J Neurophysiol 109: 924–931, 2013.
First published November 21, 2012; doi:10.1152/jn.00755.2012.

The role of envelope shape in the localization of multiple sound sources and echoes in the barn owl

Caitlin S. Baxter, Brian S. Nelson, and Terry T. Takahashi
Institute of Neuroscience, University of Oregon, Eugene, Oregon

Submitted 28 August 2012; accepted in final form 20 November 2012

Baxter CS, Nelson BS, Takahashi TT. The role of envelope shape in the localization of multiple sound sources and echoes in the barn owl. *J Neurophysiol* 109: 924–931, 2013. First published November 21, 2012; doi:10.1152/jn.00755.2012.—Echoes and sounds of independent origin often obscure sounds of interest, but echoes can go undetected under natural listening conditions, a perception called the precedence effect. How does the auditory system distinguish between echoes and independent sources? To investigate, we presented two broadband noises to barn owls (*Tyto alba*) while varying the similarity of the sounds' envelopes. The carriers of the noises were identical except for a 2- or 3-ms delay. Their onsets and offsets were also synchronized. In owls, sound localization is guided by neural activity on a topographic map of auditory space. When there are two sources concomitantly emitting sounds with overlapping amplitude spectra, space map neurons discharge when the stimulus in their receptive field is louder than the one outside it and when the averaged amplitudes of both sounds are rising. A model incorporating these features calculated the strengths of the two sources' representations on the map (B. S. Nelson and T. T. Takahashi; *Neuron* 67: 643–655, 2010). The target localized by the owls could be predicted from the model's output. The model also explained why the echo is not localized at short delays: when envelopes are similar, peaks in the leading sound mask corresponding peaks in the echo, weakening the echo's space map representation. When the envelopes are dissimilar, there are few or no corresponding peaks, and the owl localizes whichever source is predicted by the model to be less masked. Thus the precedence effect in the owl is a by-product of a mechanism for representing multiple sound sources on its map.

inferior colliculus; precedence effect; cocktail party effect; amplitude modulation

IN NATURE, SOUNDS OF INTEREST may be obscured by their own reflections and sounds from independently emitting sources. When the interfering sound is a reflection and the delay between the direct and reflected sounds is short, the sound arriving directly from the source dominates perception, and the reflection may not be perceived as a separate acoustical event. This phenomenon, "localization dominance," is a component of a set of psychoacoustical phenomena called the "precedence effect" (Blauert 1997; Gardener 1968; Litovsky et al. 1999; Wallach et al. 1949).

A recent study in the barn owl (*Tyto alba*) showed that localization dominance requires deep amplitude modulations (AMs) when the delay between the onsets of the leading and lagging sounds is removed (Nelson and Takahashi 2010). Without these AMs, the owl localized the two sources with equal frequency, in spite of the fact that the lagging carrier was a delayed copy of the lead. This suggests that the temporal relationship between the sounds of the pair is determined not from the carrier, but from the ongoing disparities between corresponding envelope features. The behavioral observations were consistent with an "envelope model" (below) that explained localization dominance in the owl, based on the selectivity of neurons in its auditory space map for binaural cues (Brainard et al. 1992; Euston and Takahashi 2002; Knudsen and Konishi 1978; Spezio and Takahashi 2003) and their tendency to discharge to upward AMs (Keller and Takahashi 2005; Nelson and Takahashi 2010).

Unlike independent sounds, reflections resemble direct sounds to varying degrees. Therefore, a reflected sound and an independent competing sound are, in principle, on a continuum of similarity to the sound of interest. Does the auditory system process echoes and sounds of independent origin differently? How does it differentiate between them?

We examined owls' head saccades to a pair of concomitant noise bursts while manipulating the similarity of the pair's envelopes to each other. As in Nelson and Takahashi (2010), the carriers were identical, except for a 2- to 3-ms delay. The envelope model computed the relative strengths of each sound's representation on the owl's auditory space map. The computed strengths predicted the probability with which the owl would localize one or the other source. Envelope pairs that caused the owl to localize the leading source tended to be similar to one another. This is consistent with the idea that a peak in the leading sound overlaps the rising edge of the corresponding peak in the lagging sound, preventing neurons from discharging to the lagging sound. Without corresponding peaks, the arbitrary shapes of the envelopes determined which of the two would evoke a stronger representation on the map, and the owl simply localized the sound with the stronger representation. Our results suggest that a direct sound with an echo and a pair of independent sounds are thus represented on the space map according to the same principles. The owl's auditory system does not appear to process echoes and sounds of independent origin differently.

MATERIALS AND METHODS

All experiments were carried out under a protocol approved by the Institutional Animal Care and Use Committee of the University of Oregon. The three barn owls used in this study [C, 5 yr old (at time of last experiment); S, 4 yr old; T, 5 yr old] were hand raised from hatching and kept in a captive-breeding colony under a permit from the US Fish and Wildlife Service.

Stimulus synthesis and presentation. The stimuli consisted of pairs of noise-modulated noises (40-ms duration; 2.5-ms liner ramps) presented from different locations in frontal space, generated by use of a custom MATLAB script (MathWorks). The carriers comprised repro-

Address for reprint requests and other correspondence: B. S. Nelson, Institute of Neuroscience, Univ. of Oregon, Eugene, OR 97403 (e-mail: bsnelson@uoregon.edu).

ducible, broadband noise bursts filtered between 2 and 10 kHz, the range relevant to the barn owl for sound localization. To synthesize the envelopes, we generated noises with random phase and energy below 150 Hz, near the upper cutoff of the owl's modulation transfer function (Dent et al. 2002; Keller and Takahashi 2000). These noises were transformed into the time domain by an inverse Fourier transform. The waveform was then scaled and DC shifted so that the minimal value was 0 and the maximal value was 1. A catalog of 50 carriers and 50 envelopes was thus generated.

To generate a lead/lag pair, an envelope and a carrier were selected at random (with replacement) from the catalog and duplicated. One copy of the carrier and of the envelope was delayed by 2 or 3 ms and designated the "lagging" carrier and envelope. The 2- or 3-ms segment at the start of the leading waveforms and the segment at the end of the lagging waveforms were removed, synchronizing the onsets and offsets of all the waveforms (synchronizing envelope; dashed lines Figs. 1, *A*, *C*) (Dizon and Colburn 2006; Zurek 1980). The delay, ∂, however, is present in the ongoing structure of the lead and lag carriers and envelopes. The envelopes of the leading and lagging sounds are designated, respectively, Env(t) and Env($t - \partial$). The leading and lagging carriers are designated Car(t) and Car($t - \partial$).

To vary the degree of their similarity, each envelope of a lead/lag pair (but not their carriers) was multiplied by a statistically independent, "decorrelating" envelope, $d_1(t)$ and $d_2(t)$, weighted by a scalar, which we termed the "decorrelating index" (DI). The envelope of the sounds to be presented from the leading and lagging speakers, S(t) and S($t - \partial$) respectively, are thus

$$S(t) = DI \cdot d_1(t) + (1 - DI) \cdot Env(t) \quad (1)$$

$$S(t - \partial) = DI \cdot d_2(t) + (1 - DI) \cdot Env(t - \partial) \quad (2)$$

The decorrelating envelopes, $d_1(t)$ and $d_2(t)$, were generated as described above for the original envelopes. The decorrelating index, DI, ranged from 0 to 1 in increments of 0.25. When DI = 0, S(t) and S($t - \partial$) were maximally correlated; when DI = 1, they were minimally correlated. An example of a partially correlated envelope pair generated in this manner is shown in Fig. 1*C*.

The degree of similarity of each envelope pair was quantified by computing the maximal value of the cross-covariance between the leading and lagging envelopes, ρ_{max}. To compute the cross-covariance, the mean of the envelopes was subtracted and the two waveforms were cross-correlated. Because of the arbitrary shapes of the envelopes used in this study, a single value of DI generated a distribution of ρ_{max} values whose means are related to DI. Because of the removal of the lead- and lag-alone segments, ρ_{max} approached but never reached 1, even when DI = 0. Figure 1, *B* and *D*, plots the cross-covariances of the envelope pairs in Figs. 1, *A* and *C*, respectively.

Finally, the carriers, $Car_{lead}(t)$ and $Car_{lag}(t - \partial)$, and corresponding envelopes, S(t) and S($t - \partial$), were multiplied, resulting in a pair of modulated noise bursts. The stimuli were converted to analog (TDT RX8) at a rate of 48,828 samples/s, amplified, and presented from loudspeakers. The sound pressure levels of the two noises (SPL re: 20 μPa), measured at the perch, were roved together in 1-dB increments between 27 and 33 dB across trials.

Note that as S(t) and S($t - \partial$) are made less similar, the terms "lead" and "lag" become poorly defined when referring to the envelopes. However, since the temporal relationship is still defined for the carriers, the terms "lead" and "lag" are retained when referring to the stimuli.

Equipment. Trials were conducted in a darkened, double-walled anechoic chamber (Industrial Acoustics IAC; 4.5 m × 3.9 m × 2.7 m). Sounds were presented from one or a pair of 10 dome tweeter(s) (2.9 cm, Morel MDT-39) mounted on an array located 1.5 m from the owl's perch. Speakers were calibrated weekly to ensure consistent output. A green LED (2.9 mm; λ = 568 nm), placed at the center of the array, served as the target that the owl had to fixate (within a 2.5° radius) to start a trial.

A custom-built, head-mounted, magnetic search-coil system (Remmel Labs) was used to measure the owl's head movements in azimuth and elevation. The output voltages of the search coils were digitized (1,000 samples/s; TDT RX8) during a saccade and stored on computer mass-storage media. Saccade trajectories and velocities were computed from the stored traces. Before each session, the coil was calibrated to ensure ± 2° accuracy. The subjects were monitored continuously throughout test sessions with an infrared camera and light source (Canon Ci-20R; IR-20W).

Behavior. Three barn owls used in this study were initially trained to make head saccades toward single sound sources. Because a barn owl's eyes are nearly immobile, unlike those of primates, its head-aim gives us a measure of the owl's localization.

During a testing session, a bird was tethered to a perch mounted in the center of the anechoic chamber, facing the speaker array. All lights in the owl's visible spectrum were extinguished in the chamber during testing, except for the LED at the center of the speaker array that the owl was required to fixate to initiate a trial.

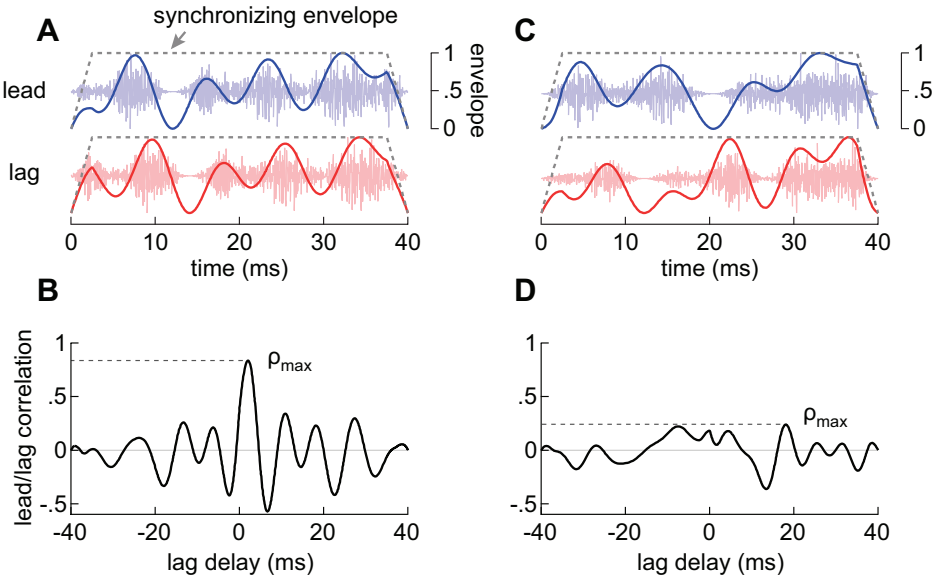

Fig. 1. Examples of the lead/lag pairs. The envelopes are plotted as thick lines and the carriers represented as faint, thin lines underneath. The lead source is plotted in blue and the lag in red. The stimuli were simultaneously gated (ramps shown as a gray dotted outline) to exclude onset and offset disparities, retaining only the portion of the stimuli that overlapped in time. In *A*, the envelopes of the lead/lag pair were 100% correlated except for the missing onset/offset disparities. The cross-covariance function of the envelopes shown in *A* is shown in *B*, with the peak ρ value (ρ_{max}) indicated by the dotted line. For the stimulus pair shown in *A*, ρ_{max} is high (0.84), and corresponding envelope features in the leading and lagging envelopes are obvious. *C* shows a lead/lag pair having a low ρ_{max} (0.24) with their cross-covariance plot shown in *D*. For the pair shown in *C*, corresponding envelope features are difficult to discern.

In a given session, two-thirds of the trials consisted of sound presentations from a single speaker chosen randomly on each trial. The stimuli in the remaining trials consisted of lead/lag stimuli from a pair of speakers, the positions of which were diametrically opposed across the center of the speaker array, i.e., an imaginary line between the two chosen speakers includes the central LED. When the speakers of a pair are diametrically opposed, the owl tends to choose one speaker over the other (Nelson and Takahashi 2008) instead of turning first toward one source, then toward the other in a given trial (Spitzer and Takahashi 2006). This speaker arrangement also makes it easier to discriminate between a lead-directed saccade and a lag-directed one. The speaker pair was randomly chosen for each trial from the five possible pairs. Whether a saccade was lead or lag directed was determined by measuring its polar angle, starting from the location where the saccade began to each subsequent point that was sampled along its trajectory, weighted by head-velocity (Nelson and Takahashi 2010). Figure 2 shows an example of a head saccade for a two-source trial in which the owl made a head saccade to the lag source. The speaker array was changed every four to five sessions to ensure that the birds did not memorize the speaker locations. The gray circles in Fig. 2 show the entire set of locations that the 10 speakers occupied.

Owls were rewarded with a piece of mouse dispensed from a remotely controlled feeder activated by the computer or by the experimenter from outside the anechoic chamber (Nelson and Takahashi 2008, 2010). Trials were excluded in which the owl moved but did not localize any part of the speaker array, for example, if the owl looked at the feeder or chamber door, began to groom itself, or performed a threat display. Trials with no movement whatsoever were also discarded. This left at least 17 acceptable trials per DI level per delay for each bird, with a mean sample size of 34.8 trials per DI for the 2-ms delay and 25.8 trials for the 3-ms delay between all birds. Each session continued until the bird had earned at least 20 rewards or became satiated. No bird was run more than once a day. Because the 2-ms and 3-ms data sets were statistically indistinguishable, the data were pooled.

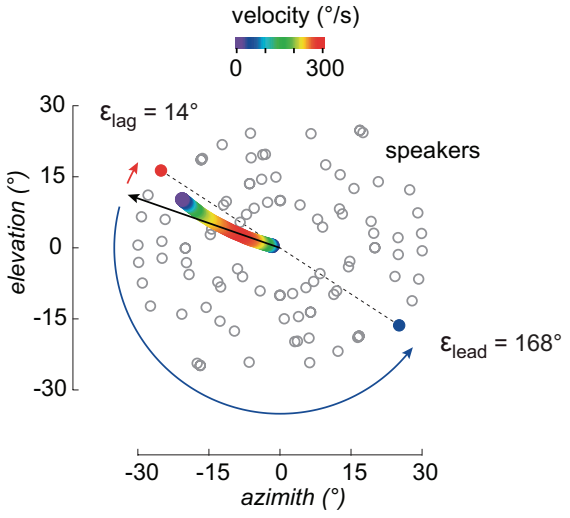

Fig. 2. Speaker array and head saccade. The gray circles mark the locations (in polar coordinates) from which leading and lagging sounds were presented. The horizontal axis is azimuth (Az) and the vertical axis is elevation (El). The trajectory of a head saccade from the central fixation point to a target speaker (red circle; −15° Az, 15° El) is plotted in a color gradient that represents velocity. In this trial, the bird localized the lag source (red circle) over the lead (blue circle). The quantities ϵ_{lead} and ϵ_{lag} are the errors, in degrees, of the bird's head saccade trajectory from the lead or lag source. The errors of the bird's head saccade trajectories to the lead and lag sources are also represented as blue or red arrows, respectively. The trajectory of the saccade was determined to be toward the source with the smaller error. The black arrow represents the velocity-weighted vector of the saccade.

Envelope model. The external nucleus of the barn owl's inferior colliculus contains a map of frontal auditory space, composed of neurons with spatial receptive fields (SRFs). The SRFs are the result of their selectivity to frequency-specific interaural time and level differences [ITD(f) and ILD(f)] that are generated by a stimulus in the neuron's preferred area of space. The envelope model is based on the observations that these space map neurons discharge when the binaural cues match those of their SRFs while the amplitude of that stimulus is rising. The model computed the responses of neurons in the space map based on these two principles, and, as we show below, predicted which of two targets the owl would localize. The envelope model is explained here as it pertains to the current stimuli. For more details, the reader is referred to the earlier paper (Nelson and Takahashi 2010).

When two broadband sounds with overlapping magnitude spectra are presented simultaneously from two speakers in the free field, ITD(f) and ILD(f) are vector sums of the values for each source's location, at a given instant (Blauert 1997; Keller and Takahashi 2005; Nelson and Takahashi 2010; Snow 1954; Takahashi and Keller 1994). The resultant vector in each frequency band is weighted by the amplitude of each source's signal. If the noises from the two sources are statistically independent of one another (i.e., uncorrelated), then, within any given frequency band at any given instant, the amplitude is likely to be higher for one source than the other. The ITD and ILD for that frequency band will approximate the values of the higher amplitude source. At another instant, the other source's amplitude may become higher. Over the course of the stimuli, the cues will spend roughly equal amounts of time assuming the values corresponding to the loci of the two sources. In the owl's space map, two uncorrelated noises evoke two separate foci of activity (Keller and Takahashi 1996a, 2005; Takahashi and Keller 1994). In the behavioral experiments below, the carriers of the leading and lagging noises were identical, except for the delay (∂) of 2 or 3 ms. In such case, a segment of the lagging noise is not aligned temporally with its corresponding segment in the leading noise because the corresponding segment in the leading sound occurred ∂ ms earlier (Fig. 1*A*). Thus a delay causes the two carriers to be uncorrelated, on a short time scale. In the barn owl, this time scale is about 0.1 ms (100 μs), so a 2- to 3-ms delay is sufficient to cause decorrelation and for the owl's space map to represent the two sources as separate foci of activity (Keller and Takahashi 1996a). Importantly, a pair of identical-but-delayed (by 2 or 3 ms) noises lacking deep AMs fails to evoke localization dominance; in other words, owls are equally likely to localize one source or the other on a given trial (Nelson and Takahashi 2010). We next explain the necessity for deep AMs and how they can bias the owl's perception of one or the other target.

If the source of one of the two independent noises, which we refer to as the "target," is located in a space map neuron's SRF, the neuron discharges when two conditions are simultaneously met (Keller and Takahashi 2005; Nelson and Takahashi 2010). The first condition is that the amplitude of the target is higher than that of the other sound of the pair, the "masker," outside the neuron's SRF. As noted above, the frequency-specific binaural cues will approach those of the sound with the higher amplitude. Thus, when the target's amplitude is higher, the ILDs and ITDs attain values similar to those generated at the cell's SRF. The second condition is that the envelope of the target must also be increasing. Like cells in a variety of auditory structures of many species, cells in the owl's space map tend to respond to rising AMs. These conditions and their relationship to the stimuli of the present study are illustrated in Fig. 3, which helps to explain the envelope model.

Figure 3*A* plots the sensitivity of the typical space map neuron against the difference, ΔE, between the amplitudes of the target and masker at a given moment. The function was estimated from the responses of isolated space map neurons to various values of ΔE obtained in an earlier study (Nelson and Takahashi 2010; see caption for equation). The values along the ordinate represent the probability

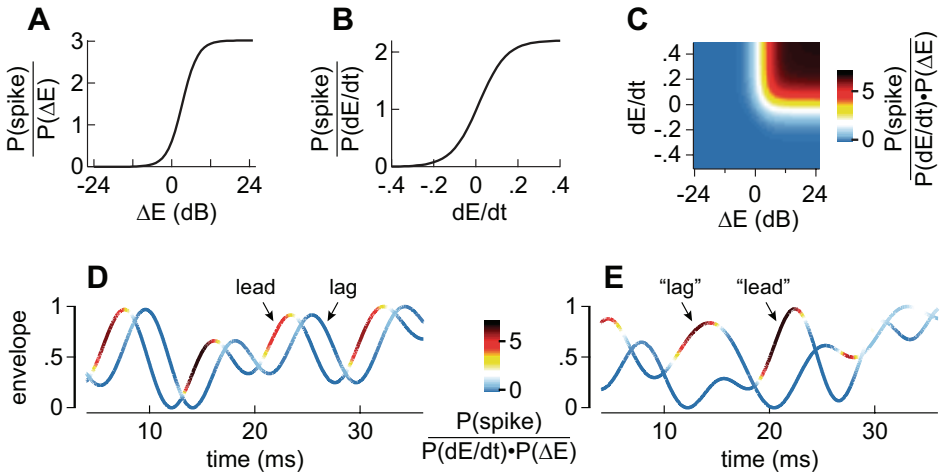

Fig. 3. Model description. *A*: the sensitivity of neurons of the owl's auditory space map to the difference between the amplitudes of the target and masker, ΔE. The function shown was estimated from the average probability of discharge across 28 isolated neurons obtained in an earlier study divided by the probability of occurrence of the various values of ΔE over the course of each stimulus (Nelson and Takahashi 2010). The sigmoid is described by the following equation: $P(\text{spike}) = 3.021/[1+e^{-(\Delta E - 3.25/2.384)}]$. *B*: the sensitivity of space map neurons is plotted against the derivative of the average envelopes of the 2 sounds, dE/d*t* (see MATERIALS AND METHODS). As in *A*, the probability of firing was divided by the probability of occurrence of the various values of dE/d*t*. The function shown is described thus: $P(\text{spike}) = 2.208/[1+e^{-(dE/dt - 0.014/0.0653)}]$. *C*: combined probabilities shown in *A* and *B*. The color value shown in the i^{th} row and j^{th} column of Fig. 3*C* is the product of the spiking probability at the i^{th} value of the derivative sigmoid (Fig. 3*B*) and the probability at the j^{th} value of the difference sigmoid, and ranges from low spiking probability (blue) to high spiking probability (red/black) (Fig. 3*C*; also see Fig. 6*E* of Nelson and Takahashi 2010). For example, the color value (blue) in the lower left corner is a product of the spiking probability for ΔE = −24 dB and dE/d*t* = −0.4. Neurons are most likely to discharge when both the ΔE and dE/d*t* of the target are positive. Conjunctions of these positive differences and derivatives are most abundant in the upper right quadrant of *C*, where both ΔE and dE/d*t* are positive. *D*: a lead/lag pair with highly correlated envelopes ($\rho_{max} = 0.84$). The probability of discharge is projected onto the envelopes as the color gradient from *plot C*, indicating the periods in which neurons are likely to spike. Because the envelopes in *D* are highly correlated, conjunctions causing high spiking probability occur preferentially for the leading stimulus. *E*: an envelope pair with minimal covariance ($\rho_{max} = 0.24$). Because of the low ρ_{max}, there are few corresponding envelope features, and the incidence of conjunctions are nearly equal for the lead and the lag (note that "lead" and "lag" are defined only by the carriers, which are not shown). For example, at roughly the 20-ms time point, ΔE favors the lead and the dE/d*t* of the leading envelope is positive as it rises. The model therefore predicts a stronger response to the leading sound than to the lagging sound at that time point. By contrast, at ≈12 ms, ΔE and dE/d*t* favor the lagging sound and the lag is predicted to elicit more spikes at that time point.

of firing divided by the probability of occurrence of various ΔE values over the course of the stimulus (see Fig. 5*C* of Nelson and Takahashi 2010). Neuronal activity is very low when ΔE <0, i.e., when the target has a lower amplitude than the masker, and increases sigmoidally for positive ΔE values. Similarly, Fig. 3*B* plots neuronal sensitivity to the derivative, dE/d*t*, of the average amplitude of the target and masker envelopes [$(Env_{target} + Env_{masker})/2$]. This function was also derived from single unit data obtained in the earlier study (see Fig. 5*F* of Nelson and Takahashi 2010). Neuronal spiking increases when the envelope of the target is increasing in amplitude. Spiking probability against both ΔE and dE/d*t*, based on single-unit recordings in the barn owl's space map, is plotted in Fig. 3*C* and shows that the highest spiking probability for space map neurons is found when both ΔE and dE/d*t* are positive (Nelson and Takahashi 2010).

Figure 3*D* shows the envelopes of a highly correlated lead/lag pair ($\rho_{max} = 0.84$). The probability of the leading and lagging stimuli to evoke spikes is shown as a color gradient from low spiking probability (blue) to high spiking probability (red/black). The leading stimulus contains several instances of high spiking probability, whereas the lagging stimulus contains no instances of high spiking probability. At a high envelope covariance and a delay of 2 or 3 ms, the lead will tend to have a greater spiking probability according to the model because it overlaps the lag in time, eliminating almost all of the instances during which the lag has a higher amplitude while its amplitude is rising. This is demonstrated in Fig. 3*D*: when the lag envelope is increasing in amplitude; e.g., at $t \approx 23$ ms, the lead has a higher envelope amplitude than the lag. During the instances when the lag envelope does have a higher amplitude than the lead, e.g., at $t \approx 28$ ms, the lag is no longer increasing in amplitude. Therefore, although each criterion may be found individually in the lag stimulus, there are few if any instances in which the two coincide. The result is that the lag fails to evoke a strong neural response. Increasing the delay between the lead and lag stimuli decreases the consistency with which the envelope peaks in the leading sound overlap and obscure the corresponding peaks in the lagging sound, thereby strengthening the representation of the lagging sound on the space map (Nelson and Takahashi 2010).

Figure 3*E* shows a pair of envelopes with a lower level of covariance ($\rho_{max} = 0.24$). The occlusion of the rising edges of the peaks in the lagging envelope by the leading envelope becomes less consistent as the covariance level diminishes, and, as a result, the lag begins to evoke stronger activity on the space map. When ρ_{max} is low, the leading and lagging stimuli occlude one another in nearly equal proportions on average.

The model predicted the relative discharge rate of the space map neurons to the leading sound, RP_{lead}:

$$RP_{lead} = P_{lead}/(P_{lead} + P_{lag})$$

where P_{lead} and P_{lag} are the probabilities of firing in space map neurons that represent the leading and lagging sources obtained from Fig. 3*C*.

RESULTS

Behavioral predictions of the envelope model. The envelope model predicts the strengths of the representations of a sound pair on the space map. Assuming that the sound with the stronger representation on the space map is more likely to elicit a saccade, this model should predict the sound of a pair toward which the owl is more likely to turn.

Results are shown in Fig. 4*A*, which plots, for each bird, the proportion of saccades to the lag against the relative probability of discharge to the lead, RP_{lead}. RP_{lead}- and lag-directed

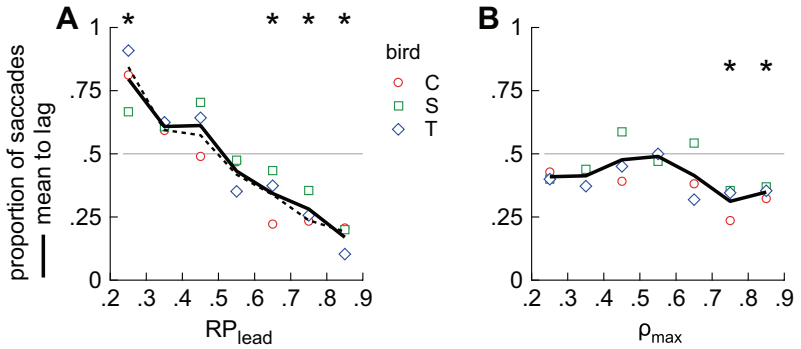

Fig. 4. *A*: Saccades plotted against spike probability. The proportion of saccades to the lag source was plotted against the probability of spikes to the lead source predicted by the envelope model. Relative probability to lead (RP_{lead}) was defined as $P_{lead}/(P_{lead} + P_{lag})$, where P_{lead} and P_{lag} represent the probability of neuronal discharge to the leading and lagging sounds, respectively. The data for each bird are plotted separately (C, red; S, green; T, blue), and the mean response for all 3 birds to the lag source is shown in black (linear regression parameters: $r = -0.987$, slope $= -1.035 \pm 0.076$, y-intercept $= 1.03 \pm 0.04$). Asterisks indicate values of ρ_{max} yielding statistically significant ($P < 0.05$; contingency table analysis) preferential localization of one source over the other. These data suggest a strong relationship between RP_{lead} and saccades to the lag or lead. The dashed line shows the mean response to the lag for all trials with $\rho_{max} \leq 0.7$ (linear regression parameters: $r = -0.976$, slope $= -1.037 \pm 0.103$, y-intercept $= 1.027 \pm 0.060$). Removal of those trials with $\rho_{max} > 0.7$ (see Fig. 5*B*) has little effect on the relationship between saccades and RP_{lead}. *B*: saccades plotted against envelope cross-covariance. The proportion of saccades to the lagging source out of the total number of saccades [Lag/(Lead + Lag)] is plotted against ρ_{max}. As in *A*, the data for each bird are plotted separately in red (C), green (S), and blue (T). The cross-bird mean of lag-directed saccades is shown in black. The proportion of saccades to the lag is lower for lead/lag pairs with high ρ_{max}, indicating localization dominance and a preference for the leading source. Statistically significant localization dominance is indicated by asterisks.

saccades are strongly and inversely correlated ($r = -0.987$, slope $= -1.035 \pm 0.076$, y-intercept $= 1.03 \pm 0.04$). The asterisks above the plot indicate the values of RP_{lead} at which saccades to the lag were significantly different from 0.5 ($P < 0.05$, contingency table analysis). When the model predicted that the two stimuli would evoke equal activity (0.5 on the abscissa), the owls turned their heads toward each of the two stimuli with equal frequency; however, when $RP_{lead} > 0.6$, the owls preferentially turned toward the leading stimulus, exhibiting localization dominance.

Importantly, when RP_{lead} was less than 0.3, the owls preferentially localized the lagging sound, even though the delay, defined in the carriers of the lead and lag, was only 2 or 3 ms, values at which one would have expected the preferential localization of the lead if localization dominance depended on the carrier.

The term "localization dominance" has traditionally meant the preferential localization of the leading source. Our observation suggests that the lagging source can also dominate.

Localization dominance and the role of envelope similarity. The relationship between the similarity of the lead and lag envelopes, ρ_{max}, and the proportion of lag-directed saccades is shown in Fig. 4*B* for each bird (C, red; S, green; T, blue; mean of birds, black line). Values of ρ_{max} for which the proportion of saccades to the lag was significantly less than 0.5 are marked by asterisks ($P < 0.05$; contingency table analysis). As shown, the number of lead- and lag-directed saccades is not significantly different from 50% when $\rho max < 0.7$. Above 0.7, the birds preferentially localized the leading sound (i.e., % lag-directed saccades < 0.5; asterisks, $P < 0.05$; contingency table analysis). Thus localization dominance depends on the similarities of the lead and lag sounds' envelopes.

One parsimonious and more proximal explanation is that RP_{lead} tends to favor the leading sound when the envelopes of the two sounds bear some resemblance to one another. Therefore, we analyzed the relationship between RP_{lead} and the similarity of envelopes, ρ_{max}. The ρ_{max} value for each lead/lag pair is plotted against its corresponding RP_{lead} value, color coded to indicate whether the pair elicited a lead-directed (blue) or lag-directed (red) saccade, in Fig. 5*A*. Plotted thus, the reason for the dependency of localization dominance on envelope similarity becomes evident. There is a dense cluster of blue points in the upper right, indicating that when ρ_{max} is high, RP_{lead} tends to be high, and correspondingly, the saccades are toward the leading sound.

The inset to the right of Fig. 5*A* shows the number of lead-directed (blue) and lag-directed (red) saccades plotted against ρ_{max}. Above $\rho_{max} = 0.7$, the lead-directed saccades predominate. Asterisks indicate statistically significant differences between the numbers of lead- and lag-directed saccades. Below 0.7, the number of lead- and lag-directed saccades is roughly equal. Note that even when $\rho_{max} > 0.7$, the difference in the proportions of lead- and lag-directed saccades is modest. The scatterplot in Fig. 5*A* shows that, for some envelope pairs with high ρ_{max} values, RP_{lead} actually favors the lagging sound, and the owls turned toward the lag as predicted by RP_{lead}.

The relationship between RP_{lead} and the relative number of saccades, shown in Fig. 4*A* and the upper inset of Fig. 5*A*, could be an artifact of the abundance of lead-directed saccades when $\rho_{max} > 0.7$. Figure 5*B* replots the same data, having removed those trials in which $\rho_{max} > 0.7$. The upper inset shows that the relationship between saccades and RP_{lead} persists. Specifically, lead-directed saccades continue to predominate when $RP_{lead} > 0.5$, and lag-directed saccades predominate when $RP_{lead} < 0.5$. If the saccades to the leading sound or lagging sound are plotted in terms of proportions, we arrive at the dashed black line shown in Fig. 4*A*, which continues to show the strong linear relationship ($r = -0.976$, slope $= -1.037 \pm 0.103$, y-intercept $= 1.027 \pm 0.060$).

The inset above the scatterplot (Fig. 5*A*) graphs the number of lead- and lag-directed saccades against RP_{lead} and shows that the predicted activity on the space map can favor either sound source. Correspondingly, to the left of the vertical line at $RP_{lead} = 0.5$ (Fig. 5*A*), the number of lag-directed saccades (red line) exceeds the number of lead-directed saccades (blue line). The reverse is true to the right of that line.

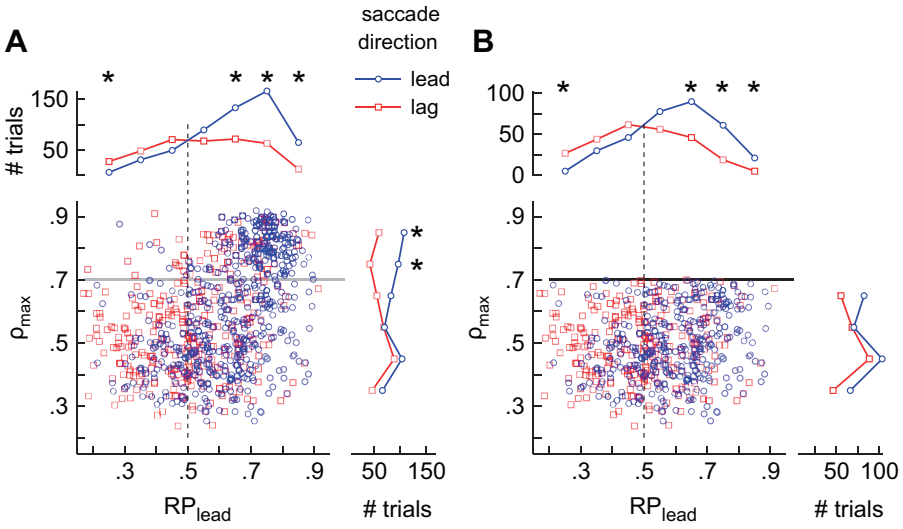

Fig. 5. Relationship Between RP_{lead} and ρ_{max}. *A*: scatterplot of ρ_{max} vs. RP_{lead}. Each trial is represented by an open circle, with blue circles for stimulus pairs that evoked lead-directed saccades and red circles for pairs that evoked lag-directed saccades. The value of RP_{lead} where the lead and lag stimuli are equally likely to elicit spikes is marked by a dashed line, and $\rho_{max} = 0.7$ is marked by a gray horizontal line. The numbers of lead- and lag-directed trials are plotted alongside the scatterplot, according to the trials' ρ_{max} values (*right*) or P_{lead} values (*top*). Bins where either the lead or lag source was preferentially localized are marked by asterisks. Note that there is a dense cluster of lead-directed trials above $\rho_{max} = 0.7$, indicating that a high covariance level increases the probability of response to the lead source. *B*: RP_{lead} and ρ_{max} are plotted as in *A*, except that trials that used stimulus pairs with $\rho_{max} > 0.7$ (black horizontal line) have been removed. Their removal has minimal effect on the RP_{lead} values for which statistically significant preferential localization of one source over the other was observed.

DISCUSSION

By examining envelope structure, we were able to predict which of two concurrent, broadband noise sources an owl would localize. The noise carriers were identical except for a 2- or 3-ms delay, which is enough to decorrelate the noise carriers and allow the owl's auditory system to represent and localize two separate sources (Keller and Takahashi 1996a, 2005; Nelson and Takahashi 2010). We showed that envelope structure serves to weight the two neural representations according to the relative incidences of the conjunction of positive ΔE and dE/dt for arbitrary envelope pairs. In turn, owls were found to localize the source predicted to have the stronger neural representation.

Analysis of the stimuli with the model demonstrated that when envelopes of the sounds in a pair are similar ($\rho_{max} > 0.7$), the relative incidence of the conjunction of positive differences and derivatives tends to favor the leading sound. This is due to the occlusion of the rising edges of peaks in the lag by corresponding peaks in the lead, which causes neurons representing the lagging source's location to respond weakly. When two sounds are dissimilar, an envelope peak in one sound does not have a corresponding peak in the other, so one sound is just as likely to obscure the other. In nature, the similarity between an acoustical reflection and its direct sound is graded. Our results suggest that the strength of the activity on the owl's auditory space map reflects this continuum. In other words, echoes are not treated in any special way, for example by a mechanism devoted to their suppression. Our explanation of the precedence effect is similar to that that of Tollin (1998) and of Trahiotis and Hartung (Hartung and Trahiotis 2001; Trahiotis and Hartung 2002), who explained the precedence effect observed with clicks without recourse to inhibition. The ability to predict the neural representation and localization of multiple sound sources from stimulus structure is a step toward identifying acoustical environments that may prove challenging for auditory stream segregation. We also note that the envelope model links the ability of neurons to respond to a temporal feature, i.e., upward AMs, with the ability to localize sounds in acoustical clutter. This raises the possibility that a loss of temporal acuity, e.g., in presbyacusis (Walton 2010), is related to a loss in spatial acuity (Abel et al. 2000; Dobreva et al. 2011). Although further studies are obviously necessary to establish causality, the ideas behind the envelope model do not depend on owl-specific features. Their applicability and relevance to human hearing is testable, provided that the frequencies used by humans for sound localization and the temporal responses of the peripheral filters at those frequencies are incorporated.

Relationship to the precedence effect. In nature, sounds of interest may be obscured by their reflections and by sounds from independently emitting sources. Unlike independent sounds, reflections that arrive after a short delay are less noticeable under everyday listening conditions, which would suggest that echoes are processed differently. Mechanisms for the suppression of echoes, such as lateral inhibition, have therefore been suggested (Gaik 1993; Lindemann 1986a,b; Pecka et al. 2007; Yin 1994; Zurek 1980).

Yet, noting that echoes are rarely perfect copies of the direct sound, others have asked what features two sounds must share to be considered a direct sound and its reflection (Blauert and Divenyi 1988; Divenyi 1992; Divenyi and Blauert 1987; Shinn-Cunningham et al. 1995). Human listeners experience the precedence effect even when the leading and lagging noise carriers are uncorrelated (Blauert and Divenyi 1988; Divenyi and Blauert 1987) and when narrowband carriers differ by well over a critical bandwidth (Divenyi 1992; Shinn-Cunningham et al. 1995), suggesting that the carriers do not need to be similar. The present results in the owl implicate the envelope instead of the carrier (also Dizon and Colburn 2006; Nelson and Takahashi 2010). The carriers of the leading and lagging sounds used in the present study were identical, so the lead/lag delay could, in theory, have been determined from the carrier regardless of the envelope similarity. Nonetheless, localization dominance failed when the lead and lag envelopes were dissimilar or when the envelopes were not modulated, suggesting that it is the envelope, not the carrier, that must be similar. In the studies with human subjects (Divenyi 1992; Shinn-Cunningham et al. 1995), the delay between the onsets of the sounds was left intact and is presumed to have provided an envelope cue for establishing the lead/lag relationship.

Limitations and generalizability of the envelope model. The envelope model was developed to predict the representation of

two concomitantly active sound sources with overlapping amplitude spectra. In this study, the envelope model was used to predict which of two sources an owl would localize. Although the model was relatively successful in predicting saccade behavior in this challenging stimulus configuration, its performance in other acoustical conditions remains to be explored.

The sound pairs used in the present study were presented at equal amplitudes and with synchronized onsets and offsets; however, the envelope model can be generalized to the more natural condition in which neither of these experimental manipulations are present. First, the amplitudes of reflected sounds are typically lower than that of the sound of interest. Independent interfering noises may also be of lower amplitude, for example, when the masking source is farther away. Under such conditions, ΔE will consistently favor the direct sound or the sound of interest, and their representations on the space map will be stronger. This is consistent with observations in human listeners that the effect of a lead/lag delay on the perceived location of a sound source can be compensated by the relative amplitude of the sources (Snow 1954).

Second, the onsets and offsets of sounds in nature are often (though not always) preserved and perceptible. Onsets and offsets are corresponding envelope features from which the temporal relationship between two stimuli may be determined. The onset, in particular, may have a stronger influence on localization than the envelope features during the superposed portion of the sound pair. Since the leading sound is briefly present alone, all frequency-specific binaural cues are spatially coherent and correspond to the location of the leading source for that brief time period. Furthermore, space map neurons would be expected to fire a strong burst at the onset of the leading sound that settles to a lower, steady-state level as adaptation sets in (Keller and Takahashi 2000; Nelson and Takahashi 2008, 2010). This onset burst may contribute more spikes to the space map's representation of the leading sound than a similar peak during the superposed segment, thereby favoring the space map's representation of the leading sound even if the leading and lagging sounds are equally represented during the superposed segment. Of course, the lagging sound is present by itself at the end of the stimulus pair, but this "lag-alone" segment would be expected to contribute less than the lead-alone segment, because the neurons tuned to the lagging-source's location would respond at the lower steady-state level.

The present study involved a localization task. Spatial hearing can also be evaluated by estimating the minimal audible angle (MAA), which quantifies the ability to discriminate changes of sound-source position. Using a pair of identical noise bursts, one of which was delayed relative to the other, Spitzer and colleagues (2003) showed that the barn owl's MAA for the lagging sound source was considerably coarser than that for the leading sound source. This phenomenon, which may be analogous to "lag-discrimination suppression" described for human listeners (Litovsky and Shinn-Cunningham 2001; Shinn-Cunningham et al. 1993), is consistent with the weaker representation of the lagging source observed in recordings from space map neurons (Keller and Takahashi 1996b; Nelson and Takahashi 2010; Spitzer et al. 2004) and predicted by the envelope model when the leading and lagging envelopes are similar ($\rho_{max} > 0.7$). As the envelopes become less similar, the model predicts that the sound with the lagging carrier may actually evoke more spikes in space map neurons than that with the leading carrier. In this case, it is possible that the MAA for the sound with the leading carrier would be coarser than that of the sound with the lagging carrier.

Finally, how would the envelope model perform under conditions with more than a single interfering sound, whether they are independent sounds or reflections? Devore and colleagues (2009) have shown that in reverberant environments, which have reflections from multiple surfaces as well as higher order reflections, the location of the direct sound can be determined from the localization cues during the brief period before reflections arrive and distort the cues corresponding to the location of the direct sound. Although our study removed the lead (and lag) alone segments, the same principle is applicable and is captured by the envelope model: a particular sound, whether it is the sound of interest or the masker, will have binaural cues that are most spatially coherent when the amplitude of that sound is highest. Over the course of the stimulus, one source or another will have an amplitude that is momentarily higher than those of the others and will be most localizable. These segments of the sound stream when one of the sounds momentarily stands out from amidst the others can happen during the course of the stimulus (present study; Nelson and Takahashi 2010), and not only at the beginning of the sound (Devore et al. 2009). If such "glimpses" are more consistently available for one of the sound sources, it should be preferentially localized by the owl, according to the envelope model.

ACKNOWLEDGMENTS

We thank Dr. C. H. Keller, Dr. E. A. Whitchurch, Dr. C. E. Carr, and J. M. Donovan for helpful discussions throughout this project.

Present address for C. S. Baxter: Department of Biology, University of Maryland, College Park, MD 20742.

GRANTS

This study was supported by a grant from the National Institute of Deafness and Communication Disorders (R01 DC003925 and AARA Supplement R01 DC003925-11S1).

DISCLOSURES

Dr. C. E. Carr, Editor, *Journal of Neurophysiology*, is C. Baxter's doctoral advisor at the University of Maryland, College Park.

AUTHOR CONTRIBUTIONS

C.S.B. and B.S.N.: equal contributions in data collection, analysis, writing/editing of manuscript. T.T.T.: writing/editing of manuscript.

REFERENCES

Abel SM, Giguere C, Consoli A, Papsin BC. The effect of aging on horizontal plane sound localization. *J Acoust Soc Am* 108: 743–752, 2000.

Blauert J. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT, 1997.

Blauert J, Divenyi PL. Spectral selectivity in binaural contralateral inhibition. *Acustica* 66: 267–274, 1988.

Brainard MS, Knudsen EI, Esterly SD. Neural derivation of sound source location: resolution of spatial ambiguities in binaural cues. *J Acoust Soc Am* 91: 1015–1027, 1992.

Dent ML, Klump GM, Schwenzfeier C. Temporal modulation transfer functions in the barn owl (Tyto alba). *J Comp Physiol A* 187: 937–943, 2002.

Devore S, Ihlefeld A, Hancock K, Shinn-Cunningham B, Delgutte B. Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron* 62: 123–134, 2009.

Divenyi PL. Binaural suppression of nonechoes. *J Acoust Soc Am* 91: 1078–1084, 1992.

Divenyi PL, Blauert J. On creating a precedent for binaural patterns: when is an echo an echo? In: *Auditory Processing of Complex Sounds*, edited by Yost WA and Watson CS. London: Psychology Press, 1987, p. 344.

Dizon RM, Colburn HS. The influence of spectral, temporal, and interaural stimulus variations on the precedence effect. *J Acoust Soc Am* 119: 2947–2964, 2006.

Dobreva MS, O'Neill WE, Paige GD. Influence of aging on human sound localization. *J Neurophysiol* 105: 2471–2486, 2011.

Euston DR, Takahashi TT. From spectrum to space: the contribution of level difference cues to spatial receptive fields in the barn owl inferior colliculus. *J Neurosci* 22: 284–293, 2002.

Gaik W. Combined evaluation of interaural time and intensity differences: psychoacoustic results and computer modeling. *J Acoust Soc Am* 94: 98–110, 1993.

Gardener MB. Historical background of the Haas and/or precedence effect. *J Acoust Soc Am* 43: 1243–1248, 1968.

Hartung K, Trahiotis C. Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. *J Acoust Soc Am* 110: 1505–1513, 2001.

Keller CH, Takahashi TT. 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, 1996a.

Keller CH, Takahashi TT. Localization and identification of concurrent sounds in the owl's auditory space map. *J Neurosci* 25: 10446–10461, 2005.

Keller CH, Takahashi TT. Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J Neurophysiol* 84: 2638–2650, 2000.

Keller CH, Takahashi TT. Responses to simulated echoes by neurons in the barn owl's auditory space map. *J Comp Physiol A* 178: 499–512, 1996b.

Knudsen EI, Konishi M. A neural map of auditory space in the owl. *Science* 200: 795–797, 1978.

Lindemann W. Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *J Acoust Soc Am* 80: 1608–1622, 1986a.

Lindemann W. Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front. *J Acoust Soc Am* 80: 1623–1630, 1986b.

Litovsky RY, Colburn HS, Yost WA, Guzman SJ. The precedence effect. *J Acoust Soc Am* 106: 1633–1654, 1999.

Litovsky RY, Shinn-Cunningham BG. Investigation of the relationship among three common measures of precedence: fusion, localization dominance, and discrimination suppression. *J Acoust Soc Am* 109: 346–358, 2001.

Nelson BS, Takahashi TT. Independence of echo-threshold and echo-delay in the barn owl. *PLoS One* 3: e3598, 2008.

Nelson BS, Takahashi TT. Spatial hearing in echoic environments: the role of the envelope in owls. *Neuron* 67: 643–655, 2010.

Pecka M, Zahn TP, Saunier-Rebori B, Siveke I, Felmy F, Wiegrebe L, Klug A, Pollak GD, Grothe B. Inhibiting the inhibition: a neuronal network for sound localization in reverberant environments. *J Neurosci* 27: 1782–1790, 2007.

Shinn-Cunningham BG, Zurek PM, Durlach NI. Adjustment and discrimination measurements of the precedence effect. *J Acoust Soc Am* 93: 2923–2932, 1993.

Shinn-Cunningham BG, Zurek PM, Durlach NI, Clifton RK. Cross-frequency interactions in the precedence effect. *J Acoust Soc Am* 98: 164–171, 1995.

Snow W. The effects of arrival time on stereophonic localization. *J Acoust Soc Am* 26: 1071–1074, 1954.

Spezio ML, Takahashi TT. 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, 2003.

Spitzer MW, Bala AD, Takahashi TT. Auditory spatial discrimination by barn owls in simulated echoic conditions. *J Acoust Soc Am* 113: 1631–1645, 2003.

Spitzer MW, Bala AD, Takahashi TT. A neuronal correlate of the precedence effect is associated with spatial selectivity in the barn owl's auditory midbrain. *J Neurophysiol* 92: 2051–2070, 2004.

Spitzer MW, Takahashi TT. Sound localization by barn owls in a simulated echoic environment. *J Neurophysiol* 95: 3571–3584, 2006.

Takahashi TT, Keller CH. Representation of multiple sound sources in the owl's auditory space map. *J Neurosci* 14: 4780–4793, 1994.

Tollin DJ. Computational model of the lateralisation of clicks and their echoes. In: *Proceedings of the NATO Advanced Study Institute on Computational Hearing*, edited by Greenberg S and Slaney M. Il Ciocco, Italy: NATO Scientific and Environmental Affairs Division, 1998, p. 77–82.

Trahiotis C, Hartung K. Peripheral auditory processing, the precedence effect and responses of single units in the inferior colliculus. *Hear Res* 168: 55–59, 2002.

Wallach H, Newman E, Rosenzweig M. The precedence effect in sound localization. *Am J Psychol* 62: 315–336, 1949.

Walton JP. Timing is everything: temporal processing deficits in the aged auditory brainstem. *Hear Res* 264: 63–69, 2010.

Yin TC. Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J Neurosci* 14: 5170–5186, 1994.

Zurek PM. The precedence effect and its possible role in the avoidance of interaural ambiguities. *J Acoust Soc Am* 67: 953–964, 1980.