# Effect of spectral content and learning on auditory distance perception

**Norbert Kopčo[1,2], Dávid Čeljuska[1], Miroslav Puszta[1], Michal Raček[1] a Martin Sarnovský[1]**

[1]Department of Cybernetics and AI, Technical University, Košice, Slovakia
[2]Hearing Research Center and Dept. of Cognitive and Neural Systems, Boston University, Boston, MA, USA
email: kopco@bu.edu

*Abstract: We studied the importance of the stimulus spectrum and of the listener's experience on the human ability to judge distance of auditory objects. The study was performed in a virtual auditory environment (VAE), created by convolution of the stimuli with non-individual Head-Related Transfer Functions (HRTFs) measured in a regular classroom. The listeners' task was to judge the distance of sounds originating at distances of 0.15 to 2 meters directly ahead or to the right of the listener's head. Three types of stimuli were presented: lowpass-filtered, highpass-filtered, and broadband stimuli. Each listener performed four blocks of measurements, two on one day and two on a later day, with separation of at least 12 hours. No significant effect of spectrum was found, presumably because the non-individual HRTFs do not provide accurate cues to the listeners, making the task very difficult even with broadband stimuli. However, a systematic learning effect was observed in all listeners: while their performance did not change across the blocks performed on the same day, it improved between the blocks performed on different days. Thus, it can be hypothesized that a process similar to memory consolidation is observed that leads to improved performance and that requires a long break between the measurement sessions.*

*Keywords: auditory distance perception, virtual auditory environment, learning, memory consolidation*

## 1 Introduction

The human ability to judge the distance of sound sources depends on many factors. In real anechoic environment, humans can accurately judge the distance of nearby sounds (within 1 m of the listener) [1] but not of distant sounds. And, in this environment, the observed performance is better for medial sources (ahead of the listener) than for lateral sources (on the side of the listener). This pattern of performance is presumably observed because the main distance cue in the anechoic space is the interaural level difference (ILD). The ILD cue is available only for lateral sources [2] and it is frequency dependent, changing more with distance at high frequencies than at low frequencies. In reverberant space, distance perception is better than in anechoic space [3], and it is less influenced by the sound source azimuth. Therefore, it can be hypothesized that the effect of the stimulus spectrum on the accuracy of distance perception will also be weaker than in anechoic space.

Another effect of reverberation in distance perception is that the accuracy of distance judgments improves with experience, often over a course of hours and days of practice [4]. The exact details of how this effect occurs are not clear. However, it is known that the learning is general in the sense that performance improves even if the listeners' position in the room changes during the experiment [5]. Still, how exactly this learning occurs, what are the critical characteristics necessary for it to occur, and what is the time course of this learning, is not well understood.

The present study measured the accuracy of distance judgments as a function of the stimulus spectrum in two directions: ahead of the listener (medially) and to the right of the listener (laterally). It was expected that the dependence of performance on the stimulus spectrum will be observed for the lateral sources but not for the medial sources. Also, it was expected that some learning will be observed. To get some estimate of the time course of this learning effect, listeners performed two experimental sessions on two different days. Each session consisted from two measurement blocks performed with a short, half hour break between them. The main question tested by this arrangement was whether the amount of learning will be the same independent of the amount of break provided between the blocks, i.e., whether the within-session improvement in performance will be the same as the across-session improvement.

# 2 Experimental methods and procedures

Four male undergraduate students (all authors except NK) participated in the study. The subjects reported to have normal hearing.

The HRTFs used in this study were measured on a human listener (different from the participating subjects) in a regular empty quiet classroom 9×5 meters (for a description of the measurement procedures, see for example [6]) and they included the impulse response of the room. The simulated sound source positions were logarithmically spaced from 15 to 200 cms ahead or to the right of the listener (see Figure 1). The HRTFs were convolved with 300-ms noise bursts to generate the

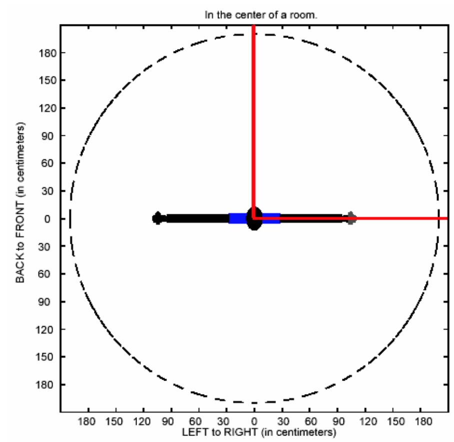


**Figure 1 Screen shot from the experiment shows a top view of a listener in the room. The red line marks the positions from which the sound source was simulated. Subject used a mouse to click on the perceived location.**

stimuli. The stimuli were bandpass-filtered to be lowpass (300-3000 Hz), highpass (3000-5700 Hz) or broadband (300-5700 Hz). The stimuli were RMS-normalized and for every presentation the level was randomly roved by ±5 dB to exclude the overall level as a distance cue. The pre-generated stimuli were played back via an Echo Darla 20 sound card to Sennheiser 570 circumaural headphones.

One trial (or measurement) consisted of the presentation of one stimulus, after which the listener indicated the perceived location of the stimulus on a computer screen.

The experiment was divided into runs, each run consisting of 45 trials with fixed stimulus spectrum and direction and with randomly changed stimulus distance. One block consisted of 6 such runs, two for each combination of azimuth and spectrum. This gives 270 trials per block and 5 trials per block for each combination of distance, azimuth and spectrum. Each subject performed four such blocks. The breaks between the blocks 1 and 2 and between the blocks 3 and 4 were approximately 30 minutes long, the break between blocks 2 and 3 was at least 12 hours long.

The data were analyzed by computing the square of the correlation coefficient ($r^2$) between the log of the actual distance and the log of the perceived distance.

# 3 Results

**Effect of signal spectrum and source azimuth**

First, the results are studied as a function of the source azimuth and spectrum. Figure 2 plots the individual subjects' correlation coefficients (small symbols) as

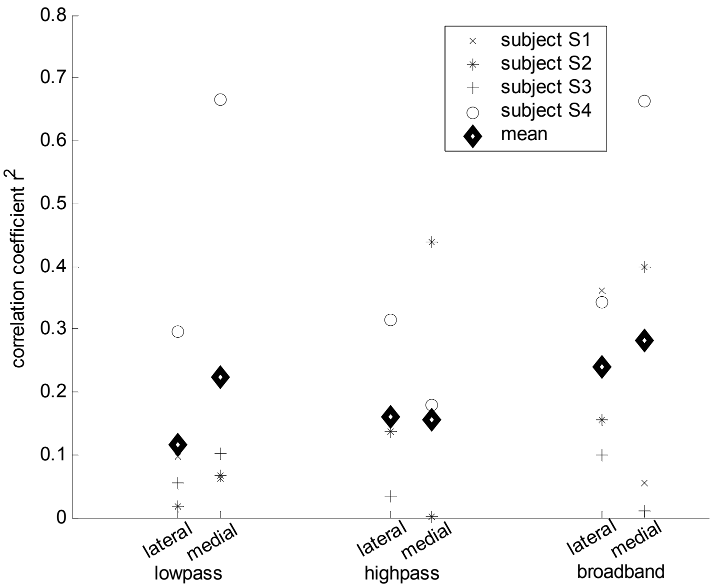


**Figure 2 Distance perception as a function of sound source azimuth and stimulus spectrum. Plotted is the square of the correlation coefficient ($r^2$) between the log of the actual distance and the log of the perceived distance. Individual data are shown by small symbols. Thick symbols are across-subject means. Data plotted separately for each combination of azimuth (lateral, medial) and spectrum (lowpass, highpass, broadband)**

well as the across-subject average of the $r^2$ (thick symbols) for all studied combinations of sound source azimuth and signal spectrum. Overall, the observed performance is very bad in all conditions, with the best average $r^2$ reaching only 0.3 (broadband medial condition), compared to $r^2$ of up to 0.7 in previous VAS studies [7] and $r^2$ of up to 0.9 in real echoic environment [3]. Therefore, it is not surprising that the effects of signal spectrum and of signal azimuth are very weak. The effect of signal spectrum follows the expected pattern (medial performance better than lateral) only in two out of three conditions (lowpass and broadband). The effect of spectrum is also in the expected direction, i.e., the broadband data are better than the lowpass or the highpass data. However, both these effects are negligible compared to the differences in performance among the subjects, which is much larger than in previous studies.

**Effect of experience**

Figure 3 plots the overall performance ($r^2$) of the subjects as a function of the experimental block and collapsed across the source azimuth and stimulus

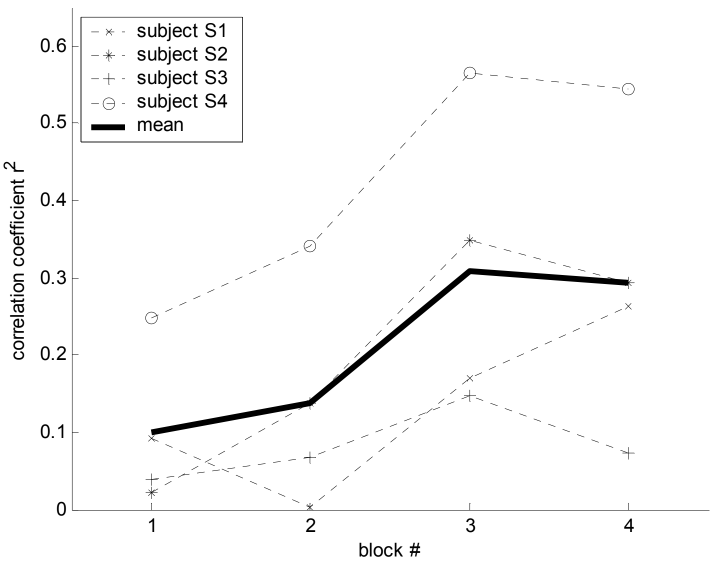


**Figure 3 Effect of experience on the accuracy of perceived distance. Plotted is the $r^2$ as a function of the experimental block. Subjects performed blocks 1 and 2 on one day and blocks 3 and 4 on another day. Individual data are shown by small symbols. Thick symbols are across-subject means.**

spectrum. The figure shows that each of the subjects improved their performance from the first to the last block (on average the increase in $r^2$ is approximately 0.15). Closer inspection shows that essentially all of this improvement happened between the blocks 2 and 3 and that the improvement between the blocks number 1 and 2 or 3 and 4 was small or none. This trend is visible in all the subjects, although the across-subject differences in individual measurements are still as large as in the previous analysis (Figure 2). The Student's t-test shows that the difference between the mean performance in the 2nd and the 3rd block is statistically significant ($t(3)=-5.18$, $p=0.0139$).

# 4 Discussion and conclusion

The results of this study do not support the hypothesis that auditory distance perception in reverberant environments is strongly dependent on the spectral content of the stimulus. However, the main reason for this observation might be in that non-individual HRTFs were used to create virtual auditory environment. Also, the observed large inter-subject differences might be a consequence of the fact that the chosen HRTFs fitted more some subjects and less other subjects. However, even for the best-performing subject (S1, symbols “o” in Figure 2) there is no clear effect of the stimulus spectrum, suggesting that no effect of stimulus spectrum would be observed even if the simulation was better than in this experiment. In order to draw reliable conclusions, the experiment must be replicated using individually measured HRTFs to improve the quality of the simulated environment.

On the other hand, the learning effect observed in this study is strong and reliable. The most interesting aspect of the learning effect observed is that it occurs only if the listener takes a long break (more than 12 hours, including sleep) between the blocks, while no learning is observed when the break is only 0.5 hour. In the previous studies that reported this kind of learning effect, the listeners participated in the experiment over a period of several days and weeks. Therefore, the current results suggest that long pauses between sessions are a critical aspect of this learning. The observed temporal profile of learning is reminiscent of that of the memory mechanism of consolidation, which is associated with the transfer of knowledge from short-term to long-term memory [8]. Therefore, it can be hypothesized that the learning mechanisms observed in this study are similar to the mechanisms of memory consolidation.

However, it is still not clear what gets learned in this process or what acoustic characteristics are important for distance perception in general.

# 5 References


[1] Brungardt, DS (1998) "Near-Field Auditory Localization". Unpublished Ph.D. Thesis, MIT

[2] Shinn-Cunningham, BG, SG Santarelli, and N Kopčo (2000) "Tori of confusion: Binaural cues for sources within reach of a listener," J Acoust Soc Am, 107(3), 1627-1636, 2000.

[3] Santarelli, S. (2001) "Auditory Localization of Nearby Sources in Anechoic and Reverberant Environments" Ph.D. Thesis, Boston University

[4] 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.

[5] Kopčo, N. (2003) "Spatial hearing, auditory sensitivity and pattern recognition in noisy environments," Unpublished Ph.D. Thesis, Boston University

[6] Kopčo N and BG Shinn-Cunningham (2003) Spatial unmasking of nearby pure-tone targets in a simulated anechoic environment, J Acoust

Soc Am, in press.
[7] Shinn-Cunningham, B.G., Santarelli, S., and Kopčo, N. (2000). "Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural versus monaural cues." Presented at the MidWinter meeting of the Association for Research in Otolaryngology, St. Petersburg, Florida
[8] Squire LR "Memory and Brain". Oxford University Press, 1987.