Auditory Localization in Rooms: Acoustic Analysis and Behavior

Norbert Kopčo[a] and Barbara Shinn-Cunningham[a,b]

[a]Hearing Research Center, Department of Cognitive and Neural Systems
and [b]Department of Biomedical Engineering
Boston University, 677 Beacon St., Boston, MA 02215

Abstract

In an ordinary room, reverberation and echoes in the signals reaching a listener's ears influence auditory localization performance. The energy of the echoes and reverberation depends on the position of the listener in the room as well as on the position of the sound source relative to the listener. In this paper, the effects of echoes and reverberation are quantified through analysis of reverberant Head-Related Transfer Functions (HRTFs) measured in an ordinary classroom. HRTFs were measured for several human listeners and a KEMAR acoustic manikin at four different listener positions in the room and multiple source positions relative to the listener. Azimuthal localization performance was also measured for several listeners in the room as a function of listener position. Compared to the acoustic cues it was found to be less sensitive to a change in room location. The only similarity was found between the magnitude of frequency-to-frequency variations in basic localization cues and the variability in localization performance, demonstrating that localization accuracy decreases with increasing reverberant energy.

Introduction

In a room, the ability of human listeners to localize sounds is influenced by echoes and reverberation (which are henceforth collectively referred to as "reverberation," for brevity; Santarelli, 2000). The effect of reverberation can be both beneficial and detrimental, improving distance perception and degrading azimuthal localization. However, the pattern of reverberation differs from room to room as well as from position to position within a given room. For a listener in the center of a room, most reflective surfaces are relatively far from the listener and reflections are diffuse for all source positions. On the other hand, when the listener is close to a wall, prominent early reflections arise whose magnitude and timing depend on the location of the source relative to the walls and to the listener.

The goal of this study is to analyze how localization cues in the signals reaching a listener's ears are influenced by reverberation and to evaluate whether acoustic effects can account for how localization performance varies with a listener's position in a room. A set of head-related transfer functions (HRTFs; see Santarelli, 2000) was measured for a manikin (KEMAR) located at different positions in a classroom. The effect of reverberation on interaural differences and spectral magnitude is evaluated by computing how these cues vary with source position relative to the listener and listener location relative to the room. Results are compared to behavioral localization results (Kopčo, Brown, and Shinn-Cunningham, 2001) for similar configurations of source and listener in the room.

Methods

Acoustic analysis

HRTFs were measured for a KEMAR manikin located at the four positions in a classroom (Center, Back, Ear, Corner) (T_{60}≈700 ms) shown in Figure 1.

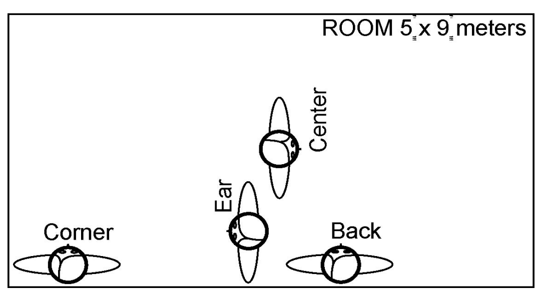

Fig. 1 Positions of KEMAR at which HRTFs were measured.

HRTFs were measured for sources in KEMAR's right front quadrant at all combinations of azimuths from 0° to 90° (15° steps) and distances 0.15, 0.40, and 1 m (for sources in the horizontal plane containing the ears). Responses to Maximum-Length Sequences (e.g., see Zahorik, 2000) were measured to estimate a 750-ms-long head-related impulse response (HRIR; 44.1 kHz sampling rate). Stimuli were presented from a PC computer using a TDT system, Crown amplifier, and a Bose cube speaker. Knowles Electret microphones mounted in earplugs in KEMAR's ear canals were fed back to the TDT to make blocked-meatus recordings. The magnitude spectrum of the measurement system was relatively flat (within 10 dB) between 300 Hz – 12 kHz range. The dynamic range was at least 40 dB at all the frequencies. HRTFs from the center-room position were time-windowed using a cosine-squared onset/offset window (1 ms) to obtain pseudo-anechoic HRTFs against which other measurements are compared.

Interaural level differences (ILDs) were computed as the difference between the left and right ear HRTF RMS energy between 2000 – 5000 Hz. ILD variability was computed as the mean absolute value of the frequency-to-frequency difference in the ILD (using a frequency step of 1 Hz). Interaural time differences (ITDs) were estimated from the interaural delay producing the maximum peak in the cross-correlation of the left- and right-ear HRIRs bandpass-filtered from 200 – 2000 Hz.

Localization experiment

Subjects were asked to localize sound sources when in the same room locations used for KEMAR measurements (Kopčo et al., 2001). Six normal-hearing subjects pointed to the perceived source location (five 150ms-long pink-noise bursts) presented from random locations between 0° – 90° azimuth and 0.15 – 1 m distance in the horizontal plane containing the ears. Each subject performed 300 trials in each room location. The (signed) mean error (*re.* actual source position) and standard deviation in response was computed from these results.

Results

Effect of reverberation on spectral cues

Figure 2 compares HRTF magnitude spectra at the four extreme source positions with KEMAR in the center of the room for anechoic and reverberant conditions. Reverberation adds frequency-to-frequency variability to magnitude spectra. This variability grows with source distance and is greatest at high frequencies. Variability increases with source azimuth for the ear contralateral to the source position and decreases with azimuth for the ipsilateral ear. Reverberation also fills in high-frequency notches, particularly at the far ear.

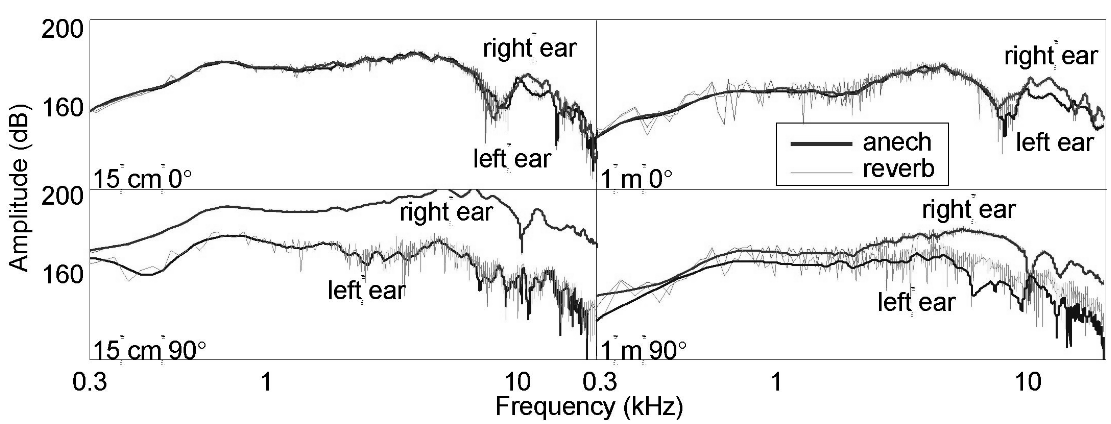

Fig. 2 Anechoic and reverberant magnitude spectra at four source positions with KEMAR in center of room.

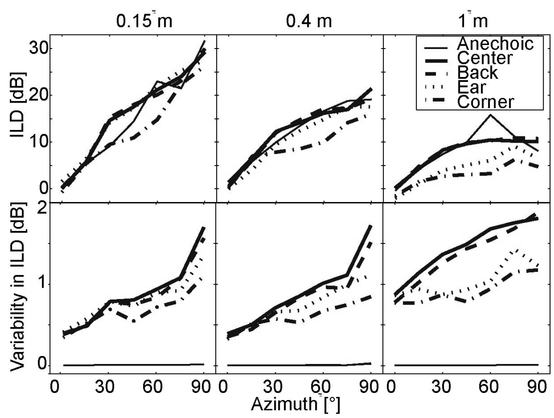

Fig. 3 ILDs and cross-freqency variability in ILDs at 4 room locations as a function of source azimuth.

Effect of reverberation on ILDs

Figure 3 shows the ILD for different room locations and source positions. ILD magnitudes tend to decrease with reverberation, particularly for distant sources and conditions in which there is asymmetry in early reflections (Ear and Corner conditions). The frequency-to-frequency variability in the ILD (which is essentially zero in the anechoic condition) tends to increase with distance and is greatest for the Center condition. For room locations with early reflections, ILD variations are smoother and more systematic with frequency.

Effect of reverberation on ITDs

Within the biologically-plausible range (±0.8 ms), the ITD of the cross-correlation peak is roughly independent of source distance and room position (Fig. 4). However, in reverberant conditions, the magnitude of this peak value decreases dramatically with distance and with the number of nearby walls. In addition, in the Corner and Ear conditions, a secondary peak (outside the biologically-plausible range of ITDs) can be of equal or larger magnitude than the primary peak in the cross-correlation.

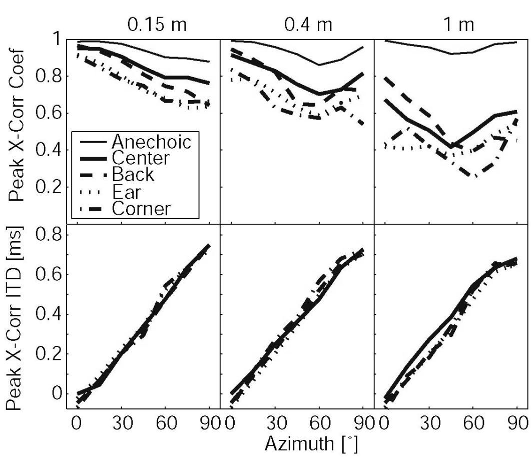

Fig. 4 The peak value in the cross-correlation function within +/- 1 ms range and the corresponding ITD.

Figure 5 illustrates that, as with the ILD, reverberation causes frequency-to-frequency variation in ITD. In the Center and Back conditions, this variation is essentially random around the “true” (anechoic) ITD. In the other conditions, the departures are more significant due to the early, asymmetric, strong reflections.

Predictions vs. localization performance

Acoustic analysis shows that all localization cues in the signals reaching a listener are influenced by reverberation in a manner that depends on room position. To the extent that these cues determine spatial auditory perception, localization performance should also be influenced in a way that varies with listener location. Figure 6 summarizes behavioral results from a localization experiment performed in the

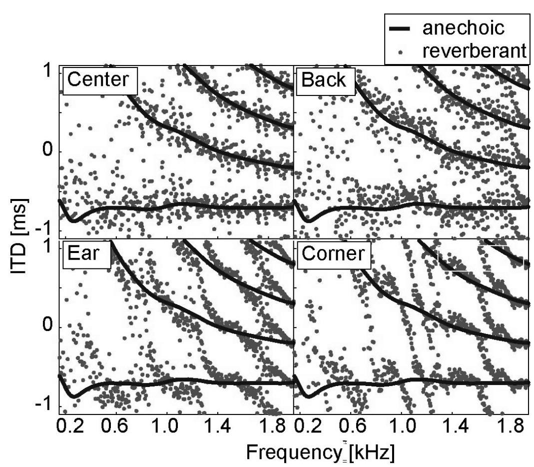

Fig. 5 ITDs as a function of frequency in the anechoic and reverberant conditions for source at 90° 1 m.

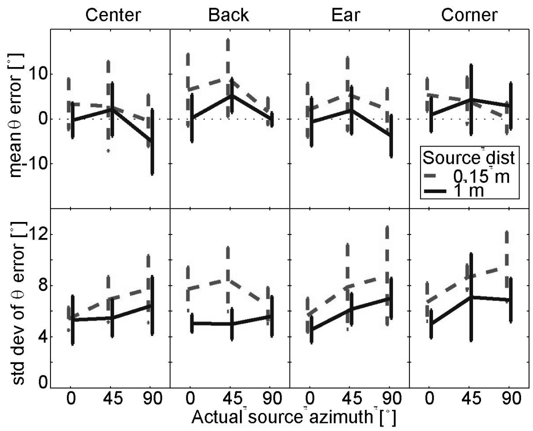

Fig. 6 Across-subject mean and std. dev. of the response error, i.e., the difference between perceived and actual source azimuth.

room in which acoustic measurements were taken (Kopčo et al., 2001). Two small but statistically-significant trends were observed. 1) Azimuthal perception in the Back and Corner positions was biased towards the median plane (approximately 3.5°). 2) The variance in perceived azimuth was smallest for listeners in the Center condition, greatest in the Corner condition, and intermediate for the other two conditions (bottom row of Fig. 6).

The azimuthal bias is difficult to explain from results of the acoustic analysis. Acoustically, Ear and Corner conditions are most similar and most influenced by reverberation, but bias is only significant for Back and Corner locations.

On the other hand, the increase in the azimuthal response variance is consistent with both ILD variability and ITD decorrelation, which are greatest for the Ear and Corner conditions. This explanation cannot account for changes in bias with distance: the variability in acoustic parameters increases with distance while variance in perceived azimuth decreases with distance. The decrease in response bias with distance may be partially explained by the measurement method. If one assumes that response variability is constant in x-y-z coordinates, the same error translates to larger angular errors for nearby sources.

Summary and discussion

Acoustic analysis shows that the effect of reverberation on localization cues varies dramatically with listener position in a room. On the other hand, effects of room position on localization performance are modest, at best. Some of this apparent discrepancy may be resolved by considering how acoustic cues change over time (as the current analysis evaluates only the expected value of the cues, ignoring variation in these cues over time). In fact, such dynamics are known to be perceptually important (cf. the "precedence effect"); for instance, the localization cues available at the onset of the stimulus will be much less distorted by reverberation than this first-order steady-state analysis suggests. Further, listeners may crudely estimate the effect of reverberation on the received stimuli and adjust the computation of source position accordingly. Future analysis will incorporate physiologically-based models of auditory processing (e.g., Colburn, 1977) to predict how basic localization cues in reverberant signals may be extracted by the brain.

References

Colburn, HS (1977): *"Theory of binaural interaction based on auditory-nerve data. II: Detection of tones in noise." Journal of the Acoustical Society of America 64:525-533.*

Kopčo, N, T Brown, and BG Shinn-Cunningham (2001): *"Effect of listener location on localization cues and localization performance in a reverberant room," Presented at the MidWinter meeting of the Association for Research in Otolaryngology, Feb 3-7, 2001, St. Petersburg, Florida*

Santarelli, S (2001): *"Auditory Localization of Nearby Sources in Anechoic and Reverberant Environments," Unpublished Ph.D. Thesis, Boston University*

Zahorik, P (2001): *"Limitations in using Golay codes for head-related transfer function measurement." of the Acoustical Society of America, 107, 1793-1796.*

Acknowledgements

Work supported by AFOSR and Alfred P. Sloan Foundation.