

Object localization in cluttered acoustical environments

T. T. Takahashi · C. H. Keller · B. S. Nelson ·
M. W. Spitzer · A. D. S. Bala · E. A. Whitchurch

Received: 1 November 2007 / Accepted: 15 March 2008
© Springer-Verlag 2008

Abstract In nature, sounds from objects of interest arrive at the ears accompanied by sound waves from other actively emitting objects and by reflections off of nearby surfaces. Despite the fact that all of these waveforms sum at the eardrums, humans with normal hearing effortlessly segregate one sound source from another. Our laboratory is investigating the neural basis of this perceptual feat, often called the “cocktail party effect”, using the barn owl as an animal model. The barn owl, renowned for its ability to localize sounds and its spatiotopic representation of auditory space, is an established model for spatial hearing. Here, we briefly review the neural basis of sound-localization of a single sound source in an anechoic environment and then generalize the ideas developed therein to cases in which there are multiple, concomitant sound sources and acoustical reflection.

1 Introduction

In the natural auditory environment, object localization entails the segregation of interesting sounds from echoes and

other sources of background noise. Although the literature on the psychoacoustics of sound source segregation is rich, as is that on the neural mechanism of sound localization of single sources, efforts to understand the neural mechanism of source segregation in cluttered environments has only recently been addressed. Below, we review our understanding of this process, with emphasis on work in the barn owl (*Tyto alba*), an established animal model of spatial hearing. We will briefly review the neural mechanisms of sound localization of single sources, and then extend the ideas to the more complex situations in which there are two independent sources with overlapping spectra and in which there is an acoustical reflection, i.e., an echo. In each case, we seek to understand the physics of the signal that impinges on the eardrum, how neurons in the auditory system respond to these signals, and how the neurons’ responses affect behavior.

The animal model that we focus on, the barn owl, is a nocturnal bird-of-prey that can hunt by hearing alone, guided by activity on a topographic representation of frontal space in the external nucleus of its inferior colliculus (ICx) (Payne 1971; Konishi 1973b; Knudsen and Konishi 1978). Lesions of this auditory space map lead to scotoma-like deficits in sound localization, and microstimulation of the space-map evokes a rapid head saccade to that area of space represented at the point of stimulation (duLac and Knudsen 1990; Wagner 1993). Moreover, we have recently found that the spatial acuity of neurons in the space map can account for the owl’s ability to discriminate between the location of two sound sources, a quantity known as the minimal audible angle (MAA) (Bala et al. 2003, 2007).

We therefore treat this space map as a neural display of the acoustic environment, and by inferring the activity evoked across it under various acoustical conditions, we attempt to predict the owl’s behavior and compare behavioral observations with predictions based on neural recordings. Although

This work was supported by grants from the National Institutes of Deafness and Communication Disorders (RO1 DC003925, F32 DC008267) and the National Institutes of General Medical Sciences (T32-GM007257).

T. T. Takahashi (✉) · C. H. Keller · B. S. Nelson · M. W. Spitzer ·
A. D. S. Bala · E. A. Whitchurch
Institute of Neuroscience, University of Oregon,
Eugene, OR 97405, USA
e-mail: terry@uoregon.edu

Present Address:

M. W. Spitzer
School of Biomedical Sciences, The University of Queensland,
St. Lucia, QLD 4072, Australia
e-mail: m.spitzer@uq.edu.au

our immediate concern is the neural mechanisms that underlie object localization in cluttered acoustical environments, the ideas are relevant to a major question of sensory physiology, i.e., the link between neural activity and perception.

2 Localization of single sound sources

The barn owl relies on binaural differences in the level and arrival-times of sounds (ILD and ITD) to localize sounds, and neurons of the space-map in ICx owe their spatial receptive fields (RFs) to the computation of ITD and ILD within parallel brainstem pathways (Konishi 1973a; Moiseff and Konishi 1983; Takahashi and Konishi 1988a,b; Poganiatz et al. 2001; Poganiatz and Wagner 2001). The spatiotemporal pattern of activity across the owl's space map and its orienting behavior ultimately depends upon these physical cues, so understanding them is critical.

In the owl, ITD varies with the sound source's azimuth for the entire range of frequencies used for sound localization (3–10 kHz) (Konishi 1973a), due to the ears' separation along the horizontal axis. ITD remains relatively, but not completely, constant across frequencies (Keller et al. 1998). Unlike ITD, ILD varies considerably with frequency, so that a single location in space is specified by a spectrum of ILDs, ($ILD(f)$). From a spatial perspective, ILD varies with the source's azimuth at the frequencies below 3 kHz, but as frequency increases, the axis along which ILD changes becomes increasingly vertical, so that ILD varies with elevation. This vertical asymmetry in the directionalities of the two ears is due, presumably, to their anatomical asymmetry.

Thus, a location in space is specified by $ITD(f)$ and $ILD(f)$ (Knudsen et al. 1979; Moiseff 1989a,b; Brainard et al. 1992; Keller et al. 1998), and neurons in the space map respond preferentially to sounds from a given location in space because they are tuned to the frequency-specific values of ITD ($ITD(f)$) and ILD ($ILD(f)$) (Olsen et al. 1989; Euston and Takahashi 2002; Spezio and Takahashi 2003). A given neuron's response to a sound from an arbitrary location thus depends upon the extent to which $ITD(f)$ and $ILD(f)$ for that location matches the neuron's preferred values. We next build on this principle to examine the neural representation of two, simultaneously-active sound sources having overlapping spectra.

3 Object localization in clutter

In nature, sound waves at the eardrums are composites of waves arriving from multiple, actively-emitting sources as well as passive reflections from nearby surfaces. Yet, we are able to hear out and attend selectively to a single source in this clutter. What are the cues that are available at the eardrum to

execute this feat, sometimes called the “cocktail party effect” (Cherry 1953; Sayers and Cherry 1957).

When two sounds with overlapping magnitude spectra are presented simultaneously from two sources, the frequency-specific binaural cues at a given instant become vector sums of the cues corresponding to each source's location (Takahashi and Keller 1994; Blauert 1997). This is illustrated for two sources symmetrically placed about the midline, emitting sinusoids of identical frequency, amplitude, and phase angles (Fig. 1a). The sinusoid received at each eardrum from each source is depicted as vectors, or *phasors* that rotate (counterclockwise) in the complex plane (e.g., Fig. 1b). A phasor's angular velocity, magnitude, and angle (re: real axis) at the signal's start represent a sinusoid's frequency, amplitude, and starting phase, respectively. All of these quantities are equal at the two speakers in this case.

The situation is more complicated at each eardrum (Fig. 1b). The sinusoid from the left speaker (green) in the left eardrum is shown as a green phasor. The sinusoid in the left ear from the right speaker (red phasor) lags a bit behind (recall that rotation is counterclockwise) because the distance to the left ear from the right source is a bit longer than from the left source. At the right ear, the phasors representing left and right sources are precisely reversed relative to those in the left ear because of the symmetrical placement of the speakers about the midline. The sinusoids sum as vectors, such that the resultant at the left and right ears (solid blue phasors) is identical in angle (phase, ϕ) and length (magnitude); consequently, the left–right differences in magnitudes (ILD) and phase angles (ITD) are zero. In other words, despite the fact that there are two speakers on either side of the midline, the superposition of the signals at the eardrums result in a single phantom source located at the midline (Fig. 1a; blue speaker; $LS = RS$).

Figure 1c shows the case in which the sinusoid from right source has a higher amplitude than that from the left source. In this case, the vector summation causes the resultant in the right ear to be further advanced (further counterclockwise) than in the left ear, i.e., the ITD is biased toward the higher-amplitude source, and the phantom source is perceived as coming from the right (Fig. 1a; blue speaker; $RS > LS$).

When the two speakers emit *uncorrelated* broadband sources, the amplitude is likely to be higher for one source than the other within any given frequency band at any given instant. The ITD and ILD for that frequency band will approximate the values of the higher-amplitude source. In other spectral bands, the situation would be reversed. Over the course of the stimulus, the cues will spend roughly equal amounts of time assuming the values corresponding to the loci of the two sources. As a result, there should be enough information to represent two separate sources.

We can confirm that this is the case by recording from the owl's auditory space map in the ICx, the neurons of which

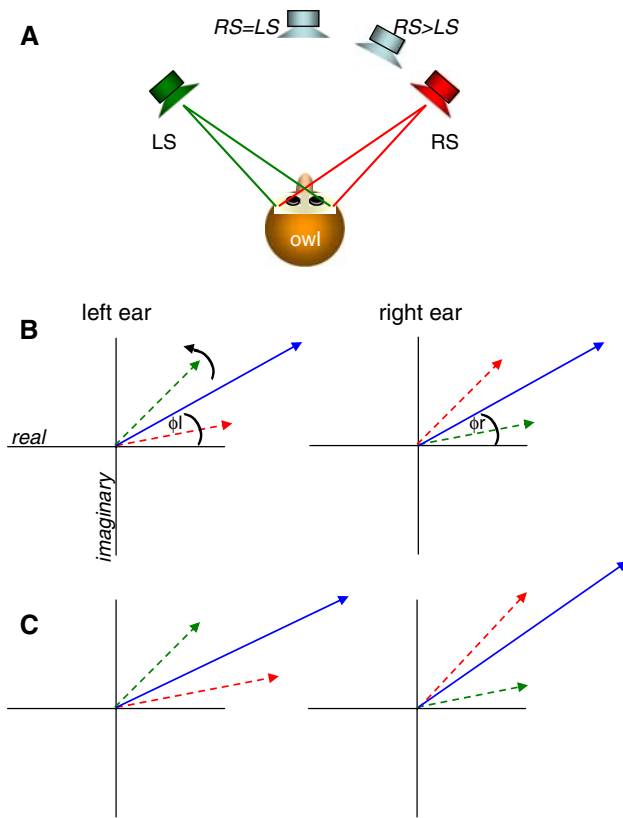


Fig. 1 Sound field with multiple sources. **a** Cartoon of an owl facing two loudspeakers simultaneously emitting sinusoidal waveforms. **b, c** Phasor representations of the sinusoids at the left and right ears. The sinusoids from the left (*green*) and right (*red*) speakers at each ear are depicted as vectors (phasors) of corresponding colors that rotate in the complex plane. The direction of rotation is counter-clockwise by convention. **1b** represents the case in which the two speakers' sinusoids are equal in frequency (rate of rotation), magnitude (length), and phase (angle re: real [horizontal] axis). In each ear, the *red* and *green* phasors summate to generate the resultant shown in *blue*. The interaural difference in the lengths and phase angles of these resultants represents, the ILD and ITD, respectively, which are zero in the case shown in **1b**. The owl should perceive a single sound-source directly in front (*blue* speaker in **a**; $R = L$). **1c** represents a case in which the amplitude of the sinusoid from the right speaker (*red*) is higher than that from the left speaker. In each ear, the *red* phasors are correspondingly longer. Vector summation produces a resultant that is more advanced (more counter-clockwise) in the right ear than in the left. The owl should perceive a sound source biased to the right, i.e., in the direction of the higher-amplitude source (*blue* speaker in **a**; $R > L$)

are tuned to the frequency-specific ITDs and ILDs at their spatial receptive fields (SRF). Figure 2 shows the response of a single space-map neuron when two sources of uncorrelated noises, diagonally separated by 42° , are passed through the cell's SRF. The cell's SRF, which was measured with a single source, is shown in Figure 2a. Figure 2b, which plots of spike count against the sound-pair's mean position generates two foci of high activity (red), indicating that the space-map neuron resolves the two sources (Takahashi and Keller 1994; Keller and Takahashi 1996a, 2005) and fires when each

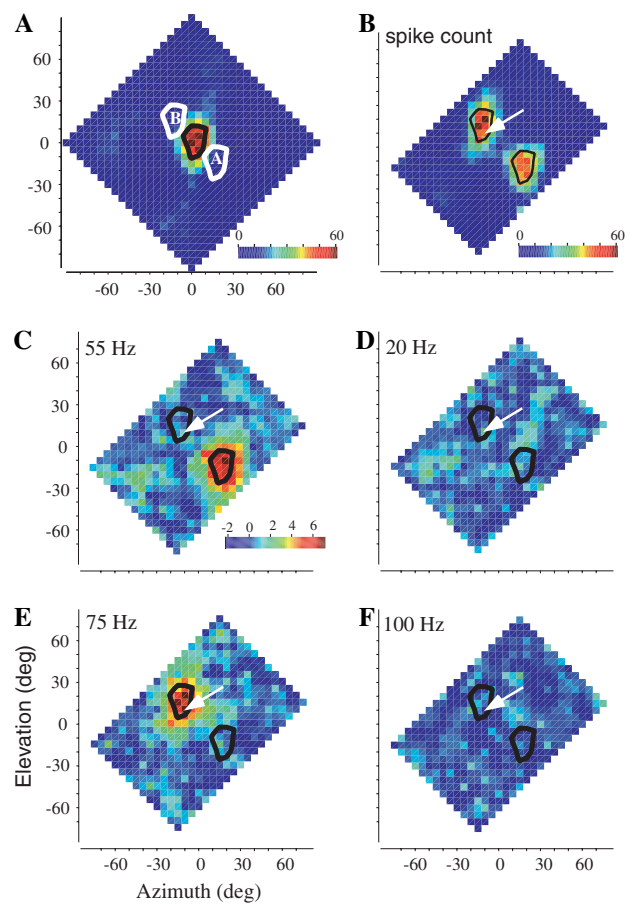


Fig. 2 Neural representation of two simultaneous sources. **a** Shows the response a space specific neuron to a broadband noise (100ms) placed at different locations in virtual frontal space (see Keller et al. 1998). In the *diamond-shaped plot*, the top corner is above the owl's head and the left corner is to the owl's left. The *pseudo-color* represents spike counts (*red*, high firing rate). The cell's spatial receptive field (SRF) is directly in front of the bird. The two *white contours*, labeled B and A, depict two diagonally-separated (by 42°) virtual sound-sources that were placed at different locations in frontal space in subsequent experiments (**b–f**). The sounds from sources A and B were uncorrelated broadband noises (100 ms) that were sinusoidally amplitude-modulated at 55 and 75 Hz respectively. **b** The firing rate evoked by sources A and B as they were swept through the cell's SRF. The neuron discharges as each source is swept through its SRF, generating two foci of activity when spike counts are plotted against the center of the two sources. (**c–f**) represent the degree of response synchrony to the 55, 20, 75, and 100 Hz components. The neuron fires at the appropriate rates when sources A and B are placed in the SRF. The synchrony to the difference frequency, 20 Hz, or an arbitrary frequency, 100 Hz, is near random. In other words, the cell is capable of spatially resolving the two sources and can convey each source's amplitude modulations. Details can be found in the report by (Keller and Takahashi 2005)

source falls in its SRF but not when the two sources flank the SRF.

The owl's localization behavior parallels the neural responses just described: when owls trained to point their heads toward single sound sources are presented with two, simultaneously emitting sources of broadband noise, they saccade

to one source or the other with equal frequency, or, in some trials, they make double saccades first toward one source, then the other (Smedstad 2003).

While the cues to resolve two simultaneous sources with overlapping amplitude spectra are available in the “acoustical array”, and the space-map’s neurons are capable of resolving them, can the auditory space also convey information about the nature of the sounds from each location? After all, knowing that there are two sources out there is only part of the problem—a human listener must also know, for example, that one person’s speech is coming from the left and another person’s speech is coming from the right. How well can cells of the owl’s space map convey information about both location and sound identity?

We now consider two sources, A and B, that emit uncorrelated noises that are sinusoidally amplitude-modulated (SAM) at 55 and 75 Hz, respectively (white contours; Fig. 2a). Given the importance of a signal’s amplitude modulation (or envelope) in deciphering communication sounds we use the ability of the cells to synchronize their firing to the envelope as an operational definition of the cell’s ability to convey information about sound identity (Drullman 1995; Shannon et al. 1995; Wright et al. 1997; Keller and Takahashi 2005). Theoretically, the binaural cues will reach source A’s ITD(f) and ILD(f) values 55 times per second, and source B’s, 75 times per second. Thus, in the space-map, the neurons that represent sources A and B should fire in synchrony at, 50 and 75 Hz respectively, preserving information about AM rates and loci.

Results are shown in Figs. 2c–f which plots the synchrony of the cell’s firing to a variety of modulation rates. The red colors indicate high synchrony. Figure 2c and d shows the cell fires in synchrony to the 55 Hz (Fig. 2c) and 75 Hz (Fig. 2d) components at the lower right and upper left foci respectively, thereby appropriately signaling the modulation rates at each location. Figures 2d and f show that the cells do not synchronize their firing to the difference frequency 20 Hz (Fig. 2e), or an arbitrary frequency, 100 Hz (Fig. 2f).

In contrast to the idea that object identity and location are separately represented in auditory cortex (Romanski et al. 1999), the reasoning above suggest that “what” and “where” are inseparable due to the physics of acoustical superposition, and the neural recordings suggest that the neurons of the space map are capable of responding to these physical cues so as to represent both “what” and “where” (Keller and Takahashi 2005). The same acoustical principles, described above for the owl, have also been suggested to play a role in the segregation of speech from the background (Roman et al. 2003; Faller and Merimaa 2004). It is interesting to note that when there are multiple sources, the binaural cues are drawn to the values of the loudest source. The barn owl’s ears are both directed directly ahead, so that sounds arriving from frontal sources would tend to be loudest and thus

capture the binaural cues (Keller et al. 1998). The well-known head saccades of the owl may be an orienting behavior that brings sound sources of interest into this acoustical “fovea”.

It is also interesting that with two sources, the AMs drive the neurons at the appropriate rates, indirectly, by changing the binaural cues. With a single sound source, AMs directly drive synchronous neural activity. Earlier studies showed that the space-map neurons were more sensitive to modulations of binaural cues than to modulations of amplitude (Takahashi and Keller 1992). We are thus led to the rather counterintuitive prediction that the perception of amplitude modulations is enhanced, at least in owls, when there is a second source, or “masker”, whose waveforms can interact with those of the “target” to generate dynamic binaural cues to which the space-map cells are very sensitive.

4 Object localization amidst echoes

The discussion above addressed the localization of two sound sources whose waveforms were independent of one another, much like to people talking simultaneously and independently of one another. Another prevalent form of acoustical interference is the echo, or more precisely, acoustical reflections. In this case, the two sounds are not independent; the reflected waveform is a delayed and corrupted version of the sound coming directly from the source. When the sound waves are longer than the time delay between the direct and reflected sounds (“lead/lag delay”), then the direct and reflected sounds overlap in time and acoustical interference would be expected. Nevertheless, humans effortlessly carry on conversations in reflective environments such as hallways, scarcely conscious of the reflected sounds that accompany the sounds arriving directly from the source. This ability to segregate direct sounds from reflected sounds is a special case of the cocktail party effect and is often called the “precedence effect” (Wallach et al. 1949). The literature on the precedence effect in humans is rich and has been thoroughly reviewed (e.g., Litovsky et al. 1999). We focus below on recent findings.

A subject’s perceptual experiences under experimental precedence conditions depend upon the lead/lag delay. When two identical sounds are presented simultaneously from two loci, humans perceive a single, fused, phantom target midway between the two actual transducers. As the delay is increased from 0 ms to about 1 ms, the phantom target appears to move proportionately closer to the leading target. This phenomenon is termed “summing localization” (as cited in Blauert 1997 pp 204–206) and can be explained in large part by the vector summation of the signals from each source (Fig. 1). Behavioral and neurophysiological evidence for summing localization has also been reported in cats and owls (Yin 1994; Keller and Takahashi 1996a; Populin and Yin 1998; Tollin and Yin 2003; Tollin et al. 2004). In an earlier study,

an untrained barn owl was placed on a perch in a sound-isolating room and the natural head saccades elicited by identical clicks presented from a pair of loudspeakers positioned 45° on either side of the midline were recorded. When the clicks were presented simultaneously, the owls typically aimed their head to a point midway between the two speakers, presumably localizing the phantom target generated by the process of superposition (Keller and Takahashi 1996b). Although delays less than 1 ms were not tested, studies of the cat cited above have shown that summing localization is observed with delays up to 0.4 ms.

If the delay between leading and lagging sources is increased to the 1–5 ms range, humans perceive a single fused phantom target closer to the leading source. This is termed “localization dominance” or “the law of the first wavefront” (Wallach et al. 1949; Haas 1951; Litovsky et al. 1999). At the same time, spatial information regarding the lagging source is degraded. The latter phenomenon, termed “lag discrimination suppression”, is manifested as a decrease in sensitivity to changes in the lagging source’s position (Zurek 1980; Yost and Soderquist 1984; Saberri and Perrott 1990; Freyman et al. 1991; Shinn-Cunningham et al. 1993; Yang and Grantham 1997).

Evidence for localization dominance has been observed in the cat and owl (Keller and Takahashi 1996b; Tollin and Yin 2003; Tollin et al. 2004; Spitzer and Takahashi 2006). The results from an owl trained to saccade toward sound-sources are shown in Fig. 3 (Spitzer and Takahashi 2006). The head aim at the end of a saccade (ordinate) is plotted against the position of the leading sound source (abscissa; solid black line). The location of the lagging sound is indicated by the blue and red lines for trials in which the lagging sound was placed to the left (blue) or right (red) of the lead. Trials with single sources are indicated by black symbols and those with a lead/lag pair are indicated by blue and red symbols for cases in which the lagging sound was placed, respectively, to the left (blue) or right (red) of the lead. In all trials, the stimuli consisted of 25 ms noise bursts; for two-source trials, the lagging stimulus was a delayed replica of the lead.

Figure 3a shows the head-aim for a lead/lag delay of 3 ms. At this short delay, saccades for two-source (lead + lag) trials are always in the direction of the leading source. Comparing the black symbols (single-source trials) to the colored ones, the precision of the saccades in two-source trials rival those of single-source trials demonstrating a powerful localization dominance. This precision is remarkable because with 25-ms sounds and a delay of 3 ms, the leading and lagging sounds largely overlap, generating complex binaural cues (Fig. 1). Figure 3b shows the trials in which the delay was 20 ms. In marked contrast to the trials with 3-ms delays, the saccades are now biased in the direction of the lagging sound, indicating that localization dominance has weakened considerably at these long delays. Figure 3c, which plots the proportion

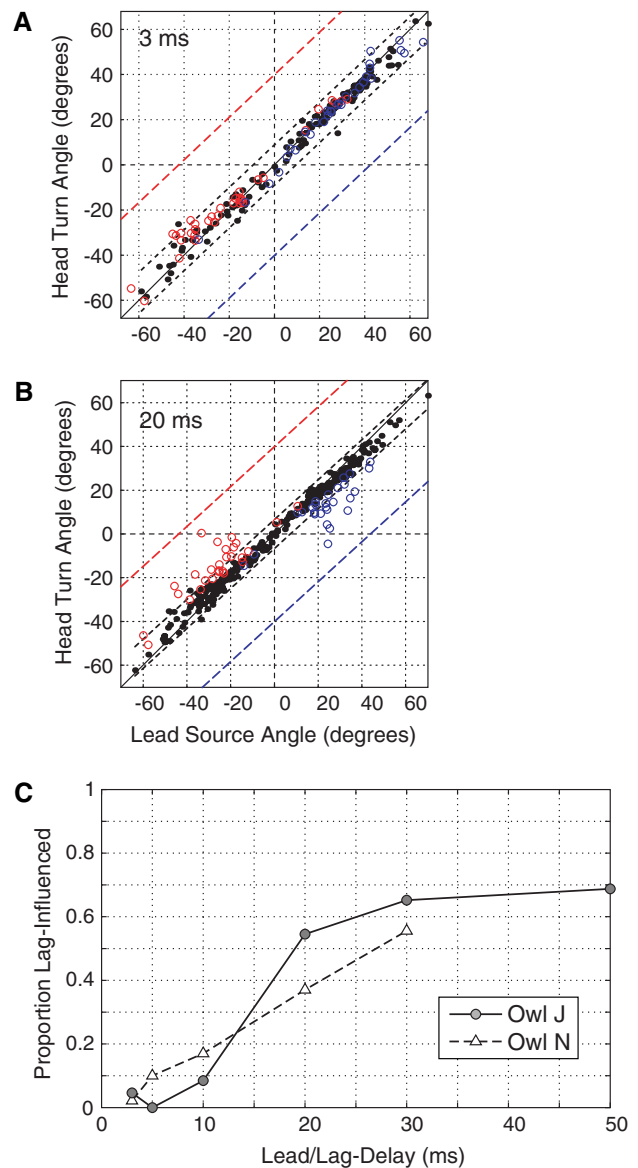


Fig. 3 Head saccades under precedence-effect conditions. In **a** and **b**, the aim of the owl’s head at the end of a saccade is plotted against the position of a single sound source (*black dots*) or the position of the leading source in trials with a lead/lag pair (*red and blue dots*). The position of the lagging source, which was always 40E to the left or right of the leading source are shown by the *blue* (lag to the left of lead) and *red* (lag to the right of lead) lines. In the experiments depicted in **a**, the lead/lag delay was 3 ms, and the owl turned its head in the direction of the leading source with an accuracy and precision rivaling those to single sources. In the experiments depicted in **B**, the delay was 20ms, and the owl’s saccades are now biased toward the lagging source, suggesting that localization dominance has weakened considerably. **c** Plots the number saccades to the lagging source against the lead/lag delay. The number of lag directed saccades increases as the delay increases, with an inflection at a delay of about 10 ms. Further details can be found in the report by Spitzer and Takahashi (2006)

of saccades biased toward the lagging stimulus against the delay, shows that localization dominance begins to weaken around 10 ms.

What neural mechanisms govern these behavioral phenomena? Why does the lagging sound's salience depend on the echo delay? These questions can be addressed by examining the neural responses to lead/lag pairs. There are a growing number of neurophysiological studies of the precedence effect in a variety of species and auditory structures (Yin 1994; Fitzpatrick et al. 1995; Keller and Takahashi 1996a,b; Litovsky 1998; Litovsky and Yin 1998a,b; Fitzpatrick et al. 1999; Mickey and Middlebrooks 2001; Litovsky and Delgutte 2002; Spitzer et al. 2004; Tollin et al. 2004), and nearly all of these studies report that a cell's response to the lag source is weaker than its response to the lead source or to a single source and that this response recovers at longer delays. Such suppression at short delays is a possible candidate for the decrease in sensitivity to the binaural cues of later arriving sounds experienced by human listeners (Zurek 1980; Stecker and Hafter 2002).

The relationship between the recovery of neural responses and the delay at which the lag source exerts its influence, or "echo threshold" is naturally of interest. In the IC of unanesthetized rabbits, the median delay value that elicits a half-maximal discharge is about 6 ms, which is in agreement with behavioral observations in the humans. Tollin et al. (2004) recording from the IC of awake, behaving cats confirmed these shorter recovery times and showed that the neural responses have recovered to 65–75% at behaviorally determined echo thresholds (Tollin and Yin 2003).

5 Future directions

Much of the knowledge on the precedence effect has come from studies that used stimuli with very short durations, such as clicks. Such brief stimuli have the pragmatic advantage in that lead and lag sounds and the neural responses they evoke can be separated in time (see, however, Hartung and Trahiotis 2001; Trahiotis and Hartung 2002). In nature, however, direct sounds and their reflections *do* overlap, raising a new set of questions that should be addressed if we are to understand spatial hearing in the real world.

One such question concerns the echo threshold. When the durations of lead and lag sounds exceed the echo delay, one can parse the stimulus into three segments: a time segment with only the leading sound wave, which is equal in length to the echo delay; the segment of time when both lead and lag overlap; and finally, a time segment with only the lag source, which would also be equal to the echo delay (if the stimuli are of equal durations). Which part of the stimulus determines the echo threshold? On the one hand, a neural response evoked during the superposed segment, e.g., at the onset of the lagging sound, if strong enough, may make the echo detectable and localizable. Alternatively, echo-threshold could be reached if the neural response during the lag-alone segment

becomes strong enough. The recovery of the firing rate during the lag-alone segment better reflects the manner in which the proportion of lag-directed saccades increases with echo delay, leading us to favor this segment, at least tentatively (Spitzer et al. 2004; Spitzer and Takahashi 2006).

Another question that long-duration sounds pose has to do with localization dominance during the overlap segment. During the overlapped period, the binaural cues are consistent with two simultaneous sounds despite the fact that the two sounds are perfectly correlated if correlations are computed over their entire lengths. This is because during the period of overlap, the non-zero echo delay causes different *local* segments of the lead and lag sounds to be present at any given moment and these local segments would be uncorrelated. (Takahashi and Keller 1994; Keller and Takahashi 1996a, 2005). The concepts described above regarding two concomitant, uncorrelated sounds therefore apply (e.g., Fig. 2); in other words, there are sufficient *acoustical* cues to resolve two separate stimuli. At the same time, however, because the leading and lagging waveforms are delayed copies of one another, there would also be sufficient information to discern their temporal relationships in their fine structures or envelopes. Do the auditory centers above the inferior colliculus process such delayed copies of sounds, which would mimic reflections, differently than two independent and uncorrelated sounds?

Asked in a more general way, is the precedence effect a special case of the cocktail party effect or is it an altogether different phenomenon processed within a dedicated pathway? A pathway specialized for the processing of echoes exists, of course, in echolocating bats. In animals that do not actively echolocate, this pathway may serve to inform the listener about the location of reflective surfaces or to provide cues to the distance to a sound source (Fitzpatrick et al. 1995; Clifton and Freyman 1997; Bronkhorst and Houtgast 1999; Fitzpatrick et al. 1999; Burger and Pollak 2001; Pecka et al. 2007).

These are but a few of the questions that need to be addressed if future studies to understand spatial hearing in cluttered naturalistic environments.

References

- Bala AD, Spitzer MW, Takahashi TT (2003) Prediction of auditory spatial acuity from neural images on the owl's auditory space map. *Nature* 424:771–774
- Bala AD, Spitzer MW, Takahashi TT (2007) Auditory spatial acuity approximates the resolving power of space-specific neurons. *PLoS ONE* 2:e675
- Blauert J (1997) Spatial hearing. The psychophysics of human sound localization. MIT, Cambridge
- Brainard MS, Knudsen EI, Esterly SD (1992) Neural derivation of sound source location: resolution of spatial ambiguities in binaural cues. *J Acoust Soc Am* 91:1015–1027

- Bronkhorst AW, Houtgast T (1999) Auditory distance perception in rooms. *Nature* 397:517–520
- Burger RM, Pollak GD (2001) Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats. *J Neurosci* 21:4830–4843
- Cherry E (1953) Some experiments on the recognition of speech with one and with two ears. *J Acoust Soc Am* 25:975–979
- Clifton RK, Freyman RL (1997) The precedence effect: beyond echo suppression. In: Gilkey RH, Anderson TR Binaural and spatial hearing in real and virtual environments. Erlbaum, Mahwah
- Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am* 97:585–592
- duLac S, Knudsen EI (1990) Neural maps of head movement vector and speed in the optic tectum of the barn owl. *J Neurophysiol* 63:131–146
- Euston DR, Takahashi TT (2002) From spectrum to space: the contribution of level difference cues to spatial receptive fields in the barn owl inferior colliculus. *J Neurosci* 22:264–293
- Faller C, Merimaa J (2004) Source localization in complex listening situations: selection of binaural cues based on interaural coherence. *J Acoust Soc Am* 116:3075–3089
- Fitzpatrick DC, Kuwada S, Batra R, Trahiotis C (1995) Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol* 74:2469–2486
- Fitzpatrick DC, Kuwada S, Kim DO, Parham K, Batra R (1999) Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex. *J Acoust Soc Am* 106:3460–3472
- Freyman RL, Clifton RK, Litovsky RY (1991) Dynamic processes in the precedence effect. *J Acoust Soc Am* 90:874–884
- Haas H (1951) Über den einfluss eines Einfachechos auf die Hörsamkeit von Sprache. *Acustica* 1:49–58
- Hartung K, Trahiotis C (2001) Peripheral auditory processing and investigations of the “precedence effect” which utilize successive transient stimuli. *J Acoust Soc Am* 110:1505–1513
- Keller CH, Takahashi TT (1996a) 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
- Keller CH, Takahashi TT (1996b) Responses to simulated echoes by neurons in the barn owl’s auditory space map. *J Comp Physiol A* 178:499–512
- Keller CH, Takahashi TT (2005) Localization and identification of concurrent sounds in the owl’s auditory space map. *J Neurosci* 25:10446–10461
- Keller CH, Hartung K, Takahashi TT (1998) Head-related transfer functions of the barn owl: measurement and neural responses. *Hear Res* 118:13–34
- Knudsen EI, Konishi M (1978) Space and frequency are represented separately in auditory midbrain of the owl. *J Neurophysiol* 41:870–884
- Knudsen EI, Blasdel GG, Konishi M (1979) Mechanisms of sound localization in the barn owl (*Tyto alba*). *J Comp Physiol A* 133:13–21
- Konishi M (1973a) Locatable and nonlocatable acoustic signals for barn owls. *Am Nat* 107:775–785
- Konishi M (1973b) How the owl tracks its prey. *Am Sci* 61:414–424
- Litovsky RY (1998) Physiological studies of the precedence effect in the inferior colliculus of the kitten. *J Acoust Soc Am* 103:3139–3152
- Litovsky RY, Yin TC (1998a) Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms. *J Neurophysiol* 80:1302–1316
- Litovsky RY, Yin TC (1998b) Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. *J Neurophysiol* 80:1285–1301
- Litovsky RY, Delgutte B (2002) Neural correlates of the precedence effect in the inferior colliculus: effect of localization cues. *J Neurophysiol* 87:976–994
- Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. *J Acoust Soc Am* 106:1633–1654
- Mickey BJ, Middlebrooks JC (2001) Responses of auditory cortical neurons to pairs of sounds: correlates of fusion and localization. *J Neurophysiol* 86:1333–1350
- Moiseff A (1989a) Bi-coordinate sound localization by the barn owl. *J Comp Physiol A* 164:637–644
- Moiseff A (1989b) Binaural disparity cues available to the barn owl for sound localization. *J Comp Physiol A* 164:629–636
- Moiseff A, Konishi M (1983) Binaural characteristics of units in the owl’s brainstem auditory pathway: precursors of restricted spatial receptive fields. *J Neurosci* 3:2553–2562
- Olsen JF, Knudsen EI, Esterly SD (1989) Neural maps of interaural time and intensity differences in the optic tectum of the barn owl. *J Neurosci* 9:2591–2605
- Payne RS (1971) Acoustic location of prey by barn owls (*Tyto alba*). *J Exp Biol* 54:535–573
- Pecka M, Zahn TP, Saunier-Rebori B, Siveke I, Felmy F, Wiegrebe L, Klug A, Pollak GD, Grothe B (2007) Inhibiting the inhibition: a neuronal network for sound localization in reverberant environments. *J Neurosci* 27:1782–1790
- Poganiatz I, Wagner H (2001) Sound-localization experiments with barn owls in virtual space: influence of broadband interaural level difference on head-turning behavior. *J Comp Physiol A* 187:225–233
- Poganiatz I, Nelken I, Wagner H (2001) Sound-localization experiments with barn owls in virtual space: influence of interaural time difference on head-turning behavior. *J Assoc Res Otolaryngol* 2: 1–21
- Populin LC, Yin TC (1998) Behavioral studies of sound localization in the cat. *J Neurosci* 18:2147–2160
- Roman N, Wang D, Brown GJ (2003) Speech segregation based on sound localization. *J Acoust Soc Am* 114:2236–2252
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 2:1131–1136
- Saberi K, Perrott Dr (1990) Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. *J Acoust Soc Am* 87:1732–1737
- Sayers B, Cherry E (1957) Mechanism of binaural fusion in the hearing of speech. *J Acoust Soc Am* 29:973–987
- Shannon RV, Zeng FG, Kamath V, Wyganski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304
- Shinn-Cunningham BG, Zurek PM, Durlach NI (1993) Adjustment and discrimination measurements of the precedence effect. *J Acoust Soc Am* 93:2923–2932
- Smedstad HB (2003) Effects of degraded sound-localization cues on neuronal spatial resolution. *Biology*. University of Oregon, Eugene 54
- Spezio ML, Takahashi TT (2003) Frequency-specific interaural level difference tuning predicts spatial response patterns of space-specific neurons in the barn owl inferior colliculus. *J Neurosci* 23:4677–4688
- Spitzer MW, Takahashi TT (2006) Sound localization by barn owls in a simulated echoic environment. *J Neurophysiol* 95:3571–3584
- Spitzer MW, Bala AD, Takahashi TT (2004) A neuronal correlate of the precedence effect is associated with spatial selectivity in the barn owl’s midbrain. *J Neurophysiol* 92:2051–2070
- Stecker GC, Hafter ER (2002) Temporal weighting in sound localization. *J Acoust Soc Am* 112:1046–1057

- Takahashi TT, Konishi M (1988a) Projections of nucleus angularis and nucleus laminaris to the lateral lemniscal nuclear complex of the barn owl. *J Comp Neurol* 274:212–238
- Takahashi TT, Konishi M (1988b) Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. *J Comp Neurol* 274:190–211
- Takahashi TT, Keller CH (1992) Simulated motion enhances neuronal selectivity for a sound localization cue in background noise. *J Neurosci* 12:4381–4390
- Takahashi TT, Keller CH (1994) Representation of multiple sound sources in the owl's auditory space map. *J Neurosci* 14:4780–4793
- Tollin DJ, Yin TC (2003) Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. *J Neurophysiol* 90:2149–2162
- Tollin DJ, Populin LC, Yin TC (2004) Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J Neurophysiol* 92:3286–3297
- Trahiotis C, Hartung K (2002) Peripheral auditory processing, the precedence effect and responses of single units in the inferior colliculus. *Hear Res* 168:55–59
- Wagner H (1993) Sound-localization deficits induced by lesions in the barn owl's auditory space map. *J Neurosci* 13:371–386
- Wallach H, Newman E, Rosenzweig M (1949) The precedence effect in sound localization. *Am J Psychol* 62:315–336
- Wright BA, Lombardino LJ, King WM, Puranik CS, Leonard CM, Merzenich MM (1997) Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature* 387:176–178
- Yang X, Grantham DW (1997) Cross-spectral and temporal factors in the precedence effect: discrimination suppression of the lag sound in free-field. *J Acoust Soc Am* 102:2973–2983
- Yin TC (1994) Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J Neurosci* 14:5170–5186
- Yost WA, Soderquist DR (1984) The precedence effect: Revisited. *J Acoust Soc Am* 76:1377–1383
- Zurek PM (1980) The precedence effect and its possible role in the avoidance of interaural ambiguities. *J Acoust Soc Am* 67:952–964