

Neural and behavioral sensitivities to azimuth degrade with distance in reverberant environments

Sasha Devore¹, Antje Ihlefeld², Barbara G. Shinn-Cunningham² and Bertrand Delgutte¹

¹ Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA
sashad@mit.edu, Bertrand_Delgutte@meei.harvard.edu

² Hearing Research Center, Boston University, Boston, MA {ihlefeld, shinn}@bu.edu

1 Introduction

Reverberation poses a challenge to sound localization in rooms. In an anechoic space, the only energy reaching a listener's ears arrives directly from the sound source. In reverberant environments, however, acoustic reflections interfere with the direct sound and distort the ongoing directional cues, leading to fluctuations in interaural time and level differences (ITD and ILD) over the course of the stimulus (Shinn-Cunningham et al. 2005). These effects become more severe as the distance from sound source to listener increases, which causes the ratio of direct to reverberant energy (D/R) to decrease (Hartmann et al., 2005; Shinn-Cunningham et al., 2005).

Few neurophysiological and psychophysical studies have systematically examined sensitivity to sound source azimuth as a function of D/R (Rakerd and Hartmann, 2005). Here, we report the results of two closely-integrated studies aimed at characterizing the influence of acoustic reflections like those present in typical classrooms on both the directional sensitivity of auditory neurons and the localization performance of human listeners. We used low-frequency stimuli to emphasize ITDs, which are the most important binaural cue for sounds containing low-frequency energy (MacPherson and Middlebrooks, 2002; Wightman and Kistler, 1992). We find that reverberation reduces the directional sensitivity of low-frequency, ITD-sensitive neurons in the cat inferior colliculus (IC), and that this degradation becomes more severe with decreasing D/R (increasing distance). We show parallel degradations in human sensitivity to the azimuth of low-frequency noise.

2 Single-unit neurophysiology

2.1 Methods

Methods for recording from low-frequency, ITD-sensitive neurons in the IC of anesthetized cats were as described by Hancock and Delgutte (2004). We focused on measuring neural responses as a function of source azimuth in simulated rooms.

Binaural room impulse responses (BRIRs) were simulated using the room-image method (Allen and Berkeley, 1979) for a pair of receivers corresponding to the left and right ears of a cat in the center of a simulated reverberant room (11 x 13 x 3 m). We did not include a model of the head in the simulations, so that the resulting BRIRs contained ITD but essentially no ILD cues. BRIRs were calculated for azimuths spanning the frontal hemifield (-90° to 90°) at distances of 1m and 3m with respect to the midpoint of the receivers. Anechoic impulse responses were created by time-windowing the direct wavefront from the 1m reverberant BRIRs. The direct-to-reverberant energy ratio (D/R) was calculated as the ratio of the energy in the direct sound (time-windowed from the BRIR) to the energy of the remaining impulse response. An overall D/R was determined for each source distance by averaging across all azimuths. Virtual room stimuli were created by convolving the BRIRs with a reproducible 400-ms broadband noise burst. The first 400 ms of the resulting signals were presented to dial-in-urethane anesthetized cats over calibrated, closed acoustic systems.

Neural responses were measured as a function of source azimuth for each virtual room condition (anechoic, 1m, and 3m). We typically used 11 azimuths (15° steps) or, occasionally, 7 azimuths (30° steps). The noise stimulus was repeated 16 times at each azimuth, with random order across azimuths. We computed the average firing rate by counting the number of action potentials over the stimulus duration.

2.2 Results

We measured neural responses as a function of azimuth for 25 IC units from 7 cats. Rate-azimuth curves for two typical units are shown in Fig. 1A. Neural rate responses depend strongly on source azimuth in the anechoic condition ($D/R = \infty$ dB), with a preference for contralateral azimuths. Reverberation reduces the range of firing rates over all source azimuths (a “demodulation”), although rates still vary systematically with azimuth. To quantify these observations, we define the *relative response range* as the range of firing rates for a given room condition normalized by the range of firing rates in the anechoic condition. The relative range for the anechoic condition is 1, by definition. For most neurons, the relative ranges in the reverberant conditions are less than 1, indicating a compression of the rate-azimuths curves; furthermore, relative range decreases with decreasing D/R (Fig. 1B).

Neural sensitivity to azimuth depends not only on the response range, but also on the variability in responses at each azimuth. To quantify sensitivity, we computed the mutual information (MI) between the source azimuth and the neural

spike counts for each neuron and room condition. MI characterizes the precision with which the source azimuth can be estimated from the neural spike counts without making additional assumptions about the neural code. To compare across experiments using different numbers of stimuli, MI was normalized by the stimulus entropy to get the relative information transfer (rIT). Figure 2 shows that the rIT systematically decreases with decreasing D/R.

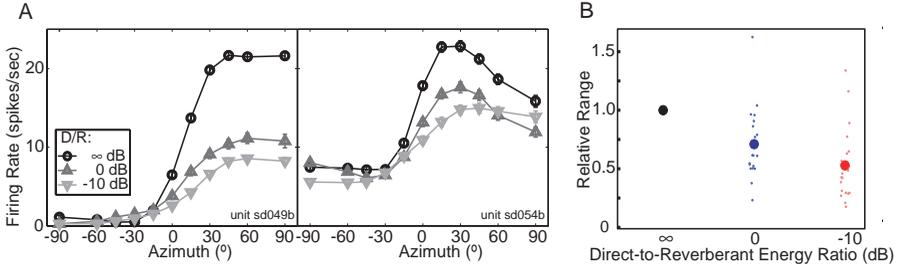


Figure 1. (A) Mean neural response rate (\pm s.e.m.) as a function of source azimuth for two neurons in the cat IC. (B) Mean (large filled symbols) relative response range as a function of D/R. Small symbols indicate relative response range for individual units.

2.3 Discussion

Reverberation leads to a decrease in the directional sensitivity of low-frequency ITD-sensitive neurons, as seen both by a reduction in the relative response range and a decrease in rIT between source azimuth and neural firing rate. Moreover, the degradation in sensitivity is largest at the smallest D/R.

It is tempting to attribute the reduction in response range with decreasing D/R to a decrease in the interaural correlation of the input signals, since the rate responses of cat IC neurons to interaurally-delayed broadband noise become compressed when the noise is statistically decorrelated (Yin et al., 1987). However, a cross-correlation model of binaural processing in the IC (Hancock and Delgutte, 2004) predicts an even larger degradation in directional sensitivity than we actually observed (see reply to comment below), indicating that the responses of IC neurons may be more robust to reverberation than cross-correlator models of binaural processing. Further studies are needed to identify the neural mechanisms underlying such robust ITD sensitivity.

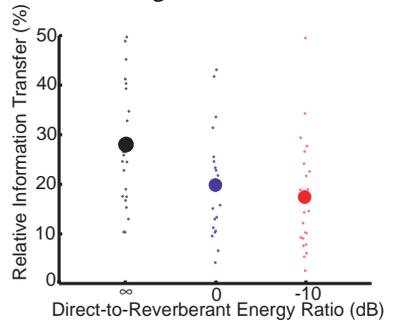


Figure 2. Mean (large filled symbols) rIT as a function of D/R. Small symbols show rIT for individual units.

3 Human psychophysics

3.1 Methods

Seven paid subjects with normal hearing participated in the experiment. Binaural room impulse responses (BRIRs; see Shinn-Cunningham et al. 2005 for details) were measured using KEMAR in a small room (T60 of approximately 485 ms; 3.3m x 5.8m x 2.6m). BRIRs were used to simulate sources from 52 locations (all combinations of 13 azimuths, evenly spaced from -90° to 90° , and four distances {15, 40, 100 and 170 cm}, in the horizontal plane containing the ears).

Simulated source azimuth and distance were varied randomly from trial to trial by convolving low-pass noise (500Hz-1kHz, 20ms \sin^2 ramp, 200ms duration) with the appropriate pair of BRIRs. Signals were presented through ER-1 insert earphones. Subjects, seated in a sound-treated booth, were asked to indicate the perceived source location on a continuous scale ranging from -90° to 90° using a graphical user interface. Each subject performed two training sessions with pseudo-anechoic simulations during which they received correct-answer feedback. Following training, subjects completed 12 sessions in the simulated room conditions without feedback. For the results reported here, subjects completed 45 trials for each azimuth and distance pair.

For each listener, mean response was calculated for each stimulus location. From the resulting confusion matrices (probability of responding each azimuth, given which azimuth was presented), the mutual information (MI) and relative information transfer (rIT) between source and response were computed.

3.2 Results

Figure 3 shows the across-subject average response angle as a function of source azimuth (error bars show the across-subject standard error). At each distance, perceived source location varies monotonically, but nonlinearly, with stimulus azimuth such that the change in perceived location decreases with increasing source laterality. In addition, for each azimuth, mean response laterality decreases with increasing source distance. However, within-subject response variability is nearly independent of source distance (Fig.3, inset).

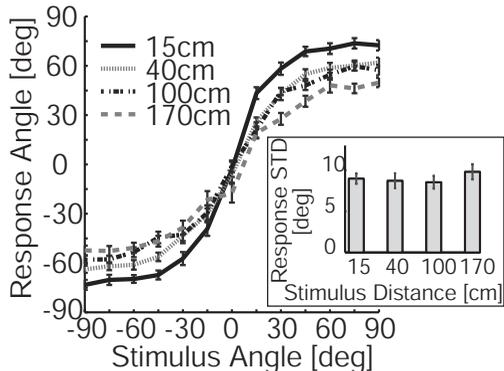


Figure 3. Localization judgments as a function of stimulus azimuth for different distances. Inset shows the across-subject and across-azimuth average of the standard deviations in responses at each distance.

Figure 4 shows that rIT and unnormalized MI between stimulus and response decrease with increasing source distance (and therefore decreasing D/R, shown at the top of the plot).

3.3 Discussion

Both localization accuracy and rIT decrease with increasing distance. Although the mean ITDs in our reverberant BRIRs decrease somewhat with increasing distance (not shown), the magnitude of these changes cannot account for the observed effects of source distance on mean localization judgments or on sensitivity to stimulus azimuth. Acoustical analyses show that, with increasing distance in a room, 1) variability in both ITD and ILD increases, particularly for lateral sources, even though the mean ITD does not change substantially and 2) the magnitude of ILD cues decreases (Shinn-Cunningham and Kawakyu 2003, Hartmann et al. 2005; Shinn-Cunningham et al., 2005). Decreases in the size of ILDs with increasing distance may partly explain why perceived source laterality decreases with increasing distance. However, it is somewhat surprising that response variability does not increase with distance, as the variability in ITD and ILD cues increases.

As measured by rIT, behavioral sensitivity to azimuth decreases with decreasing D/R. Theoretically, such a decrease could be due to either an increase in response variability or a reduction in the range of responses. Given that response variability is nearly independent of D/R, the reduction in rIT with decreasing D/R must be due to the compression of the range of perceived azimuths

Robust sound localization ability in reverberant environments is often ascribed to the precedence effect, in which the direct wavefront dominates localization judgments (Litovsky et al., 1999). Our results show that late-arriving reflections do degrade perceptual sensitivity to source azimuth.

4 General discussion

There are parallels in the effects of reverberation on neural responses and human localization performance. Physiologically, decreasing D/R leads to decreases in both the relative neural response range and rIT between stimulus azimuth and neural firing rate. Behaviorally, decreasing D/R leads to both a reduction in the range of reported azimuths and a decrease in rIT between stimulus and perceived azimuths. In interpreting these parallels, methodological differences between the neurophysiological and psychophysical studies must be kept in mind. While the BRIRs used in the physiological study contained only ITD cues, the BRIRs used in the psychophysical study contained ILDs as well. Although ITD is usually the

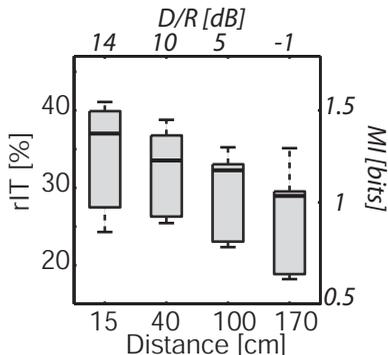


Figure 4. rIT (left axis) and MI (right axis) vs. source distance (D/R along top axis).

dominant cue for the low-frequency stimuli used in our study (Wightman and Kistler, 1992), ILD cues do carry some weight (MacPherson and Middlebrooks, 2002), and may be particularly important in our psychophysical experiments where large ILD cues are present at all frequencies for the short distances examined.

Nonetheless, we can ask whether our behavioral results can be explained by current models of sound localization based on ITD. By making explicit assumptions about the neural code for sound localization, we can relate the activity of single neurons to perceptual judgments of azimuth and therefore test whether these codes can account for the observed behavioral results. Here, we evaluate a simple 2-channel hemispheric difference model of binaural processing (Hancock, 2006; Marquardt and McAlpine, 2001; van Bergeijk, 1962). The model's key assumption is that the neural code for azimuth is derived by comparing the total neural activity in the population of ITD-sensitive neurons across the two inferior colliculi. Since we lack an explicit analytic expression relating D/R to the firing rate of IC neurons, we estimated the total activity in one IC empirically by summing the normalized rate-azimuth curves across our sample of 25

neurons. Assuming symmetry in the nervous system, the activity in the opposite IC is obtained by reflecting the activation signal across the midline (Fig. 5A). The hemispheric difference signal is then the difference in total activity between the two hemispheres. Figure 5B shows the hemispheric difference signal for each of our three room conditions in the physiological experiments. Assuming a monotonic relationship between the hemispheric difference signal and perceived azimuth, the model predictions are qualitatively consistent with the behavioral results (compare Figures 3 and 5B). The hemispheric difference signals become increasingly compressed with decreasing D/R, as do listener judgments of azimuth. Thus, the effects of reverberation on perceived source azimuth are qualitatively consistent with a neural code whereby perceived azimuth is monotonically related to the difference in summed neural activity from both sides of the midbrain.

Our results establish the usefulness of combined neurophysiological and psychophysical studies of sound localization in reverberant environments as a tool for testing neural models of binaural sensitivity. Future efforts will be aimed at testing whether other models of sound localization e.g., the Jeffress model (Jeffress, 1948) can also predict observed behaviors. Further progress in identifying neural codes for sound localization at the midbrain level will require quantitative descriptions of the effects of reverberation on ITD sensitivity.

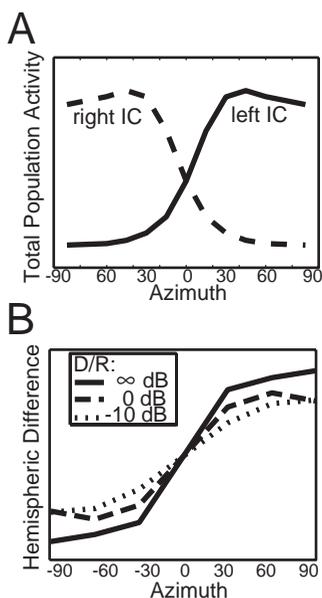


Figure 5. (A) Total neural activity in the left and right IC as a function of stimulus azimuth. (B) Hemispheric difference signal as a function of stimulus azimuth.

6 Acknowledgements

This research was supported by NIH grants DC 02258 (BD), 05209 (BD) and 05778 (BGSC), as well as AFOSR grant FA9550-04-1-0260 (BGSC). The authors thank Connie Miller for surgical assistance and Justin Kiggins's help in collecting the psychophysical data.

References

- Allen, J. B. and Berkley, D. A. (1979) Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.* 65, 943-950.
- Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J. Acoust. Soc. Amer.*, 54, 1458-1470.
- Hancock, K.E. 2006. A physiologically-based population rate code for interaural time differences (ITDs) predicts bandwidth-dependent lateralization (this volume).
- Hancock, KE, Delgutte, B. 2004. A physiologically based model of interaural time difference discrimination. *J Neurosci* 24, 7110-7117.
- Hartmann, WM, B Rakerd, and A Koller (2005). Binaural coherence in rooms, *Acta Acustica* 91, 451-462.
- Jeffress, LA. (1948). A place theory of sound localization. *J Comp Physiol Psych* 41,35-39.
- Litovsky, RY, HS Colburn, WA Yost, SJ Guzman. (1999). The precedence effect, *J Acoust Soc Am* 106, 1633-1654.
- Macpherson, EA and JC Middlebrooks (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *J Acoust Soc Am* 111: 2219-2236.
- Marquardt, T., McAlpine, D. 2001. Simulation of binaural unmasking using just four binaural channels, *Assoc. Res. Otolaryngology Abstracts*.
- Rakerd, B and Hartmann WM (2005). Localization of noise in a reverberant environment. In: *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. D. Pressnitzer, A. de Cheveigné, S. McAdams, and L. Collet (Eds). Springer: 414-423.
- Shinn-Cunningham, BG, and K Kawakyu. (2003). Neural representation of source direction in reverberant space. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (New Pfalz, New York)*, 79-82.
- Shinn-Cunningham, BG, N Kopco, TJ Martin. (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses, *J. Acoust. Soc. Am.* 117, 3100-3115.
- van Bergeijk, W. 1962. Variation on a theme of von Békésy: a model of binaural interaction. *J. Acoust. Soc. Am.* 34, 1431-1437.
- Wightman, FL and DJ Kistler (1992). The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am.* 91: 1648-1661.
- Yin, TC, JC Chan, LH Carney (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J Neurophys.*, 58: 562-583.

Comment by Hartmann

It appears from your data that the onset was included in the stimuli in a way that allows the precedence effect to operate. If the onset had been masked you would have found that the variance increases as the separation between source and receiver increases.

Reply

The precedence effect undoubtedly influenced subject's judgments of lateral source angle by allowing listeners to weigh spatial cues in the onset of the signal more heavily than those in the steady state. Indeed, the results you presented at the last ISH (Rakerd and Hartmann, 2005) suggest that the variance of listener's judgments increases with increasing distance when the onset is masked. However, this study was done in a room with a rather extreme reverberation time (4s). The distribution of spatial cues in the ongoing stimulus in such an environment is much more variable than that in everyday spaces such as classrooms (Ihlefeld and Shinn-Cunningham, unpublished observation). Some recent results by Dizon and Colburn (2006) are relevant. They found robust localization dominance, even in the absence of an onset containing the direct sound alone, when listeners were presented with an ongoing segment of the sum of a single direct source and a single delayed reflection. Such an ongoing precedence effect could allow good localization in modest reverberation even when spatial information in the onset is eliminated. So long as the mean IPD does not depend on the distance between source and listener and the time window over which the auditory system averages binaural cues is long compared to the rate of IPD fluctuations, the response variability should not increase if the onsets are removed.

Dizon, R.M. and H.S. Colburn (2006). "The influence of spectral, temporal, and interaural stimulus variations on the precedence effect," *J. Acoust. Soc. Am.* 119, 2947.

Hartmann, William M.; Rakerd, Brad; Koller, Aaron (2005). "Binaural coherence in rooms," *Acta Acustica* 91, 451.

Rakerd, B and WM Hartmann. Localization of noise in a reverberant environment. In: *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. D. Pressnitzer, A. de Cheveigné, S. McAdams, and L. Collet (Eds). Springer, 2005: 414-423.

Comment by Kollmeier

A straightforward analogy of adding (uncorrelated) reverberation is the addition of uncorrelated noise. Does your data show any evidence that the auditory system treats reverberation differently than expected when using the anechoic condition added with (interaurally uncorrelated) noise at the appropriate signal-to-noise ratio? If so, this would provide hints to an interesting de-reverberation process.

Reply

This is an excellent point which we did not have space to explore in the paper. We are investigating whether a physiologically-based cross-correlation model of ITD processing (Hancock and Delgutte, 2004) can account for neural responses in reverberation (Devore and Delgutte, 2006). Since the model is based on the average interaural correlation over the entire stimulus duration, this approach is equivalent to choosing the signal-to-noise ratio of uncorrelated and anechoic noise sources to match the interaural cross-correlation coefficient of our reverberant stimuli.

The model parameters were first fit to the anechoic data for each neuron, and then held constant to predict neural rate responses as a function of azimuth for the reverberant conditions. Figure 1 shows the range (max-min) of predicted neural response functions plotted against the range of measured neural responses. The model accurately fits the anechoic data (black circles). However, in the reverberant conditions, the model often predicts *more compression* than what was actually observed, suggesting that the auditory system may treat reverberation differently than statistical decorrelation. We are actively investigating the mechanisms that give rise to this robust directional sensitivity in reverberation (see reply to Hohmann's question).

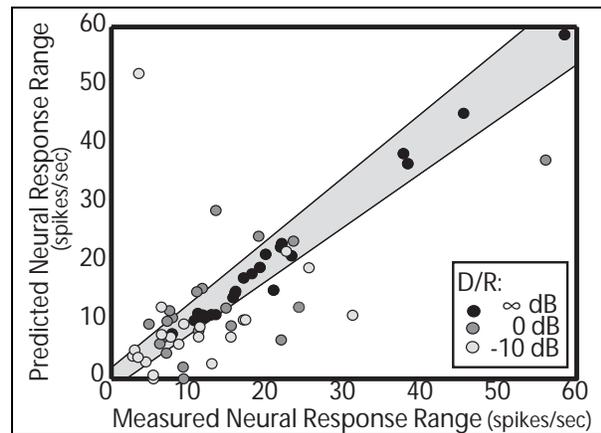


Figure 1. Predicted neural response range plotted against measured neural response range. Gray shaded region depicts ± 1 std from the line $y=x$.

Devore, S. and B. Delgutte (2006). Robustness to reverberation of directionally-sensitive neurons in the inferior colliculus. Computational and Systems Neuroscience 2006, Salt Lake City, Utah.

Hancock, K.E. and B. Delgutte (2004) "A Physiologically Based Model of ITD Discrimination". J. Neuroscience, 24(32), 7710-17.

Comment by Hohmann

You nicely showed the IPD statistics broadening with increasing reverberation. However, you did not show a relation of this effect with the changes in unit responses. Would it be possible to estimate a unit's response from its IPD tuning curve and the IPD statistics? If this is not the case and given that your assumption on how IPD is represented in the auditory system is basically correct, this would suggest that the response is more than an ergodic stochastic process. Answering this question would be interesting to further assess localization models that utilize statistics of interaural parameters, e.g., Nix and Hohmann, JASA 2006.

Johannes Nix and Volker Hohmann (2006) Sound source localization in real sound fields based on empirical statistics of interaural parameters. *J. Acoust. Soc. Am.* 119, 463-479

Response to Comment by Hohmann

The short stimulus (frozen 400-ms noise) used in the present physiological experiments provides insufficient data for an accurate characterization of short-term changes in neuronal firing rate. However, we are currently investigating the relationship between short-term interaural statistics of the stimulus and the instantaneous firing rates of ITD-sensitive neurons in the inferior colliculus (IC). The firing rates of these neurons are undoubtedly modulated, not only by changes in short-term interaural correlation (Joris, 2006), but also by other factors such as monaural temporal envelopes (Lane and Delgutte, 2006) and interactions between frequency bands. Identifying and characterizing these factors (along with the distribution of short-term IPDs) may be necessary to accurately predict the temporal discharge patterns of IC neurons in reverberant conditions.

It is worth noting that, in our psychophysical data, listener's judgments do not directly follow the IPD statistics: The mean response azimuth varies with distance, even though the mean IPD does not, and the response variability is stable with distance, though the variance of the IPD increases.

Joris, P.X., B. van de Sande, A. Recio-Spinoso, and M. van der Heijden (2006). Auditory Midbrain and Nerve Responses to Sinusoidal Variations in Interaural Correlation. *J. Neurosci.*, 26(1), 279-89.

Lane, C.C., and Delgutte, B. Neural correlates and mechanisms of spatial release from masking: Single-unit and population responses in the inferior colliculus. *J Neurophysiol.* 2005; 94: 1180-1198.