Acoustic cues underlying auditory distance in barn owls

DUCK O. KIM¹, ANDREW MOISEFF², J. BRADLEY TURNER³ & JUSTIN GULL³

¹Department of Neuroscience, University of Connecticut Health Center, Farmington, ²Department of Physiology & Neurobiology, University of Connecticut, Storrs and ³Division of Otolaryngology, Department of Surgery, University of Connecticut Health Center, Farmington, CT, USA

Abstract

Conclusions. We conclude that: (1) among several cues examined, the monaural cue of direct-to-reverberant (D/R) ratio in the ipsilateral ear provides the most information about sound-source distance; (2) interaural level difference (ILD) provides less information about sound-source distance; and (3) a comprehensive theory of three-dimensional auditory localization must incorporate the fact that all of the major acoustic cues change with distance. Objective. Neural mechanisms underlying auditory localization of distance are poorly understood. The present study was an initial step toward filling this gap in knowledge. Materials and methods. The binaural room impulse responses of adult barn owls were measured. The sound source was placed at various distances (up to 80 cm) and azimuths (0°/90°/180°) relative to the owl's head, with the elevation kept at 0°. Results. We determined the value of each cue for a 3–10 kHz band, and found that: (1) D/R ratio of signal amplitudes provided the most information about sound-source distance; (2) the ipsilateral D/R ratio represented distance more clearly than the contralateral or binaural-average D/R ratios; (3) ILD of direct signals increased with decreasing distance under certain conditions; (3) interaural time difference (ITD) of direct signals increased with decreasing distance at 90° azimuth; and (4) the spectral patterns of ILD and the monaural direct signals changed with distance in complex ways.

Keywords: Localization, sound, acoustic, hearing

Introduction

Humans and animals localize objects using all senses. Localizing sounds is important for basic functions such as escaping from threat, capturing prey, and communication. Localizing a sound in three-dimensional space requires the processing of azimuth, elevation, and distance. Mechanisms for localizing sounds in azimuth and elevation have been studied extensively. These studies have indicated that interaural time difference (ITD), interaural level difference (ILD), and spectral pattern are important cues for localization of azimuth and elevation [1–5]. In contrast, the third dimension of sound localization, distance, has received far less attention [6–10]. Our knowledge of neural mechanisms underlying auditory-distance processing is almost non-existent. The present study is an initial step toward filling this gap of knowledge.

We chose to address this topic using the barn owl. It is a good animal model for sound localization studies, partly because two-dimensional maps of auditory azimuth and elevation have been found in the midbrain of the barn owl [11]. In a barn owl, ILD changes with elevation because its two ears have asymmetrical vertical orientations; ITD and ILD are the main cues for azimuth and elevation, respectively [4].

The goals of this study were to address the following in the barn owl. What are potential acoustic cues that can provide information about auditory distance? How does each cue change with distance?

Materials and methods

Procedures of this study were approved by the Animal Care Committee of the University of...
Connecticut Health Center. Binaural room impulse responses (BRIRs) of three adult barn owls were measured with miniature microphones (Knowles model FG-3629, 2.6 mm in diameter) embedded in ear molds placed at the entrance to each ear canal. The blocked-meatus recording method offered: (1) avoidance of standing-wave interference that would be present in an unoccluded ear canal due to wave reflections from the tympanum, (2) a good signal-to-noise ratio, and (3) repeatable placements of the microphones in the intended locations without requiring external devices to hold the microphones [12,13].

Each owl was anesthetized with isoflurane and placed near the center of an acoustic chamber (double-walled with inner dimensions of 200 × 220 × 220 cm; Indust. Acoust. Co., Bronx, NY, USA). The anesthetic gas was delivered to the owl through an endotracheal tube connected to a pair of long tubes connecting between the owl and a respirator. The respirator was located outside the acoustic chamber. Our stimulus was a maximum length sequence (MLS) [14] with a period of 84 ms sampled at 49 kHz. The MLS stimulus of a finite length provided a constant amplitude spectrum across frequency. We presented the stimulus, acquired 2 channels of acoustic signals continuously from the 2 ears, and averaged the signals over 512 repetitions of the stimulus period, thus requiring a total of 43 s for stimulation and data acquisition at each sound location. We obtained BRIRs of the two ears by taking discrete inverse Fourier transforms of the transfer functions. A digital processor (Tucker Davis Tech. RP2) performed the simultaneous tasks of continuous stimulation and acquisition at the above sample rate.

The sound source was a compression driver (Selenium DT150 Super Tweeter) that produced a constant sound pressure level within ±6 dB over a 1–20 kHz range. It was well suited for this study of the barn owl because the above frequency range included a range of the barn owl’s low hearing threshold, 1–10 kHz [15]. The driver was fitted with a custom rigid conical coupler with an opening 6.2 mm in diameter. The sound emitted from a 6.2 mm opening was expected to behave as a point source for frequencies below 18 kHz based on the following. The upper limit frequency, f, for a point-source behavior of an opening with a radius, R, is related by: \( f = C/(6*R) \); where C is sound velocity [16]. We used a stimulus level of 75 dB SPL re 20 \( \mu Pa \) (rms level with the A-filter weighting) at 5 cm from the opening of the conical coupler.

We recorded BRIRs under two conditions of the walls of the acoustic chamber: (1) the walls were lined with sound-absorbent foam panels (Sonex, 3", Ilbruck), and (2) removable sound-reflective boards were positioned in front of the Sonex panels.

Our definitions of polarities of the coordinates, ILD and ITD, are as follows. Zero azimuth and elevation correspond to the straight-ahead direction relative to the animal’s head, positive azimuth to the right side, positive ILD to the right ear being more intense, and positive ITD to the right ear leading.

**Results**

The BRIRs of a barn owl are shown in Figure 1. The direct component of each impulse response was obtained by applying a Blackman window to the impulse response over a time range of ±2 ms re onset. (Note: A range of negative time was equivalent to the range at the tail end of the period because both the stimulus and the response were periodic.) The direct components of the impulse responses in the two acoustic environments were essentially the same. The reverberant component of each impulse response was obtained by applying a Blackman window over a time range of 2–52 ms re onset. The reverberant signal contained larger and more numerous echoes in the reflective environment (upper panel) than in the absorbent environment (lower panel).

![Figure 1. Monaural signals of binaural room impulse responses (BRIRs) for the left ear of an adult barn owl. The sound source was located at: distance =20 cm, azimuth = right 90°, and elevation = 0°. The responses were recorded in an acoustic chamber with the walls covered with sound-reflective boards (upper panel) or with sound-absorbent foams (Sonex, Ilbruck; lower panel).](image-url)
We chose 52 ms as the end of the reverberant signal because, around this point, the signal power level decayed about 60 dB relative to the onset merging into the noise floor.

Potential cues for auditory distance include: ILD, ITD, direct signal level, direct-to-reverberant (D/R) ratio of signal amplitudes, and spectral pattern [6–10]. To compute ILD, we defined binaural ratio spectrum (BRS) to be a ratio of the right-ear spectrum divided by the left-ear spectrum. ILD corresponds to level of BRS in dB. A barn owl’s ILD spectral patterns of the direct signals in the reflective environment for various distances at three different azimuths are shown in Figure 2, where ILD value is represented by color. For 0° azimuth (top), ILD was mostly restricted near 0 dB (mostly green) except for frequencies near 10 kHz. For azimuths of 45° and 90°, in contrast, ILD changed much with frequency and distance in complex ways exhibiting peaks (red) and troughs (blue).

The barn owl’s sound localization is accurate in a frequency range of 3–10 kHz [17]. Accordingly, we averaged ILD over this frequency range, and show these ILD values as a function of a logarithmic scale of distance for three azimuths in Figure 3A. The values of ILD were increased with decreasing distance, becoming as large as 23 dB when azimuth was 45° or 90°. When azimuth was 0°, ILD was little affected by distance remaining at a small value.

We employed a novel method of determining ITD over a specified bandwidth of frequencies, 3–10 kHz in this case. We defined binaural cross spectrum (BCS) to be a product between the right-ear spectrum and a complex-conjugate of the left-ear spectrum. Inverse Fourier transform of BCS is mathematically equal to a cross-correlation function between the two ears’ impulse responses. We defined ITD to be the center of gravity of the square of the above cross-correlation function. The latter is mathematically equal to energy-weighted group delay (EGD) in the frequency domain [18]. We applied a Blackman window to the BCS over 3–10 kHz and determined ITD as EGD in the windowed BCS using the above equality between the center of gravity and EGD. Advantages of this method of determining ITD are indicated in the Discussion section.

The results of this ITD analysis (Figure 3B) indicated that ITD increased with azimuth, as expected, and that ITD also changed somewhat with distance. The amounts of ITD change were greater at 90° and 45° azimuths than at 0° azimuth.

The dashed lines in Figure 3B correspond to linear regression lines.

As direct-signal level increases with decreasing sound-source distance, it is a potential cue for distance. However, its usefulness is limited because a listener must have a prior knowledge of the intensity of a sound source itself in order to use direct-signal level as a cue for distance. This information is generally not available to a listener. Another cue, D/R ratio, does not require a listener’s knowledge of the source intensity. This is an important advantage of D/R as a potential cue for auditory distance.

To obtain the D/R ratio, we applied a Blackman window to each ear’s D/R spectrum over a 3–10 kHz frequency range and obtained level (in dB) of a weighted average of the magnitude of the windowed spectrum. The D/R ratio of the ipsilateral (right) ear (Figure 3C) exhibited the clearest representation of sound-source distance at all azimuths examined in the sense that D/R ratio increased with decreasing distance. (Note: At 0° azimuth, the right and left ears are referred as the ‘ipsilateral’ and ‘contralateral’ ears, respectively, for convenience.) The dashed lines correspond to linear regression lines of D/R versus log-distance. The ipsilateral D/R (Figure 3C) represented distance more clearly than the contralateral D/R (Figure 3D) in the sense that the former changed more steeply and consistently with distance. The regression line slopes for the ipsilateral ear were −6.8 and −8.1 dB/doubling of distance for 45° and 90° azimuths, respectively (Figure 3C). The corresponding contralateral ear slopes were shallower, −2.0 and −2.2 dB/doubling (Figure 3D). At 0° azimuth, the regression line slopes of the two ears

Figure 2. Color-coded interaural level difference (ILD) spectra for the direct signals of a barn owl (no. 1) in the reflective environment for various distances. The three panels represent three azimuths (0°; right 45° and right 90°); elevation =0°. The distance scale is logarithmic.
were similar: −4.1 and −3.4 dB/doubling for the right and left ears, respectively.

For a lateralized sound location (away from 0° azimuth), a consequence of the fact that the ipsilateral D/R provides more information about distance than the contralateral D/R is that the monaural ipsilateral D/R provides more information about distance than the binaural average D/R (that can be inferred from the data of the two separate ears).

We made acoustical measurements of the above set of cues with two other barn owls in the sound-reflective environment and with one of the owls in the sound-absorbent environment. These results from three owls are summarized in Table I. The results from three owls in the sound-reflective environment were consistent in that: (1) monaural ipsilateral D/R ratio provided the clearest representation of sound-source distance; and (2) changes of the various response measures as a function of distance and azimuth described in Table I were largely similar.

In the sound-absorbent environment, the results from one owl (no. 1) also indicate that the monaural ipsilateral D/R ratio provided the clearest representation of sound-source distance. However, the slope of the monaural D/R ratio versus log-distance tended to be slightly shallower in the absorbent environment than in the reflective environment.

Discussion

To the best of our knowledge, the present study is the first to describe the dependence of acoustic cues in an experimental animal species upon sound-source distance. The main findings of this study are that, among several acoustic cues examined, the monaural cue of D/R ratio in the ipsilateral ear provides the most information about the sound-source distance and that a comprehensive theory of three-dimensional auditory localization must incorporate the fact that all of the major acoustic cues change with distance.

We introduced in this study a new analysis procedure to determine ITD in the frequency domain over a specified bandwidth (see the text related to Figure 3B in the Results section). The standard concept of phase delay between the two ears corresponds to interaural phase difference (in units of cycles) divided by frequency \[3\]. One may consider taking a scalar average of the phase delays over a specified bandwidth of frequencies to obtain ITD associated with the bandwidth. We suggest that the present measure of ITD is more informative because it reflects not only the phase but also the amplitude characteristics of the signals, whereas the scalar average of phase delay reflects only the phase characteristics ignoring the amplitude characteristics. Another advantage of the present procedure of computing ITD is that it is designed to work over a specified bandwidth of frequencies while remaining in the frequency domain without having to convert the bandpassed signals from the frequency domain to the time domain. If one wishes to apply the usual concept of obtaining ITD from cross-correlation of two ears’ impulse responses \[9,19\] to bandpassed...
<table>
<thead>
<tr>
<th>Environment</th>
<th>Sound-reflective</th>
<th>Sound-absorbent</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owl no.</td>
<td>D/R slope, ipsilateral</td>
<td>D/R slope, contralateral</td>
<td>ILD (dB)</td>
</tr>
<tr>
<td></td>
<td>(dB/doubling)</td>
<td>(dB/doubling)</td>
<td></td>
</tr>
<tr>
<td>0°</td>
<td>4.1</td>
<td>6.8</td>
<td>5.4</td>
</tr>
<tr>
<td>45°</td>
<td>4.2</td>
<td>7.0</td>
<td>6.8</td>
</tr>
<tr>
<td>90°</td>
<td>4.2</td>
<td>7.0</td>
<td>6.8</td>
</tr>
<tr>
<td>Mean</td>
<td>4.2</td>
<td>7.0</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Measurements were made with three owls in a sound-reflective environment and with one of the three owls in a sound-absorbent environment. The slopes of D/R versus log-distance were determined from linear regression lines over a distance range of 5–80 cm. D/R ratio, direct-to-reverberant ratio; ILD, interaural level difference; ITD, interaural time difference; AILD, among-ear level difference; ITD, interaural time difference.
ear. Besides testing these predictions behaviorally, future studies can also test them neurophysiologically by recording responses of barn owl’s auditory midbrain neurons to virtual sound fields containing normal and altered acoustical cues. As a part of a future physiological study, the present finding demonstrating potential effectiveness of a monaural cue highlights the need to examine the responses of neurons in the monaural pathways regarding their sensitivities to auditory distance. This aspect may be easy to overlook because the prevailing concepts of sound localization are dominated by studies of binaural cues (ITD and ILD).

The present observations were made in an acoustic environment with relatively small dimensions, $200 \times 220 \times 220$ cm. The concept of a role played by monaural ipsilateral D/R ratio for distance perception is expected to be also applicable for environments of larger dimensions on the basis of observations in humans [8–10]. Future studies of animals should ascertain whether that is the case.

Acknowledgements

This paper was presented at CORLAS 2007. The study was supported by grants from the Deafness Research Foundation and from the Department of Neuroscience, University of Connecticut Health Center. We thank Drs T. Takahashi and C. Keller, University of Oregon, and Dr E. Knudsen, Stanford University, for providing us with barn owls, and Dr Shigeuki Kowada for stimulating discussions and encouragement.

References