Effects of preceding exposure on distance perception and distance learning in varying and fixed virtual environments

Myroslav Fedorenko[1,2] and Norbert Kopčo[1]

[1]Institute of Computer Science, P. J. Šafárik University, Košice, Slovakia

[2]M. Dragomanov Ukrainian State University, Kyiv, Ukraine

Background: Previous research showed that listeners calibrate to the acoustic environment when judging distance in real reverberant rooms, resulting in gradually improving performance even without explicit training. In contrast, similar enhancements are not observed in anechoic environments, suggesting that the improvements are due to tuning to specific reverberation-related cues in a given room. The mechanisms underlying this "room learning" are not well understood. A previous study [Schoolmaster, Kopčo, & Shinn-Cunningham, J Acoust Soc Am 113, 2285, 2003], examined 1) how consistency in simulated room presentation affects performance, and 2) whether previous exposure to consistent vs. inconsistent simulation influences the ability to interpret the distance cues in different rooms. It showed that distance perception is more accurate when the simulated environment is consistent than randomly chosen from trial to trial. Here, we further analyze the data to examine how they depend on type of simulated room (anechoic vs. reverberant) and the direction of stimulus presentation (frontal vs. lateral). And, we develop a model that proposes how subjects combine available distance cues in different contexts.

1. Introduction

Level-independent distance perception possible for nearby sources(< 1 m)

- in **anechoic** space: at locations **away from midline**, using interaural level difference (**ILD**; e.g., Brungart et al., 1998)
- in **reverberation**: in **all directions**, using reverberation-related cues like direct-to-reverberant energy ratio (**DRR** + ILD; e.g., Kopco et al., 2012)

Weights used by listeners to combine DRR/ILD/other cues depend on **context** of previously presented stimuli (Doreswamy et al., 2019).

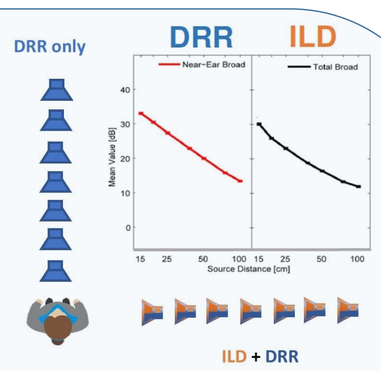


Fig.1.1. Direction of ILD and DRR cues.

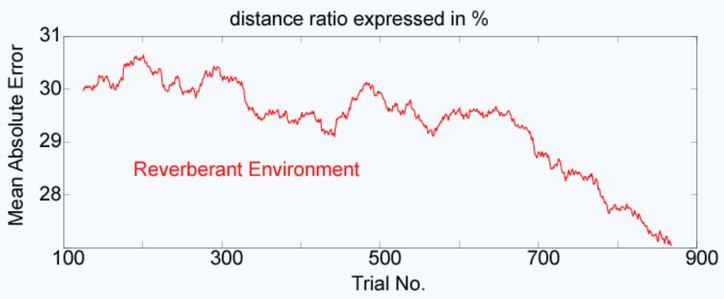


Fig.1.2. Distance perception in reverberant environment

In **reverberation** (but not in anechoic space), distance perception improves **spontaneously, without feedback or** any **training**, just by listeners actively performing the task in sessions with duration of **several hours** (even if split over multiple days). (Shinn-Cunningham, 2000; Santarelli, 2000)

This **spontaneous learning** in a fixed room can strongly depend on **availability of cues** (e.g., level vs. DRR), especially **during initial exposure** to a given room (Hladek et al., 2013).

2. Current study

In **virtual** and **mixed reality**, the presented **environments** can **change rapidly**.

- ☆ How does **consistency of simulated environment** affect **distance perception** and the **spontaneous learning** processes?
- ☆ How does **varying** the environment from **trial to trial** (vs keeping the environment consistent) influence distance perception? E.g. when listeners performs the task in 3 different virtual environments: Will they be able to concurrently maintain/tune to **3 separate model rooms**, or will they create **1 combined model**?
- ☆ Does **initial exposure to in/consistent rooms** affect performance in both consistent and inconsistent contexts? E.g., if starting in consistent rooms means that listeners will learn characteristics of each room, will it transfer to better performance in inconsistent rooms in a later session?
- ☆ Is distance perception and spontaneous learning influenced by the **early reflections** when listener is near **the corner** of a room?

3. Experiment in Virtual Environment

Stimuli:
- five 150-ms-long pink noise bursts (30-ms gaps);
- roved by 15 dB (to eliminate level cue);
- 9 distances (15 to 170 cm, log spaced);
- 2 directions (medial and lateral);

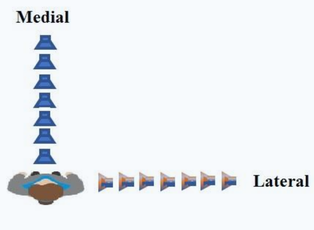
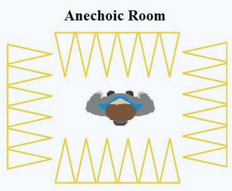


Fig.3.1. Sound direction.

Room conditions:
- 3 virtual environments simulated using **individually** measured **BRIRs**;
- **anechoic, center, and corner** of a midsize classroom;



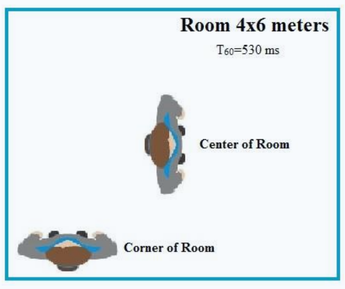


Fig.3.2. Room conditions: Anechoic, Corner and Center of the room.

One trial:
- subject **informed** about **room condition**;
- simulated source **presented** over headphones;
- subject **indicated heard position** by a mouse click on screen.

Each subject performed two blocks (**contexts**): **FIXED** and **MIXED**:
- block consisting of 6 **sessions**, each containing 8 **runs**;
- each **run** had 45 trials which **held direction fixed**, only varying distance;
- **FIXED** sessions: simulated room fixed within a **block**;
- **MIXED** sessions: simulated room **selected randomly on each trial**.

Two subject groups:
- initFixed group (4 subjects): **FIXED** block followed by **MIXED**;
- initMixed group (4 subjects): **MIXED** block followed by **FIXED**.

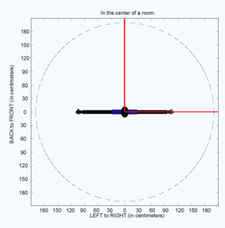


Fig.3.3. Screen shots from the experiment. Subject used a mouse to click perceived location.

4. Group starting with FIXED context

Fig.4.1. Across-subject mean and std. error in perceived distance as a function of simulated distance for each room condition and source for the initFixed group.

Fixed Room Context

InitFixed Group

Block 1: room FIXED within a block.

Level-independent distance perception:
- better for lateral than medial sources;
- better in room (Center/Corner) than anechoic;
- slightly better in Center vs Corner.

Nearby targets overestimated.

Distant targets underestimated in Anechoic, overestimated in Corner (edge effect), accurately judged in Center.

Mixed (vs. Fixed) Room Context

InitFixed Group

Block 2: room varying from trial-to-trial in block

Worse performance in all rooms & directions.

Bias induced by **mixed context** in all rooms, independent of direction:
- in Anechoic, responses shifted closer,
- in Corner and Center, responses shifted further away.

5. Group starting with MIXED context

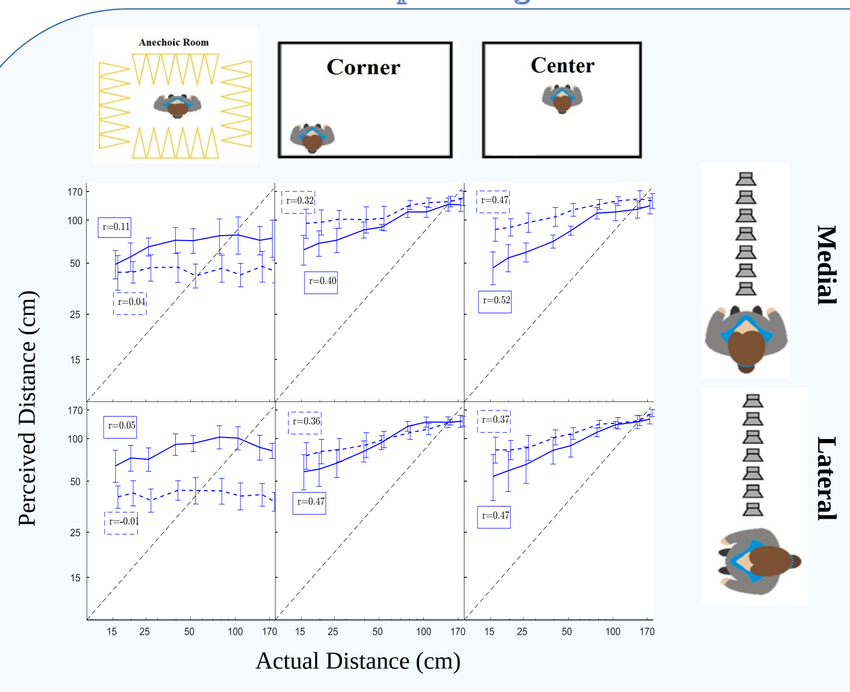


Fig.5.1. Across-subject mean and std. error in perceived distance as a function of simulated distance for each room condition and source for the initMixed group.

initMixed Group

Block 1: MIXED
Block 2: FIXED

Overall performance worse, specially for nearby & lateral sources

Effect of Mixed vs Fixed context:
- similar to initFixed gr.;
- weaker mainly because the Fixed condition is worse.

Bias effects not visible in corr. coef. *r* (e.g., Anechoic)

Again:
Anechoic < Corner < Center

6. Summary of Results using Correlation Coefficient *r*

Performance tends to be better for:
- initFixed group;
- Fixed context;
- lateral direction;
- room (Center>Corner>Anecho).

Complex dependences between factors (4-way interaction: p = 0.034):

- initMixed group: no effect of direction

- initFixed group: effect of context (Mixed – Fixed):
  - varies with room and direction
  - is largest for **Corner Medial**

Not only the current context, but also initial/preceding context affects performance.

Early reflections modulate effect of initial context for med sources.

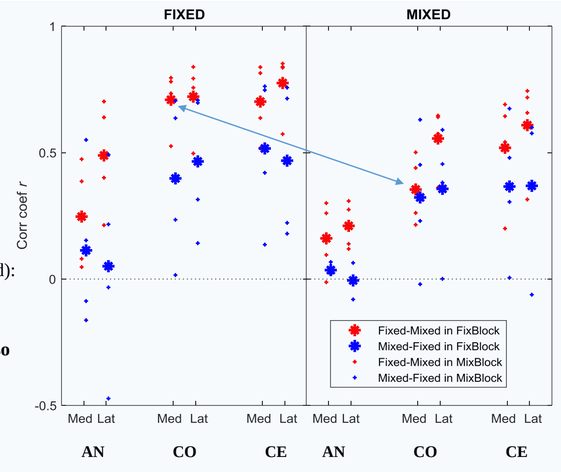


Fig.6.1. Correlation coefficient between log10 of actual and log10 of responded distance.

7. Discussion and Conclusions

1. Fixed rooms:

   Performance better for room than Anechoic and, sometimes, for lateral than medial sources. → **Both DRR and ILD cues are used by listeners when available.**

   Performance slightly worse in Corner than Center.
   → **Early reflections in Corner are detrimental for distance judgments.**

2. Context effect:

   Mixing rooms from trial induces biases: underestimation in Anechoic and overestimation in Center/Corner. → **Listeners cannot separately process distance information from different rooms on trial-by-trial basis.**
   Biases consistent with listeners creating a single DRR-to-distance mapping in Mixed context, since in such 1 combined room model:
   - Anechoic ~ very large DRR → percepts biased closer,
   - Center/Corner ~ smaller DRR → percepts biased away from listener.

3. Initial/Preceding context:

   Starting in **Mixed** context tends to cause more deterioration re. **starting Fixed**.
   However, the effect is complex:

   - initMixed group performed equally for lateral and medial sources.
     → **If starting in a Mixed context, listeners did not benefit from ILD cue for lateral sources in the Mixed or in the Fixed context, even though in the Fixed session (performed as 2nd) all the cues were consistent.**

   - initFixed group can benefit from ILD, but the effect of context (Mixed vs Fixed) varied with room and direction (largest for Corner Medial).
     → **How the cues are combined and weighted depends on the current context, the initial context, which cues (ILD/DRR) are available, as well as on early reflections.**

   Factors that determine these complex interactions need further examination.

References

- Kolarik AJ, Moore BC, Zahorik P, Cirstea S, Pardhan S. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. Atten Percept Psychophys. 2016 Feb;78(2):373-95. doi: 10.3758/s13414-015-1015-1. PMID: 26590050; PMCID: PMC4744263.
- D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source.," *J. Audio Eng. Soc.*, vol. 49, pp. 904–916, 2001.
- J. M. Kates, K. H. Arehart, R. K. Muralimanohar, and K. Sommerfeldt, "Externalization of remote microphone signals using a structural binaural model of the head and pinna," *J. Acoust. Soc. Am.*, vol. 143, no. 5, pp. 2666–2677, May 2018, doi: 10.1121/1.5032326.
- W. M. Hartmann and A. Wittenberg, "On the externalization of sound images," *J. Acoust. Soc. Am.*, vol. 99, no. 6, pp. 3678–3688, 1996.
- A. Kulkarni and H. S. Colburn, "Role of spectral detail in sound-source localization," *Nature*, vol. 396, no. 24/31 December, pp. 747–749, 1998.
- J. Catic, S. Santurette, and J. Buchholtz, "The effect of interaural-level-difference fluctuations on the externalization of sound," *J Acoust Soc Am*, vol. 134, no. 2, pp. 1232–1241, 2013.
- P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, "Auditory distance perception in humans: A summary of past and present research," *ACTA Acust. UNITED WITH Acust.*, vol. 91, no. 3, pp. 409–420, 2005.
- B. G. Shinn-Cunningham, S. Santarelli, and N. Kopco, "Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural vs. monaural cues," in *23rd mid-Winter meeting of the Association for Research in Otolaryngology*, 2000, p. 88.
- A. W. Bronkhorst and T. Houtgast, "Auditory distance perception in rooms," *Nature*, vol. 397, no. 11 February, pp. 517–520, 1999.
- P. Zahorik, "Estimating sound source distance with and without vision," *Optom. Vis. Sci.*, vol. 78, no. 5, pp. 270–275, 2001.
- Barb Shinn-Cunningham and Matt Schoolmaster contributed to data collection and initial analysis. (Schoolmaster, Kopčo, & Shinn-Cunningham, J Acoust Soc Am 113, 2285, 2003)

Perception and Cognition Laboratory

PCL.UPJS.SK

[Work supported by Horizon Europe HORIZON-MSCA-2022-SE-01 grant N° 101129903 and VEGA 1/0350/22]