Abstract 103: LOCALIZATION OF NEAR-FIELD SOURCES IN A REVERBERANT ROOM

Scott Santarelli, Norbert Kopco, and Barbara Shinn-Cunningham

Hearing Research Center and Department of Cognitive and Neural Systems, Boston University

0. ABSTRACT

Subjects were asked to point to the position of sound sources within one meter of the head while seated in a medium-sized, echoic classroom. These localization results are compared to results from a previous study that used an identical procedure (but a different set of subjects) in anechoic space. Overall, localization in the reverberant room is worse than observed in the anechoic conditions (a surprising result, given that, for sources near the head, the reverberant energy is weak relative to the direct sound reaching the ears).

In order to understand these results more fully, gross interaural time difference (ITD) and interaural level difference (ILD) cues were estimated for near-field sources in anechoic space using a simple model. This analysis is especially interesting since iso-ITD and iso-ILD surfaces in the near field differ from the "cone-of-confusion" (iso-binaural cue) surfaces that occur for sources relatively far from the head. To analyze response errors for each subject, the estimated ITD and ILD values are found for both source and response positions. From this analysis, localization errors are estimated in units of binaural cues (ITD and ILD) and non-binaural cues. This analysis implies that reverberation does not affect interaural time and interaural intensity localization errors in the same way. It also appears that there are consistent subject differences in both binaural abilities and in non-binaural localization abilities, and that these differences capture much of the observed intersubject variability.

Project supported by AFOSR grant F49620-98-1-0108. We wish to thank Douglas Brungart for providing us with his raw data as well as

1. BACKGROUND

Most previous studies of spatial hearing focus on sources

- relatively far from the head
- varying in direction only (ignoring distance) *-or-*
  varying in distance only (ignoring direction)
- in anechoic space *-or-* under headphones

Brungart (1997)

- 3-d localization of anechoic sources near the head

Current study

- same experimental setup and protocol in reverberant space.

HYPOTHESES

Performance should not be dramatically affected by reverberation.

For sources relatively close to the head, the amount of direct energy from the source is large compared to the amount of reverberant energy (i.e., very similar to an anechoic room).

Reverberation should affect spectral cues more than binaural cues.

Reverberation causes frequency-dependent changes in acoustic cues.

Spectral cues depend on the energy in narrow frequency ranges that will be distorted in different ways for different source locations.

Binaural cues should be relatively more robust since information can be integrated across all available frequencies.

COORDINATE SYSTEMS

For sources relatively far from the head, the set of locations leading to the same binaural cues form "cones of confusion." Subjects often make directional errors that fall near the same cone of confusion as the source.

Most localization studies analyze data in terms of the azimuth and elevation. Researchers label obvious cone-of-confusion errors as "reversals;" however, this ignores many possible cone of confusion errors.

Some researchers (e.g., Duda, 1997) have proposed using the coordinate system shown below (FIGURE 1). This coordinate system separates cone of confusion errors (along θ) from other errors (ϕ, r). Our analysis is similar (see 2. DOUGHNUTS OF CONFUSION, below).

Duda & Martens (1998) & Brungart (1998) have argued that the large interaural level differences occuring for sources near the head provide distance information.

However, near-field ILDs vary with both distance and direction (see FIGURE 2). Analysis using radial distance (r) as a coordinate confounds ITD and ILD performance.

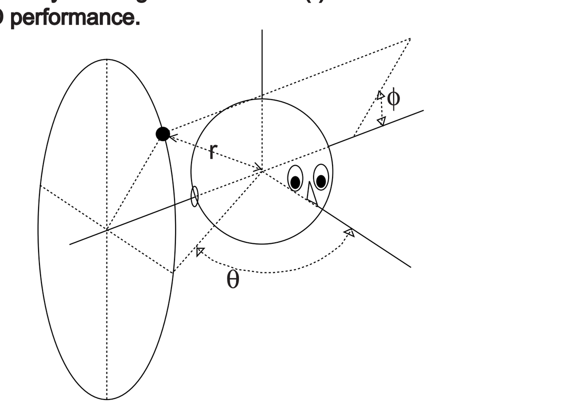

FIGURE 1: Alternate coordinate system (from Duda, 1998, p. 52). The three coordinates are angle of cone centered on the interaural axis (θ), angle from horizontal plane (ϕ), and radial distance (r).

2. DOUGHNUTS OF CONFUSION

PERCEPTUALLY-BASED COORDINATES

We analyzed our data using a spatial representation based on the perceptually-relevant cues of

- interaural time differences (ITDs)
- gross interaural level differences (ILDs) and
- other, mainly spectral cues.

In the near field, ILDs occur due to both

- head shadow effects (as in the far field) and
- difference in the path lengths from source to left and right ears.

ISO-BINAURAL SURFACES

To a first order approximation (using point receivers in free space), the near-field ILD component is constant when the ratio of the distances from source to ears is constant.

The set of all point-source positions leading to the same pathlength-component of the ILD (the iso-ILD surface) is a sphere whose center falls on the interaural axis (see left side of Figure 2, below).

Sources near midline fall on iso-ILD spheres of nearly infinite radius with centers nearly infinitely far from the head.

As the point source approaches one ear, the iso-ILD sphere degenerates to a point (a sphere of radius 0 at the ear).

Iso-ITD sufaces form the familiar "cones of confusion" that vary primarily with the aimuth of the source (see right side of Figure 2).

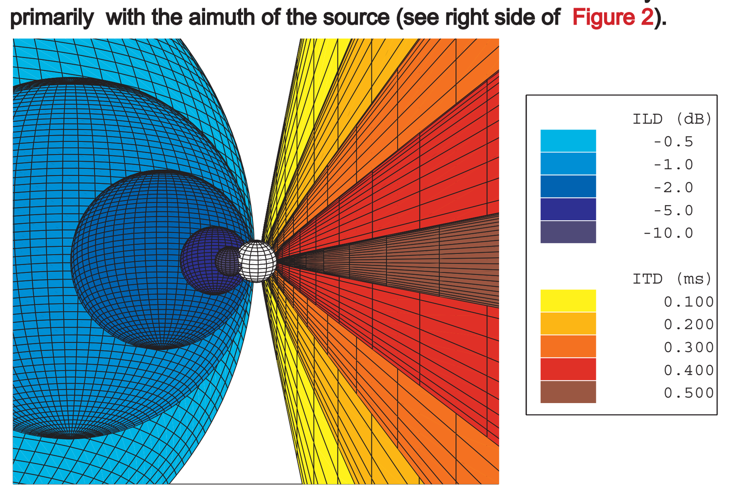

FIGURE 2: Iso-ILD (left) and Iso-ITD (right) surfaces for a point-receiver model (head in white).

CIRCLES OF CONFUSION

The set of locations with the same gross binaural cues is the intersection of an ILD sphere and an ITD cone of confusion, forming a circle perpendicular to and centered on the interaural axis.

The ITD, ILD, and angle along this circle of confusion can uniquely represent any position in space. The circles of confusion are identical to the circle angle coordinate discussed in Duda (1997; see FIGURE 1). ITD units are used in our scheme, but otherwise, the ITD coordinate is very close to the cone-of-confusion angle proposed in FIGURE 1. The major difference in our coordinate scheme is the ILD coordinate, shown on the left of FIGURE 2.

By comparing the ITD, ILD, and circle of confusion angle of source and response, one can see how accurately subjects use gross ITD, near-field ILD, and any monaural (spectral) cues.

VOLUMES OF UNCERTAINTY

Our ability to judge binaural cues is not perfect. We can discriminate

- ITDs within about 10 ms of the source ITD and
- ILDs within about 0.5 dB of the source ILD.

From gross binaural cues, subjects should be able to determine near-field source position within a volume delineated by iso-ILD spheres within one JND of the source ILD and cones of confusion within one JND of the source ITD.

We call these doughnut-shaped volumes of rotation (symmetric about the interaural axis) doughnuts of confusion (see Figure 3, below).

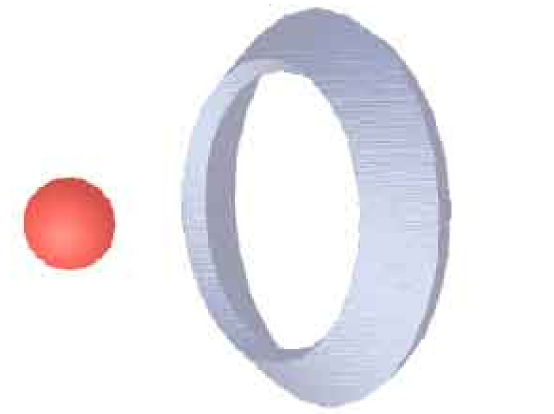

FIGURE 3: Cartoon showing (ITD 400 µs, ILD 2 dB) *doughnut of confusion* (gray) and head (pink).

CAVEATS

This analysis ignores head and pinnae effects.

For a spherical head, one would still predict doughnuts of confusion; however, the value of the ILD within a given doughnut-volume would vary with frequency. For a real head, interaural cues will vary along our "circles of confusion," (see Duda, 1997). We are currently measuring how gross ILD, ITD, and spectrum vary with source position in our reverberant room to evaluate our simple assumptions.

A fixed distance error corresponds to different size errors in units of ITD, ILD, and angle.

An error of 2 cm corresponds to a larger errors for sources near the head compared to far from the head.

This is also true of a spherical coordinate representation. However, in our scheme, nonuniformity in the ILD coordinate is even more pronounced. Of course, this may be reflected in human performance, if ILD cues are used in near-field listening situations.

3. METHODS

SUBJECTS

- seven total: two female, five male (22 – 44 years of age)
- six with normal hearing; one with marginal high-frequency loss (R6)

REVERBERANT ROOM

- 14' x 20' rectangular classroom with carpeted floor and hard walls
- reverberation time R_{60} approximately 250 ms

TEST CONDITIONS

- ROOM: subject in center of room (facing short wall)
- BOARD: 8'x4' tiled board (parallel median plane) 10" from left ear

STIMULI

- five 150-ms long pink noise bursts separated by 30 ms silence
- random locations in 1 m diameter hemisphere to right of subject
- level equalized (to overcome distance effects), + additional 15 dB rove

EQUIPMENT

- wooden chair with attached head rest
- PC sound card, Crown D-75A amplifier, point source (D. Brungart)
- Polhemus Isotraks on point source and response wand
- mirror on plastic easel (to allow subjects to view responses)

TRIAL PROCEDURE

At the start of each trial, subjects closed their eyes and a random location was chosen by the computer. The experimenter then positioned the point source (within 15 deg direction and 6 cm distance of the chosen location as measured by the electromagnetic tracker). The computer presented one of five random noise bursts after recording the location of the source. The subject opened their eyes and positioned the response wand at the heard location. The computer then recorded this location.

OVERALL PROCEDURE

At the start of each session, the subject's head location was measured for calibration. Each session lasted 1 - 1.5 hours, and consisted of multiple 50-trial blocks separated by 5-minute rest intervals. Each subject received 200 practice trials (in one session) prior to testing. Each subject performed 1000 test trials/condition (roughly 10 hrs/condition). All subjects performed the ROOM condition first; five of seven subjects then performed the BOARD condition.

BRUNGART PROCEDURE

- ANECHOIC listening condition

4. ANALYSIS

In Brungart's study and in the current study, the size and distribution of response errors varies with location. Thus, we must normalize the data before results are compared. In the next few sections, plots show source versus response location for ITD, ILD, & angle. Data was first binned for each coordinate. The percentage of responses falling into each response bin was then calculated for each source bin.

Plots are organized as 24 x 24 grids:

- columns represent source position
- rows represent response position
- gray-scale represents % responses in the row (for that column).

Perfect performance would result in a black diagonal line. Any spread of data around the diagonal represents noise in the responses. Any systematic deviation from the diagonal represents bias in the responses.

5. RESULTS: ITDS

Data are plotted on 24 x 24 grids (FIGURE 4). Legend (below right) shows how to interpret each plot. Here, each grid is 20 µs wide, with entries ranging from 0 µs to 480+ µs. In each of the next few sections, data from each subject are presented individually. Recall that R1-R5 performed both ROOM and BOARD conditions (data aligned vertically).

ITD results are comparable for all three conditions. All subjects appear to rely on ITD, with varying degrees of accuracy (spread of data varies from subject to subject).

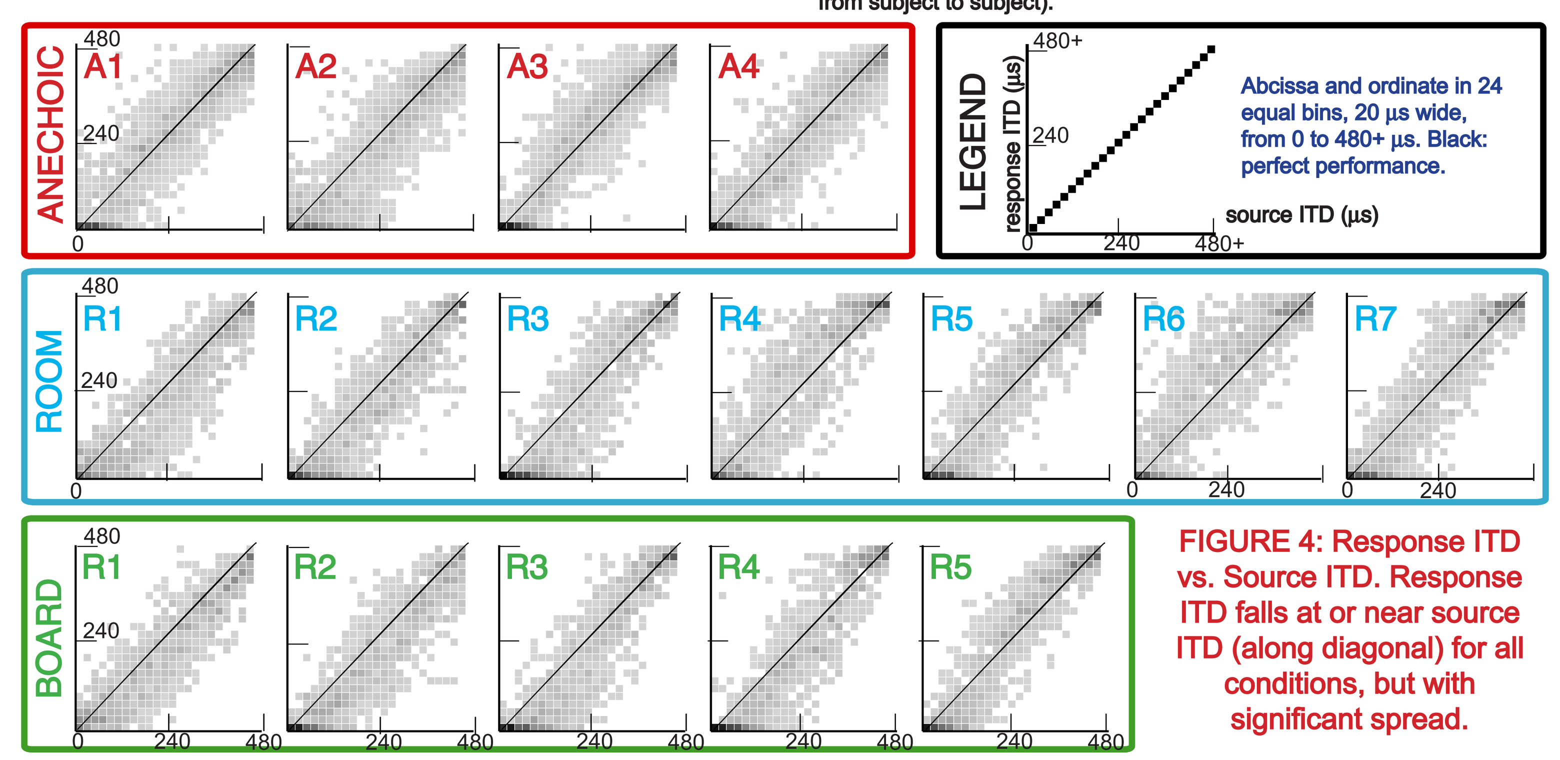

FIGURE 4: Response ITD vs. Source ITD. Response ITD falls at or near source ITD (along diagonal) for all conditions, but with significant spread.

6. RESULTS: CIRCLE OF CONFUSION ANGLE

Plots (FIGURE 5) show the angles of source and response along the circle of confusion. Bins are 15 deg wide to span 360 deg. Straight ahead is 0 deg; up is 90 deg; behind is 180 deg; and down is 270 deg. Common up/down and front/back reversals fall along dashed lines (see LEGEND).

Data show large intersubject differences. For instance, subject A1 has a tendency to confuse sources in front and down with sources in back and down (front to back reversals). Other ANECHOIC subjects show consistent responses (small spread of data) near, but slightly below, the diagonal.

Many of the subjects in the reverberant conditions show greater spread in their responses and more reversals than the ANECHOIC listeners. In particular, the best ROOM/BOARD results (R1 or R2) are worse than the best ANECHOIC subjects. R6, whose results are extremely poor, suffers from marginal high-frequency hearing loss.

The effect of adding a single hard reflective surface near the head is very idiosyncratic. For most subjects (R1, R2, R3, and R5), ROOM and BOARD conditions are similar. On the other hand, R4 showed front to back confusions in the ROOM condition. In the BOARD condition, results are reversed: R4 showed a large number of back to front confusions.

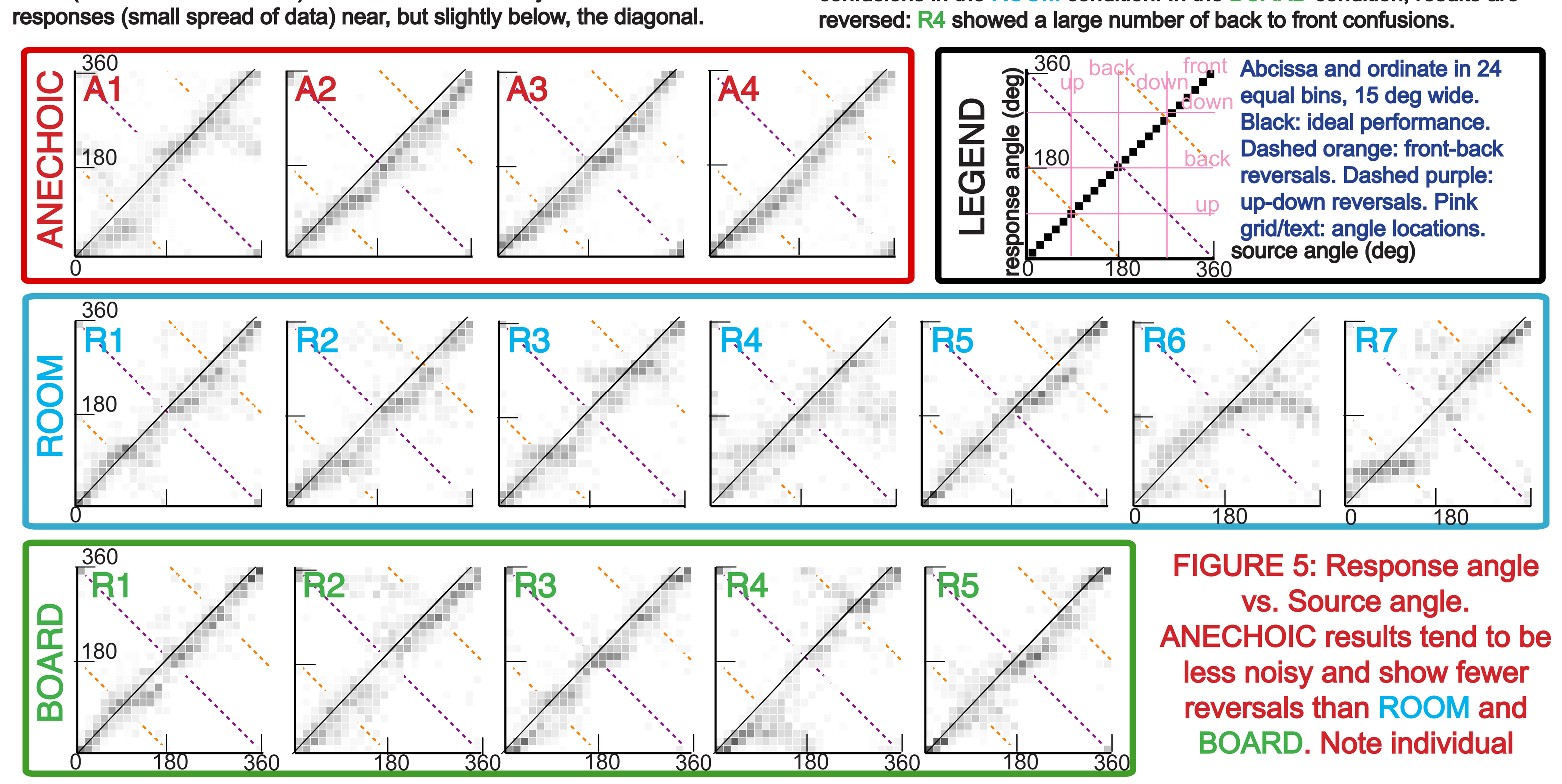

FIGURE 5: Response angle vs. Source angle. ANECHOIC results tend to be less noisy and show fewer reversals than ROOM and BOARD. Note individual

7. RESULTS: ILDS

ILD results are plotted in 0.5 dB bins ranging from 0 to 12 dB (Figure 6). Results show interesting differences across the three conditions.

ANECHOIC results tend to fall below the diagonal: subjects tend to underestimate the ILD of the source (responses are too far from the head and/or too close to the median plane).

ROOM results tend to fall on the diagonal for small ILDs (sources far from the head and/or near the median plane). For ILDs greater than 3-4 dB in the ROOM condition, most subjects overestimate ILD (responses tend to be too close to the head and/or too near the interaural axis).

BOARD results look similar to ROOM results; however, subject by subject comparison shows that ILD error is somewhat smaller. Since BOARD results were gathered after ROOM results, this slight decrease in bias may reflect learning (see 9. LEARNING) rather than acoustic effects.

The pattern of results in the reverberant conditions suggests that subjects are not using ILD cues appropriately or consistently, at least at large ILDs (near the head). Given that source level was, on average, lower when sources were near the head (to remove intensity as a cue for source distance), this failure may indicate that the direct-to-reverberant energy ratio was used instead of ILD by subjects in the reverberant conditions. See 8. RESULTS: DISTANCE for further discussion of this possibility.

Differences in ILD results are summarized in FIGURE 7, which plots the average signed error of the ILD judgements for each subject (computed by averaging the errors plotted in FIGURE 6 over all bins). Bias is negative for all four subjects in the ANECHOIC condition and is greater or equal to zero for all seven subjects in the ROOM condition. In the BOARD condition, bias decreased for four of the five subjects.

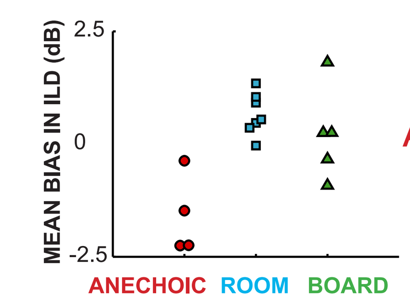

FIGURE 7: Mean ILD bias in three conditions. ILD is underestimated in ANECHOIC condition but overestimated in ROOM and BOARD conditions (symbols displaced horizontally for clarity).

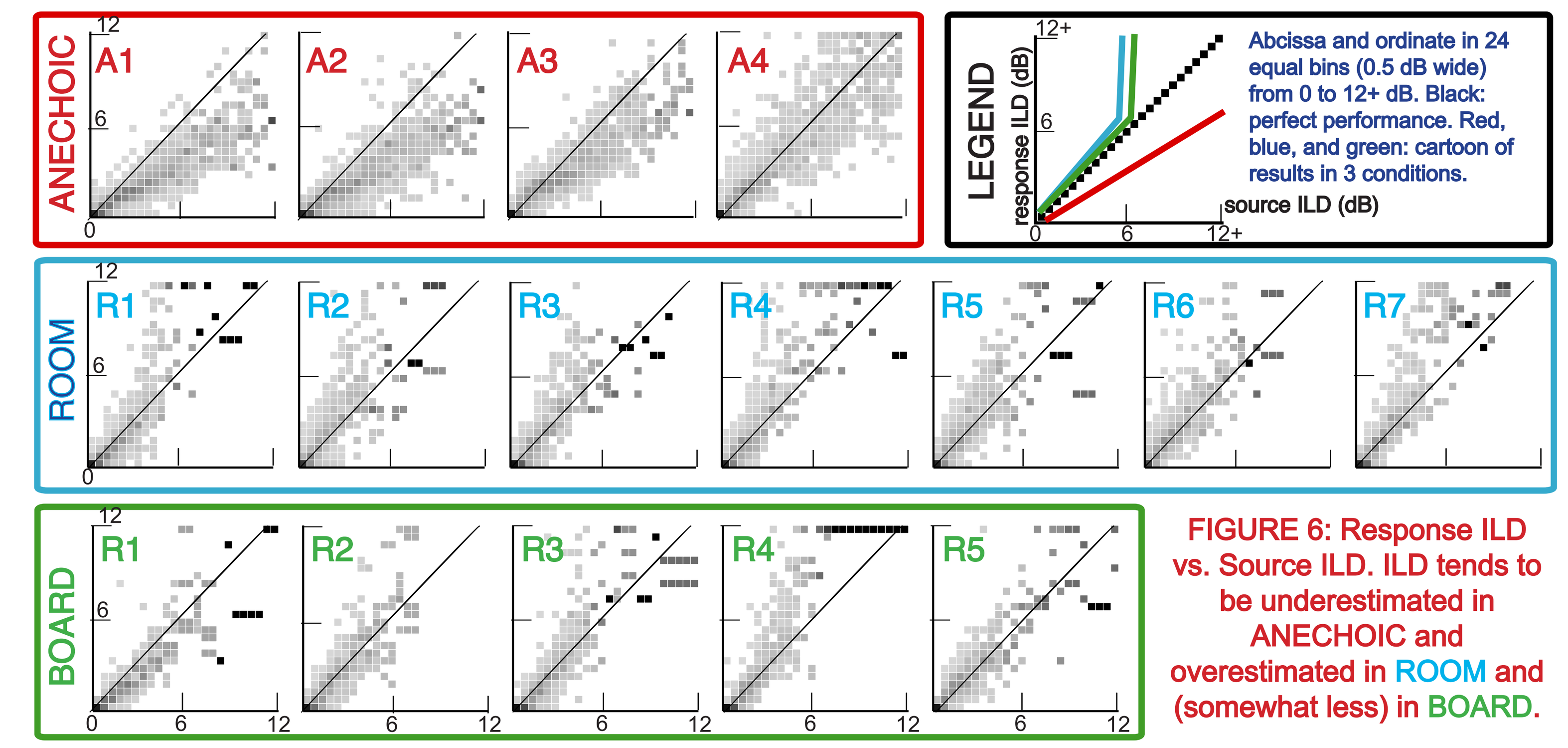

FIGURE 6: Response ILD vs. Source ILD. ILD tends to be underestimated in ANECHOIC and overestimated in ROOM and (somewhat less) in BOARD.

8. RESULTS/DISCUSSION: DISTANCE

ILD cues are weak near the median plane (i.e., the rate of change of ILD with postiion is small). Brungart (1998) found that ANECHOIC subjects are poor at judging source distance when sources are near the median plane. This pattern implies that ANECHOIC subjects attend ILD cues and that their performance suffers without this cue (near the median plane).

In the current ROOM and BOARD ILD results, subjects are not far off in response ILD when source ILD is small (generally, when sources are far from the head), but overestimate ILD when source ILD is large. This might occur if 1) subjects do not use ILD *per se*, but use some other cue for judging distance (e.g., the direct-to-reverberant energy ratio; see Mershon et al., 1989; Butler et al., 1980), and 2) this other cue is inaccurate for sources near the head (subject judgements are biased for close sources).

In these experiments, overall level at the ears was roughly equalized and then roved an additional 15 dB to ensure that direct sound level was not a useful cue. As a result, the absolute power emitted from near sources was less than from far sources. Thus, for near sources in the reverberant room, reflective energy might fall near threshold, effectively creating an infinite direct-to-reverberant energy ratio. If this ratio was used to judge distance (and ILD were ignored), ROOM and BOARD subjects (unlike ANECHOIC subjects) would tend to underestimate distance of near sources (see FIGURES 6 & 7), but their distance accuracy would be independent of source azimuth.

Figure 8 shows the strength of the correlation between source and response distance for the three conditions. The top panel [a) side] examines sources within 7.5 deg of the interaural axis. The bottom panel [b) front/back] examines sources within 7.5 deg of the median plane.

For sources to the side (where ILD is a useful cue), all subjects are consistent in distance judgements (high correlations). For sources near the median plane, performance of ANECHOIC subjects deteriorates to chance, but performance of ROOM and BOARD subjects does not change.

These results indicated that subjects in the reverberant conditions do not rely on ILD (which should be a robust binaural cue), but ignore it in favor of some other cue, possibly the direct-to-reverberant energy ratio.

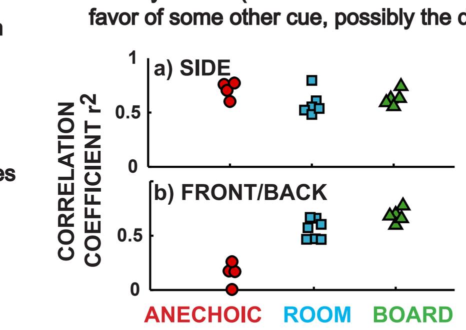

FIGURE 8: Correlation between source and response distance for sources a) to the side and b) near the median plane. Loss of ILD cues in ANECHOIC results in low correlation, but not in ROOM and BOARD.

9. LEARNING

Absolute errors in all dimensions tended to be smaller in the BOARD condition than in the ROOM (summary data not shown). To see if these changes were due to learning, a 2-way ANOVA analysis compared unsigned errors in ITD, ILD, and angle on the first and last 100 (of 1000) trials in each condition (with subject as a second factor).

In the ROOM condition, significant decreases were found in all three errors ($p < 0.005$) between the first and last 100 trials. For all cues, subject errors varied significantly. The interaction of these two factors was significant for all three cues (i.e., the decrease in error varied from subject to subject).

In the BOARD condition, there was no significant change between the first and last 100 trials for any cues. Subject differences were significant for all cues. The interaction term was not significant for any of the cues.

Finally, comparison of the last 100 ROOM trials with the first 100 BOARD trials showed no significant effect of condition for the ILD or angle cues. ITD errors increased significantly (but slightly) with the addition of the BOARD. Subject differences were once again significant, but the interaction between subject and condition were not significant for any of the cues.

Taken together, these results indicate that 1) subjects continue to improve on the task even after 200 practice trials (i.e., during the ROOM condition), 2) there is only a slight deleterious effect of adding the wallboard, primarily on ITD judgements, and 3) subjects vary in their abilities to use different

10. CONCLUSIONS

Analysis of localization performance using a perceptual-based coordinate system gives insight into what cues are important.

Subjects in all conditions use ITD cues consistently.

In reverberant conditions, what should be reliable binaural cues (ILD cues) in the near field are ignored in favor of some other cue(s), like the ratio of direct to reverberant energy. As a result, subjects in ROOM and WALL conditions condition show large localization bias for near sources (in ILD units) compared to subjects in the ANECHOIC condition. However, distance judgements are more uniform as a function of azimuth for the ROOM and WALL conditions than for the ANECHOIC condition.

Subject abilities vary greatly in all dimensions.

In general, the addition of a strong reflection that arrives close in time to the initial wavefront may cause 1) an immediate slight decrease in ITD performance, 2) idiosyncratic changes in circle of confusion judgements (e.g., R4), and 3) little effect on ILD judgements.

Subject performance continues to improve over hundreds of trials when listening and responding in a reverberant environment.

11. REFERENCES

Brungart, D. S. (1998). Near-field auditory localization. Unpublished Ph.D. dissertation, Electrical Engineering and Computer Science. Cambridge, MA, Massachusetts Institute of Technology.

Butler, R., E. Levy, & W. Neff (1980). "Apparent distance of sounds recorded in echoic and anechoic chambers." Journal of Experimental Psychology: Human Perception and Performance 6: 745-750.

Duda, R. O. (1997). Elevation dependence of the interaural transfer function. Binaural and Spatial Hearing in Real and Virtual Environments. R. Gilkey and T. Anderson. New York, Erlbaum: 49-76.

Duda, R. O. & W. L. Martens (1998). "Range dependence of the response of a spherical head model." Journal of the Acoustical Society of America 104(5): 3048-3058.

Mershon, D. H., W. L. Ballenger, et al. (1989). "Effects of room reflectance and background noise on perceived auditory distance." Perception 18: 403-416.