# Perception of scrambled reflections

Fritz Menzer[1], Berhard U. Seeber[1]
[1] *Audio Information Processing, Technische Universität München, Deutschland, Email: fritz.menzer@tum.de, seeber@tum.de*

## Introduction

While sound reflections are an essential feature of the reverberation caused by rooms, they are usually not perceived individually, and rather contribute to an overall spatial impression. From a physical point of view, conclusions regarding the shape and size of a room can be drawn from the timing and directions of arrival of individual reflections. However, the mechanisms employed by the human auditory system for interpreting the reverberation characteristics are not known, and the question arises whether the auditory system uses a geometrical representation of the room in this process.

Within the context of the precedence effect, many studies have shown that the human brain is able to suppress simple reflections, i.e. echoes arriving before the so-called echo threshold are not heard as distinct auditory events. This echo threshold can be increased by a repeated exposure to sound containing reflections [1], meaning that the auditory system adapts to the reflection pattern. Many studies related to the precedence effect have been performed using very simple stimuli where only a direct sound and a single reflection is present and where the source signal consists of clicks or short noise bursts. However, as was shown by Djelani and Blauert [2], precedence effect experiments can also be performed using speech as a source signal and using more complex reflection patterns. Nevertheless, precedence effect experiments usually rely on an unnatural change in the reflection pattern, such as changing the direction of arrival of a reflection by 90°, or mirroring a room about the median plane of the listener [2].

In this study, a different paradigm was used, where a realistic scene is presented and the test subjects have to detect natural changes in the reflection patterns. Using this paradigm, two cases are compared: in the first case, a room was simulated using the image source model and head-related impulse responses corresponding to the directions of arrival of the individual reflections. In the second case, the directions of arrival of the reflections were randomly modified to make a geometrical interpretation impossible.

Compared to other studies, here a very detailed simulation of a room was used (using an image source model simulating tens of thousands of reflections) and instead of clicks, speech was used as the source signal, i.e. the signals used in this study are realistic with respect to both the modeled acoustic environment and the naturalness of the source signal.

The primary goal of this study was to examine if any differences in the capability to detect changes in the reflection pattern can be observed when the directions of arrival of the reflections are scrambled such that the resulting reflection pattern would be extremely unlikely to occur in reality, compared to the geometrically correct case. The scrambling process was designed in a way that it maintains reflection timings, T60 and DRR.

In the following sections, the experimental setup and the method for generating the stimuli are explained, followed by information about the test subjects. Preliminary results for the main experiment and a pilot study for a discrimination experiment, as well as the conclusion section are found at the end of this paper.

## Experimental Setup

The experiments presented in this paper were conducted in a sound-proof booth using headphone-based virtual acoustics. Binaural room impulse responses were generated using the image source model [3] implemented in the Simulated Open Field Environment [4], which was configured to use an HRTF set selected individually for each subject from the AUDIS catalogue [6] using a subjective selection method [5]. Source signals were convolved with the simulated binaural room impulse responses and were played back at a sampling rate of 44100 Hz using Sennheiser HD600 headphones. The test subjects gave feedback by pressing keys on a computer keyboard (space for the detection experiment and up/down arrow keys for the discrimination experiment).

## Method Overview

In the two experiments presented here, test subjects listened to a speech signal convolved with a binaural room impulse response (BRIR). In the main experiment, the BRIR was replaced by another BRIR after a randomly chosen time between 3 and 6 sec. The subjects had to indicate that they heard a change in the reverberation characteristics of the room by pressing a key within 1.5 sec from the change of the impulse response.

The simulation of the BRIRs was performed by using the image source model [3] implemented in SOFE [4] for a rectangular 5 m-by-9 m-by-4.5 m room and image sources within a 200 m-by-216 m-by-180 m cube, a listener position (1.500 m,1.800 m,1.600 m) and and 6 different source positions illustrated in Figure 1. While in the first part of the stimulus, the source position was always position 1, in the second part it was any of the positions 1 to 6 (a repetition of position 1 means corresponds to a “no change” condition).

The source signal used to generate the stimuli was taken from track 50 of the EBU SQAM CD [7]. It consists of

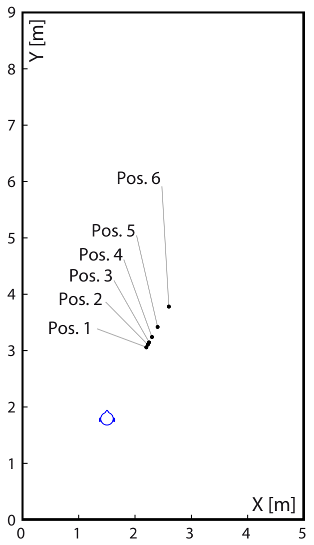

**Figure 1:** Source and listener positions used in this experiment. The simulated room was 5 m wide and 9 m long, with walls corresponding the the rectangular frame of the illustration.

a recording of the sentence "In the course of a December tour in Yorkshire, I rode for a long distance in one of the public coaches, on the day preceding Christmas" spoken by a male speaker. This sentence was split at a randomly chosen point between 3 and 6 sec after its beginning. The two resulting parts overlap by 2 ms with a sinusoidal crossfade, as illustrated in Figure 2. The first part was convolved with a binaural room impulse response for position 1, while the second part was convolved with a binaural room impulse response for one of the positions 1 to 6. The signal was stopped 1.5 sec after the split point, i.e. the signal presented to the subject finished at the same time as the valid time interval for giving the response. This method saved time and also gave the subjects a feedback on when a change that was missed had happened.

## Impulse Response Scrambling

To create physically implausible room impulse responses, scrambled versions of the binaural room impulse responses were obtained by changing the directions of arrival of the individual reflections in the impulse response by randomly chosen angles. This means that the timing of the reflections, as well as their overall energy, is conserved. The binaural room impulse response is obtained by convolving the individual reflection responses (containing the wall absorption filters, etc.) by different head-related impulse responses than for the geometrically correct version. This procedure preserves the reverberation time T60 and the direct-to-reverberant ratio. For each reflection, the angle of arrival was always modified by the same amount across all 6 different positions used in this experiment. From one position to another the difference in the impulse response was comparable to the difference in the geometrically correct simulation.

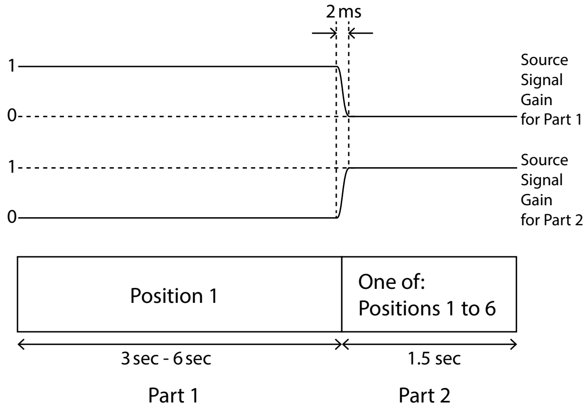

**Figure 2:** Envelopes used for the stimuli generation. By multplication with the two envelopes shown above, a speech signal was split into two parts overlapping by 2 ms with a sinusoidal crossfade. The first part was convolved with a binaural room impulse response corresponding to position 1, whereas the second part was convolved with a room impulse response for one of the positions 1 to 6. The stimulus was obtained by adding the two convolved signals.

## Tested Conditions and Stimulus Presentation

In total, 24 conditions were tested: 6 positions for the second part of the stimulus, each with all four combinations of geometrically correct and scrambled image source positions and normal and inversed channel assignments for the binaural output signal (where the inversed channel assignment corresponds to a situation where the room and source positions have been mirrored about the median plane of the listener).

The generated signals were divided into blocks of 23.2 ms (1024 samples at 44100 Hz sampling rate) and played back using playrec (http://www.playrec.co.uk/). Each time playrec indicated that a block had finished playing, the last key pressed on the keyboard was polled and in case the space bar was pressed during a 1.5 sec time window after the change occurred, the playback was stopped and this trial was considered a success (i.e. the change was detected). If the subject did not press the space bar, or pressed it before the change happened, the trial was considered a failure (no change detected). The trials where no change occurred were used only to determine a "false alarm" rate.

The experiment was split into 10 blocks, each of which took approximately 2 min to complete. Out of the 24 conditions, some conditions were tested only once every two blocks, leaving 18 trials per block. In total, the experiment consists of 180 trials. At the end of each block the test subjects could choose to either start the next block or to take a break. Most subjects split the experiment into two sessions of 5 blocks, separated by a short break. Before starting the test, the subjects could familiarize with the stimuli and the task by doing 2 blocks of 18 trials identical to those in the actual listening test. The results of these 2 blocks were discarded.

## Subjects

The preliminary results shown here are from six subjects (5 male, 1 female, aged 24-49) which participated in this experiment. All reported normal hearing.

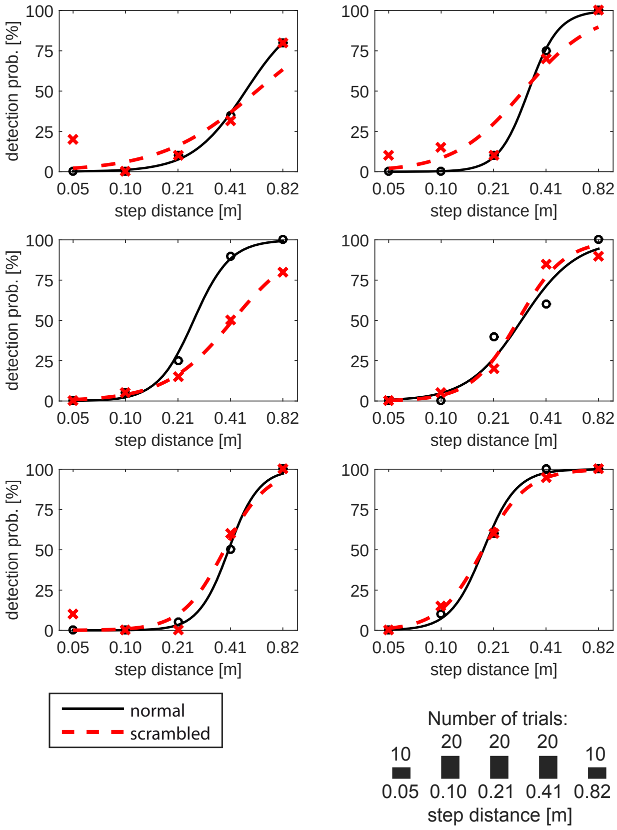

**Figure 3:** Individual results for 6 subjects. Black circles mark measured results for the geometrically correct condition, red crosses mark measured results for the scrambled condition. Black lines and red dashed lines show the fitted psychometric functions for the goemetrically correct and scrambled conditions, respectively.

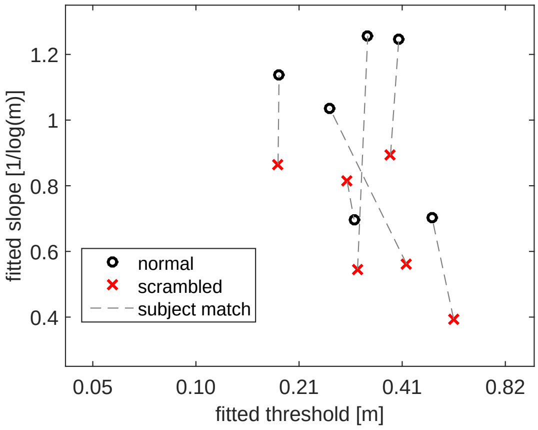

**Figure 4:** Threshold vs. slope of fitted psychometric functions for the different subjects.

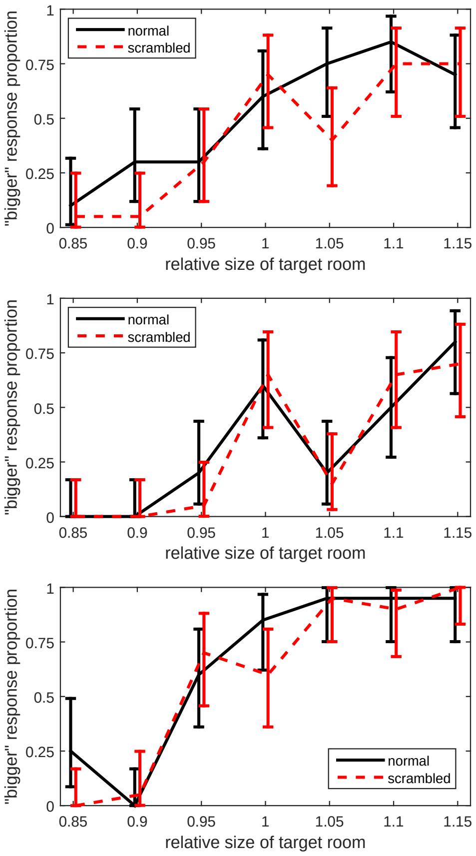

**Figure 5:** Results of the pilot study for a room size discrimination experiment. The horizontal axis represents the relative room size and the vertical axis the likelyhood to perceive the target room to be bigger.

## Results

Preliminary results are shown in Figure 3. The probability of detecting a change in the reverberation characteristic is shown as a function of the position of the source in the second part of the stimulus. For the statistical analysis, the trials for both normal and inversed channel assignments are merged, resulting in 10 to 20 trials per condition. For each subject, logistic psychometric functions were fitted and the parameters of these functions (slope and 50% threshold) are shown in Figure 4. While no siginificant difference in the thresholds could be observed, the slopes are significantly flatter in the scrambled condition (a one-sided T-test shows that the difference is significant on a 2% level).

## Discrimination Experiment Pilot Study

In order to study the observed differences between geometrically correct and scrambled rooms from a different angle, and in the hope to enable experiment paradigms that prove the existence of a build-up effect, a discrimination experiment was set up where the subjects had to characterize a change in room size. The results of the pilot study for this experiment are shown in Figure 5. As no clear differences between the geometrically correct and the scrambled conditions could be observed and fitting a psychometric function did not seem appropriate given the data from the pilot study, this experiment paradigm was not further pursued. A possible explanation for these results is that a change in room size may be detected by observing only the timing of reflections: if the room size increases, all reflections arrive later (and vice versa a smaller room leads to earlier reflections). Since the scrambling does not change the reflection timing, the similar performance for geometrically correct and scrambled conditions can be explained. Furthermore, as Djelani and Blauert [2] have noted already, extended speech signals are not suitable for testing build-up effects.

## Conclusions

An experiment was conducted where subjects had to detect changes in reverberation patterns in two conditions: a room simulated using geometrically correct image source positions and a room where the image source positions have been scrambled, i.e. randomly rotated around the center of the listener's head. For the scrambled condition, flatter psychometric functions were observed than for the geometrically correct condition. A discrimination experiment using a similar paradigm where changes in the reverberation pattern due to changes in room size had to be discriminated did not lead to conclusive results, which may be explained by the fact that scrambing does not affect reflection timings.

## Acknowledgement

This study was funded by BMBF project 01 GQ 1004A.

## References

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

[2] T. Djelani, J. Blauert: Investigations into the build-up and breakdown of the precedence effect. acta acustica / ACUSTICA **87** (2001) 253-261.

[3] J.B. Allen, D.A. Berkley. Image method for efficiently simulating small room acoustics. The Journal of the Acoustical Society of America **65(4)** (1979) 943-950.

[4] B. U. Seeber, S. Kerber, E. R. Hafter: A system to simulate and reproduce audio-visual environments for spatial hearing research. Hearing Research **260** (2010) 1-10.

[5] B.U. Seeber, H. Fastl: Subjective selection of non-individual head-related transfer functions. Proceedings of the 2003 International Conference on Auditory Display (2003) 259-262.

[6] J. Blauert et al.: The AUDIS catalog of human HRTFs. 16th International Congress of Acoustics (1998).

[7] European Broadcasting Union: Sound Quality Assessment Material recordings for subjective tests - Users' handbook for the EBU SQAM CD. Geneva (2008).