Neural bases of an auditory illusion and its elimination in owls

Kourosh Saberi[1], Yoshifumi Takahashi[1,2], Haleh Farahbod[1] and Masakazu Konishi[1]

[1] *Division of Biology 216-76, California Institute of Technology, Pasadena, California 91125, USA*

[2] *Present address: Osaka University Graduate School of Medicine, Department of Otolaryngology and Sensory Organ Surgery, 2-2 Yamadaoka, Suita, Osaka, Japan*

Correspondence should be addressed to K.S. (kourosh@caltech.edu)

Humans and owls localize sounds by detecting the arrival time disparity between the ears. Both species determine the interaural time difference by finding the delay necessary to match the leading signal with the lagging one. This method produces ambiguity with periodic signals, because the two signals can be matched by delaying either one or the other. As predicted, owls localized periodic signals in illusory directions, whereas they always perceived the real source when signal bandwidth exceeded a certain value. This bandwidth also enabled higher-order auditory neurons to discriminate between real and illusory sources.

Here we report similarities between humans and owls in the manner in which they perceive and deal with illusions in sound localization. Sound propagating from its source takes a longer time to reach one ear than the other ear, resulting in an interaural time difference (ITD). Because this varies as a function of sound incidence angle, it can be used to encode the direction of sound sources in the horizontal plane. Humans and owls use the same methods of computing the ITD. Human psychophysical studies led to hypotheses regarding mechanisms of ITD detection[1–3]. These theories involve binaural cross-correlation: using axonal delay-lines of different lengths to create every possible biologically relevant delay, the brain measures the ITD by determining the delay required to match the leading signal to the lagging signal. The degree of match between the two signals is assessed by neurons that fire maximally when impulses from the two sides coincide. Neural circuits that carry out these computations have been found in the auditory systems of mammals, birds and reptiles[4–9].

This theory of binaural correlation can predict how humans and animals localize a sound given the physical properties of the signal. In addition to the ITD, these properties include frequency, bandwidth and signal-to-noise ratio[2,10–12]. The present paper concerns the role of signal bandwidth in sound localization. As narrowband signals such as tones are periodic, cross-correlation of these sounds produces more than one solution. For example, when a tone arriving at one ear leads the same tone in the other by a quarter of its period, the two signals can be matched by delaying either the leading side by a quarter period or the lagging side by three-quarters of a period. A broadband signal is unambiguously localized, because left and right signals can be matched for all component frequencies only once by delaying the leading side by the ITD. Given ambiguous signals that generate both real and illusory sound sources, humans perceive the source closer to the midline of the head[2,10]. Owls also localize illusory sound sources at loci that can be predicted from the tonal period and ITD[13]. Here we demonstrate that owls resolve this ambiguity in localization if signal bandwidth is increased. Analysis of this behavioral phenomenon in owls allowed the study of its neural basis. We show that the same range of signal bandwidths resolves ambiguity in sound localization by owls at both behavioral and neuronal levels.

RESULTS

Behavior

The owl's task was to turn its head in the perceived direction of a sound source. The relationship between ITDs and head-turning angles is known[14]. For example, an ITD of 50 μs induces a head turn of 20°, 100 μs, 40° and 150 μs, 50°. Cross-correlation theory can predict the direction of both the real and phantom sources. For example, a speaker at +20° (right of 0°) broadcasting a 5 kHz (period, 200 μs) tone creates a phantom source at –50° (left side). Cross-correlation of the left and right ear tones yields maxima at 50 μs, 50 + 200 μs and 50–200 μs. Of these time lags, 50 μs represents the direction of the real source and a lag of –150 μs indicates a phantom source at –50°, whereas a maximum at +250 μs is outside the owl's physiological range.

Figure 1 illustrates the responses of one owl to sounds from a target loudspeaker at one of four locations or from headphones delivering sounds with ITDs simulating these positions. If narrowband signals with a center frequency of 5 kHz were presented from 20° (top right), the owl consistently responded as if it perceived a phantom source at –50° (opposite hemifield). Likewise, a real source at –20° produced head-turns toward +50°. The owl always correctly turned its head toward the real source at +20° when signal bandwidth exceeded 3 kHz. However, the owl consistently localized sources located at +50° or –50° at all bandwidths (left). When choosing between real and phantom sources, the other owl consistently preferred the smaller angle. This owl localized a phantom source at +20° when narrowband sound was presented from –50°; conversely, it localized a phantom source at –20° when the real source was at +50°. Directions of phantom sources were predictable as they were for the first owl. This owl always correctly localized real sources at either + 20° or –20°.

Neurophysiology

The discharge of a single tectal neuron varied as a function of both stimulus bandwidth and ITD (Fig. 2). All test signals were

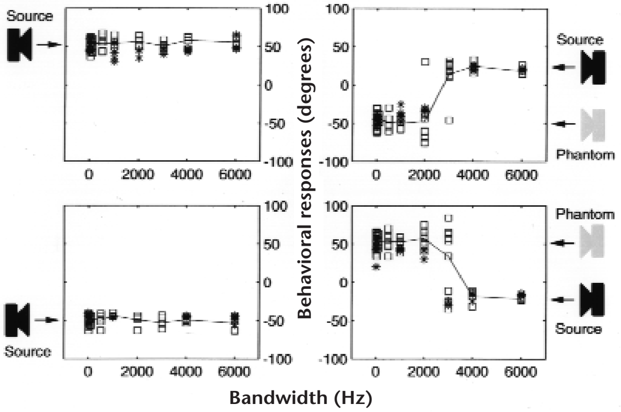

Fig. 1. Localization of real and illusory sound sources by an owl. Ordinate shows head-orienting angles of one owl to sounds presented through either a loudspeaker (*) or headphones (☐). Negative and positive values denote left and right of the midline, respectively. This owl appeared to perceive an illusory sound source only for the conditions shown in the right panels. Right panels show responses to the source at either +20° (top) or –20° (bottom); these same panels show responses to headphone-delivered sounds with ITDs of +50 (top) and –50 µs (bottom). Left panels show data for loudspeakers located at +50° (top) and –50° (bottom), as well as the equivalent headphone conditions (+150 and –150 µs, top and bottom, respectively). Note a change from phantom to true-source positions when the stimulus bandwidth exceeded 3 kHz (right).

delivered by earphones and contained a center-frequency of 7 kHz (period, ~143 µs). This neuron responded to the real ITD and its phantoms, –93 µs (= 50 – 143 µs) and +193 µs (= 143 + 50 µs). When bandwidth was narrow (0–1 kHz), the neuron responded equally strongly to both the real ITD and its phantoms. As bandwidth increased, responses to the real time difference became larger than those to time differences corresponding to phantoms. Note that when bandwidth exceeded 3 kHz, the neuron's response to the true ITD was clearly identifiable.

To quantitatively analyze the effects of signal bandwidth, we recorded from 11 tectal neurons. Spike rates in response to the real time difference and its phantoms are shown as a function of bandwidth for six of these neurons (Fig. 3). The response magnitudes are color coded such that 'hot' colors indicate greater responses than 'cool' colors. Responses to the real time difference rose above responses to phantoms as bandwidth increased. Responses of all 11 neurons to the real ITD (MP, main peak) increased relative to their responses to phantoms as bandwidth increased (SP, side peaks; Fig. 4). Equal responses to the real ITD and its phantoms are designated 0%, whereas 100% means that the neurons responded only to the real ITD and not to its phantoms; increases below 50% are not statistically significant (see legend). This increase exceeded 50% at a bandwidth of 3 kHz (see also refs. 15, 16); this was the same as the bandwidth at which owls ceased to localize illusory sources. The dashed curve shows the distribution of normalized mean responses, and the solid curve indicates a theoretical relationship between ITD and bandwidth predicted by a cross-correlation model of binaural interaction described in earlier papers[17–19].

DISCUSSION

We have shown previously that owls perceive phantom sources in localizing narrowband sounds[13]. In this paper, we show that owls always localize the real source of a broadband signal. One owl always localized the farther source, whether real or phantom, whereas the second owl preferred the nearer source. We can explain these differences by assuming that the owls selected one of two perceived images, one real and one illusory in the hemifield contralateral to the real source. Owls can easily select a single natural acoustic target to strike, as natural signals used for prey capture are broadband[20]. Our experiments with narrowband signals presented two targets simultaneously. Under these conditions, each owl developed its own strategy, choosing either the larger or the smaller angle in each case.

Phantom images originate from the temporal periodicity contained in narrowband signals; cross correlation of these signals gives rise to multiple maxima separated by an integer multiple of the signal period, that is, ITD ± nT, where n is an integer and T is the period[3,21,22]. In both mammals and owls, the auditory system carries out binaural cross-correlation initially in separate frequency bands[5,7,23]. The ability of human

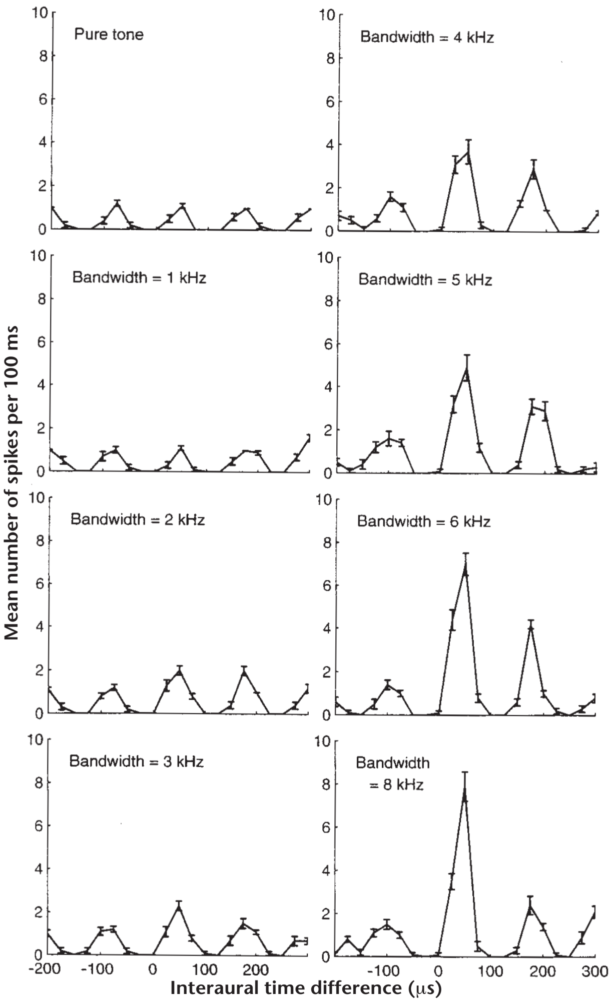

Fig. 2. Response of a tectal neuron to ITD varies with signal bandwidth. For bandwidths narrower than 3 kHz, the neuron responded equally to both the true (at 60 µs) and phantom ITDs. As bandwidth increased further, responses to the true ITD stood out above those to phantom delays.

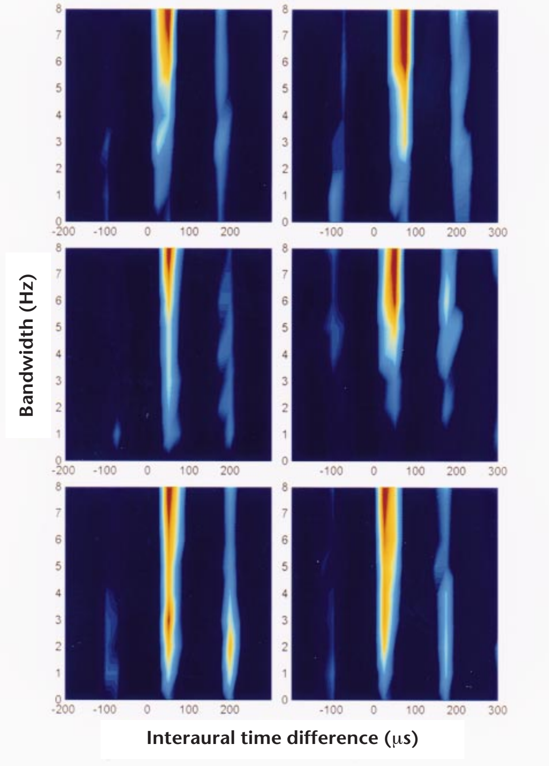

Fig. 3. Responses of tectal neurons as a function of bandwidth and ITD. This display combines the effects of bandwidth and ITD on neural responses. Regions of 'hotter' colors correspond to greater neural activity. Each panel shows results from one neuron. The middle-right panel shows data combined from the eight panels of Fig. 2.

beings to localize broadband sounds unambiguously is due to frequency convergence[2]. However, the exact nature of the convergence is a subject of debate[24,25]. In owls, the results of cross-correlation in separate frequency bands are also brought together in the external nucleus of the inferior colliculus[26,27]. This frequency convergence is responsible for the owl's ability to distinguish the real sound source from its phantoms in broadband signals. The neuronal response to the real ITD (or source) increases relative to the responses to the phantoms, because neurons tuned to different frequencies but to the same ITD 'add' their outputs. This process allows responses to the real ITD, but not to phantoms, because the ± nT terms do not coincide across different frequencies. In addition, the owl's auditory system uses inhibition to reduce the neuronal response to phantom ITDs[28].

The bandwidth necessary for discriminating between real and phantom sources is the same at both the behavioral and neuronal levels. Successful attempts to correlate neuronal response properties with perceptual performance abound. However, examples in which perception and neuronal responses are compared in the same species are rare in auditory research. The present paper shows that such combined approaches are possible when appropriate animals and behaviors are chosen.

METHODS

Behavior. Behavioral experiments used two tame owls. The behavioral measure of sound localization was the natural head-turning response of owls. We trained the owl to orient toward a sound source at zero degrees (directly in front) before turning the head in the direction of the test signal from either headphones or a loudspeaker (2 inch Quam) on a semicircular track with a radius of 1 m. All tests were done in total darkness in a large echo-free soundproof room. We observed and recorded the behavior of owls with an infrared video camera and a recorder. We obtained head-turning angles by measuring the direction of the beak on a 19-inch television monitor by frame-by-frame playback of videotapes. The resolution of this method was 4°. Owls were rewarded only when they made a rapid response toward a specific direction, regardless of whether that direction was the same as the source. Responses that were slow or that drifted near the terminating angle were not rewarded.

The headphones were designed to fit snugly in the owl's ear canal and consisted of a Knowles ED-1914 receiver as a sound source and a Knowles BF-1743 damped coupling assembly for smoothing the frequency response of the receiver. The headphones were held by an aluminum bar bolted on a post that had been previously cemented on the skull under anesthesia (ketamine, 10 mg per kg). Stimuli of 200 ms in duration were presented at a sampling rate of 40 kHz through digital-to-analog converters. We created an ITD by delivering a stimulus waveform to one ear and its time-shifted copy to the other ear. Although owls do not use temporal cues in stimulus envelopes for sound localization[14], the beginning and end of the left and right envelopes were simultaneous and gradual with 20-ms rise/fall ramps. We synthesized signals containing different bandwidths with the same center frequency of 5 kHz (period, 200 μs). This period length allowed us to create two sound sources within the range of ITDs detected by the owl. On each trial, one of four loudspeaker positions (–50°, –20°, +20°, +50°) or one of four ITDs (–150, –50, +50, +150 μs) was randomly paired with one bandwidth. Because many combinations of speaker locations or ITDs and bandwidths were possible, some bandwidths in the free-field condition were not tested. The sound level was 20 dB above the owl's threshold for all stimuli. Interaural level differences that the owl uses for localization in the vertical plane were kept at zero for all headphone experiments.

Neurophysiological methods. Two owls were used for neurophysiology. The protocol for the care and treatment of owls was approved by the Institute's animal care committee (Protocol # 1053-7). Owls were anesthetized with intramuscular injections of ketamine (20 mg per kg per h, Phoenix Pharmaceutical) and diazepam (0.5 mg per kg per h, Steris Laboratories). Owls were kept under anesthesia throughout the experiment and the

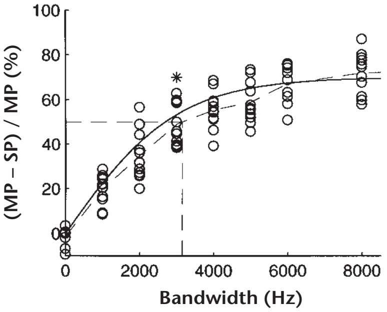

Fig. 4. Relative amplitudes of responses by tectal neurons to real and phantom ITDs. Data are from 11 tectal neurons. Dashed curve is the mean data from all neurons. A plot of discharge rate against ITD is a periodic curve consisting of equally spaced peaks. The response to the true time difference (the neuron's best ITD) is called the main peak (MP) and those to other time differences are called the side peaks (SP). The differences in discharge rate between the main and side peaks were calculated as a percentage change in average discharge rate at two nearest side peaks compared to the main peak [100% × (MP – SP)/MP]. A cross-correlation model predicts the relationship between MP and SP as shown by solid curve. Dashed horizontal and vertical lines mark the 50% reduction point. Asterisk shows the smallest bandwidth for which responses significantly differed from those to zero-bandwidth stimuli ($t_{10} = 3.70$, $p < 0.01$). For each of 11 neurons, the difference (D) between discharge rate at main peak and largest side peak was first obtained. A *t*-test showed that mean D becomes significantly greater than zero only for bandwidths exceeding 3 kHz.

edges of the skin wounds were doused once per hour with 0.2 ml 2% lidocaine HCl (Astra Pharmaceutical). We fixed an inverted T-shaped stainless steel post on the skull with dental cement. This post held the owl's head such that the flat roof of the upper mandible was tilted down by 90° with reference to the horizontal bar of the stereotaxic frame. After each experimental session which lasted several hours, the craniotomy was cleaned with an antibacterial agent, Chlorhexiderm (0.05% chlorhexidine gluconate, DVM Pharmaceuticals) and resealed with a small plastic sheet and dental cement (Hygenic, Perm Reline & Repair Resin Type II, Class I), and the scalp wounds were sutured shut. After the surgery, owls were observed in small cages in a separate recovery room. When they had recovered well enough to fly, they were returned to their living cages.

Sound stimuli were delivered by the same type of headphone used for behavioral experiments; in addition, a Knowles 1939 microphone was attached for monitoring sound pressure levels in the ear canal. The microphone was initially calibrated against a 12-mm B&K microphone in free field. Neurophysiological experiments employed the same set of stimuli that were used for headphone experiments except that the stimulus duration was 100 ms, and rise and fall ramps were 5 ms. A standard value of 20 dB above a neuron's threshold was chosen as the stimulus sound level for all stimuli.

We recorded from neurons of the owl's optic tectum, the last stage in the processing of cues for sound localization before information reaches the motor map for control of head turning[29]. These neurons were chosen for study as they have the right combination of physiological properties: they are sensitive to both ITDs and a broad range of frequencies. The optic lobe was accessed through a hole made in the exoccipital bone at the back of the head. This allowed direct access to the bony cover of the optic lobe through an air-filled cavity under the skull. Electrodes were inserted through holes made in this bony cover. Both tectal and midbrain auditory neurons could be recorded with ease. We isolated single tectal neurons with parylene-insulated tungsten microelectrodes (A-M systems; shaft thickness, 250 µm; 5 MΩ impedance at 1 kHz). We confirmed the identity of tectal neurons by both their responses to visual stimuli and the histological locations of electrolytic lesions as visualized in cresyl-violet-stained sections. Neural waveforms were amplified, bandpass filtered (1–10 kHz) and simultaneously monitored with an oscilloscope (Tektronix, 5110) and an audio amplifier (Grass, AM8). TTL pulses representing the timing of neural impulses were sent to the computer for on-line data display and storage. Data display consisted of plots of discharge rate against frequency, interaural time and level differences.

ACKNOWLEDGEMENTS

This work was supported by NIH (National Institutes of Deafness and Other Communicative Disorders) and a Human Frontier Science Program Fellowship (Y.T.). We thank John C. Middlebrooks, Terry Takahashi and Ervin R. Hafter for commenting on an earlier draft of this paper. We also thank Roian Egnor for comments and assistance during the course of this research.

RECEIVED 26 MARCH; ACCEPTED 14 MAY 1999

1. Jeffress, L. A. A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41, 35–39 (1948).
2. Jeffress, L. A. in *Binaural Signal Detection: Vector Theory. Foundations Of Modern Auditory Theory* Vol. 2 (ed. Tobias, J. V.) 349–368 (Academic, New York, 1972).
3. Sayers, B. M. & Cherry, E. C. Mechanism of binaural fusion in the hearing of speech. *J. Acoust. Soc. Am.* 29, 973–987 (1957).
4. Carr, C. E. & Konishi, M. Axonal delay-lines for time measurement in the owls brain-stem. *Proc. Natl. Acad. Sci. USA* 85, 8311–8315 (1988).
5. Carr, C. E. & Konishi, M. A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.* 10, 3227–3246 (1990).
6. Overholt, E. M. & Rubel, E. W. & Hyson, R. L. 1992. A circuit for coding interaural time differences in the chick brainstem. *J. Neurosci.* 12, 1698–1708 (1992).
7. Yin, T. C. & Chan, J. C. Interaural time sensitivity in medial superior olive of cat. *J. Neurophysiol.* 64, 465–488 (1990).
8. Smith, P. H., Joris, P. X. & Yin, T. C. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J. Comp. Neurol.* 331, 245–260 (1993).
9. Oertel, D. The role of timing in the brainstem auditory nuclei of vertebrates. *Annu. Rev. Neurosci.* 61, 497–516 (1999).
10. Stern, R. M., Zeiberg, A. S. & Trahiotis, C. Lateralization of complex binaural stimuli: a weighted-image model. *J. Acoust. Soc. Am.* 84, 156–165 (1988).
11. Colburn, H. S. Theory of binaural interaction based on auditory-nerve data: I. General strategy and preliminary results on interaural discrimination. *J. Acoust. Soc. Am.* 54, 1458–1470 (1973).
12. Colburn, H. S. in *Auditory Computation* (eds. Hawkins, H., McMullen, T., Fay, R. & Popper, A. N.) 332–400 (Springer, New York, 1995).
13. Saberi, K., Farahbod, H. & Konishi, M. How do owls localize interaurally phase-ambiguous sounds? *Proc. Natl. Acad. Sci. USA* 95, 6465–6468 (1998).
14. Moiseff, A. & Konishi, M. Neuronal and behavioral sensitivity to binaural time differences in the owl. *J. Neurosci.* 1, 40–48 (1981).
15. Mazer, J. How the owl resolves auditory coding ambiguity. *Proc. Natl. Acad. Sci. USA* 95, 10932–10937 (1998).
16. Mazer, J. *Integration of Parallel Processing Streams in the Inferior Colliculus of the Barn Owl.* Thesis. Caltech (1995).
17. Saberi, K. *et al.* Effects of interaural decorrelation on neural and behavioral detection of spatial cues. *Neuron* 21, 789–798 (1998).
18. Saberi, K. An auditory illusion predicted from a weighted cross-correlation model of binaural interaction. *Psychol. Rev.* 103, 137–142 (1996).
19. Saberi, K. Modeling interaural-delay sensitivity to frequency modulation at high frequencies. *J. Acoust. Soc. Am.* 103, 2551–2564 (1998).
20. Konishi, M. How the owl tracks its prey. *Am. Sci.* 61, 414–424 (1973).
21. Hafter, E. R. & Shelton, B. R. Counterintuitive reversals in lateralization using rectangularly modulated noise. *J. Acoust. Soc. Am.* 90, 1901–1907 (1991).
22. Grantham, D. W. in *Spatial Hearing and Related Phenomena. Handbook of Perception and Cognition: Hearing* (ed. Moore, B. C.) 297–345 (Academic, San Diego, 1995).
23. Takahashi, T. T. & Konishi, M. Selectivity for interaural time difference in the owl's midbrain. *J. Neurosci.* 6, 3413–3422 (1986).
24. Shackleton, T. M., Meddis, R. & Hewitt, M. J. Across-frequency integration in model of lateralization. *J. Acoust. Soc. Am.* 91, 2276–2279 (1992).
25. Trahiotis, C. & Stern, R. M. Across-frequency integration in lateralization of complex binaural stimuli. *J. Acoust. Soc. Am.* 96, 3804–3806 (1994).
26. Wagner, H., Takahashi, T. & Konishi, M. Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J. Neurosci.* 7, 3105–3116 (1987).
27. Keller, C. H. & Takahashi, T. Binaural cross-correlation predicts the responses of neurons in the owl's auditory space map under conditions simulating summing localization. *J. Neurosci.* 16, 4300–4309 (1996).
28. Mori, K. Across-frequency nonlinear inhibition by GABA in processing of ITD. *Hear. Res.* 111, 22–30 (1997).
29. duLac, S. & Knudsen, E. I. Neural maps of head movement vector and speed in the optic tectum of the barn owl. *J. Neurophysiol.* 63, 131–146 (1990).