The role of spectral composition of sounds on the localization of sound sources by cats

Daniel J. Tollin, Janet L. Ruhland, and Tom C. T. Yin

Tollin DJ, Ruhland JL, Yin TC. The role of spectral composition of sounds on the localization of sound sources by cats. J Neurophysiol 109: 1658–1668, 2013. First published December 28, 2012; doi:10.1152/jn.00358.2012.—Sound localization along the azimuthal dimension depends on interaural time and level disparities, whereas localization in elevation depends on broadband power spectra resulting from the filtering properties of the head and pinnae. We trained cats with their heads unrestrained, using operant conditioning to indicate the apparent locations of sounds via gaze shift. Targets consisted of broadband (BB), high-pass (HP), or low-pass (LP) noise, tones from 0.5 to 14 kHz, and 1/6 octave narrow-band (NB) noise with center frequencies ranging from 6 to 16 kHz. For each sound type, localization performance was summarized by the slope of the regression relating actual gaze shift to desired gaze shift. Overall localization accuracy for BB noise was comparable in azimuth and in elevation but was markedly better in azimuth than in elevation for sounds with limited spectra. Gaze shifts to targets in azimuth were most accurate to BB, less accurate to HP, and NB sounds, and considerably less accurate for tones. In elevation, cats were most accurate in localizing BB, somewhat less accurate to HP, and less yet to LP noise (although still with slopes >0.60), but they localized NB noise much worse and were unable to localize tones. Deterioration of localization as bandwidth narrows is consistent with the hypothesis that spectral information is critical for sound localization in elevation. For NB noise or tones in elevation, unlike humans, most cats did not have unique responses at different frequencies, and some appeared to respond with a “default” location at all frequencies.

interaural time difference; interaural level difference; spectral cues; sound localization; minimum audible angle

THERE ARE THREE ACOUSTIC CUES to the spatial location of sound sources: interaural time differences (ITDs), interaural level differences (ILDs), and spectral cues (Tollin and Yin 2009). The two interaural cues, ITDs and ILDs, are important for localization of sound in the horizontal, or azimuthal, dimension at low and high frequencies, respectively. Sound localization in the vertical, or elevational, dimension is based on patterns of the broadband power spectra at each ear that result from the direction-dependent acoustic filtering properties of the head and pinnae.

Numerous psychophysical experiments in humans have established the importance of ITDs (Rayleigh 1907; Stevens and Newman 1936), ILDs (Durlach and Colburn 1978), and spectral cues (Batteau 1967; Gardner and Gardner 1973) for localizing sound sources of different spectral compositions. These experiments have had the advantage that the instructions to the human subjects are relatively clear cut (i.e., the subjects understand what the behavioral task is). However, detailed studies of neurophysiological mechanisms in human subjects are naturally problematic, and noninvasive functional imaging or evoked-potential techniques do not allow for fine-grained analysis of neural mechanisms needed to test specific hypotheses. Conversely, virtually all physiological experiments have been derived from single-unit recordings in animals, where the associated psychophysical correlates are more challenging and in many cases lacking all together. To help close this gap, we have developed a behavioral sound localization task for cats that utilizes an ethologically natural orienting response (Populin and Yin 1998; Tollin et al. 2005).

The role of the interaural cues in sound localization has been widely studied and forms the basis for the Duplex theory, which posits that ITDs are used for the localization of low-frequency sounds whereas ILDs are used for higher frequencies. There is general support for the Duplex theory for pure tone stimuli in both human (Durlach and Colburn 1978; Rayleigh 1907; Stevens and Newman 1936) and animal subjects (Casseday and Neff 1973; Martin and Webster 1987); however, there is evidence that human subjects are also sensitive to the ITDs of the envelope of high-frequency waveforms (Bernstein and Trahiotis 1985; Henning 1974).

The importance of spectral cues for localization of sound in azimuth and elevation has been demonstrated by manipulating the spectrum of a stimulus while studying localization performance. Localization of pure tones or narrow-band noise by human subjects is accurate in the azimuthal dimension but very inaccurate in elevation. The perceived elevation of a sound source often correlates more with the frequency spectra of the stimulus rather than source location (Frens and van Opstal 1995; Middlebrooks 1992; Roffler and Butler 1968), which may be related to spectral peaks (Middlebrooks 1992) or “boosted bands” (Blauert 1969/1970) in the head-related transfer functions (HRTFs) at those elevations. Behavioral studies of sound localization in azimuth and elevation in animals where the frequency spectrum of the stimulus was varied include measurements of minimum audible angle (MAA; see below) with tones and walking to or orienting the head to tones or bandpass-filtered noises. In agreement with the role of spectral cues for vertical localization, Martin and Webster (1987) found that most cats could not perform the MAA task with pure tones from speakers on the midsagittal plane. Also, Huang and May (1996a) found localization of tones difficult in both azimuth and elevation, although Jenkins and Merzenich’s...
(1984) cats were able to walk to the source of tonal stimuli (azimuthal targets only). Problems in using a task in which the animals must walk to speakers are that the potential target locations are specified, vertical localization cannot be tested, and it is difficult to do many trials. Finally, using a gaze-orienting behavioral response, Tollin and Yin (2003) provided evidence that cats do use spectral notches for sound localization in elevation.

Psychophysical studies of sound localization usually fall into two broad categories: those measuring absolute localization and those measuring relative localization (Moore et al. 2008). Most relative measures involve determination of the MAA, which requires a subject to discriminate a change in the location of a sound source in space and to measure the minimum discriminable speaker separation (Casseday and Neff 1973; Mills 1958). Absolute localization requires the subject to estimate in some way (by gaze, pointing with the head or fingers, giving the angle of deviation, or walking to the target, among others) the spatial position of the sound source (Jenkins and Masterton 1982; Makous and Middlebrooks 1990; Wightman and Kistler 1992). Although it is usually assumed that absolute and relative measures of localization are closely related, there are well-known discrepancies between the two measures that cast doubt on this assumption (Heffner et al. 2005; Moore et al. 2008). In animals most studies have used MAA tasks to probe sound localization acuity (Casseday and Neff 1973; Heffner and Heffner 1988) or to study the effects on MAA following lesions to various parts of the auditory system (Heffner and Heffner 1984, 1990; Heffner and Masterton 1975; Kelly 1980).

To provide an estimate of the cat’s absolute sound localization ability, and to relate the animal studies more closely to human psychophysics, in this study we trained cats to look at visual and acoustic targets and measured gaze movements to targets varying in azimuth and elevation. We compared localization to broadband noise with various spectrally impoverished stimuli (pure tones, low-pass, high-pass, and bandpass-filtered noise). Following results from human psychophysical studies, we hypothesized that localization of broadband (BB) noise by cats would be superior to localization of high-pass (HP) or low-pass (LP) noise. Moreover, given the importance of high-frequency (>5 kHz) spectral cues for localization in elevation, we expected cats to localize HP noise better than LP noise for targets in elevation and to have great difficulty in localizing narrow-band (NB) noise and pure tones in elevation. In addition, we expected that cats would localize tones and NB noise at idiosyncratic elevations depending on frequency, as in some human studies. We found deterioration of localization to targets in elevation as bandwidth narrows, consistent with the hypothesis that spectral information is critical for sound localization in elevation, but no evidence for frequency-dependent vertical localization of tones.

Methods

Subjects and surgery. All surgical and experimental procedures were approved by the University of Wisconsin Animal Care and Use Committee and complied with the guidelines of the National Institutes of Health. The experiment was conducted over a period of months, utilizing two different, yet comparable, sound-attenuating chambers. In five adult female domestic cats, we implanted a stainless steel head post and fine wire coils (AS632; Cooner Wire, Chatsworth, CA; or S170012A7-FEP; Alan Baird Industries, Ho-Ho-Kus, NJ) in each eye under aseptic surgical conditions. Anesthesia was induced with an intramuscular injection of ketamine (20 mg/kg) and maintained throughout the surgery by inhalation of isoflurane (1–2% in 1 l/min O2) via a tracheal cannula. Postoperative analgesia was provided by ketoprofen (20 mg/kg) once a day for 3 days, and penicillin or cephalaxin monohydrate was given for 7 days as an antibiotic.

Experimental apparatus and stimuli. Many of our methods and materials have been described previously (Populin and Yin 1998; Tollin et al. 2005). All experiments were conducted in either a double-walled (2.2 × 2.5 × 2.5 m) or single-walled (2.4 × 2.2 × 2.0 m) dimly illuminated (or dark) sound-attenuating chamber (IAC, Bronx, NY). All walls and major pieces of equipment were covered with sound-absorbing acoustic foam (10.2 cm; Sonex; Ilbruck, Minneapolis, MN) to minimize acoustic reflections. The magnetic search coil (CNC Engineering, Seattle, WA) technique (Fuchs and Robinson 1966) was used to measure the positions of the eyes, and the analog output of the coil system was saved to disk by sampling at 500 Hz.

Targets in these experiments consisted of acoustic or visual stimuli presented from 1 of 19 different locations in the frontal hemifield (±40 or ±45° in azimuth, ±30° or ±23° to +18° in elevation, depending on which chamber was being used) distributed along two arcs, the horizontal and the vertical meridian. The radius of the arc was either 62 or 80 cm, depending on which sound chamber was used. Visual stimuli consisted of a 2.0-mm-diameter red (λ max = 635 nm) light-emitting diode (LED) located at the center of each speaker. Acoustic stimuli were delivered from 1 of 19 Morel Acoustics speakers (model MDT20) with similar frequency-response characteristics. The speakers themselves were hidden from view behind a black translucent cloth through which illuminated LEDs could be easily seen and sounds heard. The acoustic stimuli were generated digitally using a custom-built digital stimulus system or a Tucker-Davis (Alachua, FL) stimulus presentation system and custom software and were delivered by a 32-channel multiplexer at a rate of 100 kHz. The outputs of the multiplexer were connected to 19 individual audio amplifiers and attenuators that drove each speaker.

Acoustic stimuli consisted of broadband noise [nominally flat (∼±5 dB) out to 25 kHz, but because of the speaker characteristics, the low-frequency end began to roll off with a 3-dB cutoff frequency of ∼1.5 kHz], HP or LP noise (with a lower or upper cutoff frequency of 5 kHz to include or exclude, respectively, the high-frequency spectral cues in the HRTF) of 1-s duration, 1/6 octave NB noise with center frequencies of 6, 8, 12, 16 kHz of 100-ms duration, and pure tones (0.5, 1, 2.5, 6, 8, 10, 12, 14 kHz of 500-ms duration). The noise and tone stimuli had 7- and 4-ms rise/fall ramps, respectively. Sound level was 48–58 dB SPL, reverberant ±2 dB.

During initial training, the heads of the cats were restrained in the center of the coils comprising the magnetic search coil system (see Fig. 1 of Populin and Yin 1998). After the cats learned the task, the heads were freed, but a body restraint helped to maintain the position of the head within the center of the coil system. All aspects of the experiments including selection of the visual or acoustic stimuli, the location of the target speaker and/or LED, the acquisition of the eye position, and determination and delivery of reward were under computer control.

Eye coil calibration. As described in prior publications from our laboratory (Populin and Yin 1998; Tollin et al. 2005), the eye coils were calibrated behaviorally, relying on the cats’ instinct to look at the LEDs when they were suddenly illuminated in the darkened chamber. When all LEDs were suddenly illuminated, the output of the coil system was sampled and stored. This was repeated two to three times for each of the LED locations. The vertical and horizontal components of the voltage from the coil for these final eye positions were plotted against vertical and horizontal target position, respectively, and separately fit with linear functions, which were then used by the data collection software to convert the voltage output of the coil system to degrees of visual angle. Within the spatial range of this experiment (≥45° in both azimuth and elevation), the voltage output of the coil system was...
system and the location of the target were well fit by a linear function. In all cases the correlation coefficients exceeded 0.95.

Psychophysical procedure and training. The cats were on a controlled-access diet for 5–6 days per week. Water was always available. During the psychophysical task, the cats earned food rewards, which were small dollops of pureed cat food that were delivered via a peristaltic pump after each “successful” trial. Although the cats typically worked until they were satiated during each day of testing, their weights were monitored daily and we ensured that they were maintained within 15% of the original weight.

All cats were initially trained using operant conditioning with their heads restrained (Populin and Yin 1998). Subsequently, the cats were trained to make gaze saccades (head unrestrained) to the location of lights and sounds. The cats were rewarded under computer control if they maintained their eye position for a period of time after the gaze saccade in the vicinity of the target, as determined by a square electronic acceptance window centered on the target location. During training, the acceptance windows for the visual trials were gradually decreased to $\frac{\pi}{6} - 8^\circ$, but the windows for the auditory noise trials remained large ($\frac{\pi}{8} - 12^\circ$). The windows for the tone trials were even narrower.

Table 1. Mean gain and $\delta$ for BB, HP, LP, and NB noise

<table>
<thead>
<tr>
<th></th>
<th>BB Noise</th>
<th>HP Noise</th>
<th>LP Noise</th>
<th>NB, 6 kHz</th>
<th>NB, 8 kHz</th>
<th>NB, 12 kHz</th>
<th>NB, 16 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>0.91 (934)</td>
<td>0.85 (1,110)</td>
<td>0.62 (1,075)</td>
<td>0.25 (120)</td>
<td>0.17 (241)</td>
<td>0.35 (120)</td>
<td>0.14 (245)</td>
</tr>
<tr>
<td>Azimuth</td>
<td>0.92 (934)</td>
<td>0.88 (1,110)</td>
<td>0.85 (1,075)</td>
<td>0.54 (120)</td>
<td>0.67 (241)</td>
<td>0.90 (120)</td>
<td>0.74 (245)</td>
</tr>
<tr>
<td><strong>Cat 18</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>0.93 (2,529)</td>
<td>0.73 (1,007)</td>
<td>0.53 (981)</td>
<td>0.12 (117)</td>
<td>0.12 (229)</td>
<td>0.49 (80)</td>
<td>0.18 (82)</td>
</tr>
<tr>
<td>Azimuth</td>
<td>0.95 (2,529)</td>
<td>0.77 (1,007)</td>
<td>0.79 (981)</td>
<td>0.20 (117)</td>
<td>0.23 (229)</td>
<td>0.44 (80)</td>
<td>0.44 (82)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\delta$, degrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
</tr>
<tr>
<td>Azimuth</td>
</tr>
<tr>
<td><strong>Cat 18</strong></td>
</tr>
<tr>
<td>Elevation</td>
</tr>
<tr>
<td>Azimuth</td>
</tr>
</tbody>
</table>

Values are gain (with number of trials given in parentheses) and precision ($\delta$) for broadband (BB), high-pass (HP), low-pass (LP), and narrow-band (NB) noise.

Fig. 1. Localization of long-duration broadband (BB), high-pass (HP), and low-pass (LP) noise targets. A: final gaze position (small open symbols) for stimuli presented from 12 target locations (corresponding large symbols) at ($\pm 45^\circ$, 0°), ($\pm 32^\circ$, 0°), ($\pm 18^\circ$, 0°), ($\pm 9^\circ$, 0°), (0°, 18°), (0°, 9°), (0°, −14°), and (0°, −23°). Central fixation LED is shown as +. B: accuracy of the vertical (gaze shift elevation, top) and horizontal (gaze shift azimuth, bottom) components of the saccades. Each point corresponds to a single trial. The motor error (abscissa) is the horizontal or vertical component of the distance between the initial gaze position on each trial and the actual position of the target. The gaze shift amplitude (ordinate) is the corresponding horizontal or vertical component of the response to that target position from the initial gaze position. Red line is the linear regression of saccade amplitude component and the motor error. Gain is the slope of the regression line and represents localization accuracy ($gain = 1$ corresponds to perfect localization accuracy); $\delta$ is the residual error after regression and is an indication of response precision or consistency; $n$ is the number of trials. Data are from cat 18.
larger, ±15°, to encourage the cats to look to the apparent location of the sound source rather than training them to look to a particular location in space to get a reward (see Populin and Yin 1998). However, all data were analyzed, even if the cat did not get the trial “correct” during testing. In every session the stimulus type (broadband, LP, HP, and BP noise or tones), target position, and trial type (saccade or fixation) were randomized so that the cat could not predict what the next stimulus would be.

All data presented in this report were collected in the saccade psychophysical task. Here, the cat was initially required to fixate an LED presented from straight ahead (0°, 0°) and to maintain gaze fixation within the acceptance window for a variable period of time. If the cat satisfied this initial fixation condition, then simultaneously the fixation LED was extinguished and an acoustic or visual target was presented from 1 of the 19 locations in the frontal hemifield (approximately ±45°). The cat was then required to make a gaze saccade to the apparent location of the target and maintain fixation at that location for another 600–1,000 ms. If during this time the cat’s gaze position was within the specified acceptance window, the cat was given a food reward. To deliver food rewards, a small lightweight (40 g) aluminum headset-like apparatus was attached to the head post, which held the feeding tube near the mouth. This feeding system moved with the head and allowed the cat to receive the rewards near its mouth without hindering its head movements.

Analysis of final gaze position. The key dependent variable in this experiment was the final horizontal and vertical gaze position at the completion of the saccadic shift to the apparent location of the target (Populin and Yin 1998). We used a velocity criterion to determine the “end of fixation,” or when the gaze was no longer stationary, by determining the time at which the magnitude of the velocity trace exceeded 2 SD of the mean velocity computed during the initial fixation. During initial fixation, the gaze was expected to be constant and the velocity close to zero. The final gaze position was determined by the position at the time of “return to fixation,” which was computed as the time at which the magnitude of the velocity trace returned to within 2 SD of the baseline mean velocity. If corrective movements were made within 200 ms of the end of the initial saccade, the final position was determined from the return to fixation of the corrective saccade.

To quantify saccade accuracy and precision, the initial gaze motor error and the final gaze shift were computed for each trial, separately for horizontal and vertical components. The gaze motor error was defined as the difference between the target-in-space position and the initial gaze position, which describes the magnitude of the desired gaze shift needed to acquire the target position perfectly given the

![Fig. 2. Localization of BB noise compared with narrow-band (NB) noise with center frequencies of 6, 8, 12, and 16 kHz for cat 17. Same format as Fig. 1. In A, targets in elevation (top) have been separated from targets in azimuth (bottom) for ease of presentation.](http://jn.physiology.org/doi/10.1152/jn.00358.2012)
initial gaze position. If the gaze is initially directed at exactly (0°, 0°), the motor error is equivalent to the absolute position of the target in space. In our experiments, since the cats were always required to fixate the LED at (0°, 0°) before the saccade, the gaze motor error is usually nearly equivalent to the position of the target in space. The gaze shift was defined as the angular magnitude and direction of the gaze saccade. To obtain a quantitative measure of the localization performance across all target locations, a linear function was fit to the plots of gaze vs. motor error, separately for the horizontal and vertical components of the target locations. This procedure was performed for each of the stimuli. The coefficients of the fits are indicators of localization performance. For example, the slope of the response-target localization function, which we shall refer to as “gain,” indicates the accuracy with which the cats localized the targets (Frens and van Opstal 1995). A gain of 1.0 indicates that, on average, across all trials and all target positions, the cats located the targets to their actual positions (unless of course if the line does not cross the origin), whereas gains <1.0 indicate that the target locations were underestimated. Standard statistical bootstrapping techniques (Elfron and Tibshirani 1986) were used to obtain an estimate of the 95% confidence intervals of the gain. Here, for a given stimulus configuration, 1,000 synthetic data sets, containing exactly the same number of trials as the empirical data set, were created by randomly sampling with replacement localization data from individual trials from the empirical data sets for each cat separately. As described above, the horizontal and vertical components of the behavioral responses were analyzed separately. A linear function was fit to each synthetic data set, resulting in 1,000 measurements of the gain, from which the 95% confidence interval was obtained. Using the empirical data set, we also computed the standard deviation of the residuals of the fitted function, which represents the distribution of behavioral responses about the mean gain. This latter value gives a numerical estimate of the precision (or consistency) of the localization responses, which we refer to as δ.

RESULTS

These experiments were designed to examine the ability of cats to localize filtered noise and pure tones via gaze shift. The results and statistical analyses were based on the localization performance of five adult female cats. Two cats were tested with BB and filtered noise; three cats were tested with BB and tones. Several of our cats have been subjects in previous publications (cats 17, 18, and 21), but all of the data reported in this article are new and were obtained in different experimental sessions with different stimuli. The performances of all three cats to BB noise were similar and confirm the reliability of our behavioral measures. Overall, 15,119 trials were analyzed.

Auditory localization of BB noise. To study the effect of variations in sound spectra on localization, we compared the accuracy and precision of gaze shifts to filtered noise and tone targets with those obtained from localization of 1,000-ms-duration BB noise bursts. Figure 1A, left, shows the horizontal and vertical final gaze positions plotted as scatter plots for noise targets; the target positions are indicated by the large symbols. The final gaze position for each trial is indicated by a single small symbol. The responses to all 12 positions show little scatter (good precision) and are located near each target, indicating high accuracy. Figure 1B shows scatter plots of gaze motor error (i.e., the angular distance between the initial position of the gaze and the actual target) vs. actual gaze displacement (i.e., the angular distance and direction that the gaze moved) for the same data as in Fig. 1A. This cat has high accuracy, with gains of 0.95 and 0.93 for localization in azimuth and elevation, respectively, and high precision with a δ value of 2.4° for localization in azimuth and 2.8° in elevation. All five cats had good to excellent accuracy (gains for azimuth and elevation between 0.78 and 0.95) and precision (δ ranging between 2.4° and 4.6°) to BB noise (see Tables 1 and 2). The variability in the behavioral performances as measured by the gain and δ in Tables 1 and 2 were typical of cats in our laboratory working with the head free (see Tollin et al. 2005).

Localization of HP, LP, and NB noise. Figure 1A, middle and right, shows scatter plots of the final gaze position to long-duration HP and LP noise targets, respectively, for the same cat. The cutoff frequency for both noise stimuli was 5 kHz, which is approximately the lower boundary for the midfrequency spectral notches in the cat’s HRTF (Musicant et al. 1990; Rice et al. 1992; Tollin and Koka 2009). Thus the HP noise contains most of the first-notch information hypothesized to be necessary for localization of sounds in elevation, but the LP noise does not. The scatter plots show that there was still good horizontal localization, although not as good as for BB noise. However, vertical localization was clearly affected by the decreased bandwidth. Figure 1B, middle and right, quantifies the comparable moderate decline in accuracy (gains of 0.77 and 0.79) and precision (δ of 3.4° and 3.3°) to targets in azimuth for both HP and LP noise. Performance for the vertical targets was more severely affected by the filtering, with more severe decline for LP (gain of 0.53) than HP (gain of 0.73) noise. All declines were statistically significant with respect to responses to BB noise, using 95% confidence intervals obtained with bootstrapping methods. The differential effect of spectral filtering on the LP noise, compared with the HP, is consistent with the role of the midfrequency spectral notches on vertical localization. These effects of HP and LP noise...
stimuli on localization were similar in both cats tested (Table 1). The effect of restricting spectral information on localization is shown in Fig. 2 for another cat tested with filtered noise targets, in this case to BB noise and four NB noise targets. For clarity, the scatter plots in Fig. 2A are shown separately for the horizontal and vertical targets. Gaze shifts and associated computations of localization gain and precision (δ values) are shown in Fig. 2B. Table 1 shows the localization gains and δ values for cats 17 and 18 to noise and filtered-noise targets.

The accuracy of localization deteriorated dramatically for NB targets in elevation: in Fig. 2B the gains ranged from 0.14 to 0.35. The other cat tested showed a similar decrease in accuracy in elevation for NB, with gains between 0.12 and 0.49 (Table 1, cat 18). By contrast, accuracy in azimuth remained moderately high for the data shown in Fig. 2B, with gains ranging from 0.54 to 0.90, but precision declined compared with BB noise for all four of the NB noise targets. However, the results in azimuth were not as consistent as in elevation: cat 18 was much less accurate but more precise (smaller δ value) localizing NB noise to targets in azimuth than cat 17 (Table 1). Overall, gaze shifts to targets in azimuth were most accurate to BB but less accurate for HP, LP, and NB sounds. Precision usually declined for NB noises for both azimuth and elevation in both cats (Table 1).

Figure 3 summarizes the performance for both cats. The gains are plotted separately for elevational (open symbols) and azimuthal (closed symbols) components of the responses as a function of target type. For sound sources varying in elevation, localization was much more sensitive to spectral content: there was a steady decline in accuracy from BB noise (mean gain for 2 cats of 0.92) to HP noise (mean gain of 0.79) to LP noise (mean gain of 0.58) to NB noise (mean of 0.23). All declines were

![Figure 3](http://jn.physiology.org/)

**Fig. 3.** Performance for both cats. The gains are plotted separately for elevational (open symbols) and azimuthal (closed symbols) components of the responses as a function of target type. For sound sources varying in elevation, localization was much more sensitive to spectral content: there was a steady decline in accuracy from BB noise (mean gain for 2 cats of 0.92) to HP noise (mean gain of 0.79) to LP noise (mean gain of 0.58) to NB noise (mean of 0.23). All declines were

![Figure 4](http://jn.physiology.org/)

**Fig. 4.** Localization of BB noise (left) compared with tones of 1, 6, 10, and 14 kHz for cat G03. Same format as Fig. 2. In B, bottom row, the value b indicates the y-intercept of the regression line.
statistically significant at $P < 0.05$ with respect to responses to BB noise except for accuracy of cat 17 to HP noise in elevation, which was not significantly decreased compared with BB noise. Since the NB noises were shorter (100 ms) in duration than the BB noises (1 s), we considered the possibility that the decreased accuracy to NB noise could in part be due to duration rather than spectra. We have systematically analyzed in a separate study the effect of duration on accuracy and found in three other cats the mean difference in gain from 100 ms to 1,000 ms in duration was 0.12, corresponding to only an ~13% reduction in accuracy, for both azimuthal and elevational targets (Ruhland et al. 2005). These effects are minimal compared with the much larger changes shown in Figs. 2 and 3 for BB and NB noise.

The better performance for both cats at the NB noise centered on 12 kHz is interesting since the spectral notches in the HRTFs for the locations tested (from +18° to −23° on the midsagittal plane) fall between about 12 and 9 kHz (Musicant et al. 1990; Rice et al. 1992; Tollin and Koka 2009). This suggests that the NB noise centered on 12 kHz may have provided some information about the midfrequency notch to aid in vertical localization.

### Localization of pure tone targets

With the poor performance for localizing NB noise in elevation shown in Fig. 2 and Table 1, we expected performance to be even worse to pure tone stimuli. Figure 4 shows localization of another cat to pure tones and BB noise for comparison. Although this cat (cat G03) could localize in both azimuth and elevation BB noise targets accurately (gains of 0.86 in both dimensions) and precisely (δ of 4.4° and 3.5°, respectively), she had great difficulty localizing tones in both azimuth and elevation. For all tones from 0.5 to 14 kHz placed vertically, she responded by looking to a default position centered around 10° left and 0° elevation, resulting in gains ranging between −0.01 and 0.05. For tones placed horizontally, she again responded by saccading to the left but with a weak azimuthal bias, i.e., looking farther left for the leftmost target and more toward the center for targets to the right. The result was that cat G03’s elevational gains were all essentially zero (<0.05) and the azimuthal gain values were very poor, ranging from 0.22 to 0.35, i.e., with a small positive slope. The intercepts of the regression lines were all very negative (from −18.2° to −11.5°), reflecting the leftward bias of her localizations.

Table 2 shows the localization gains and δ values for two other cats to BB noise and pure tone targets, whereas Fig. 5 summarizes the performance for all three cats. The most striking result was that none of the cats were able to localize tones in elevation (Table 2 and Fig. 5B) (gains ranging from −0.02 to 0.17 with a mean of 0.05), but they could accurately and precisely localize BB noise targets in elevation (mean gain of 0.85; Fig. 5 and Table 2). Although their performance in localizing BB noises was comparable, the three subjects varied in their ability to localize pure tone targets in azimuth (Fig. 5A). Each of the three cats had stable or only slightly declining gains as tone frequency increased. One subject (cat 28) had gains in azimuth for low-frequency pure tones that were even higher than those she had for BB noise.

Cats do not localize the source of pure tone targets by frequency. Figure 6 plots the means and standard deviations of response locations to tone targets (A) and NB noise targets (B) that varied in elevation. Separate symbols represent tone targets of different frequencies. For NB noise or tones in elevation, the cats’ localization responses did not vary systematically with different frequency. Instead, particularly for certain cats (cats G03 and 18), they appeared to localize the tones to a “default” location, similar for all frequencies but differing between cats. Of course in our experiment we cannot determine whether the cats actually perceived these stimuli to be located at the default position or whether they could not localize the sound and instead saccaded to the default position in hopes of getting a reward.

### Latency

We used the end of fixation as a measure of response latency for the gaze shifts. Figure 7 plots latencies as a function of gaze motor error in azimuth (A) or elevation (B) for BB noise (left) and four different tone frequencies for three different cats (cat 21, G03, and 28). By plotting data for these different conditions in this unconventional way with dot rasters, we can see the variability across stimuli, direction, amplitude, and subjects. Plotting the data against gaze motor error rather than target location has the advantage that the jitter in initial fixation of the central LED prevents all of the data points from falling in a line, thereby enhancing

<table>
<thead>
<tr>
<th>Cat</th>
<th>BB Noise</th>
<th>0.5 kHz</th>
<th>1 kHz</th>
<th>2.5 kHz</th>
<th>6 kHz</th>
<th>8 kHz</th>
<th>10 kHz</th>
<th>12 kHz</th>
<th>14 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elevation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cat 21</td>
<td>0.83 (656)</td>
<td>0.04 (302)</td>
<td>0.06 (271)</td>
<td>0.03 (283)</td>
<td>0.02 (293)</td>
<td>0.00 (146)</td>
<td>0.03 (186)</td>
<td>−0.02 (61)</td>
<td>0.03 (98)</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>0.81 (656)</td>
<td>0.64 (302)</td>
<td>0.75 (271)</td>
<td>0.69 (283)</td>
<td>0.56 (293)</td>
<td>0.56 (146)</td>
<td>0.61 (186)</td>
<td>0.62 (61)</td>
</tr>
<tr>
<td>Cat 28</td>
<td>0.87 (549)</td>
<td>0.01 (257)</td>
<td>0.11 (340)</td>
<td>0.08 (315)</td>
<td>0.07 (318)</td>
<td>0.08 (130)</td>
<td>0.17 (98)</td>
<td>0.13 (145)</td>
<td>0.17 (78)</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>0.78 (549)</td>
<td>0.85 (257)</td>
<td>1.00 (340)</td>
<td>0.85 (315)</td>
<td>0.70 (318)</td>
<td>0.77 (130)</td>
<td>0.70 (98)</td>
<td>0.73 (145)</td>
</tr>
<tr>
<td>Cat G03</td>
<td>0.86 (418)</td>
<td>0.00 (129)</td>
<td>−0.01 (170)</td>
<td>0.02 (171)</td>
<td>−0.01 (178)</td>
<td>0.05 (184)</td>
<td>0.05 (158)</td>
<td>0.01 (168)</td>
<td>0.04 (147)</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>0.86 (418)</td>
<td>0.30 (129)</td>
<td>0.29 (170)</td>
<td>0.22 (171)</td>
<td>0.22 (178)</td>
<td>0.22 (184)</td>
<td>0.35 (158)</td>
<td>0.36 (168)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cat</th>
<th>BB Noise</th>
<th>0.5 kHz</th>
<th>1 kHz</th>
<th>2.5 kHz</th>
<th>6 kHz</th>
<th>8 kHz</th>
<th>10 kHz</th>
<th>12 kHz</th>
<th>14 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elevation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cat 21</td>
<td>3.4</td>
<td>6.5</td>
<td>5.8</td>
<td>6.2</td>
<td>5.7</td>
<td>5.1</td>
<td>5.6</td>
<td>5.3</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>2.9</td>
<td>18.3</td>
<td>15.2</td>
<td>11.6</td>
<td>16.7</td>
<td>14.2</td>
<td>16.0</td>
<td>16.3</td>
</tr>
<tr>
<td>Cat 28</td>
<td>4.6</td>
<td>7.7</td>
<td>7.9</td>
<td>7.6</td>
<td>7.1</td>
<td>7.5</td>
<td>7.6</td>
<td>7.8</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>4.1</td>
<td>19.4</td>
<td>18.3</td>
<td>11.8</td>
<td>15.6</td>
<td>11.2</td>
<td>9.1</td>
<td>9.2</td>
</tr>
<tr>
<td>Cat G03</td>
<td>3.5</td>
<td>4.3</td>
<td>4.6</td>
<td>4.3</td>
<td>4.1</td>
<td>6.2</td>
<td>5.6</td>
<td>7.1</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>4.4</td>
<td>7.0</td>
<td>7.1</td>
<td>6.0</td>
<td>7.7</td>
<td>7.4</td>
<td>9.1</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Values are gain (with number of trials given in parentheses) and δ for BB noise and pure tones.
visibility of individual data points. The red horizontal dashed lines indicate the means of the latency distributions. For all three cats, the mean latency to BB noise was shorter than for any of the tone stimuli regardless of direction (azimuth or elevation) or amplitude of gaze shift. With a few exceptions, latencies to BB noise targets to the right but not the left (Fig. 4) was reflected in the latency distributions. For all three cats: the latency distributions to tones in most cases were wide and broadly distributed from 80 to 300 ms (mean ± SD: 188 ± 107 ms), whereas the BB noise latencies were qualitatively shorter and more tightly distributed than for tones. The disparities in the mean latencies and the general shapes of the distributions suggest fundamental differences in the neural processing leading to gaze shifts for noise and tones. In particular, the uniform dispersion of latencies to tones is unusual (e.g., see May et al. 2009). Our data are consistent with the hypothesis that the longer and greater variability of latencies to tones reflects the cats’ uncertainty of the source location (Luce 1986; May et al. 2009; Tollin and Yin 2003).

There were also some idiosyncratic behaviors evident in the latency distributions. Cat G03’s difficulty in localizing BB noise targets to the right but not the left (Fig. 4) was reflected in the wide range of latencies to the right but not the left, although this bias in the latency distribution was also present for BB noise where there was no bias in performance. Cat 21 had shorter latencies for distal targets varying in elevation than for proximal ones.

DISCUSSION

We have used an absolute sound localization task to study the role of spectral composition in sound localization performance in the cat. Our results demonstrate that absolute localization by gaze shift of spectrally impoverished stimuli (tones or NB noise) is severely degraded in elevation and less so in azimuth compared with localization of BB noise. More moderate losses were seen for HP or LP noise with the cutoff frequency at 5 kHz. Latency of the gaze shift was also longer and more variable for the stimuli that were difficult to localize.

Absolute vs. relative localization. Most previous studies of the ability of cats to localize noise of varying spectral content and/or pure tones have utilized relative rather than absolute measures of sound localization (see Moore et al. 2008 for review). The differences between absolute and relative behavioral paradigms can complicate the comparison of behavioral performance across studies. For example, in human psychophysical studies, some researchers have inferred (e.g., Grantham et al. 2003; Heffner and Heffner 1988; Perrott 1984; Perrett et al. 1987), and others have directly suggested (Hartmann and Rackerd 1989; Makous and Middlebrooks 1990; Recanzone et al. 1998), that measures of auditory spatial acuity, as measured by the MAA, are specifically related to the precision of absolute sound localization estimates, and thus that the former can be used as a proxy for the latter and that the two can be compared. However, Moore et al. (2008) showed that psychophysical measures of acuity (e.g., the MAA) were occasionally very poor indicators of absolute localization precision, especially when sound sources were more difficult to localize accurately, such as NB noise and tone stimuli (see also Heffner et al. 2005), and argued that to compare the precision (δ) in absolute behavioral paradigms to estimates of acuity, the MAA must be adjusted by the localization accuracy (gain), which can be done by dividing δ by the gain.

When the gain is near 1.0, then the MAA can be approximated by the precision δ. This condition is approximately met for all of our subjects for BB noise stimuli, where the mean gain for targets in elevation was 0.88 and that in azimuth was 0.86. In azimuth the mean precision (and unadjusted estimated MAA) for the 5 subjects was 3.7°, whereas the mean gain-adjusted MAA (δ/gain) was 4.3°, which compares with the MAAs for BB noise in azimuth reported in psychophysical studies of cats of 5.9° (Casseday and Neff 1973), 6.0° (Heffner and Heffner 1988), 4.1° (Huang and May 1996a), and 3.4° (Martin and Webster 1987).

Localization of targets in elevation. Turning to localization of sources in elevation, studies in humans (Butler 1999; Frens and Van Opstal 1995; Middlebrooks 1992; Roffler and Butler 1968) have sometimes found that the perception of the vertical location of NB noise or tones was influenced more by the frequency composition of the stimuli, rather than the actual location of the target. It has been hypothesized that specific tone frequencies and NB noise with certain center frequencies are associated with the perception of certain locations in space, because these frequencies and the resulting frequency spectra associated with the stimulus correlated best with, and were thus more predictive of, the acoustic spectral shape cues that result from filtering by the pinnae for sounds originating from those locations (Butler 1999; Middlebrooks 1992; Roffler and Butler 1968).

In other words, the perceived spatial locations of NB
sounds are illusions that do not match the location from which they are delivered, but rather match the location from which the stimulus spectra most closely matches the acoustical spectral shape cues (Middlebrooks 1992). Note, however, that this illusion does not always produce robust perceptions. For example, in one study subjects reported that tones were never as easy to localize in elevation as from targets in azimuth (Roffler and Butler 1968). The cats in the present study could not accurately localize tone targets in elevation, and at least some of them tended to localize tones from targets in elevation at a default and idiosyncratic location. It appears that cats do not experience the illusion as robustly as humans when localizing NB noise and tone targets in elevation (however, see Xu et al 1999).

In a similar absolute sound localization study, Huang and May (1996b) also reported that their cats could not localize tone targets in azimuth or elevation. These researchers used the position of the head as the indicator of where the cat localized the sound in space. They did not want to shape the localization behavior in any way by rewarding approximate attempts, e.g., rewards did not depend on turning the head in the proximity of the target but on maintaining a paw press through the presentation of the sound and through the subsequent presentation of a light at the location of the sound source. Despite being unable to localize tones in an absolute sense (i.e., the cats were unable to report the location of the sources of the tones), their cats were able to localize noise targets in azimuth and elevation quite well, although the rewards did not depend on correct response either. Similarly, Martin and Webster (1987) reported that three of four cats tested could not discriminate changes in sound source elevation for tones of 2, 8, or 16 kHz even though they could easily discriminate azimuthal changes with those frequencies (1 cat required an −20° separation in elevation to reach threshold). Therefore, their results may simply indicate that noise was easier and more naturally to be localized than tones when the subjects were not motivated.

At least two prior studies have attempted to measure the MAA for BB noise in cats for sources varying in elevation. In four cats Martin and Webster (1987) reported that the MAA for noise measured about the midline was <3.8°; that is, their animals were still performing above chance for the smallest sound source separation. Huang and May (1996a) reported that if the psychometric functions of their cats were extrapolated down to chance, then vertical plane MAAs for BB noise would be estimated to 3.6°. In the present study, using precision to estimate MAA, our five cats produced a mean vertical plane MAA for BB noise of 3.7° ± 1.0° (SD), virtually identical to the MAAs measured by Martin and Webster (1987) and Huang and May (1996a) using traditional relative sound localization paradigms. This latter result further confirms our absolute sound localization data as well as supporting the validity of the gain-adjusted precision method of Moore et al. (2008) to estimate localization acuity from absolute sound localization data.

Finally, we point to other studies and behavioral methods that may be useful in future studies to help reconcile differences between sound localization performance measured in relative (i.e., MAA) and absolute paradigms. Nodal et al. (2008), for example, specifically compared acoustic orientation and approach-to-target responses of ferrets to noise targets. They concluded that the neural processing stage at which localization errors occur is common to both head orientation and approach-to-target spatial responses; however, this study did not address the contributors to localization precision or acuity. A similar conclusion was reached by Tollin et al.
(2010), who demonstrated that the spatial orienting of the pinnae by cats was predictive of the ultimate gaze position even though the pinnae moved with extraordinarily short latencies relative to the gaze shifts; the short latencies of the pinnae movements could only be accounted for via brain stem mechanisms. The mobile pinna of the cat compared with the human may also provide an explanation for differences in localizing tones in elevation. However, studies of the effects of lesions of auditory cortex on localization of noise have suggested that different neural circuits are responsible for orientation of the head toward a sound source and the ability to associate a sound with a location in space (Beitel and Kaas 1993; Heffner and Heffner 1990; Lomber and Malhotra 2008; Malhotra and Lomber 2007; Thompson and Masterton 1978); however, none of these studies addressed the neural contributors of spatial precision or acuity.

In summary, we have found our sound localization task based on the unrestrained head gaze shifts to be a reliable and valid method to study the auditory cues responsible for sound localization. We conclude that for cats, like humans, localization in elevation, but not so much in azimuth, is severely degraded as the stimulus bandwidth is reduced, suggesting that vertical localization depends on the broadband power spectra at each ear resulting from the filtering properties of the head and pinnae at high frequencies. Unlike humans, however, the cats do not seem to show frequency-dependent biases in vertical judgments. These results can now be used to inform neurophysiological studies of these phenomena.

Fig. 7. A: latency of gaze saccades to BB noise and tone stimuli in azimuth. B: latency of gaze saccades to the same stimuli in elevation. The short red lines indicate the mean latency at each condition.
GRANTS
This work was supported by National Institute of Deafness and Other Communicative Disorders Grants DC-00116 and DC-02840 (to T. C. T. Yin) and DC-011555 (to D. J. Tollin).

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS
D.J.T. and T.C.T.Y. conception and design of research; D.J.T. and J.L.R. performed experiments; D.J.T. and J.L.R. analyzed data; D.J.T. and T.C.T.Y. interpreted results of experiments; D.J.T. and J.L.R. prepared figures; D.J.T. and J.L.R. drafted manuscript; D.J.T., J.L.R., and T.C.T.Y. edited and revised manuscript; D.J.T., J.L.R., and T.C.T.Y. approved final version of manuscript.

REFERENCES