME DIFFERENCES, given by the difference in arrival times of waveform features at the two ears, are useful localization cues only for long wavelengths. In (**a**), the signal comes from the right, and waveform features such as the peak numbered 1 arrive at the right ear before arriving at the left. Because the wavelength is greater than twice the head diameter, no confusion is caused by other peaks of the waveform, such as peaks 0 or 2. In (**b**), the signal again comes from the right, but the wavelength is shorter than twice the head diameter. As a result, every feature of cycle 2 arriving at the right ear is immediately preceded by a corresponding feature from cycle 1 at the left ear. The listener naturally concludes that the source is on the left, contrary to fact.

as waves that come from the forward direction are boosted near 3000 Hz. The most dramatic effects occur above 4000 Hz: In this region, the wavelength is less than 10 cm and details of the head, especially the outer ears, or pinnae, become significant scatterers. Above 6000 Hz, the ATF for different individuals becomes strikingly individualistic, but there are a few features that are found rather generally. In most cases, there is a valley-and-peak structure that tends to move to higher frequencies as the elevation of the source increases from below to above the head. For example, figure 4 shows the spectrum for sources in front, in back, and directly overhead, measured inside the ear of a Knowles Electronics Manikin for Acoustic Research (KEMAR). The peak near 7000 Hz is thought to be a particularly prominent cue for a source overhead.

The direction-dependent filtering by the anatomy, used by listeners to resolve front–back confusion and to determine elevation, is also a necessary component of externalization. Experiments further show that getting the ATF correct with virtual reality techniques is *sufficient* to externalize the image. But there is an obvious problem in the application of the ATF. *A priori*, there is no way that a listener can know if a spectrally prominent feature comes from direction-dependent filtering or whether it is part of the original source spectrum. For instance, a signal with a strong peak near 7000 Hz may not necessarily come from above—it might just come from a source that happens to have a lot of power near 7000 Hz.

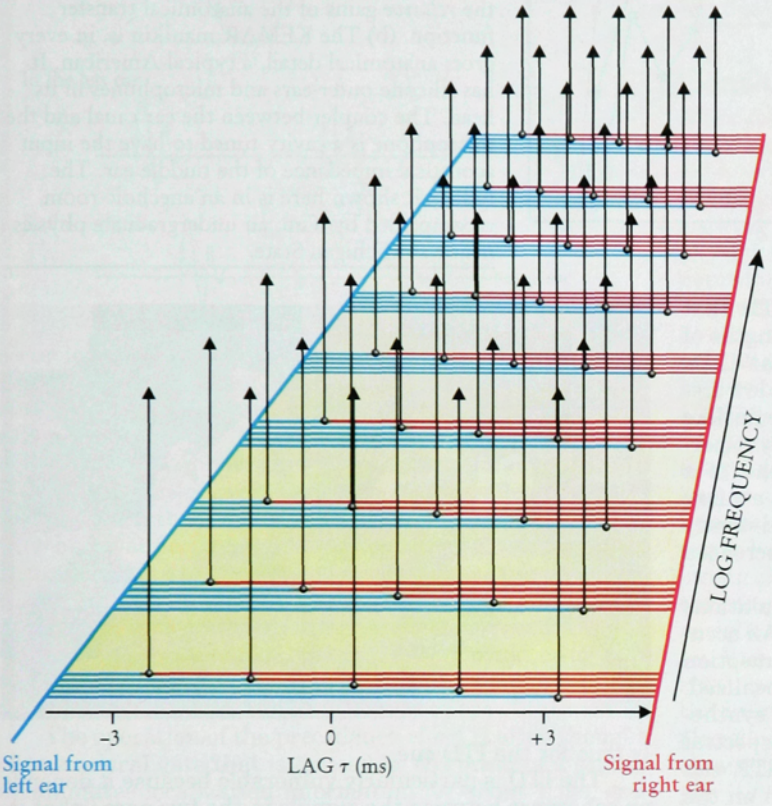
Confusion of this kind between the source spectrum and the ATF immediately appears with narrow-band sources such as pure tones or noise bands having a bandwidth of a few semitones. When a listener is asked to say whether a narrow-band sound comes from directly in front, in back, or overhead, the answer will depend entirely on the frequency of the sound—the true location of the sound source is irrelevant.[5] Thus, for narrow-band sounds, the confusion between source spectrum and location is complete. The listener can solve this localization problem only by turning the head so that the source is no longer in the midsagittal plane. In an interesting variation on this theme, Frederic Wightman and Doris Kistler at the University of Wisconsin—Madison have shown that it is not enough if the source itself moves—the listener will still be confused about front and back. The confusion can be resolved, though, if the listener is in control of the source motion.[6]

Fortunately, most sounds of the everyday world are

The Binaural Cross-Correlation Model

In 1948, Lloyd Jeffress proposed that the auditory system processes interaural time differences by using a network of neural delay lines terminating in e–e neurons.[10] An e–e neuron is like an AND gate, responding only if excitation is present on both of two inputs (hence the name "e–e"). According to the Jeffress model, one input comes from the left ear and the other from the right. Inputs are delayed by neural delay lines so that different e–e cells experience a coincidence for different arrival times at the two ears.

An illustration of how the network is imagined to work is shown in the figure. An array of e–e cells is distributed along two axes: frequency and neural internal delay. The frequency axis is needed because binaural processing takes place in tuned channels. These channels represent frequency analysis—the first stage of auditory processing. Any plausible auditory model must contain such channels.



Inputs from left ear (blue) and right ear (red) proceed down neural delay lines in each channel and coincide at the e–e cells for which the neural delay τ exactly compensates for the fact that the signal started at one ear sooner than the other. For instance, if the source is off to the listener's left, then signals start along the delay lines sooner from the left side. They coincide with the corresponding signals from the right ear at neurons to the right of $\tau = 0$, that is, at a positive value of τ. The coincidence of neural signals causes the e–e neurons to send spikes to higher processing centers in the brain.

The expected value for the number of coincidences N_c at the e–e cell specified by delay τ is given in terms of the rates $P_L(t)$ and $P_R(t)$ of neural spikes from left and right ears by the convolution-like integral

$$N_c(\tau) = T_W \int_0^{T_S} dt'\, P_L(t')\, P_R(t'+\tau),$$

where T_W is the width of the neuron's coincidence window and T_S is the duration of the stimulus.[11] Thus, N_c is the cross correlation between signals in the left and right ears. Neural delay and coincidence circuits of just this kind have been found in the superior olive in the midbrain of cats.[12]

broadband and relatively benign in their spectral variation, so that listeners can both localize the source and identify it on the basis of the spectrum. It is still not entirely clear how this localization process works. Early models of the process that focused on particular spectral features (such as the peak at 7000 Hz for a source overhead) have given way, under the pressure of recent research, to models that employ the entire spectrum.

The experimental art

Most of what we know about sound localization has been learned from experiments using headphones. With headphones, the experimenter can precisely control the stimulus heard by the listener. Even experiments done on cats, birds, and rodents have these creatures wearing miniature earphones.

In the beginning, much was learned about fundamental binaural capabilities from headphone experiments with simple differences in level and arrival time for tones of various frequencies and noises of various compositions.[7] However, work on the larger question of sound localization had to await several technological developments to achieve an accurate rendering of the ATF in each ear. First were the acoustical measurements themselves, done with tiny probe microphones inserted in the listener's ear canals to within a few millimeters of the eardrums. Transfer functions measured with these microphones allowed experimenters to create accurate simulations of the real world using headphones, once the transfer functions of the microphones and headphones themselves had been compensated by inverse filtering.

Adequate filtering requires fast, dedicated digital signal processors linked to the computer that runs experiments. The motion of the listener's head can be taken into account by means of an electromagnetic head tracker. The head tracker consists of a stationary transmitter, whose three coils produce low-frequency magnetic fields, and a receiver, also with three coils, that is mounted on the listener's head. The tracker gives a reading of all six degrees of freedom in the head motion, 60 times per second. Based on the motion of the head, the controlling computer directs the fast digital processor to refilter the signals to the ears so that the auditory scene is stable and realistic. This virtual reality technology is capable of synthesizing a convincing acoustical environment. Starting with a simple monaural recording of a conversation, the experimenter can place the individual talkers in space. If the listener's head turns to face a talker, the auditory image remains constant, as it does in real life. What is most important for the psychoacoustician, this technology has opened a large new territory for controlled experiments.

Making it wrong

With headphones, the experimenter can create conditions not found in nature to try to understand the role of different localization mechanisms. For instance, by introducing

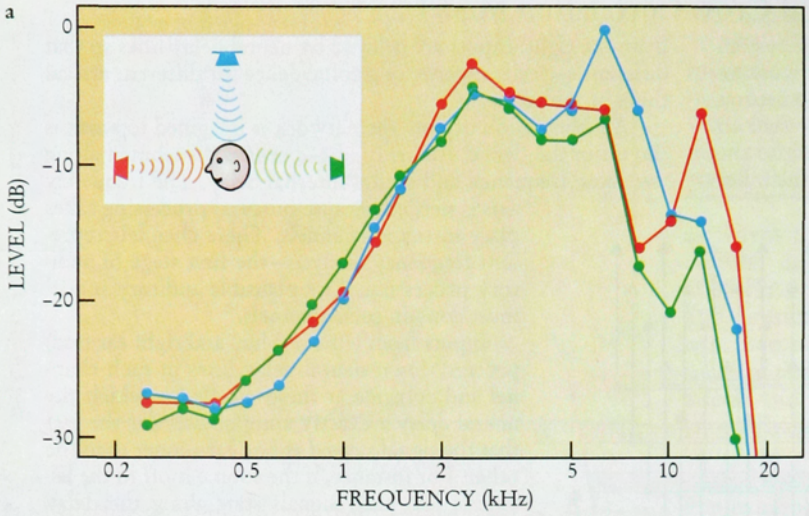


FIGURE 4. THE ANATOMICAL TRANSFER function, which incorporates the effects of secondary scatterers such as the outer ears, assists in eliminating front–back confusion. **(a)** The curves show the spectrum of a small loudspeaker as heard in the left ear of a manikin when the speaker is in front (red), overhead (blue), and in back (green). A comparison of the curves reveals the *relative* gains of the anatomical transfer function. **(b)** The KEMAR manikin is, in every gross anatomical detail, a typical American. It has silicone outer ears and microphones in its head. The coupler between the ear canal and the microphone is a cavity tuned to have the input acoustical impedance of the middle ear. The KEMAR shown here is in an anechoic room accompanied by Tim, an undergraduate physics major at Michigan State.

an ILD that points to the left opposed by an ITD that points to the right, one can study the relative strengths of these two cues. Not surprisingly, it is found that ILDs dominate at high frequency and ITDs dominate at low frequency. But perception is not limited to just pointlike localization; it also includes size and shape. Rivalry experiments such as contradictory ILDs and ITDs lead to a source image that is diffuse: The image occupies a fuzzy region within the head that a listener can consistently describe. The effect can also be measured as an increased variance in lateralization judgements.

Incorporating the ATF into headphone simulations considerably expands the menu of bizarre effects. An accurate synthesis of a broadband sound leads to perception that is like the real world: Auditory images are localized, externalized, and compact. Making errors in the synthesis, for example progressively zeroing the ITD of spectral lines while retaining the amplitude part of the ATF, can cause the image to come closer to the head, push on the face, and form a blob that creeps into the ear canal and finally enters the head. The process can be reversed by progressively restoring accurate ITD values.[8]

A wide variety of effects can occur, by accident or design, with inaccurate synthesis. There are a few general rules: Inaccuracies tend to expand the size of the image, put the images inside the head, and produce images that are in back rather than in front. Excellent accuracy is required to avoid front–back confusion. The technology permits a listener to hear the world with someone else's ears, and the usual result is an increase in confusion about front and back. Reduced accuracy often puts all source images in back, although they are nevertheless externalized. Further reduction in accuracy puts the images inside the back of the head.

Rooms and reflections

The operations of interaural level and time difference cues and of spectral cues have normally been tested with headphones or by sound localization experiments in anechoic rooms, where all the sounds travel in a straight path from the source to the listener. Most of our everyday listening, however, is done in the presence of walls, floors, ceilings, and other large objects that reflect sound waves. These reflections result in dramatic physical changes to the waveforms. It is hard to imagine how the reflected sounds, coming from all directions, can contribute anything but random variation to the cues used in localization. Therefore, it is expected that the reflections and reverberation introduced by the room are inevitably for the worse as far as sound localization is concerned. That is especially true for the ITD cue.

The ITD is particularly vulnerable because it depends on coherence between the signals in the two ears—that is, the height of the cross-correlation function, as described in the box on page 27. Reverberated sound contains no useful coherent information, and in a large room where reflected sound dominates the direct sound, the ITD becomes unreliable.

By contrast, the ILD fares better. First, as shown by headphone experiments, the binaural comparison of intensities does not care whether the signals are binaurally coherent or not. Such details of neural timing appear to be stripped away as the ILD is computed. Of course, the ILD accuracy is adversely affected by standing waves in a room, but here the second advantage of the ILD appears: Almost every reflecting surface has the property that its acoustical absorption increases with increasing frequency; as a result, the reflected power becomes relatively smaller compared to the direct power. Because the binaural neurophysiology is capable of using ILDs across the audible spectrum with equal success, it is normally to the listener's advantage to use the highest frequency information that can be heard. Experiments in highly reverberant environments find listeners doing exactly that, using cues above 8000 Hz. A statistical decision theory analysis using ILDs and ITDs measured with a manikin shows that the pattern of localization errors observed experimentally can be understood by assuming that listeners rely entirely on ILDs and not at all on ITDs. This strategy of reweighting localization cues is entirely unconscious.

The precedence effect

There is yet another strategy that listeners unconsciously employ to cope with the distorted localization cues that

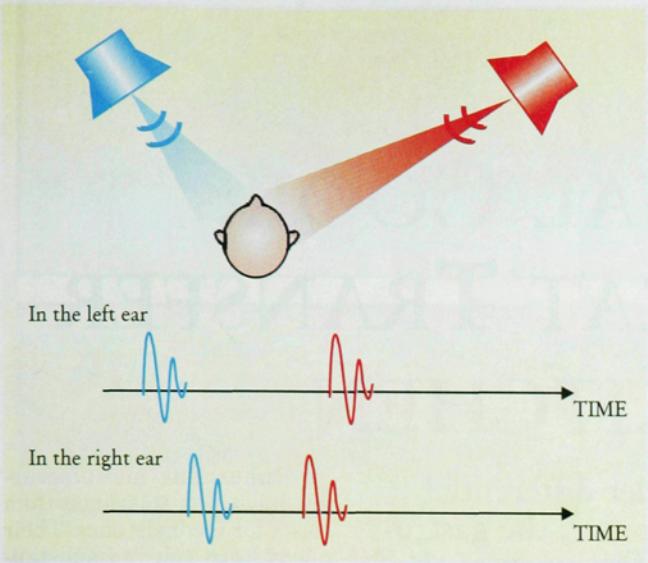


FIGURE 5. PRECEDENCE EFFECT demonstration with two loudspeakers reproducing the same pulsed wave. The pulse from the left speaker leads in the left ear by a few hundred microseconds, suggesting that the source is on the left. The pulse from the right speaker leads in the right ear by a similar amount, which provides a contradictory localization cue. Because the listener is closer to the left speaker, the left pulse arrives sooner and wins the competition—the listener perceives just one single pulse coming from the left.

occur in a room: They make their localization judgments instantly based on the earliest arriving waves in the onset of a sound. This strategy is known as the precedence effect, because the earliest arriving sound wave—the direct sound with accurate localization information—is given precedence over the subsequent reflections and reverberation that convey inaccurate information. Anyone who has wandered around a room trying to locate the source of a pure tone without hearing the onset can appreciate the value of the effect. Without the action of the precedence effect on the first arriving wave, localization is virtually impossible. There is no ITD information of any use, and, because of standing waves, the loudness of the tone is essentially unrelated to the nearness of the source.

The operation of the precedence effect is often thought of as a neural gate that is opened by the onset of a sound, accumulates localization information for about 1 ms, and then closes to shut off subsequent localization cues. This operation appears dramatically in experiments where it is to the listener's advantage to attend to the subsequent cues but the precedence effect prevents it. An alternative model regards precedence as a strong reweighting of localization cues in favor of the earliest sound, because the subsequent sound is never entirely excluded from the localization computation.

Precedence is easily demonstrated with a standard home stereo system set for monophonic reproduction, so that the same signal is sent to both loudspeakers. Standing midway between the speakers, the listener hears the sound from a forward direction. Moving half a meter closer to the left speaker causes the sound to appear to come entirely from that speaker. The analysis of this result is that each speaker sends a signal to both ears. Each speaker creates an ILD and—of particular importance—an ITD, and these cues compete, as shown in figure 5. Because of the precedence effect, the first sound (from the left speaker) wins the competition, and the listener perceives the sound as coming from the left. But although the sound appears to come from the left speaker alone, the right speaker continues to contribute loudness and a sense of spatial *extent*. This perception can be verified by suddenly unplugging the right speaker—the difference is immediately apparent. Thus, the precedence effect is restricted to the formation of a single fused image with a definite location. The precedence effect appears not to depend solely on interaural differences; it operates also on the spectral differences caused by anatomical filtering for sources in the midsagittal plane.[9]

Conclusions and conjectures

After more than a century of work, there is still much about sound localization that is not understood. It remains an active area of research in psychoacoustics and in the physiology of hearing. In recent years, there has been growing correspondence between perceptual observations, physiological data on the binaural processing system, and neural modeling. There is good reason to expect that next year we will understand sound localization better than we do this year, but it would be wrong to think that we have only to fill in the details. It is likely that next year will lead to a qualitatively improved understanding with models that employ new ideas about neural signal processing.

In this environment, it is risky to conjecture about future development, but there are trends that give clues. Just a decade ago, it was thought that much of sound localization in general, and precedence in particular, might be a direct result of interaction at early stages of the binaural system, as in the superior olive. Recent research suggests that the process is more widely distributed, with peripheral centers of the brain such as the superior olive sending information—about ILD, about ITD, about spectrum, and about arrival order—to higher centers where the incoming data are evaluated for self-consistency and plausibility, and are probably compared with information obtained visually. Therefore, sound localization is not simple; it is a large mental computation. But as the problem has become more complicated, our tools for studying it have become better. Improved psychophysical techniques for flexible synthesis of realistic stimuli, physiological experiments probing different neural regions simultaneously, faster and more precise methods of brain imaging, and more realistic computational models will one day solve this problem of how we localize sound.

The author is grateful to his colleagues Brad Rakerd, Tim McCaskey, Zachary Constan, and Joseph Gaalaas for help with this article. His work on sound localization is supported by the National Institute on Deafness and Other Communication Disorders, one of the National Institutes of Health.

References

1. J. W. Strutt (Lord Rayleigh), Phil. Mag. **3**, 456 (1877).
2. W. A. Yost, J. Acoust. Soc. Am. **70**, 397 (1981).
3. J. W. Strutt (Lord Rayleigh), Phil. Mag. **13**, 214 (1907).
4. G. F. Kuhn, J. Acoust. Soc. Am. **62**, 157 (1977).
5. J. Blauert, *Spatial Hearing*, 2nd ed., J. S. Allen, trans., MIT Press, Cambridge, Mass. (1997).
6. F. L. Wightman, D. J. Kistler, J. Acoust. Soc. Am. **105**, 2841 (1999).
7. N. I. Durlach, H. S. Colburn, in *Handbook of Perception*, vol. 4, E. Carterette, M. P. Friedman, eds., Academic, New York (1978).
8. W. M. Hartmann, A. T. Wittenberg, J. Acoust. Soc. Am. **99**, 3678 (1996).
9. R. Y. Litovsky, B. Rakerd, T. C. T. Yin, W. M. Hartmann, J. Neurophysiol. **77**, 2223 (1997).
10. L. A. Jeffress, J. Comp. Physiol. Psychol. **41**, 35 (1948).
11. R. M. Stern, H. S, Colburn, J. Acoust. Soc. Am. **64**, 127 (1978).
12. T. C. T. Yin, J. C. K. Chan, J. Neurophysiol. **64**, 465 (1990). ■