What does the ability to suppress a single reflection tell us about localisation performance in rooms?

Stefan Kerber[1], Bernhard U. Seeber[1]

[1] *MRC Institute of Hearing Research, University Park, Nottingham, NG7 2RD,*
{s.kerber; b.seeber}@ihr.mrc.ac.uk

Introduction

Localisation in rooms is characterised by the presence of multiple reflections interfering with the direct sound from source to listener. Nevertheless, due to the "precedence effect", normal hearing listeners are able to cope in such situations by apparently ignoring the later arriving reflections. In studies on the precedence effect this situation is abstracted by playing a sound (the "lead") and a single delayed reflection (the "lag") from different directions (for review see [1] and [2]). In realistic situations, however, multiple early reflections as well as late reflections might interfere with the direct sound.

In the current study two experiments were carried out: Experiment 1, the "localisation dominance task", quantified how auditory localisation is affected by presenting a single reflection additionally to a target sound. Participants were asked to indicate the perceived direction of a given lead-lag pair. In Experiment 2, room reflection patterns for the target sounds were calculated with room-simulation software. Target sounds were then played with their reflections at different direct-to-reverberant ratios and participants were asked to localise them. The aim was to compare to what degree localisation performance in rooms can be predicted from the simple lead-lag precedence effect paradigm. This will also help to answer the question to which degree later reflections affect localisation performance in real rooms.

Experiments

General methods

Both experiments were conducted in the free-field using the „Simulated Open Field Environment (SOFE)", a setup consisting of 48 loudspeakers in an anechoic room. The setup offers 36 loudspeakers in the horizontal plane, spaced in 10° steps. Additionally, twelve loudspeakers are installed at elevations of ±40°. All loudspeakers were digitally equalized to ±1.5 dB between 250 and 12000 Hz. A screen covers loudspeakers in the front, left and right of the participant and allows for video projection. A detailed description of the setup can be found in [3]. In the experiments participants were asked to position a light spot projected on the screen to the perceived direction of a sound using a trackball [4]. Target sounds were pulse trains (10 ms on, 120 ms off, 780 ms long, 2 ms Gaussian ramps on each pulse) cut out of a uniform exciting noise of 60 dB SPL. So far four normal-hearing persons (median age 31 years) participated in the experiments.

Exp. 1: Localisation dominance task

In the localisation dominance task the target sound (the „lead") was presented at +40°. It was followed by a delayed and scaled copy of itself from -40° (the „lag"). Six delays between 0 and 48 ms were administered. Scaling was done so that the level of the lag was between -10 and +10 dB relative to the lead. When participants perceived only one sound, e.g. at short delays, they pointed to it with the light spot. When lead and lag were heard separately, participants were instructed to point either to the most-left or most-right of the sounds. Eight trials were collected for each delay, lag level and instruction from each participant.

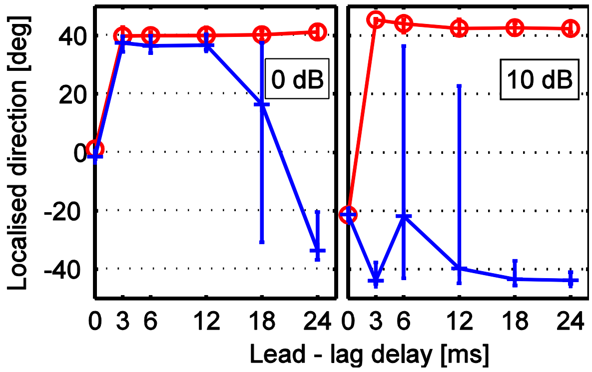

Figure 1: Results of one participant in the localisation dominance task presented as medians and quartiles as a function of lead-lag delay. Red circles depict trials where the instruction was to point to the lead, blue pluses where it was to the lag. Relative level of the lag was 0 dB (left panel) and +10 dB (right panel).

Figure 1 gives results from one participant as medians and interquartiles for relative levels of the lag of 0 and 10 dB. For equal level of lead and lag and small delay times only one sound close to the lag (+40°) is perceived. Increasing the delay leads to a split, so that for larger delays lead and lag are heard separately close to their actual positions. A split also occurs for a relative lag level of +10 dB at all delays but zero. However, high variability at short delays shows that even at this high lag level sounds were sometimes heard as a single sound at the lead. This general pattern of results holds for all participants with some variance in the level where the complete split occurs. This variance was also observed by others ([5]).

Exp. 2: Localisation in simulated rooms

Reflections of a "living room" were simulated using the mirror-image source method. Sources and reflections were mapped to the loudspeakers of the SOFE, thereby reproducing the spatio-temporal characteristics of the room reflection pattern. Using this capability, target sounds were presented in a simulated living room from eleven directions between -80° and 80°. The distance between sources and listener was fixed at 1.26 m. Again, participants indicated the origin of the sounds using the light pointer method. Reverberation was altered by changing the direct-to-reverberant ratios (DRRs) of the simulated rooms between -18 and 0 dB. This was done by keeping the direct sound part in the simulated room impulse responses constant whilst scaling the reverberant part. Following this, overall levels of the different room impulse responses were adjusted. This approach has the advantage that temporal and spatial reflection patterns were identical across the different DRRs ("rooms"). This method for calculating different DRRs also implies that the difference in DRR between two rooms matches the difference in level of their respective first reflections.

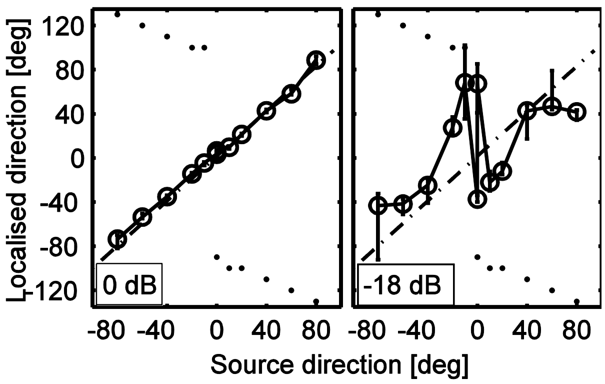

Figure 2: Medians and interquartiles for the localised direction vs. the real direction in simulated reverberation for one participant. The dashed line depicts optimum performance. Black dots give the direction of the first reflections.

Figure 2 shows results for localisation in a simulated room for the same participant as before. For a DRR of 0 dB (left panel) localisation is hardly affected by the reverb. Median responses are close to the target indicated by a dashed line. A linear fit shows a regression slope of 0.98 and the correlation between target and localised position is highly significant. Lowering the DRR to -18 dB (right panel) severely affects localisation performance: The median responses of the participant are shifted towards the direction of the first reflection (indicated by black dots), and the size of the interquartile ranges increases by a factor of six. This general pattern again holds for all participants with some variance in the DRR for which degradation starts. Note that this variation is smaller than the one observed in Experiment 1.

Comparison of results between experiments

Root-mean-square (rms) errors were calculated to compare results between both experiments (Figure 3). For the localisation dominance task this was done for all lag levels at a delay time of 3 ms and under the assumption that correct localisation would be at the direction of the lead (+40°). The delay of 3 ms was chosen because it matches approximately the delay of the first reflection in the room impulse responses. RMS-errors for localisation in rooms were calculated over all test directions.

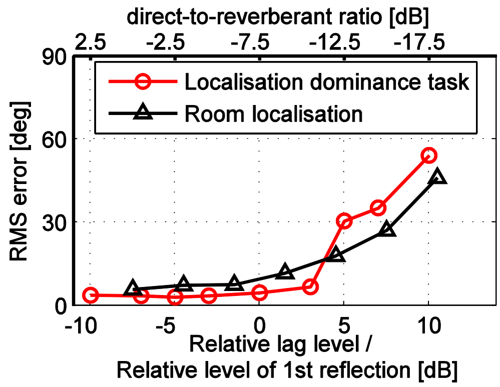

Figure 3: RMS-error in both experiments as a function of relative level of the first reflection (bottom axis) for one participant. The corresponding DRRs in Exp. 2, are given on the top axis.

Generally, rms-errors calculated for both experiments match well. Performance starts to decline at similar levels of the first reflection and rms-errors increase at a similar rate. This shows the high impact of the first reflection on localisation performance in rooms. However, for this participant, the rms-error for localisation in reverb is slightly lower than its respective counterpart in Exp. 1 for higher lag levels. This is surprising, because increasing the energy in the reflections by adding later reflections should make localisation more difficult and consequently lead to an increase in error. Also, it should lead to a decrease in correlation of the overall sound, thus making the extraction of binaural cues more difficult.

Summary and Acknowledgements

Our results so far suggest that the results of a simple lead-lag localisation task follow those of localisation in simulated rooms. A pronounced impact of the first reflection seems to be present with further research needed to clarify the exact role of later parts of the room impulse response on sound localisation.

This study was funded by Cochlear Europe Ltd. and the intramural programme of the MRC.

References

[1] Litovsky, R. Y.; Colburn, H. S.; Yost, W. A.; Guzman, S. J.: The precedence effect. J. Acoust. Soc. Am. 106(1999), 1633-1654.

[2] Blauert, J.: Spatial Hearing. The Psychophysics of Human Sound Localization. The MIT Press (1997).

[3] Seeber, B. U.; Kerber, S.; Hafter, E.: A system to simulate and reproduce audio-visual environments for spatial hearing research. Hear. Res. 260 (2010), 1-10.

[4] Seeber, B. U.: A New Method for Localization Studies. Acta Acustica – Acustica 88 (2002), 446-450.

[5] Saberi, K.; Antonio, J. V.: Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval. J. Acoust. Soc. Am. 114 (2003), 420-429.