Learning of reverberation cues for auditory distance perception in rooms

Norbert Kopčo[1], Pierre Silvera[2], Konstantin Tskhay[2], Beáta Tomoriová[1], and Aaron Seitz[2]

[1]Technical University of Košice, Košice, Slovakia [2]University of California, Riverside

1. BACKGROUND: LEARNING AND DISTANCE PERCEPTION

Auditory perception of distance is not well understood (Zahorik et al, 2005)

For familiar sounds, overall received sound pressure level (loudness) considered to be the main distance cue (Warren, 1999)

In rooms, reverberation provides distance information. Candidate cue Direct-to-Reverberant energy ratio, D/R (Bronkhorst and Houtgast, 1999).

Amount of reflected energy varies from room to room. Auditory system has to adapt in each room to correctly map D/R to source distance.

In rooms, there is a learning effect: distance performance improves with experience (Shinn-Cunningham, 2000)

Improvement occurs over course of days, suggesting that memory consolidation process occurs (Lechner et al., 1999)

Learning process can be disrupted on a short-term scale, e.g., if inconsistent D/R cues are presented (Schoolmaster et al., 2004)

2. CURRENT STUDY

Study spontaneous learning (i.e., no feedback) of distance perception in a specific room: Does room learning occur when level cue is available?

Measure distance perception in a room over several days

Two run types, differing by distance cues available in sounds:

- A: overall presentation Sound Pressure Level (Amplitude) fixed
- R: overall presentation Sound Pressure Level (Amplitude) roved from trial to trial.

Four listener groups, differing by
- test run order (always interleaved) RARARA or ARARAR
- training session order: 3 sessions/days of A followed by 3 sessions of R, or vice versa.

All sources in front of listener.

Seven sessions, performed on different days.

HYPOTHESES

H1: R-test performance will be initially worse than A-test performance. R-training, but not A-training, will improve R-test performance.

H2: R-training will also improve A-test performance.

H3: There will be short-term adaptation effects: RA-test group results will differ from AR-test group.

H4: Learning will be mainly in form of consolidation, occuring across sessions, not within session.

3. METHODS
EXPERIMENTAL PROCEDURES

29 normal hearing subjects, divided into 4 groups

Source Stimuli
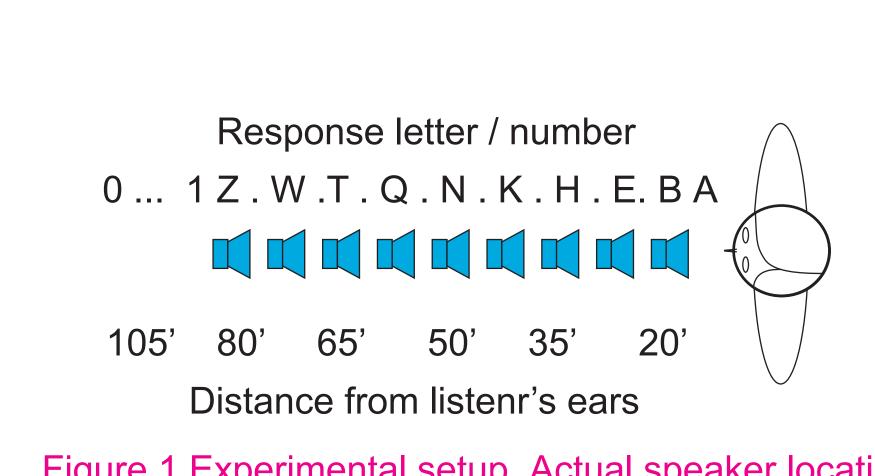- 500-ms-long, broadband noise burst

Figure 1 Experimental setup. Actual speaker locations and the letters/numbers (A-Z,1-0) used by listeners to indicate perceived distance. The nearest speaker was not used to present stimuli.

- A stimuli:
  - fixed presentation level, received level 49 - 54 dB SPLA
- R stimuli:
  - received level equalized and roved by +/- 12 dB

Source Locations (see Figure 1)
- Eight distances (nearest of 9 speakers not used)

Room
- small empty room, hard walls, carpeted floors, ceiling tails, background noise level 32 dB SPLA

One run
- subject informed about stimulus condition (A or R)
- 80 trials, each speaker used 10 times in random order
- subject indicated heard position by pressing corresponding letter/number on computer keyboard
- at end of each run, subject informed about his/her performance

Experiment
- one practice session at the end of which criterion

performance had to be achieved
- 7 experimental sessions (Fig. 2)
- session consisted of 12 runs

Testing
- first half of sessions 1, 4 and whole session 7
- R-test and A-test runs interleaved in a fixed order throughout experiment

Training
- sessions 1 (second half), 2, 3 used one type of runs (A or R)
- sessions 4 (second half), 5, 6 used the other type of runs (R or A)

Subject groups
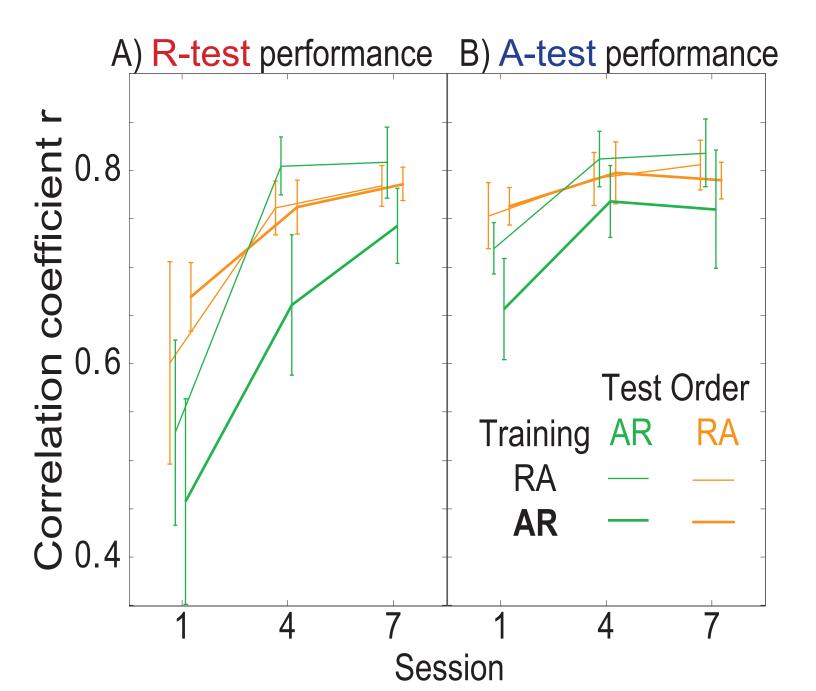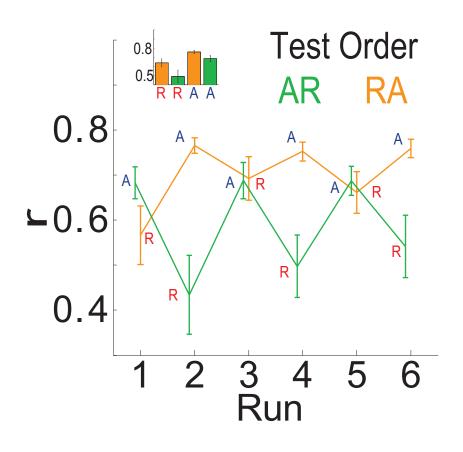- 4 groups, all combinations of:
- testing run order fixed to AR or RA, and
- training session order fixed to "sessions 1-3 R, sessions 4-6: A" or "sessions 1-3 A, sessions 4-6: R"

DATA ANALYSIS

Computed correlation coefficient r between log(response distance) & log(simulated distance) within each run.

Unless specified otherwise, graphs show across-subject means (+ standard error of the mean).

Figure 2 Ordering of test runs and training runs across sessions. Subject groups differed by order of test runs (Test = repeated pairs of RA or AR, fixed throughout experiment) and by order of training sessions (Train1=A, Train2=R) or (Train1=R, Train2=A).

4. RESULTS - TESTING PERFORMANCE

Raw test results (Figure 3):

Session 1
- R-test performance worse than A-test performance
- large inter-subject variability within groups

Session 4
- R-test performance improved
  - more than A-test performance (panel A vs B)
  - more due to R- than to A-training only in the AR-test group, equally for R- and A-training in RA-test group
  - A-test performance improved slightly for all groups

Session 7
- differences between groups and within groups very small

Initial (pre-training) session (Figure 4A):
- AR-test group worse than RA-test group
- difference between R-test and A-test larger for AR group

Final session (Figure 4B):
- difference between R-test and A-test
  - largely eliminated,
  - independent of training order

Figure 3. Performance in R-test sessions (panel A) and A-test sessions (panel B) as a function of test session number.
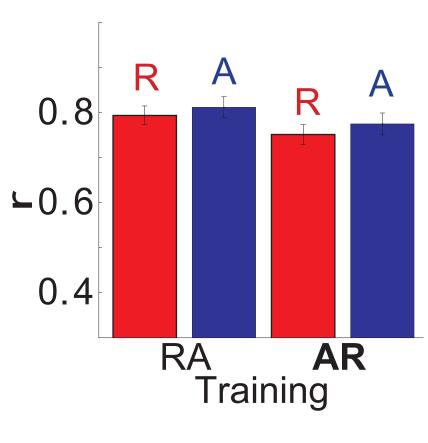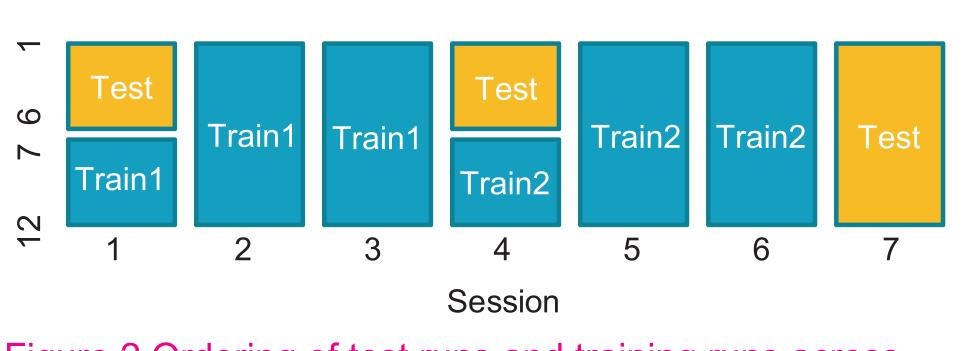

Figure 4. Detailed analysis of initial and final test performance. A) Pre-training performance in the first 6 test runs of session 1, collapsed across training order groups. Inset bar graph shows across-run average. B) Final session performance (average of runs 1-12).

> Distance judgments in rooms improve over time with both A and R types of training. Training on the reverberation distance cue is advantageous for distance judgements when overall level cue is eliminated. However, short-term factors (ordering of test runs) within initial session can eliminate the advantage.

5. RESULTS - TRAINING PERFORMANCE

Temporal profile of learning during the first three training sessions (Figure 5)

Learning on R-runs (Figure 5A):
- no improvement from session 1 to 2
- slight improvement in session 3
- large improvement in session 4

Learning on A-runs (Figure 5B):
- large improvement from session 1 to 2
- no improvement afterwards

Temporal profile of learning within training sessions (Figure 6)
- averaged across all subjects and all training sessions of the same type, no consistent temporal change within R or A sessions
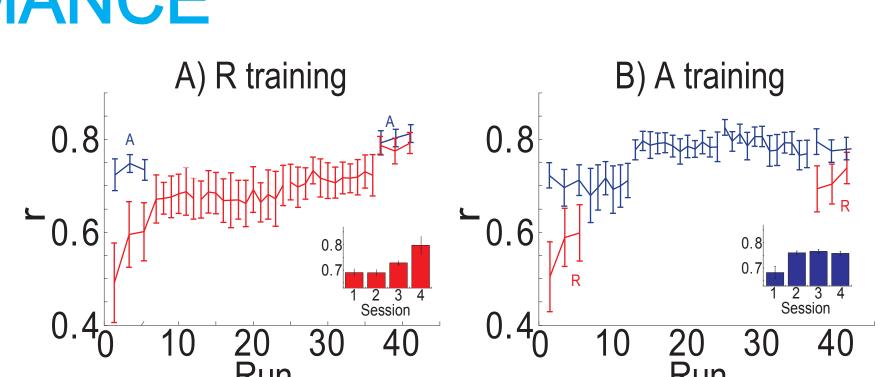

Figure 5. Temporal profile of performance during training runs in sessions 1-3 (and test runs of session 4). A) Performance of RA-training subjects on R-training runs. B) Performance of AR-training subjects on A-training runs. Inset bar graphs show across-session average performance. First three and last three bars show performance of immediately preceding/following test runs (and also performance during interleaved test runs of the opposite type).

> Almost no learning observed during training sessions.
> Most learning occurred between sessions, suggesting consolidation processes.
> In A-training, consolidation was fast, occurring after first session.
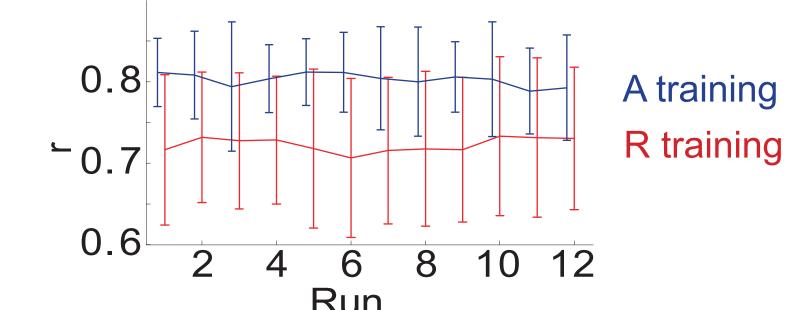> In R-training, main consolidation occurred later, after third session.

Figure 6. Temporal profile of performance during full-training sessions (2,3,5,6). Performance as a function of run number within a session averaged across all subjects.

6. CONCLUSIONS AND DISCUSSION

Auditory distance judgements in a room improve over days of no-feedback task performance.

Room learning likely contributes to improvement, both when overall level cue is available (A) and when it is not (R) available (H1 confirmed only partially).

Training on the room-related cue does not bring additional improvement in performance with the level cue available (contrary to H2).

Initial performance (as well as the learning effect) is critically influenced by whether the experiment started with level cue available (A) or not (R), confirming H3.

Specifically, if no level cue is available initially (R-test), listeners immediately improve their R-performance. If the level cue is initially available (A-test), three sessions are required to learn to use the room cue.

Alternatively, some cue switching mechanism might be involved. E.g., initial A testing might block the usage of R-cue. The blockage is overcome only after 3 days.

Consolidation is the main form of learning observed (confirming H4). Consolidation is
- fast when level cue available
- slow when only room reverberation cue available.

Discussion

These results might be particular to training without feedback:
- most likely it is (just) an improvement in consistency of cue-to-response mapping (absolute errors might grow),
- results are likely to differ if feedback is provided.

Procedural learning might have influenced performance.

Future studies will examine whether the learning is room-specific and other its properties.

7. REFERENCES

Bronkhorst, AW, Houtgast T (1999): Auditory distance perception in rooms. Nature 397, 517–520.

Lechner, H.A., Squire, L.R. and Byrne, J.H. 100 years of consolidation - remembering Müller and Pilzecker. Learning & Memory 6:77-87, 1999.

Shinn-Cunningham, BG (2000). "Learning reverberation: Considerations for spatial auditory displays," in Proceedings of the International Conference on Auditory Display, Atlanta, GA, 2-5 April 2000, 126-134.

Schoolmaster, M, N Kopčo, and BG Shinn-Cunningham (2004). "Auditory distance perception in fixed and varying simulated acoustic environments," J. Acoust. Soc. Am., (Presented at th 147th meeting of ASA, New York, NY)

Warren, R.M. (1999) Auditory Perception: A New Analysis and Synthesis. New York: Cambridge University Press.

Zahorik, P., Brungart, D.S., and Bronkhorst, A.W. (2005) Auditory Distance Perception in Humans: A Summary of Past and Present Research. Acta Acustica united with Acustica, No. 3, pp. 409-420.