LOCALIZING SOUND IN ROOMS

Barbara Shinn-Cunningham

Department of Cognitive and Neural Systems and Biomedical Engineering
Boston University, 677 Beacon St., Boston, MA 02215
Email: shinn@cns.bu.edu
Ph: 617-353-5764
FAX: 617-353-7755

INTRODUCTION

Relatively little psychoacoustic work has examined how realistic echoes and reverberation affect spatial auditory perception. Within psychoacoustics, echoes and reverberation are generally thought to 1) cause little degradation in directional perception (as suggested by studies of the "precedence effect"; e.g., see Litovsky, Colburn, Yost & Guzman, 1999) and 2) improve distance perception (by some essentially unknown mechanism; e.g., see Mershon & King, 1975).

Head-related transfer functions (HRTFs) show how the signals that reach the two ears are related to the original source signal from a specific location in space (Wightman & Kistler, 1989a; Wightman & Kistler, 1989b; Wenzel, 1992; Carlile, 1996). HRTFs have been examined in detail in anechoic space as a function of source direction and,

Figure 1: Magnitude spectrum (dB) of HRTFs in the center of a reverberant room as a function of source position relative to listener. Left and right columns show near (15 cm) and far (1 m) sources, respectively. Top, middle, and bottom rows show the lateral angle of the source relative to median plane (0ß, 45ß, and 90ß to the right, respectively).

450 ms) using a maximum-length-sequence technique. Measurements were made for individual human listeners as well as a KEMAR manikin

Figure 2: Interaural phase difference versus frequency for the same listener and source positions as in Figure 1.

Figure 3: Four listener configurations for which HRTFs were measured in a reverberant room (not to scale).

In order to systematically evaluate HRTFs as a function of listener location and orientation, HRTFs were measured for four different configurations of the listener (Kopčo & Shinn-Cunningham, 2001). Figure 3 diagrams the listener positions/orientations for which HRTFs were measured (for the same six relative source positions shown in Figures 1 and 2). Results show that the cleanest results are obtained when a listener is in the center of the room (configuration 1 in Figure 3). Spatial cues becoming increasingly degraded as the listener approaches a wall (configuration 3), are even worse when the subject has his back to the wall (configuration 2), and are most distorted when the listener is located in the corner of the room (configuration 4).

Figure 4 demonstrates how much worse the acoustic distortion can be by showing ITD as a function of frequency for the same relative source positions as in Figure 2, but for listener configuration 4 (corner of the room; note that Figs 2 and 4 use different ITD scales). The blue symbols show the ITD that would arise for anechoic HRTFs; the red symbols show the corresponding ITD for the reverberant HRTFs. For a source near to and directly in front of the listener, echoes and reverberation only marginally affect ITD; however, for all other conditions, the ITD is dramatically distorted.

Figure 4: ITD versus frequency for the same source positions as in Figure 2. Blue symbols show anechoic and red symbols reverberant results for a listener located in the corner of the room.

Taken as whole, acoustic measures suggest that directional localization performance should be degraded in a room compared to in anechoic space, and that this degradation should depend on where the listener is located in the room. Directional performance should be worst when a subject is located in the corner of the room and best when a listener is in the center of the room. In contrast, reverberation should provide source distance information. The degree to which distance perception varies with source and listener position may help in teasing out what aspects of reverberation provide distance information to the listener.

BEHAVIORAL MEASURES

Human localization performance was measured in the same room in which acoustic measures were made (Santarelli, Kopčo & Shinn-Cunningham, 1999; Santarelli, 2000; Santarelli, Kopčo & Shinn-Cunningham, 2000; Kopčo et al., 2001) using an experimental procedure essentially identical to that employed in a previous anechoic localization study (Brungart et al., 1999a). In the experiments, a human experimenter positioned a small speaker at a condition in every spatial dimension. Further examination of the data showed that listener s accuracy improved over hours of practice in the first condition but was essentially unchanged during the second condition. No similar change were seen in the previous anechoic data (reanalyzed for these trends). These results imply that subjects "adapt" to a room over time, and that whatever the subjects learn transfers from one configuration (without a board) to another (with the board in place) that is very different, acoustically.

A follow-up study was recently conducted to explore how robust these effects are (Kopčo et al., 2001). We hypothesized that with practice in a room, subjects adapt and localization improves, and that this learning transfers from one listener configuration to another; i.e., that there is some "room specific" characteristics of reverberation common across all listener positions and orientations in the room. To examine these hypotheses, two groups of listeners performed a localization task similar to that in the initial experiment. Each listener performed four sessions of localization, each from one of the four configurations shown in Figure 3. The first group performed the sessions in the order indicated in Figure 3, starting in the center of the room (configuration 1) and ending in the corner of the room (configuration 4). The second group performed the sessions in the opposite order.

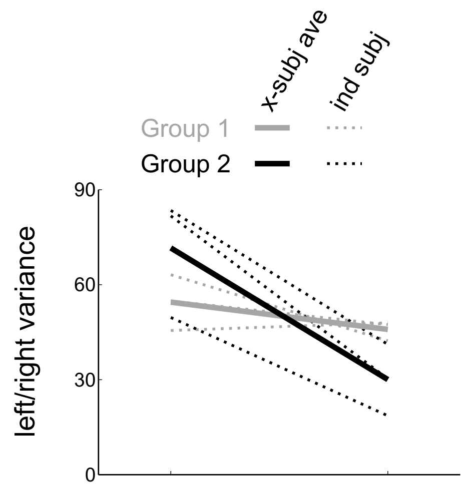

Figure 5: Response variability versus session. Solid lines show across-subject means. Dashed lines show individual subjects.

To the extent that room position affected localization accuracy, we hypothesized that performance would be best when listeners were in the center and worst when listeners were in the corner of the room. To the extent that practice in the room improved localization, performance should be better in the last session of the experiment and worst in the initial session, independent of room configuration order. If both factors influence localization accuracy, the second subject group should show the largest improvement from session one to session four, because both the acoustic and the learning effects would push the results in the same direction. In contrast, for the first subject group, who begin the experiment in the easiest acoustic setting (but without any prior experience in the room), the two effects would interact. In this case, insight into the relative importance of learning and room acoustics on localization performance could be gleaned by comparing results for the two groups.

Response variability in the left/right dimension is shown in Figure 5 for the two groups for the initial session (left) and the final session (right). Results show that the Group 2 subjects (for whom both learning and acoustic effects should cause performance to be best in session 4) show much larger changes in response variability between session 1 (the most acoustically-challenging, corner configuration) and session 4 (the room center configuration). The Group 1 subjects, who started in the easy room configuration and moved to the hardest room configuration, showed only a modest decrease in variability between sessions 1 and 4.

These results support the hypothesis that both learning and room acoustics influence localization accuracy. In addition, since the learning transfers across room configurations that are acoustically very different (and that lead to very different signals at the ears), the results suggest that with practice on the task, subjects learn some very general characteristic about the room reverberation that is similar for all room positions, independent of the exact structure of the echoes and reverberation interacting with the direct sound.

In another set of experiments (Shinn-Cunningham, Santarelli & Kopčo, 2000a), measured HRTFs were used to simulate anechoic and reverberant listening conditions under headphones. Subjects were asked to indicate the heard distance of the simulated sources for sources that were presented both binaurally and monaurally, for sources to the side (along the interaural axis) and to the front. Simulated source distances ranged from 15 cm to 1 m, the range in which ILD cues vary dramatically with distance for sources along the interaural axis. We expected to find that subjects could judge source distance accurately for binaural presentations of lateral sources because subjects in a real anechoic space have been shown to do relatively well on a similar task. Binaural and monaural presentations of reverberant simulations were used so that we could determine whether the reverberation cue for source distance arose from monaural effects (such as the direct-to-reverberant energy ratio; e.g., see Mershon & King, 1975; Bronkhorst et al., 1999) or binaural effects (such as the interaural decorrelation caused by reverberant energy, which is correlated with the direct-to-reverberant energy ratio).

Results of this study were compelling. In every anechoic condition, subject performance was near chance. In all reverberant conditions, subject performance was well above chance. Further, for lateral sources simulated with the reverberant HRTFs, monaural and binaural distance perception was essentially equal; binaural cues were irrelevant for the task. Interestingly, for medial sources, turning off one ear did affect distance judgments slightly, with subjects consistently overestimating the simulated source distance. However, we believe this bias arises because the simulated sources were heard in the wrong direction (i.e., along the interaural axis), where the pattern of reverberation varies differently with distance than it does for medial sources.

Results suggest that reverberation is an important distance cue. Even when sources are so close to the listener that there exist reliable ILD distance cues, these cues are ignored when listeners expect (are calibrated for) a reverberant listening environment. The cue for distance is probably correlated with the direct-to-reverberant energy ratio, although it is unlikely that the human auditory system can accurately compute such a ratio from the total signal reaching the ear. Further, the distance cue provided by reverberant energy is not a binaural cue, but a monaural cue; however, perceived direction (which is strongly influenced by binaural cues) affects perceived distance.

SUMMARY

Inclusion of realistic echoes and reverberation in virtual auditory environments will have a number of dramatic effects, including increasing the realism of the display (Begault, 1992b; Durlach, Rigapulos, Pang, Woods, Kulkarni, Colburn & Wenzel, 1992; Gilkey, Simpson & Weisenberger, 2001), improving distance perception (Shinn-Cunningham, 2000a), providing information about the room itself (Gilkey et al., 2001), and degrading directional accuracy, albeit slightly (Shinn-Cunningham, 2000b). Relatively little is known about which aspects of reverberation are most critical for each of these perceptual results. Further, it is likely that these different perceptual effects arise from different aspects of the reverberation. For instance, while our results hint that distance perception is driven more by monaural than binaural cues, impressions about room size depend on the amount of interaural decorrelation induced by echoes and reverberation, a binaural cue.

These results have a number of implications for the design of effective, efficient acoustic room simulators, pointing to the need to take into account how various aspects of reverberation influence perception. Further work is necessary to tease apart how reverberation influences various percepts important in virtual environments. More specifically, we must examine how accurately room reflection patterns must be simulated in a virtual environment to achieve accurate distance perception as well as realism (while some work addressed these issues, e.g., Begault,

Kopčo, N. and B. G. Shinn-Cunningham (2001). Effect of listener location on localization cues and localization performance in a reverberant room. 24th MeetingAssoc Res Otolaryng, St. Petersburg Beach, FL.
Litovsky, R. Y., H. S. Colburn, W. A. Yost and S. J. Guzman (1999). The precedence effect. Journal of the Acoustical Society of America, **106**(4): 1633-1654.