Auditory spatial discrimination by barn owls in simulated echoic conditions

Matthew W. Spitzer, a) Avinash D. S. Bala, and Terry T. Takahashi

Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403

(Received 11 December 2001; revised 20 December 2002; accepted 20 December 2002)

In humans, directional hearing in reverberant conditions is characterized by a “precedence effect,” whereby directional information conveyed by leading sounds dominates perceived location, and listeners are relatively insensitive to directional information conveyed by lagging sounds. Behavioral studies provide evidence of precedence phenomena in a wide range of species. The present study employs a discrimination paradigm, based on habituation and recovery of the pupillary dilation response, to provide quantitative measures of precedence phenomena in the barn owl. As in humans, the owl’s ability to discriminate changes in the location of lagging sources is impaired relative to that for single sources. Spatial discrimination of lead sources is also impaired, but to a lesser extent than discrimination of lagging sources. Results of a control experiment indicate that sensitivity to monaural cues cannot account for discrimination of lag source location. Thus, impairment of discrimination ability in the two-source conditions most likely reflects a reduction in sensitivity to binaural directional information. These results demonstrate a similarity of precedence effect phenomena in barn owls and humans, and provide a basis for quantitative comparison with neuronal data from the same species. © 2003 Acoustical Society of America.

PACS numbers: 43.66.Qp, 43.66.Pn, 43.66.Gf [LRB]

I. INTRODUCTION

The barn owl (Tyto alba) is a nocturnal predator that makes use of well-developed acoustic and neuronal specializations to localize and capture prey. In a natural environment, the task of localizing a target sound source is complicated by reflections that convey conflicting directional information. Nevertheless, barn owls, like many other animals, are able to localize sounds in echoic conditions. Numerous human psychophysical studies have addressed spatial hearing abilities using stimuli that simulate a direct sound source and its reflection. In general, the results indicate that directional information conveyed by the direct (leading) source dominates perception, and that listeners are relatively insensitive to directional information conveyed by reflections (lagging sources). Perceptual dominance of the leading sound is commonly referred to as the “precedence effect.” Psychophysical studies of the precedence effect have typically measured one of three basic phenomena (Litovsky et al., 1999). At short delays between leading and lagging sounds, subjects report “perceptual fusion,” meaning that the direct sound and reflection are heard as a single acoustic event (reviewed in Blauert, 1997). “Localization dominance” refers to the finding that information conveyed by the direct sound dominates perception of the location of the fused image, although the actual perceived location is also influenced to a lesser extent by the location of the lag source (Wallach et al., 1949; Shinn-Cunningham et al., 1993). Studies addressing “lag discrimination suppression” have shown that a listener’s ability to detect changes in directional attributes of a lagging source is impaired relative to that for single or leading sources (Zurek, 1980; Perrott et al., 1989; Litovsky and Macmillan, 1994). The latter two phenomena indicate that, despite the perceptual dominance of the leading source, a representation of the lagging source is retained, and that this representation contributes to perception.

Behavioral data suggest that precedence effects occur in a wide variety of species, including barn owls (Keller and Takahashi, 1996); rats (Kelly, 1974); cats (Cranford, 1982); and crickets (Wytenbach and Hoy, 1993). In contrast to human psychophysical studies, previous animal behavioral studies have used lateralization tasks, in which the ability to distinguish between left-leading and right-leading sound pairs is measured as a function of the delay between leading and lagging sources. In all species tested, lateralization performance is very good for intermediate delay values (owl: 2 to 10 ms; cat: 0.3 to 3 ms; rat: 0.5 to 4 ms) and decreases at shorter or longer delays. Thus, at intermediate delays the presence of the lag does not interfere with the ability to lateralize the lead source. At longer delays it is presumed that the lag source becomes audible as a separate event, causing a decline in performance as subjects begin to lateralize both lag and lead sources. The range of delays at which optimal performance occurs in all animals, except crickets, is similar to the optimal range (1 to 5 ms) for discrimination of the interaural time difference of a leading sound by humans (Litovsky and Shinn-Cunningham, 2001). The fact the precedence effect occurs across classes, and even phyla, indicates that it is a fundamental process in spatial hearing.

Physiological studies have documented neuronal corre-
lates of the precedence effect in a variety of animal models. Single-unit studies have addressed the processing of lead–lag sound pairs by neurons that encode spatial cues (interaural time difference, ITD) at several stages of the mammalian auditory pathway (Yin, 1994; Fitzpatrick et al., 1995; Litovsky and Yin, 1998; Fitzpatrick et al., 1999), and by space-tuned neurons in the barn owl’s auditory midbrain (Keller and Takahashi, 1996). A common finding of these studies is that the response to a stimulus at a neuron’s best spatial location or ITD is reduced when that stimulus is preceded, at an interval of a few milliseconds, by a stimulus at another location or ITD. Thus, at short delays the neuronal representation of the ITD, or spatial location, of a lagging sound is at least partially attenuated. It is unclear how such neuronal effects relate to precedence phenomena at the behavioral level, as comparable measures of neuronal and behavioral responses have yet to be obtained in the same species.

The present study was designed to provide behavioral measures of the precedence effect in barn owls that can be compared directly to similar measures derived from neuronal data. A spatial discrimination paradigm was used to facilitate such a comparison, as both behavioral and neuronal discrimination performance can be quantified using equivalent metrics derived from signal detection theory (Britten et al., 1992). An added advantage of spatial discrimination paradigms is that it is possible to address processing of both lead and lag sources explicitly (Perrott et al., 1989; Litovsky and Macmillan, 1994; Litovsky, 1997). In addition to providing a basis for comparison with physiological data, results of this study are relevant to understanding the generality, or lack thereof, of precedence phenomena in different species. This is, to our knowledge, the first study to measure spatial discrimination of lead and lag sources in any nonhuman species.

II. EXPERIMENT 1: SPATIAL DISCRIMINATION OF SINGLE, LEAD, AND LAG SOURCES

The first experiment measured the ability of owls to reliably detect changes in the azimuth of single, leading, or lagging sound sources. Previous human studies using free-field stimulus presentation have demonstrated that minimum audible angles (MAAs) for lag sources are substantially larger than those for single sources (Perrott et al., 1989; Litovsky and Macmillan, 1994; Litovsky, 1997), while lead source MAAs are intermediate between those for single and lag sources (Litovsky and Macmillan, 1994; Litovsky, 1997). A recently developed discrimination paradigm (Bala and Takahashi, 2000), based on habituation and recovery of the pupillary dilation response (PDR), was used to provide comparable measurements from barn owls.

A. Methods

Subjects were three adult barn owls from a captive breeding colony, all of which had previously served as subjects in similar experiments. Each owl had a steel headplate surgically attached to its skull. Following recovery from surgery, owls were gradually acclimatized to being restrained in the experimental apparatus in three or four sessions over several weeks (Bala and Takahashi, 2000). During experiments, owls were restrained in a cloth jacket, and supported by a metal frame. The position of the owl’s head was fixed in a constant position relative to a speaker array by attaching the headplate to a holder mounted on the supporting frame [Figs. 1, 2(A)]. Once acclimatized, owls would lay quietly in the experimental apparatus without struggling throughout test sessions with durations of 1 to 2 h. During test sessions, owls were continuously monitored using an infrared sensitive video camera.

Experiments were conducted in a double-wall, sound-isolating anechoic chamber (Industrial Acoustics Company, IAC). The inner walls, ceiling, and floor were lined with 6-in. Metadyne 150 (IAC) anechoic absorption cones leaving...
an inner dimension of 4.5×3.9×2.7 m. A wire-cable meshwork formed a suspension floor over the floor cones. The Metadyne cones and their supporting panels formed a suspension floor over the floor cones. The chamber deviated from the inverse-square law by no more than ±1.5 dB between 3 and 10 kHz measured along transects extending between opposite corners of the room. The ambient noise inside the chamber was below 18 dB SPL$_A$.

1. Acoustic stimulation

Stimuli were presented from three speakers mounted on a hoop forming an arc in the azimuthal plane at eye level (Fig. 1). For measuring discrimination of sources at 3° separations, speakers were positioned 1.98 m from the owl’s head. For all other measurements, speakers were positioned at a distance of 0.79 m. Stimulus waveforms were generated digitally on a computer and converted to analog form at 30,000 samples per second (Tucker-Davis Technologies Power DAC PD1 or Modular Instruments M308). Analog signals were attenuated with programmable attenuators (PA4, Tucker Davis Technologies, or M300, Modular Instruments), amplified (HB6, Tucker-Davis Technologies, or MacIntosh 754) and delivered to the speakers (Peerless, 2-in. cone tweeters). Stimuli were “frozen” noise segments (i.e., the same waveform was repeated on each trial) with durations of 25 ms, for owls 889 and 882, or 100 ms, for owl 896, with onsets and offsets shaped by 2.5-ms cosine ramps. Frozen noise was used to prevent false positive responses that might otherwise occur if the owls could detect the differences between random noise stimuli (see Behavioral measurements, below). Digital stimulus waveforms were generated with flat spectra from 3 to 9 kHz. The final stimulus spectrum, shaped by the transfer functions of the speakers, was flat from 3 to 9 kHz (±3 dB). Transfer functions of the different speakers did not differ by more than 2.5 dB at any frequency from 3 to 9 kHz. Stimulus sound-pressure level was 74 dB SPL$_A$, measured with a 1-in. microphone (Bruel & Kjaer model 1760) and sound-level meter (Bruel & Kjaer type 2235) at the approximate position of the owl’s head before the beginning of each test session.

2. Behavioral measurements

The acoustically evoked PDR was measured using an infrared (IR) pupillometer, consisting of an IR-emitting diode and an IR-detecting diode [Figs. 1, 2(A)]. The right eye was held open by threads taped to the eyelids, leaving the nictitating membrane free to moisten the cornea. The pupillometer was positioned such that the IR emitter illuminated the entire cornea, and the detector sampled light reflected by the ventral-nasal quadrant of the retina. The voltage output of the IR detector, thus, varied in proportion to the dilation of the pupil. Output of the IR detector was amplified, digitized at a sampling frequency of 1875 samples per second, and stored on computer disk. The magnitude of the PDR was quantified by summing the area under the IR-detector voltage trace for 2 s following stimulus onset, after first “zeroing” the trace by subtracting the voltage level measured at sound onset. Spatial discrimination was measured using a method based on habituation and recovery of the PDR described in a previous report (Bala and Takahashi, 2000). The PDR is a reflexive dilation of the pupil evoked by any salient, novel stimulus, such as a sound, in the owl’s environment. Repetition of the same sound causes the PDR to habituate. After the

FIG. 2. Measurement of acoustically evoked pupillary dilation response (PDR). (A) Positioning of the pupillometer. (B) Initial habituation of the PDR in owl 882. Pupillary responses, measured as the voltage output of the pupillometer, have progressively diminishing amplitudes on the first four habituating trials of a test session (h1 through h4). The habituating stimulus is a lead–lag pair of 25-ms noise bursts (delay=3 ms), with the lead source located at 6° azimuth. Stimulus onset occurs at time=0 s. (C) Recovery from habituation. On the first test trial (t1, overall trial number 160), the location of the lead sound is changed by 20°. The amplitude of the evoked PDR (black line) is greater than those evoked on the three preceding habituating trials (gray lines).
PDR has habituated to one sound, presentation of a different sound can evoke a “recovered” PDR, with amplitude comparable to that of the response to initial presentation of the first sound. Thus, the degree to which a second test sound recovers the PDR from habituation to the first habituating sound can be used as a measure of the owl’s ability to detect the difference between the two sounds.

Individual test sessions consisted of 300 stimulus trials with an intertrial interval varying randomly from 9 to 12 s. In a single session, a constant habituating stimulus was presented for the first 119 trials. Following this initial habituation period, a test stimulus was presented every 40 trials. Spatial discrimination was measured in three conditions (Fig. 1): (1) In the single source condition, the habituating stimulus consisted of a noise burst presented from a speaker at 6° azimuth (Fig. 1, “H”). On test trials the same noise burst was presented from another speaker, displaced to the right (Fig. 1, “T”). The azimuthal separations tested included all (owl 882) or some (owls 889 and 896) of the following: 3°, 5°, 7°, 10°, 15°, 20°. (2) In the lead source condition, the habituating stimulus consisted of a noise burst from the speaker at 6° azimuth followed, at a delay of 3 ms, by the same noise burst from a speaker at a fixed location of −20° azimuth (Fig. 1, “F”). On test trials the azimuth of the lead source was changed. Each owl was tested with all (owl 882) or some (owls 889 and 896) of the following lead source separations: 3°, 5°, 10°, 15°, 20°, 25°. (3) In the lag source condition, the habituating stimulus consisted of noise burst from the speaker at −20° azimuth followed, at a delay of 3 ms, by the same noise from the speaker at 6° azimuth. On test trials the azimuth of the lag source was varied. Each owl was tested with a subset of the following lag source separations: 3°, 5°, 10°, 15°, 20°, 25°, 30°, 35°, 40°. The speakers at positions H, T, and F were always the same, regardless of test separation, to control for the possibility that the owls could distinguish between habituating and test speakers. Using the same frozen noise segment as the stimulus on all trials insured that recovered responses were not evoked by detectable differences in the stimulus waveforms across trials, other than those caused by changing the location of a speaker. Discrimination performance was quantified using the nonparametric statistic, proportion correct \( p(c) \), calculated from empirical receiver operating characteristic (ROC) curves, as described in the Results section.

Minimum audible angles were estimated from psychometric functions obtained by plotting discrimination performance \( p(c) \) as a function of azimuthal separation between habituating and test sources. A Gauss–Newton least-squares method was used to fit the psychometric data with sigmoid functions of the form

\[
p(c) = \delta - (\delta - 0.5) \exp\left[-(x/\alpha)\beta\right], \tag{1}
\]

where \( x \) is azimuthal separation between habituating and test speakers, \( \alpha \) is a threshold parameter that determines the displacement of the function along the abscissa, \( \beta \) is a slope parameter (Britten et al., 1992). The parameter corresponding to asymptotic performance level at large azimuthal separations, \( \delta \), was made equal to the highest performance level measured for each owl under each test condition. This constraint resulted in more meaningful sigmoid fits for lag discrimination by owls 882 and 896, and otherwise had little effect on the resulting curve fits, particularly within the region of steepest slope that was used for threshold determination. This method resulted in significant fits (F test, \( p < 0.01 \)) for all conditions in all owls, with one exception (owl 896, single sources). In cases where the curve fit was significant, the MAA was estimated from the fit curve by determining the source separation corresponding to a discrimination performance level of 0.71 \( p(c) \). The single source MAA for owl 896 was conservatively estimated as the smallest source separation for which discrimination performance exceeded 0.71 \( p(c) \).

B. Results

Spatial discrimination for single, lead, and lag sources was measured using habituation and recovery of the PDR in three owls. Representative habituation and recovery sequences from a single owl (owl 882) are illustrated in Figs. 2 and 3. Initial presentation of the habituating stimulus (consisting of a 25-ms noise burst from the speaker at 6° azimuth, followed at a 3-ms delay by the same noise from the speaker at −20° azimuth) evokes a large pupillary dilation, as indicated by a large positive voltage deflection [h1 in Fig. 2(B)]. On subsequent trials (h2 through 4), the response to repetitions of the same stimulus diminishes progressively. After the magnitude of responses to the habituating stimulus has declined to a relatively stable level, a test stimulus is presented. In this example, the location of the lead sound is changed by 20°. The magnitude of the PDR evoked on this test trial [Fig. 2(C, t1)] is larger than those evoked on the preceding habituating trials (gray lines). Thus, a change in lead source location serves to recover the PDR from habituation.

Responses of the same owl from four test sessions with the same stimulus configuration show similar patterns of habituation and recovery (Fig. 3). In each session, the amplitude of responses to the habituating stimulus exhibits a decreasing trend that is most evident for the first ten trials, and continues until a relatively stable response level is obtained by trial 80, or sooner. The magnitude of responses to test stimuli (unfilled large circles) is generally larger than that to habituating stimuli (filled circles). Thus, both habituation and recovery from habituation appear to be stable. To facilitate comparison of our results with the majority of previous human and animal studies of the precedence effect, an attempt was made to use shorter stimuli. However, in initial experiments measuring single source discrimination with sound durations of 15 ms or less it was observed that, following the first test stimulus, responses to successive test stimuli habituated. For owls 889 and 882, stable recovery from habituation was obtained when stimulus duration was increased to 25 ms. For owl 896, stimulus duration had to be increased to 100 ms to achieve stable recovery.

In the test sessions shown, it is clear that the mean response magnitude for test stimuli is greater than that for habituating stimuli, but there is also considerable overlap between the distributions of response magnitudes for the two types of stimuli. The difference in mean response magnitude
for test and habituating stimuli, relative to the pooled variance of responses, can provide a measure of the owl’s ability to discriminate the test stimulus from the habituating stimulus. Because the number of test stimuli in a single session had to be limited to prevent the owl from habituating to the test stimulus, responses were pooled across multiple sessions to obtain reliable estimates of discriminability for a given test condition. Despite the overall similarity of responses, however, there were differences in mean response magnitude and variance between sessions, most likely reflecting small differences in placement of the pupillometer relative to the eye, and in retraction of the eyelids. In order to pool responses across test sessions it was necessary to first normalize responses from individual sessions. For this purpose, the amplitudes of responses to individual habituating and test trials were converted to z-scores, relative to the distribution of habituating trial responses after trial 100. Responses from the initial 100 trials of each session were discarded to insure that the calculated mean reflected the amplitude of the fully habituated PDR.

The owls’ abilities to discriminate between test and habituating stimuli were quantified by the nonparametric statistic, $p(c)$, calculated by determining the area under probability-space ROC curves generated from the pupillary responses. A nonparametric discrimination measure was used because the variance of responses to test and habituating stimuli often differed. To generate ROC curves, normalized responses were pooled from four or more test sessions with the same test condition and test azimuth separation from a single owl. ROC curves were generated by applying methods developed by Britten et al. (1992) for use with neuronal response data. The continuously distributed PDR amplitude data were treated as results of a rating experiment, in which the amplitude of the response on each trial was considered as the subject’s rating of the likelihood that a test stimulus was presented. Following this assumption, “hits” were defined as responses to test stimuli that exceeded an arbitrary response criterion, and “false alarms” were defined as responses to habituating stimuli that exceeded criterion. Probability-space ROC curves were then obtained by plotting the proportion of hits against the proportion of false alarms for each of a set of criterion values, spanning the range of measured response amplitudes, and including values above and below the extreme response amplitudes. The discrimination metric, $p(c)$, is calculated by determining the area under the ROC curve. This quantity varies from 0.5, indicating chance performance, to 1.0, indicating perfect discrimination, and is equivalent to the proportion correct in a two-alternative forced choice task (Green and Swets, 1966).

Examples of ROC curves obtained from pooled response data for single-, lead-, and lag-source discrimination by owl 889 are shown in Fig. 4. In each test condition, the ROC curve for the smallest test separation is close to the unity line (dashed), indicating discrimination performance near chance levels [$p(c)=0.5$]. The area under the ROC curves generally increases as test separation is increased, reflecting an increase in the mean amplitude of test responses relative to habituated responses. Note, however, that in the single- and lead-source conditions, maximum discrimination was achieved at the second largest test separation. In a few cases,
the ROC curves for discrimination of large test separations were clearly asymmetric (e.g., single source: 20°; lead source: 10°, 15°). This effect occurs when the variance of responses to test and habituating responses differs (Green and Swets, 1966), and was often observed at larger test separations because the variance of test response amplitudes tended to increase in proportion to the mean amplitude.

Psychometric functions for azimuthal discrimination of single, lead, and lag sources were obtained from three owls by plotting \( p(c) \) as a function of test separation (Fig. 5). For all three owls, the discrimination functions for lag sources (squares) have shallower slopes than those for single (circles) and lead sources (triangles). Consequently, for arbitrary criterion performance levels from 0.65 to 0.85 \( p(c) \), the MAAs for lag sources are greater than those for lead sources, with the magnitude of threshold elevation being dependent on criterion level. In addition, for all three owls, the MAA for single sources was less than that for lead sources. Thus, the ability to resolve sound source locations in either of the two-source conditions was impaired relative to that in the single-source condition, with the greater impairment occurring in the lag-source condition. Discrimination performance of each owl reached chance levels at the smallest test separations in the lead- and lag-source conditions, indicating that owls were unable to discriminate between the speakers used to present habituating and test stimuli.

Discrimination performance of each owl was summarized by determining MAAs for the three test conditions (Table I). MAAs were estimated by fitting the discrimination data to sigmoid curves (see Methods), and calculating the angular separation at which the fit curves achieved an arbitrary threshold value of 0.71 \( p(c) \). This threshold measure was used, in preference to the threshold parameters of the curve fits, because the individual psychometric functions for single- and lead source conditions appear to reach saturation at different \( p(c) \) values, and because we did not test sufficiently large lag source separations to be sure that the lag source performance curves reached saturation. The 0.71 \( p(c) \) criterion was chosen to be consistent with the most similar human psychophysical study (Litovsky, 1997).

An alternative, statistical criterion was also used to measure azimuthal discrimination thresholds. The statistical threshold (ST) is defined as the smallest test separation at which the mean amplitude of responses to test stimuli was significantly greater than the mean amplitude of responses to habituating stimuli \((p<0.005, \text{Mann-Whitney U-test})\). Responses to test and habituating stimuli were pooled across test sessions as in the preceding analysis. ST values obtained in this manner are shown in Table I. Within the limits of sampling resolution, results obtained with this method agree with those of the former analysis: In all cases, both methods identify transitions from subthreshold to suprathreshold performance between the same pairs of test separations.

Threshold elevation factors (TEF), defined as the ratio of MAAs in the two-source and single-source conditions (Litovsky, 1997), are also shown in Table I to facilitate comparison with results of previous human psychophysical studies. A significant curve fit could not be obtained for owl 896 in the single source condition. In this case, the MAA was estimated as the smallest azimuthal separation (3°) at which discrimination performance exceeded threshold. This value is in agreement with results of a previous study, using similar methods, in which it was shown that the owl’s MAA measured using 100-ms sounds is approximately 3° (Bala and Takahashi, 2000)

III. EXPERIMENT 2: DISCRIMINATION OF NONSPATIAL CUES ASSOCIATED WITH LEAD–LAG COMBINATIONS

Results of the preceding experiment suggest that sensitivity to the binaural cues that barn owls use to localize sounds (Moiseff, 1989) is reduced, but not eliminated, in simulated echoic conditions. This conclusion depends on an assumption that discrimination was based on the difference in perceived location of habituating and test stimuli. It is possible, however, that discrimination was based on differences in the perceived quality of sounds between test and
that were filtered using an owl’s head-related transfer functions for the appropriate lead (−20°) and lag (6° or 36°) locations. Lead–lag delay was 3 ms. The period of spectral rippling, equal to the reciprocal of the delay between lead and lag waveforms, is plotted as a function of lag source azimuth in Fig. 6(B) for habituating and test stimuli used in the lag discrimination condition of experiment 1. Changing the azimuth of the lag source results in orderly changes of the ripple period at the left (circles, solid line) and right (×’s, dotted line) ears. In humans, changes in spectral ripple period give rise to changes in perceived pitch (Yost et al., 1978). In addition to spectral rippling, changing the location of one source causes small differences in the overall shape of the power spectrum [i.e., differences in heights of neighboring peaks in Fig. 6(A)] and changes in the temporal fine structure of monaural signals. The latter effect is illustrated in Fig. 6(C). For each ear, the waveform received in the habituating condition was cross correlated with that received in each of the lag-source test conditions. The dotted and solid functions plot the maxima of these normalized cross correlograms (coherence of monaural signals against the lag source’s azimuth for the right (dotted) and left (solid) ears, respectively. Increasing the azimuthal separation of lag sources on habituating and test trials results in a orderly decrease in coherence of the monaural signals at either ear, providing additional potential cues for discrimination.

It is thus possible that the owls discriminated test trials from habituating trials by detecting changes in the quality of the sounds and not changes of their perceived locations. If this were the case, results of experiment 1 may underestimate the reduction of spatial sensitivity in echoic conditions. To control for this possibility, we measured discrimination using stimuli approximating the signals that would have been present at one ear on habituating and test trials of experiment 1, presented from a single, stationary speaker.

A. Methods

Experimental subjects and PDR measurement procedures were the same as those in experiment 1. General aspects of experimental design were also similar, so only differences will be detailed. All stimuli were presented from a single speaker at 6° azimuth. The initial habituation period consisted of 99 trials. Three different test stimuli, corresponding to different lag-source locations, were tested in each session. After 100 habituating trials, test stimuli were presented in blocks of three, with each test trial separated by nine intervening habituating trials. Test blocks were separated by 20 habituating trials. Five test blocks were presented in each session.

Stimuli similar to the waveforms resulting from the summation of lead and lag sounds at an owl’s ear in experiment 1 were generated by filtering sounds with head-related transfer functions (HRTFs) corresponding to lead and lag source locations, delaying the lag signal by 3 ms, and adding the resulting waveforms. Filtering with HRTFs replicates the transformation of both phase and amplitude spectra that occurs between a source at a given location and a position in the external ear canal immediately adjacent to the tympanum. HRTFs were measured as previously described (Keller...
et al., 1998) in an owl not used as a test subject. Using this method, we measured the ability of owls to discriminate stimuli that approximate the waveforms present at one ear on habituating and test trials in experiment 1. Stimuli corresponding to those in the lag discrimination condition were chosen, because spatial discrimination was most impaired for lag sources.

The input stimulus waveforms were the same frozen noise segments used in experiment 1. To generate the lead sound, the waveform was convolved with the time-domain representation of the HRTF for the lead location. To generate the lag sound, a 3-ms onset segment, consisting of zeros, was concatenated to the beginning of the noise, and the resulting signal was convolved with the HRTF for the lag location. The final signals, approximating the monaural signals for each pair of lead and lag locations, were generated by adding the appropriate pairs of lead and lag signals. Because the HRTFs used were measured in another owl, the resulting stimuli do not exactly replicate the monaural waveforms experienced in experiment 1. However, the relative time delays for sources at different locations, that give rise to spectral rippling, are very similar in different owls (Keller et al., 1998). Furthermore, prominent spectral notches that give rise to large individual differences in HRTFs of mammals (Middlebrooks et al., 1989; Musicant et al., 1990; Spezio et al., 2000) are not present in the HRTFs of barn owls at the locations along the frontal horizon used in this study (Keller et al., 1998). Thus, although the method used does not exactly replicate the monaural signals experienced in experiment 1, it generates habituating and test signal waveforms that differ in spectral shape and temporal correlation in a manner similar to the actual monaural waveforms in that experiment.

Owls were tested for their ability to discriminate the different monaural signals generated in the lag source discrimination condition of experiment 1. Thus, in each case the habituating stimulus was generated by filtering the lead waveform with the left- or right-ear HRTF for −20° azimuth, and the lag waveform with the same ear’s HRTF for +6° azimuth. The test stimuli were generated by filtering the lead waveform with the HRTF for −20°, and the lag waveform with the HRTFs corresponding to a subset of the lag source locations tested in experiment 1. In this manner the three owls were tested for discrimination of monaural combination equivalent to the following lag source test separations:

Owl 889...left ear: 5°, 10°, 15°, 20°, 30°
Owl 882...left ear: 10°, 20°, 30°; right ear: 5°, 10°, 20°, 30°, 34°
Owl 896...left ear: 10°, 20°, 34°; right ear: 3°, 15°, 34°.

B. Results

Discrimination performance for the summed lead–lag signals is plotted as a function of lag source azimuth separation for test stimuli in Fig. 7. Discrimination performance for left- and right-ear signals is plotted as black and gray circles, respectively. The psychometric curves for lag source discrimination from Fig. 5 are replotted for comparison (squares, dotted lines). Owls 882 and 896 were unable to discriminate any of the test signals from the habituating signal [Figs. 7(B) and (C)]. Owl 889 achieved a maximum discrimination level of 0.69 \( p(c) \) for the combination signal corresponding to a lag source separation of 10° [Fig. 7(A)]. Discrimination performance declined to chance levels at larger equivalent source separations. There was no indication that discrimination performance increased with increasing equivalent source separation for any of the subjects, as would be expected if lag-source discrimination were dependent on sensitivity to monaural cues. Furthermore, in all cases discrimination performance was at chance levels for combination signals corresponding to lag source separations that were themselves discriminable in experiment 1. The preceding results were confirmed by statistical analysis of responses, pooled across sessions. For owl 889, the mean response to the test stimulus equivalent to a 10° separation was significantly greater in amplitude than the mean response to the corresponding habituating trials at a significance level of \( p = 0.007 \) (Mann-Whitney U test). For no other stimulus in any owl was there a significant \((p < 0.05)\) difference between response amplitudes on test and habituating trials. These results indicate that sensitivity to the spectral and temporal cues available at either ear is insufficient to account for discrimination of changes in lag source location.

<table>
<thead>
<tr>
<th>Owl ID (sound duration)</th>
<th>Condition</th>
<th>Single source</th>
<th>Lead source</th>
<th>Lag source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MAA</td>
<td>ST</td>
<td>MAA</td>
</tr>
<tr>
<td>889 (25 ms)</td>
<td></td>
<td>4°</td>
<td>5°</td>
<td>8°</td>
</tr>
<tr>
<td>882 (25 ms)</td>
<td></td>
<td>8°</td>
<td>10°</td>
<td>12°</td>
</tr>
<tr>
<td>896 (100 ms)</td>
<td></td>
<td>(3°)</td>
<td>3°</td>
<td>11°</td>
</tr>
</tbody>
</table>

TABLE I. Minimum audible angles (MAA) and statistical thresholds (ST) for the three owl subjects. MAA values represent the angular separation corresponding to a discrimination performance of 0.71 \( p(c) \), obtained from the sigmoid curve fits and rounded to the nearest degree. ST values represent the smallest tested azimuthal separation at which the mean amplitude of responses to test stimuli, pooled across sessions, was significantly greater than the mean amplitude of responses to habituating stimuli. Threshold elevation factors (TEF), shown for the two-source conditions, are the ratio of the corresponding MAA to the single-source MAA. The single-source MAA and associated TEF values for owl 896 are shown in parentheses to indicate that they are estimated values (see the text).
IV. DISCUSSION

This study represents the first application of a spatial discrimination paradigm to measure the precedence effect in an animal model. The results demonstrate that the ability to detect a change in the location of a sound source is impaired in the presence of a simulated reflection. The ability to localize the simulated reflection itself is even further impaired. Results of the control experiment (experiment 2) demonstrate that discrimination of changes in source location in echoic conditions was not based on sensitivity to changes in sound quality associated with changes in the monaural signals. Thus, the impairment of spatial resolution in simulated echoic conditions most likely reflects a reduced ability to use binaural directional information.

Results for azimuthal discrimination obtained in experiment 1 are similar to those obtained in comparable human psychophysical studies. Several previous human studies have...
demonstrated lag discrimination suppression (Litovsky et al., 1999), an impaired ability to detect changes in the azimuthal location or binaural attributes of lagging sounds relative to either single- or leading-sounds. This effect has been demonstrated using either free-field (Perrott et al., 1989; Litovsky and Macmillan, 1994; Litovsky, 1997) or head-phone (Zurek, 1980; Saberi and Perrott, 1990; Shinn-Cunningham et al., 1993; Tollin and Henning, 1998; Litovsky and Shinn-Cunningham, 2001) stimulus presentation. The human study that used the most similar stimulus presentation methods and parameters to those employed in the present study also reported lead discrimination suppression (Litovsky, 1997). Furthermore, the magnitude of impairments for lead and lag sources reported in that study, as quantified by TEFs, are roughly equivalent to those measured with stimuli of the same duration (25 ms) in the present study. The similarity of the present results to those of previous human studies suggests that similar processes may give rise to the precedence effect in barn owls and humans.

A. Effect of acoustic interactions on available binaural cues

Azimuthal sound localization by barn owls is primarily dependent on sensitivity to interaural time differences (ITD; Moiseff, 1989; Poganiatz et al., 2001), which are detected via neural processes analogous to interaural cross correlation (Carr and Konishi, 1990). One factor that may contribute to impaired spatial resolution in echoic conditions is degradation of ITD cues caused by acoustic superposition of lead and lag sounds. The effect of superposition of lead and lag waveforms on the ITD cues available in experiment 1 is illustrated in Fig. 8. A single source at +20° azimuth generates a constant ITD of 49 μs. In this and the following plots the monaural acoustic signals were simulated by filtering the stimulus waveform used in experiment 1 with an owl’s head-related transfer functions. Running interaural cross correlations were constructed by calculating the cross-correlation function of the signals at the two ears in a 0.5-ms window that was advanced in 33-μs increments throughout the total stimulus duration. Running time is plotted on the x axis, ITD on the y axis (positive values indicate the right signal leads), and interaural correlation (IAC) on the z axis and color scale. (B) The combination of a lead source at −20° azimuth and a lag source at +20°, with a lag delay of 3 ms, generates a time-varying pattern of interaural correlation. The ITDs of lead and lag sources, presented in isolation, are indicated by pointers and white lines on the plot surface. Here, the monaural signals were generated by summation of lead and lag waveforms that had been filtered by the HRTFs for the appropriate source locations. (C) The effects of peripheral filter response times on the available ITD cues are illustrated by constructing the interaural cross correlation after passing the monaural signals through a bank of gamma-tone filters (Slaney, 1993), modified to match the bandwidths of barn owl auditory-nerve fibers (Köppl, 1997). At each time increment, the interaural cross-correlation function was computed separately in each frequency channel. The resulting frequency-specific cross-correlation functions were then averaged and displayed as in plot B.

FIG. 8. Running interaural cross correlograms illustrate the effects of lead–lag superposition on ITD cues available in experiment 1. (A) A single source at +20° azimuth generates a constant ITD of 49 μs. In this and the following plots the monaural acoustic signals were simulated by filtering the stimulus waveform used in experiment 1 with an owl’s head-related transfer functions. Running interaural cross correlograms were constructed by calculating the cross-correlation function of the signals at the two ears in a 0.5-ms window that was advanced in 33-μs increments throughout the total stimulus duration. Running time is plotted on the x axis, ITD on the y axis (positive values indicate the right signal leads), and interaural correlation (IAC) on the z axis and color scale. (B) The combination of a lead source at −20° azimuth and a lag source at +20°, with a lag delay of 3 ms, generates a time-varying pattern of interaural correlation. The ITDs of lead and lag sources, presented in isolation, are indicated by pointers and white lines on the plot surface. Here, the monaural signals were generated by summation of lead and lag waveforms that had been filtered by the HRTFs for the appropriate source locations. (C) The effects of peripheral filter response times on the available ITD cues are illustrated by constructing the interaural cross correlogram after passing the monaural signals through a bank of gamma-tone filters (Slaney, 1993), modified to match the bandwidths of barn owl auditory-nerve fibers (Köppl, 1997). At each time increment, the interaural cross-correlation function was computed separately in each frequency channel. The resulting frequency-specific cross-correlation functions were then averaged and displayed as in plot B.
window duration was chosen to illustrate the available ITD cues with a temporal resolution exceeding that of the barn owl’s binaural system (Wagner, 1991), while being long enough to allow the lowest frequency stimulus components to contribute to the cross-correlation function. When a leading source is added at −20° azimuth, with a lead–lag delay of 3 ms, the pattern of interaural correlation over time becomes more complex [Fig. 8(B)]. In the initial and final 3-ms stimulus segments, only one of the two sources is active, and prominent peaks of interaural correlation occur at ITDs of −49 or +49 μs, corresponding to the lead or lag source locations, respectively. These ITDs are indicated by pointers and white lines on the plot surface. In the intervening 22-ms segment, both sources are active and superposition of lead and lag waveforms at each ear results in a complex, time-varying pattern of interaural correlation. Throughout much of this overlap segment, peaks are still evident at, or near ITDs of −49 and +49 μs. However, the level of interaural correlation at each of these ITDs varies considerably over time, and peaks of interaural correlation also occur, intermittently, at other ITDs. The average pattern of interaural correlation within the initial, overlap, and final stimulus segments is illustrated in Fig. 9(A). Within the overlap segment, two peaks are evident near the ITDs of lead and lag sources. However, the maximum level of interaural correlation is substantially reduced relative to the initial and final segments. The peak level of interaural correlation within the overlap stimulus segment varied with test speaker azimuth location, and was below 0.5 at all test separations below 40° in experiment 1 [Fig. 9(B)].

The reduction of interaural correlation caused by acoustic superposition of lead and lag waveforms would be expected to contribute to the reduction of azimuthal discrimination abilities for lead and lag sources. Saberi et al. (1998) measured the effect of decreasing interaural correlation on head turns elicited by dichotic noise bursts in barn owls. The precision of head turns was unaffected as interaural correlation was decreased from 1 to about 0.5. At interaural correlation levels below 0.5 the precision of head turns declined rapidly. Thus, the reductions of interaural correlation caused by acoustic superposition in the present study would be expected to impair spatial resolution. In order to minimize this effect, the owls could have adopted a strategy of listening only to the initial stimulus segment, in the case of lead discrimination, or to the final stimulus segment, in the case of lag discrimination. However, their ability to isolate the ITD information conveyed by the initial and final segments from that conveyed by the overlap segment would be limited by the temporal resolution of the binaural pathway. Wagner (1991) demonstrated that, in a lateralization task, barn owls average ITD information within a “temporal window” with equivalent rectangular duration ranging from 3 to 42 ms. If the owls processed directional information with a temporal window of similar duration in the present study, spatial resolution would have been reduced as a result of averaging degraded ITD information conveyed by the overlap segment with the uncorrupted ITD information conveyed by either initial or final segments. This effect is illustrated in Fig. 10, which shows the time-dependent value of binaural correlation computed within two double exponential time windows of the form
\[ W(t) = (1-w) \ast \exp(-t/T_p) + w \ast \exp(-t/T_s), \]

where \( T_p \) and \( T_s \) are time constants and \( w \) is a weighting factor. The short (solid line; \( T_p = 1 \text{ ms}, T_s = 12 \text{ ms}, w = 0.83 \)) and long (dotted line; \( T_p = 3 \text{ ms}, T_s = 64 \text{ ms}, w = 0.81 \)) time windows used in this illustration were generated using parameters that provided the best fits to the behavioral data for the two owls in Wagner’s study, and had equivalent rectangular durations of 3 and 14.5 ms, respectively.

In both cases, the binaural correlation values decline as the time window center nears the temporal midpoint of the stimulus, and the maximal levels of binaural correlation occurring near stimulus onset and offset are lower than those measured using rectangular time windows to isolate the initial and final stimulus segments [Fig. 9(A)]. Thus, temporal averaging should limit the ability of owls to make use of the ITD information conveyed by the initial and final stimulus segments. The effect of acoustic superposition on ITD cues illustrated here may contribute to the impairment of discrimination performance for lead and lag sources relative to single sources. It cannot explain the difference in performance between lead-and lag-source conditions, however, because the ITD cues to the locations of both sources are equally affected.

**B. Localization dominance**

A common interpretation of results obtained in lateralization and discrimination studies of the precedence effect is that lead and lag sounds are combined, perceptually, to form a single, fused event. The perceived lateral position of the fused event reflects a sum of the ITDs of lead and lag sources, weighted so that the lead source makes the dominant contribution. As a result, the amount by which the ITD of one of the sources must be changed to produce a just-detectable change in the perceived location of the fused event is greater for the lag source than it is for the lead source. This interpretation is formalized in a model proposed by Shinn-Cunningham et al. (1993), that uses a single metric to summarize lead-source dominance. According to this model, the perceived lateral position, \( \alpha \), is a noise-corrupted weighted sum of the ITDs of the leading \( (\tau_1) \) and lagging \( (\tau_2) \) sounds

\[ \alpha = c \tau_1 + (1-c) \tau_2 + \eta, \]

where \( \eta \) is internal noise. The perceptual weighting factor, \( c \), quantifies the extent to which the laterality of the lead sound dominates the laterality of the perceptual image.

This model can be used to calculate the perceptual weighting factor from the discrimination data obtained in the present study with free-field stimulus presentation, following an approach detailed by Litovsky and Macmillan (1994). Because ITD is directly proportional to sound source azimuth, Eq. (2) can be used to derive the following equations that relate the change in perceived azimuthal location of fused images \( (\delta \alpha) \) to the change in the azimuth of either lead \( (\delta \Delta z_1) \) or lag \( (\delta \Delta z_2) \) sources:

\[ \delta \alpha = c \delta \Delta z_1, \]  
\[ \delta \alpha = (1-c) \delta \Delta z_2. \]

The discriminating, \( d' \), of a difference in azimuth is defined as the difference in means of the distributions of perceived azimuths for two locations in units of their common standard deviations. The discriminating ability of a change in azimuth for single, lead, and lag sources is thus

\[ d'_s = c \delta \Delta z, \]

\[ d'_l = c \delta \Delta z_1, \]

\[ d'_r = (1-c) \delta \Delta z_2, \]

where \( s \) is the standard deviation of \( \alpha \). We assume that the variance of perceptual judgments is equal in all three conditions. Because the threshold criterion level of discriminating ability is the same for all test conditions, the right-hand expressions of any pair of Eqs. (5) can be set equal to one another to give three equations that relate the discrimination thresholds obtained in different test conditions \( (\delta \text{single}, \delta \text{lead}, \delta \text{lag}) \) in terms of \( c \)

\[ \delta \text{single}/\delta \text{lead} = c, \]

\[ \delta \text{single}/\delta \text{lag} = 1-c, \]

\[ \delta \text{lead}/\delta \text{lag} = (1-c)/c. \]

Estimates of \( c \) calculated using Eqs. (6) and the threshold measurements from experiment 1 are listed in Table II. All but two of the nine estimates of \( c \) are greater than 0.5, consistent with the view that, overall, the lead source tended to dominate perceived laterality of a fused image. There was considerable variation in the values of \( c \) computed for each owl with different equations, and across subjects for Eqs. (6a) and (6b). The values of \( c \) computed by Eq. (6c) exhibited the least variation across subjects. For owls 889 and 896 the \( c \) values given by Eq. (6a) were lower than those given by Eq. (6c), and the values given by Eq. (6b) were higher than those given by Eq. (6c). This result is expected if some factor, in addition to the pulling effect of the stationary source, acts to reduce discrimination abilities in the two source conditions. A likely cause of the observed discrepancy is the reduction of binaural correlation resulting from acoustic superposition. For this reason, we believe that the \( c \) values given by Eq. (6c), which does not require comparison of thresholds in one- and two-source conditions, provides the best estimate of perceptual weighting in the present discrimination task.

Shinn-Cunningham et al. (1993) demonstrated that the value of \( c \) depends on several stimulus parameters, including sound duration and lag delay. The values of \( c \) calculated from our data for barn owls are lower, on average, than those obtained in studies using short duration stimuli, including 6-ms noise bursts (Litovsky and Macmillan, 1994), and
clicks (Wallach et al., 1949; Gaskell, 1983; Yost and Soderquist, 1984; Saberi and Perrott, 1990; Shinn-Cunningham et al., 1993). The barn owl c values are comparable to those calculated by Shinn-Cunningham et al. (1993) from data obtained by Zurek (1980) in a pointer matching experiment, using continuous noises as stimuli at lag delays from 3 to 9 ms. Thus, it is possible that the low c values obtained in the present study reflect similar effects of stimulus duration to those in humans.

Several mechanisms have been proposed to account for lead-source dominance in localization. Some authors have proposed that neuronal inhibitory mechanisms triggered by a leading sound suppress the neuronal representation of directional information conveyed by later-arriving sounds (Harris et al., 1963; Lindemann, 1986; Zurek, 1987). Neurophysiological data, showing that addition of a leading sound reduces the responses of binaural or space-tuned neurons to their preferred stimuli, have been interpreted as evidence for such inhibitory processes (Yin, 1994; Fitzpatrick et al., 1995; Keller and Takahashi, 1996; Litovsky and Yin, 1998; Fitzpatrick et al., 1999). Recent modeling studies, however, suggest potential alternative explanations. Tollah (1998) demonstrated that lead-source dominance could result from a temporal weighting function that emphasizes the most recent monaural inputs at the site of binaural comparison, in the absence of any onset-triggered inhibitory mechanism. In biological terms, this weighting function could reflect the physiological properties of neurons that initially detect ITDs in the mammalian medial superior olive (Goldberg and Brown, 1969; Yin and Chan, 1990) or the avian nucleus laminaris (Carr and Konishi, 1990). More recent modeling studies demonstrated that interactions between lead and lag sounds within peripheral filters, in combination with adaptation of hair cell responses, can account for several classic precedence phenomena reported in psychophysical studies (Hartung and Trahiotis, 2001) as well as the physiological results described above (Trahiotis and Hartung, 2002), without requiring any inhibitory processes. At lag delays shorter than the reciprocal of a filter’s bandwidth, the lag sound will arrive while the filter is still responding to the lead sound. Interactions between responses to the lead and lag sound result in new “internal, effective ITDs” that are dominated by the ITD of the lead sound. The behavioral results obtained in the present study are compatible with any of these proposed mechanisms with the likely exception of interactions within peripheral filters.

It is unlikely that peripheral interactions account for the impaired performance in the lead- and lag-source conditions of the present experiment, because response times of filters in the frequency range of the stimuli are too short for substantial interactions to occur. The passband of stimuli used in this experiment (3–9 kHz) was chosen to encompass the 4 to 8-kHz range over which sound localization by barn owls is most accurate (Konishi, 1973; Knudsen and Konishi, 1979). To assess the effects of peripheral filtering in this frequency range, we used a bank of gamma-tone filters (Slaney, 1993), similar to those used by Hartung and Trahiotis (2001), but with center frequencies ranging from 2670 to 9514 Hz in 1/12-octave increments and proportional bandwidth adjusted to better approximate that of barn owl auditory-nerve fibers (Köppl, 1997). The impulse response of the 2670-Hz filter reaches maximum amplitude at 1.27 ms and decays by 12 dB at a response time of 3 ms. In the range of frequencies from 4 to 8 kHz, the attenuation relative to peak of impulse response amplitudes measured at 3 ms ranged from 20 to 41 dB. Thus, unlike filters tuned to the low frequencies that humans use for ITD extraction, the ringing response of filters tuned to frequencies that barn owls use for sound localization is substantially damped at the 3-ms lag delay used in this experiment.

Nevertheless, to illustrate the effect of peripheral filtering on the “internal ITD” expected in the present experimental conditions, a running interaural cross correlogram was computed after first passing the left- and right-ear signals through the filter bank, computing the interaural cross correlogram within each frequency channel, and then averaging the results across frequencies. The resulting, “filtered” cross correlogram is shown in Fig. 8(C). Most of the major features of the unfiltered cross correlogram are preserved, despite the temporal smearing effects of the filter responses. In particular, during the overlap segment, intermittently appearing peaks corresponding to lead and lag sources are clearly evident. Thus, the damped ringing of peripheral filters at a delay of 3 ms is of insufficient amplitude to obscure the ITD cues to either lead- or lag-source locations. Note that this analysis likely overestimates the effects of peripheral filtering on azimuthal resolution, because the lower frequency bands with the longest response times are weighted equally to the higher frequencies, which make the greatest contribution to sound localization in owls. The relative insensitivity to changes in lag-source azimuth is therefore more likely the result of either a central inhibitory mechanism that partially suppresses the neuronal representation of lag-source location, asymmetric temporal weighting at the site of binaural interaction, adaptation of hair cell responses, or some combination of these mechanisms. Sufficient information is not currently available to evaluate the contribution of the latter three mechanisms to behavioral precedence phenomena in barn owls.

**C. Discrimination of nonspatial cues**

Results of experiment 2 demonstrate that barn owls are not able to detect the monaural cues associated with changes of lag-source location, including changes in rippling of the combined signal spectrum. Sensitivity to rippled noise stimuli has been assessed in several nonhuman species. In the echolocating bat, Eptesicus fuscus, sensitivity to changes in spectral ripples is sufficient to account for detection of changes in distance between two simulated target surfaces (Simmons et al., 1990). This ability presumably reflects specializations of the bat’s auditory system for target ranging and acoustical imaging along the range axis. Thresholds for detection of spectral ripples against a flat-spectrum reference condition have been measured in chinchillas (Shofner and Yost, 1995) and budgerigars (Amagai et al., 1999). In budgerigars, the only avian species to be studied, thresholds for detection of spectral ripples are similar, or slightly lower than those of humans for repetition pitches in the range from...
make use of the topographic auditory spatial representation in the barn owl’s auditory midbrain to define the neuronal representation of sound sources in simulated echoic conditions, and to relate neuronal measures of discrimination performance to the present behavioral data.

**ACKNOWLEDGMENTS**

We thank Dr. Kip Keller for providing valuable assistance in writing the computer code used for stimulus generation and data acquisition and Dr. K. Hartung for providing many insightful comments, including a firm definition of the term “suppression.” Supported by NIDCD grants DC00448 to M.W.S. and DC03925 to T.T.T.


