How do owls localize interaurally phase-ambiguous signals?

Kourosh Saberi, Haleh Farahbod, and Masakazu Konishi

This information is current as of February 2007.

Online Information & Services	High-resolution figures, a citation map, links to PubMed and Google Scholar, etc., can be found at: www.pnas.org/cgi/content/full/95/11/6465
References	This article cites 11 articles, 7 of which you can access for free at: www.pnas.org/cgi/content/full/95/11/6465#BIBL
	This article has been cited by other articles: www.pnas.org/cgi/content/full/95/11/6465#otherarticles
E-mail Alerts	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.
Rights & Permissions	To reproduce this article in part (figures, tables) or in entirety, see: www.pnas.org/misc/rightperm.shtml
Reprints	To order reprints, see: www.pnas.org/misc/reprints.shtml

Notes:

How do owls localize interaurally phase-ambiguous signals?

Kourosh Saberi*, Haleh Farahbod, and Masakazu Konishi

Division of Biology, California Institute of Technology, Pasadena, CA 91125

Contributed by Masakazu Konishi, March 24, 1998

ABSTRACT   Owls and other animals, including humans, use the difference in arrival time of sounds between the ears to determine the direction of a sound source in the horizontal plane. When an interaural time difference (ITD) is conveyed by a narrowband signal such as a tone, human beings may fail to derive the direction represented by that ITD. This is because they cannot distinguish the true ITD contained in the signal from its phase equivalents that are ITD $\pm$ nT, where T is the period of the stimulus tone and n is an integer. This uncertainty is called phase-ambiguity. All ITD-sensitive neurons in birds and mammals respond to an ITD and its phase equivalents when the ITD is contained in narrowband signals. It is not known, however, if these animals show phase-ambiguity in the localization of narrowband signals. The present work shows that barn owls (*Tyto alba*) experience phase-ambiguity in the localization of tones delivered by earphones. We used sound-induced head-turning responses to measure the sound-source directions perceived by two owls. In both owls, head-turning angles varied as a sinusoidal function of ITD. One owl always pointed to the direction represented by the smaller of the two ITDs, whereas a second owl always chose the direction represented by the larger ITD (i.e., ITD $-$ T).

Perceptual phenomena offer not only a window into the workings of the human brain, but also powerful tools for comparing the brain mechanisms of humans and animals. The phenomenon that we use in this report is binaural fusion, in which humans perceive a single auditory phantom image in the middle of their heads when listening to the same continuous broadband waveform delivered to the two ears. When the same waveform reaches the right ear before the left, the locus of the phantom image shifts toward the right ear. Broadband signals are necessary for the unambiguous perception of a single phantom image. Narrowband signals such as tones create ambiguous percepts. When an interaural time difference (ITD) is conveyed by a tone of appropriate frequency, humans may fail to localize the place represented by the true time difference (1–4). This is because the brain cannot count the number of cycles that may occur between the ears, i.e., the true time difference ITD cannot be distinguished from its phase-equivalent ITD $\pm$ nT, where T is the period of the stimulus tone and n is an integer. For example, the right ear signal leading by ITD is equivalent to the left ear signal leading by $T -$ ITD. We call this phenomenon phase-ambiguity. Auditory physiologists have long known of neurons that respond to ITD and ITD $\pm$ nT in the auditory systems of various mammals (5–8). It is, however, not known if neurons with this response property underlie the phase ambiguous perception in humans. We do not even know whether the animals that have these neurons show phase-ambiguity in sound localization, and if so, in what ways.

This work demonstrates phase-ambiguous sound localization in barn owls in which phase-ambiguous neuronal responses have been extensively studied throughout the auditory pathway (9–12). Barn owls (*Tyto alba*) are particularly suitable for this type of work, because they have been shown to localize sounds delivered by earphones by turning their heads, which provides an objective measure of localization ability.

MATERIALS AND METHODS

Surgery and Animal Care. Two owls were used in these experiments. All surgical instruments were sterilized. Owls were anesthetized with ketamine (10 mg/kg). A scalp area about 10 mm × 10 mm was incised after subcutaneous injection of a local anesthetic, lidocaine. The first layer of the skull was removed within this area with a pair of rongeurs and a stainless-steel post was cemented between the first and second layers. The whole operation lasted a few minutes. The wound was cleaned with an antibacterial agent, chlorhexidine (0.05%), and closed with dental cement and the scalp was sutured shut. After the surgery, owls were encased in a snugly fitting cylinder to prevent struggling and kicking as they recovered from anesthesia. Owls were observed in small cages in a separate recovery room until they came out of the cylinder. When owls recovered well enough to fly, they were returned to their living cages. To ensure adequate motivation for psychophysical performance, owls were maintained at 90% of their free-feeding weight. Owls were weighed daily and food intake was adjusted accordingly.

Behavioral Methods. The head-turning response did not need any training, but it had to be reinforced with food to avoid habituation to repeated presentations of the stimulus during a run of 30–50 trials. We trained the owls to feed from an apparatus that dispensed a small amount of mouse meat at a time, allowing 20–30 trials in an hour. Test sessions seldom continued beyond 2 hr. We initially used a small free-field speaker ("hoop speaker") mounted on a semicircular track to encourage the owls to localize it at various azimuthal angles. The owls were trained to initially orient to one source placed straight ahead (zero speaker) and wait for and localize a second sound (i.e., the signal) from the hoop-mounted speaker.

When the owls consistently localized both speakers, training began with the earphones. A metal bar was attached to the head post and held the left and right earphones in place. All sound stimuli were digitally synthesized in an IBM-compatible PC using Matlab and presented, after appropriate lowpass filtering, through 16-bit D/A converters (Sound Blaster; Milpitas, CA) at a rate of 40 kHz. Sound stimuli were delivered by an earphone assembly consisting of a Knowles (Itaska, IL) ED-1914 receiver as a sound source and a Knowles BF-1743 damped coupling assembly for smoothing the frequency response of the receiver. These components were encased in an aluminum cylinder 7 mm in diameter and 8.1 mm in length that fit into the external ear canal.

The localization tests in this work examined the owls' ability to localize sounds ballistically; i.e., without hearing the sound during head turning. We trained the owls with broadband noisebursts until they reliably responded by head turning to the spatial angle predicted from the ITD (13, 14). We then collected data by using tonal stimuli. The acoustic signal was a single presentation of a tone (200 msec in duration and 25 msec rise and decay times) in which the ITD was varied on each presentation. The sound pressure level for all experiments was 20 dB SPL (sound pressure level), which is about 20 dB above the hearing threshold of owls (0 dB sound pressure level between 3 and 8 kHz at the eardrum; ref. 15). The interaural level difference was always kept 0 so that the owls turned their heads only in the horizontal plane (13). At the beginning of each trial, a single broadband click was presented from the zero speaker to which the owl would orient its head. After the owl had fixated to the zero speaker, the test stimulus (pure tone) would be presented through the headphones. The owl responded to the test stimulus by turning its head and maintaining that fixation for a minimum of 1 sec, after which it was awarded with a small piece of mouse meat.

We chose 17 values of ITD per tone, leading either to the right (+) or left (−). For owl 1 we collected approximately five head-turning responses per ITD and frequency (5 or 6 kHz), and for owl 2 we collected five head-turning responses but only for one frequency (5 kHz) and four values of ITD (−150, −50, +50, and +150 μsec) that covered the relevant range of azimuths in determining the owl's general pattern of responses. The owls performed in complete darkness in an IAC (Bronx, NY) sound-proof anechoic chamber (5 × 3 × 3 m) and were monitored with an infrared video camera. We recorded behavioral sessions on video tape from above and measured head angles with a protractor on a 19-inch monitor using the owls' beak position as a reference point. The initial angle that was always at or near the zero speaker direction was subtracted from the terminal angle that was the angle at which the owl stopped. The angular resolution of this method was ≈4°.

RESULTS

We first describe the results for owl 1. Fig. 1 shows head-turning angles measured for this owl as a function of ITD. *Upper* and *Lower* show responses to a 5- and 6-kHz tone, respectively. Each symbol is one response and there is a minimum of five responses per ITD (some responses overlap and are not visible). Head-turning angles appear to vary as a sinusoidal function of ITD. To show the sinusoidal nature of the data, we fit a sine function to the data using a Matlab implementation of a multivariate Nelder-Meade simplex algorithm that minimizes the squared deviations of the function from the data. The two parameters of the fit were the amplitude and frequency of the sinusoid. For the 5-kHz tone, the best-fitting parameters were $f = 4,969.8$ Hz and $a = -51.1°$, and for the 6-kHz tone, these parameters were $f = 6,023.8$ Hz and $a = -45.5°$. The solid lines in both panels of Fig. 1 are fits to the data. In addition, 5- and 6-kHz sinusoids are plotted in *Upper* and *Lower*, respectively; however, they are not distinguishable from the statistical fits because they so closely overlap. The maximum angular responses for this owl were about 50° (right or left) in response to ITDs of −150, −50, +50 and +150 μsec. The 50° responses for ITDs of ±50 μsec are derived from ITDs of 200 − 50 and −200 + 50 where 200 μsec is the period of the waveform. Fig. 2 shows data from owl 2. In contrast to owl 1, this owl consistently responded with positive head-turning angles (≈+20°) to a positive 50-μsec ITD, and negative angles to a negative 50 μsec ITD. Note that the owl also responds with a +20° head-turning angle to −150 μsec ITD, and −20° angle to +150 μsec ITD, which are derived from 200−150 and −200 + 150 μsec ITDs.
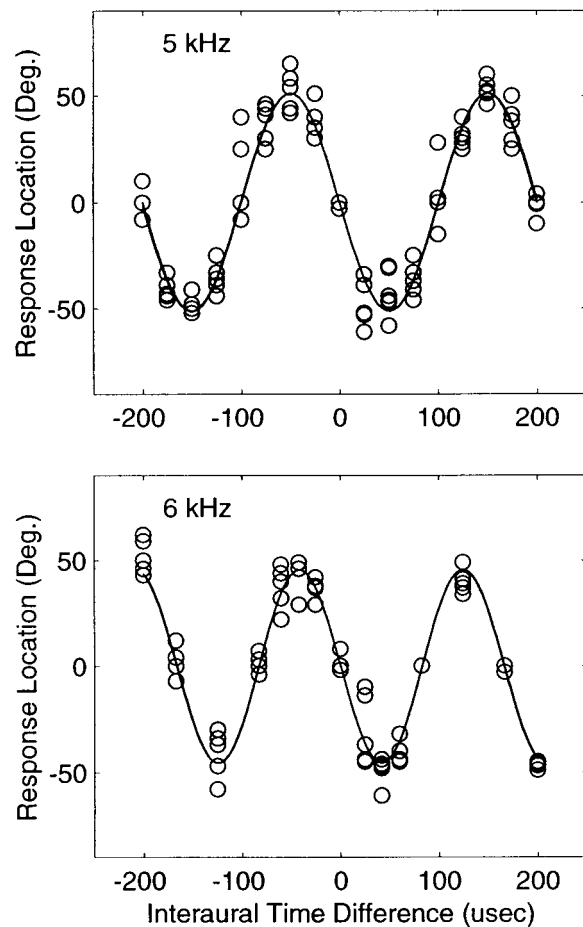

FIG. 1. Phase-ambiguous localization by owl 1. (*Upper*) Head-turning angles along the horizontal plane as a function of the ITD of a 5-kHz tone. Zero degrees is directly in front, and positive values are to the right of the owl. (*Lower*) Responses to a 6-kHz tone. Each symbol represents one head-turning response. There are a minimum of five responses per ITD. Head-turning angles vary as a sinusoidal function of ITD. The solid lines are two overlapping functions in each panel. One function is a 5-kHz sinusoid (*Upper*) and 6 kHz (*Lower*), and the other function is a sinusoidal fit to the data (4,969.8 Hz and 6,023.8 Hz in *Upper* and *Lower*, respectively).

None of the responses from either owl were greater than about 60–70° for any pure-tone stimulus. This may be understood in terms of the maximum achievable ITD in a pure tone stimulus of 5 kHz. Note that maximal ITD for this tone occurs at 100 μsec when the left and right signals are phase-shifted at the two ears by 180°. When the ITD is 200 μsec, the left and right signals are back in phase and, thus, this disparity is equivalent to an ITD of 0. A similar condition exists for a 6-kHz tone. However, as one might expect, because of its shorter period compared with a 5-kHz tone, the 6-kHz stimulus produces smaller maximum head-turning angles of only ≈45°.

DISCUSSION

Human psychoacoustical and animal neurophysiological findings led to the hypothesis that the auditory system carries out cross-correlation in the neural signals representing the sounds in the two ears to detect ITDs. In the barn owl, this computation initially takes place in separate frequency bands and its band-limited results are later combined across frequencies (9, 11). We can predict the nature of phase-ambiguity from cross correlating two 5-kHz tones (i.e., one for each ear) with the right signal leading the left side by 50 μsec (Fig. 3). Note that
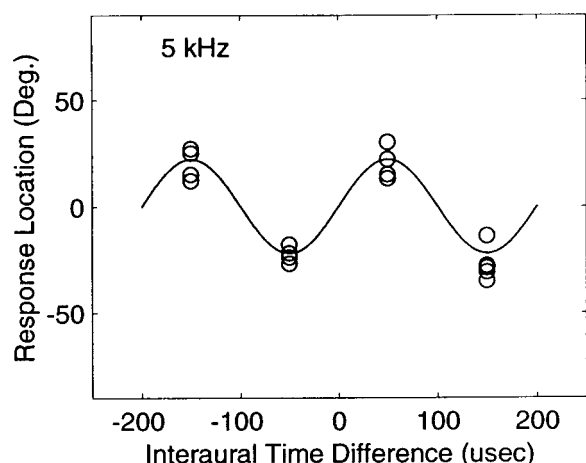
FIG. 2. Phase-ambiguous localization by owl 2. Head-turning angles are plotted against the ITD of a 5-kHz tone. The solid line is a 5-kHz sinusoidal fit. Note that the sinusoidal curves of Figs. 1 and 2 are displaced by 180°.

the two maximal values of the cross-correlation function correspond to ITDs of +50 (right side leading) and −150 μsec (left side leading). These ITD values are within the behaviorally and physiologically relevant range of ITDs (0–170 μsec) for the barn owl.

It is clear from the cross-correlation function (Fig. 3) that narrowband stimuli are phase ambiguous. The owl may resolve the ambiguity either in favor of only the smaller ITD (+50), or only the larger ITD (−150), or both, or neither (i.e., perceiving an intermediate or diffused locus). Owl 1 behaved as if it perceived a single locus that corresponded to the larger of the two ITDs, i.e., −150 μsec, thus resulting in a head turning angle of about −50°. This interpretation applies to other values of ITD as can be seen from the negative phase of the sinusoidal fits in both panels of Fig. 1. Responses of owl 2 were opposite to those of owl 1's in that this owl responded to +50 μsec ITD with a head-turning angle of +20°. This behavior, as well as the owl's responses to other ITDs, suggests that owl 2 favored the smaller of the pair of ITDs present in each test.

The cross-correlation analysis also predicts how the owl (owl 1) behaves when presented with an antiphasic signal, i.e., when the ITD in a 5-kHz tone is either −100 or +100 μsec. The cross-correlation function shows a minimum for these values of ITD. In this case, the owl's responses are more dispersed, but with an average that is near 0. When humans are presented with a low-frequency tone that is phase inverted to one ear, they hear a sound that is diffused over most of the interaural axis. Additionally, they hear two "heavier" images of equal strength, each near one ear. Of course one cannot tell what the owl precisely hears when presented with such a stimulus. However, it is clear that such a stimulus would not favor one ear over the other. Excluding any bias in the owl's behavior, one might expect either a bimodal distribution of responses, a uniform distribution of responses, or no response, i.e., no head turning or very small jitter in head position (as we have observed on several such trials). The data at ±100 μsec for the 5 kHz tone show the largest variance compared with other ITDs. If an owl responds randomly, the responses would of course be uniformly distributed with a mean of 0, and a variance that is larger than that for other ITDs. We suspect that the owl here is responding quasirandomly, occasionally not moving its head at all or very little, and sometimes selecting a random position.

Given the data reported here, there is little doubt that barn owls experience phase ambiguity in localizing narrowband signals. The owls neither fluctuated between the two targets nor localized an intermediate locus between the two targets,
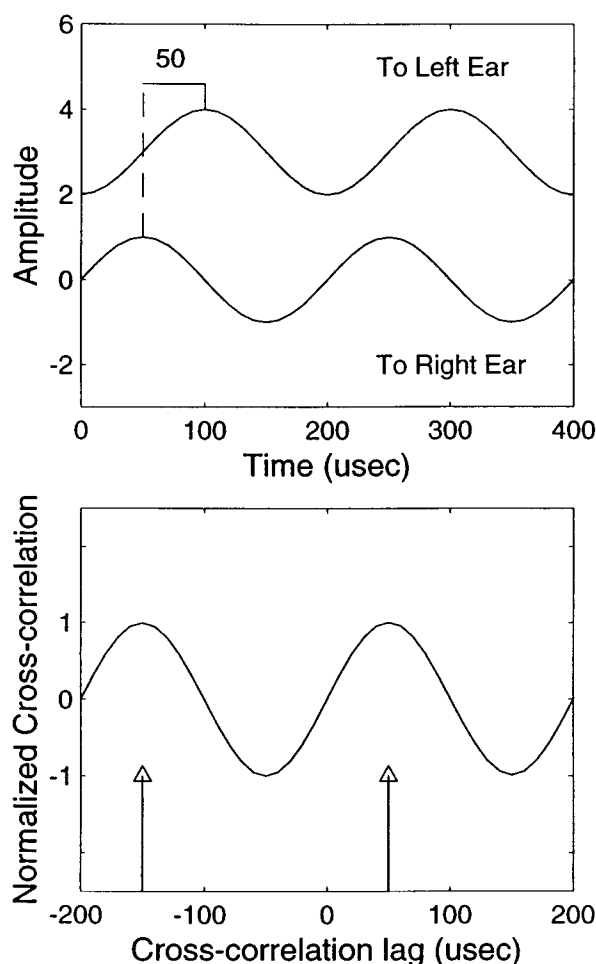

FIG. 3. Cross-correlation of narrowband signals. Top diagram shows two 5-kHz pure tones presented to the left and right ears. The waveform to the right ear is leading that to the left by ¼ period or 50 μsec. Equivalently, one may consider the left waveform leading the right by ¾ period. The two waveforms are of the same amplitude, but one has been vertically displaced in the figure for easier visual comparison. The lower panel shows the cross correlation of the waveforms to the left and right ears. The phase-ambiguous property of the waveforms results in two peaks in the cross correlation function at +50 and −150 μsec (arrows), corresponding to two directions in space. Note that the physiological range of the owl's ITD encoding is ± 170 μsec.

but instead, chose one target. In this respect, the differences between the owl and human are interesting to consider. The frequency range in which humans can use interaural phase differences to derive ITDs is much lower (highest frequency ≈1,200 Hz) than that for barn owls (highest frequency ≈9 kHz). The signal customarily used in human psychophysical studies of phase ambiguity has been a 500-Hz tone. Therefore, the ¼ period phase shift of a 500-Hz tone would produce two equally possible target loci represented by 500 and −1,500 μsec (instead of the 50 and −150 μsec for the owls at 5 kHz). Humans consistently report a single distinct source whose perceived location corresponds to the smaller of the two ITDs. The larger ITD of −1,500 μsec is outside the natural range of ITDs (600–700 μsec). This fact suggests that the perception of the near target by humans might be due to an inability to encode the large ITD. To clarify this issue, we listened to a 1-kHz tone with an ITD of 350 μsec and found, again, that only a single target is perceived even though the two ITDs 350 μsec and 650 μsec (1,000 − 350) are both roughly within the natural range of ITDs for humans. As the ITD of the 1-kHz tone was increased to values >400 μsec, a secondary faint image was

perceived on the lagging side, however, this image was much weaker than that corresponding to the leading side and is probably due to the fact that the phase shift was approaching ½ period (i.e., a disparity of 180°).

Finally, we argue here that a causal link can be established between neuronal responses and perception. Barn owls translate an ITD in a broadband signal into a single spatial direction (13, 14), whereas they experience phase-ambiguity in interpreting a time difference in a narrowband signal. Higher-order ITD-sensitive neurons such as those in the external nucleus of the inferior colliculus and the optic tectum are broadly tuned for frequency (10). When an ITD is contained in a broadband signal, these neurons respond only or most strongly to the true ITD and weakly or not all to other ITDs that are the phase-equivalents of the true ITD in a dominant spectral component. When an ITD is contained in a narrowband signal such as a tone, the same neurons show phase-ambiguity by responding equally strongly to the true ITD and its phase-equivalents. As the tone's ITD is gradually changed away from the neuron's preferred interaural delay, the neuron's response gradually decreases and reaches a minimum when the tone to one ear is out of phase relative to the other ear (i.e., ITD $\pm$ $T/2$ where T is the tone's period and ITD is the neuron's preferred interaural delay). Thus, this match between perception and neuronal response is not a coincidence, but most likely due to a cause and effect relationship.

1. Sayers, B. M. & Cherry, E. C. (1957) *J. Acoust. Soc. Am.* **29,** 973–987.
2. Jeffress, L. A. (1972) *Foundations of Modern Auditory Theory*, ed. Tobias, J. V. (Academic, New York), Vol. 2, pp. 349–368.
3. Stern, R. M., Zeiberg, A. S. & Trahiotis, C. (1988) *J. Acoust. Soc. Am.* **84,** 156–165.
4. Saberi, K. (1996) *Psychol. Rev.* **103,** 137–142.
5. Yin, T. C. T., Chan, C. K. & Carney, L. H. (1987) *J. Neurophysiol.* **58,** 562–583.
6. Yin, T. C. T. & Chan, C. K. (1990) *J. Neurophysiol.* **64,** 465–487.
7. Goldberg, J. M. & Brown, P. B. (1969) *J. Neurophysiol.* **32,** 613–636.
8. Moushegian, G., Rupert, A. L. & Whitcomb, M. A. (1964) *J. Neurophysiol.* **270,** 1174–1191.
9. Takahashi, T. & Konishi, M. (1986) *J. Neurosci.* **6,** 3413–3422.
10. Wagner, H., Takahashi, T. & Konishi, M. (1987) *J. Neurosci.* **7,** 3105–3116.
11. Carr, C. E. & Konishi, M. (1990) *J. Neurosci.* **10,** 3227–3246.
12. Mori, K. (1997) *Hearing Res.* **111,** 22–30.
13. Moiseff, A. (1989) *J. Comp. Physiol. A* **164,** 637–644.
14. Moiseff, A. & Konishi, M. (1981) *J. Neurosci.* **1,** 40–48.
15. Konishi, M. (1973) *Am. Sci.* **61,** 414–424.