Perceptual Differences for Modifications of the Elevation of Early Room Reflections

(article starts on next page)

Perceptual Differences for Modifications of the Elevation of Early Room Reflections

Leon Müller and Jens Ahrens

Chalmers University of Technology, Gothenburg, Sweden

Correspondence should be addressed to Leon Müller (`leon.mueller@chalmers.se`)

ABSTRACT

Acoustic room responses usually comprise components that propagate in non-horizontal directions. Oftentimes, audio capture and reproduction systems are not capable of maintaining such elevation information reliably hence it is important to understand their perceptual significance when auralizing rooms. This work investigates the ability of the human hearing system to distinguish between early reflections with different elevation angles by performing loudspeaker- and headphone-based listening experiments using manipulated spatial room impulse responses. The results show that changing the elevation of a strong early reflection can lead to clearly perceivable differences and factors that influence the detectability are identified. Projecting all elevated reflections of a spatial room impulse response with no very prominent ceiling reflection to the horizontal plane showed no perceivable differences.

1 Introduction

Loudspeaker-based audio reproduction setups are many times confined to the horizontal plane. When used for auralization of rooms, this has the consequence that the elevation angle of non-horizontally propagating components of the room response is altered. Similarly, recent developments in spatial audio recording techniques such as the equatorial array [1], which achieves simplification of the hardware setup at the price of projecting any captured sound field onto the horizontal plane, raise the question of what the perceptual significance of such alterations of the elevation angles is.

In the last decades, several authors investigated the perceptual effects of isolated ceiling reflections on pa-

rameters such as localization accuracy [2], auditory envelopment [3] as well as timbral [4] and spatial aspects [5]. Recently, publications such as [6] and related work by the same authors evaluated different perceptual aspects of vertical reflections based on three-dimensional loudspeaker reproduction systems. Thereby only cases with and without added ceiling reflection were compared, and it was not evaluated how a projection of the elevated reflection onto the horizontal plane affects the auditory impression. To the authors' best knowledge, none of the previously performed studies actually investigated the effects of altering the elevation of an early room reflection, which motivates the experiments described in the following. This paper is based on the work previously presented in a Master's thesis [7].
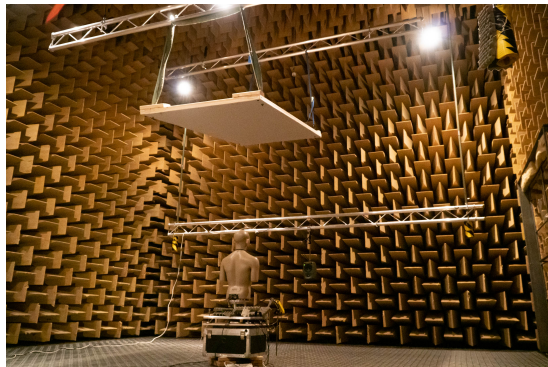
Fig. 1: Ceiling reflection measurement setup

Fig. 2: Loudspeaker setup in anechoic chamber with artificial head at participant position and smartphone mounted in front of participant

2 Methods

In order to investigate the perceptual consequences of a modification of the elevation of early reflections, spatial room impulse responses (SRIRs) of two rooms were measured, modified by adding a strong ceiling reflection, and evaluated in a listening experiment as described in the following.

Obtaining stimuli that only differ in strength and incidence angle of a single elevated reflection can be achieved by physically placing a reflecting panel at different positions in a real room and measuring the resulting spatial impulse responses. However, it was decided to instead construct such sound fields by combining measured spatial impulse responses of real rooms without pronounced ceiling reflections with a measured impulse response (IR) of a single isolated ceiling reflection. This way, the strength and timing of this extra reflection can be modified in the rendering process without requiring additional measurements.

2.1 Measurements

Spatial room impulse responses of the two different rooms *Big Hall* (T30$_{1kHz}$ = 1.43 s) and *Listening Lab* (T30$_{1kHz}$ = 0.1 s) were measured using a six-element star-shaped open microphone array with a radius of 50 mm as described in [8] and decomposed using the spatial decomposition method (SDM) [9]. Thereby, the time difference between the individual microphone signals is used to estimate the direction of arrival (DOA) for each sample of the measured signal. These DOAs are then smoothed by applying a low-pass filter and

quantized to a grid of incidence angles. This quantization step can be considered as assigning the components of the SRIR to different virtual loudspeaker positions. This processing was implemented by using a modified version of the BinauralSDM toolbox [10] with an analysis window length of 36 samples at a sampling rate of 48 kHz, a smoothing window of 16 samples and a nearest neighbor interpolation for the DOA quantization to the virtual loudspeaker positions.

For this project, the measured SRIRs were quantized to eight virtual loudspeaker positions on the horizontal plane (Table 1, Speaker Nr. 1-8) and one top speaker at 90° elevation (Table 1, Speaker Nr. 9). Additionally, a separate version of the measured SRIRs without top speaker i.e. with all room reflections projected to the horizontal plane was rendered in order to investigate the perceptional differences between a purely horizontal rendering with no elevated reflections and a rendering with an additional elevated height speaker.

In order to measure the impulse response of an isolated ceiling reflection, a single 2.40 m x 1.20 m gypsum panel was mounted in an anechoic chamber as shown in Figure 1. The plate was centered 1.3 m above the receiver position with an upwards tilt of 12°, which results in a ceiling reflection with an elevation angle of approximately 84°. The measured IR was truncated and windowed to cut off the direct path, and the original delay between direct and reflected sound (ca. 5 ms ≈ 1.7 m path-length difference) was preserved by zero padding the truncated reflection IR.

Table 1: Loudspeaker positions in anechoic chamber

Speaker Nr.	Azimuth	Elevation	r
1 (Center)	0°	0°	1.03 m
2	315°	0°	1.10 m
3	270°	0°	1.03 m
4	225°	0°	1.10 m
5	180°	0°	1.03 m
6	135°	0°	1.10 m
7	90°	0°	1.03 m
8	45°	0°	1.10 m
9 (Top)	0°	90°	1.09 m
10	0°	10°	1.05 m
11	0°	45°	1.26 m
12	90°	45°	1.26 m

2.2 Reproduction

While headphone-based reproduction methods can be considered as the most relevant for spatial audio applications they often rely on generic head-related transfer functions (HRTFs) which can result in a less accurate localization of elevated sound sources [11]. Therefore, it was decided to perform listening experiments using both a head tracked headphone-based reproduction method as well as a multichannel loudspeaker setup in an anechoic chamber. This allows for investigating the situation with both individual as well as non-individual HRTFs. Both reproduction methods are described in the following.

For the loudspeaker-based experiment, an array of 12 loudspeakers was set up in an anechoic chamber as shown in Figure 2 with speaker coordinates according to Table 1. Each speaker was calibrated in amplitude and delay for a listening position in the center of the array. In order to combine the measured SRIRs and the measured isolated ceiling reflection, the signals obtained from convolving the measured SRIRs with a source signal were played back via the loudspeakers according to the SRIR's virtual loudspeaker positions and the reflection signal was added to the loudspeaker that corresponds to the desired reflection incidence direction for each step in the experiment.

For the headphone-based experiment, the multichannel audio signals used for the loudspeaker-based experiment were dynamically convolved with non-individualized HRTFs of a KEMAR artificial head according to the loudspeaker positions and the participants' current head orientation. These binaural signals
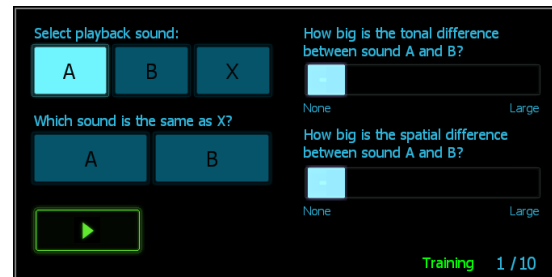
Fig. 3: Experiment user interface

were then played back via equalized AKG K-702 headphones. The overall playback volume was set to a comfortable level of approximately 64 dBA for both experiments, and the participants were instructed to avoid excessive head movements while listening to the stimuli.

2.3 Listening Experiment

2.3.1 Paradigm

In order to evaluate if participants perceive any differences between two stimuli with different reflection elevation angles, an ABX test was performed. This means that the participants listened to three sounds A, B and X of which sound X was equal to either sound A or sound B, and the task was to determine whether sound A or sound B corresponds to sound X. During the experiment, the participants had the possibility to seamlessly switch between the three sounds A, B and X, which allows for detecting even small audible differences. In addition to the ABX test, the participants reported the perceived spatial and tonal difference between sounds A and B on a continuous scale ranging from *none* to *large* for each comparison.

Each participant performed a training composed of 10 sets of stimuli to familiarize with the interface and the range of stimuli. The experiment itself comprised 50 sets of stimuli present in random order.

2.3.2 Stimuli

In order to limit the overall duration of the listening experiment, not all possible combinations of source signals, measured rooms and reflection angles were evaluated. Instead, an informal pilot experiment was performed in order to determine which comparisons

Table 2: Scenarios evaluated in listening experiment

Description	SRIR	SRIR Quantization Speaker Nrs.	Signal	Added Reflection Levels (dB)	Added Reflection Positions (Azimuth, Elevation)
Drums Without Room	-	-	Drums	-3, -6, -9, -12	$(0°, 0°)$ vs. $(0°, 10°)$
				-3, -6, -9, -12	$(0°, 0°)$ vs. $(0°, 45°)$
				-3, -6, -9, -12	$(0°, 0°)$ vs. $(90°, 45°)$
Speech Without Room	-	-	Speech	-3, -6	$(0°, 0°)$ vs. $(0°, 10°)$
				0, -3, -6	$(0°, 0°)$ vs. $(0°, 45°)$
				0, -3, -6	$(0°, 0°)$ vs. $(90°, 45°)$
Drums in Big Hall	Big Hall	1-9	Drums	0, -3, -6, -9	$(0°, 0°)$ vs. $(0°, 45°)$
				0, -3, -6, -9	$(0°, 0°)$ vs. $(90°, 45°)$
Drums in Listening Lab	Listening Lab	1-9	Drums	-3, -6, -9	$(0°, 0°)$ vs. $(0°, 45°)$
				-3, -6, -9	$(0°, 0°)$ vs. $(90°, 45°)$
Natural Rooms	Big Hall Listening Lab	1-8 vs. 1-9	Drums	-	-

might be most relevant for the actual listening experiment. This resulted in choosing the combinations listed in Table 2.

For the cases with a single elevated reflection in an otherwise anechoic environment (Table 2: Drums Without Room, Speech Without Room), the source signal was directly played back via the center speaker without convolving it with a measured SRIR and the reflection signal was played back via a single speaker corresponding to the desired reflection position.

For the scenarios of the measured rooms without any added ceiling reflection (Table 2: Natural Rooms), two different SRIR quantizations with and without a single top speaker at 90° elevation were compared to each other.

Previous research indicates that the choice of source signals influences the audibility of individual reflections as well as the binaural localization accuracy. Particularly, strong transients such as clicks or tone burst often produced a smaller localization blur than less transient stimuli such as sinusoids or speech [12]. Therefore, an anechoic drum recording was used as the primary source signal, additionally an anechoic speech recording was included for comparison. The level of the added ceiling reflection was varied between 0 dB and -12 dB relative to the signal played back via the center speaker which presented the direct sound component of the SRIRs, and reflections at 10° and 45° elevation as well as 0° and 90° azimuth were compared to their corresponding counterpart on the horizontal plane with 0° elevation.

2.3.3   Setup and User Interface

Since pilot measurements presented in [7] showed that a regular computer screen as well as a chair generate reflections that could significantly influence the perception of the presented stimuli, it was decided to make the participants stand in the anechoic chamber, and a small smartphone was mounted in front of the participants as user interface for the experiment as shown in Figure 2. The commercial app *Lemur* was used to implement a custom graphical user interface as shown in Figure 3 consisting of buttons for playback control and the ABX test as well as two sliders to report perceived spatial and tonal differences. Thereby, the participants were required to first listen to all stimuli and touch all GUI elements at least once before being able to answer. After each comparison, the GUI was reset to its default state in order to avoid user errors.

2.3.4   Participants

The loudspeaker-based experiment was performed by a group of 25 different participants consisting of students and faculty members of the division of applied acoustics as well as a small number of subjects without an academic background in acoustics. All of the participants reported to have normal hearing, 22 participants stated to have a background in acoustics and 12 participants claimed to have experience with critical listening in the context of spatial audio. The headphone-based experiment was performed by 13 participants which all reported to have normal hearing and an academic background in acoustics. Nine of these participants

stated to have experience with critical listening in the context of spatial audio. Ten of the 13 participants of the headphone-based experiment also performed the loudspeaker-based test before with at least a week of separation between both experiments.

2.3.5 Evaluation

In order to asses the results from the ABX test, the procedure described in [13] was used. Thereby, it is assumed that the participants will randomly choose either answer A or B whenever they do not perceive any difference between both sounds. This means that the number of correct identifications for a case with no perceivable difference follows a binomial distribution and the probability of randomly obtaining at least a certain number or correct identifications for a given number of trials can be estimated by calculating a *p*-value as described in [13]. From this *p*-value for a specific number of participants a 95 % confidence threshold can be derived. If the number of correct identifications exceeds this threshold, one can assume with 95% confidence that the participants heard a difference. However, it is important to notice that the ABX tests can only prove that there is an audible difference between stimuli A and B, it can not prove that no difference is perceivable.

For the loudspeaker-based experiment, this threshold lies at 17 correct identifications out of the total of 25 participant responses, for the headphone-based experiment at 9 correct identifications out of 13. This corresponds to 68 % respectively 69 % correct identifications. For convenience, it was decided to use 69 % as the 95 % confidence threshold percentages for both experiments in the following evaluation. Thereby it is important to keep in mind that even though the 95 % confidence threshold percentages of both experiments are very similar, the difference in the number of participants still leads to a lower statistical significance of the headphone-based experiment compared to the loudspeaker-based one.

In general, this ABX evaluation method only allows to conclude on the ability of the entire group of participants to differentiate between the two compared stimuli and not on the perception of individual subjects. Therefore, also the results from the perceived spatial and tonal difference slider values should be taken into account, which were evaluated without any normalization by calculating the arithmetic mean and 95 % confidence intervals from all participant responses for each condition.

3 Results

3.1 Elevation Changes on the Median Plane

Figure 4 shows the loudspeaker- and headphone-based listening test results for a comparison between a reflection with an elevation angle of 45° and 0°, respectively, at an azimuth angle of 0°. The orange and blue bars show the percentage of correct ABX identifications of each condition for the loudspeaker- and headphone-based experiment. The green areas mark the 95 % confidence region. Íf a bar reaches the green area, one can assume with at least 95 % confidence that the participants perceived a difference. Additionally, error bars show the mean perceived spatial and tonal differences as well as the corresponding 95 % confidence intervals for each condition. The error bars drawn on top of the blue bars represent the results from the loudspeaker-based experiment, the error bars drawn on top of the orange bars show the results from the headphone-based experiment.

Figure 4 (a) shows the results for the drum signal without added room, i.e. only direct sound and the ceiling reflection. Here, the results from the ABX test clearly indicate that the participants perceived a difference between the two reflection elevation angles down to a reflection level of -9 dB for the headphone-based experiment and -12 dB for the loudspeaker-based experiment. Additionally, the perceived spatial and tonal differences decrease with the reflection level.

Using speech as source signal for the same comparison did not result in clear ABX results as shown in Figure 4 (b) since only the loudspeaker-based experiment with a reflection level of -3 dB passed the 95 % confidence threshold. While the loudspeaker-based ABX results for 0 dB reflection level do not exceed the 95 % confidence threshold, the reported tonal and spatial differences are still relatively high. This can indicate that at least some participants did perceive a difference.

For the drum signal and the *Big Hall* SRIR, the ABX responses indicate perceivable differences for a loudspeaker-based reproduction for a reflection level of 0 dB and -3 dB as shown in Figure 4 (c). The ABX responses for the headphone-based reproduction are not conclusive for this scenario since they pass the 95 % confidence threshold both at 0 dB and -6 dB reflection level but not at -3 dB. Additionally, it can be observed that the reported spatial and tonal differences for this scenario are significantly higher in the
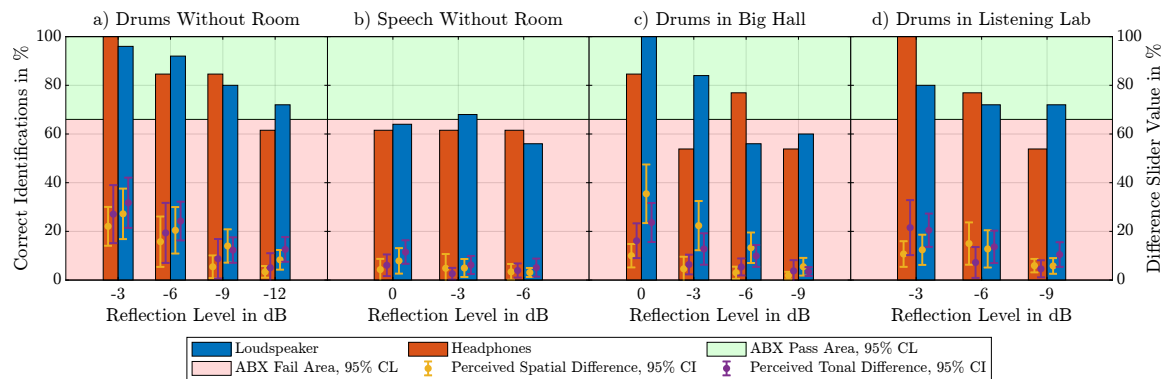
Fig. 4: Loudspeaker- and headphone-based listening test results for comparison between a reflection with an elevation angle of 45° and 0° at an azimuth angle of 0° with varying reflection levels, source stimuli and room SRIRs
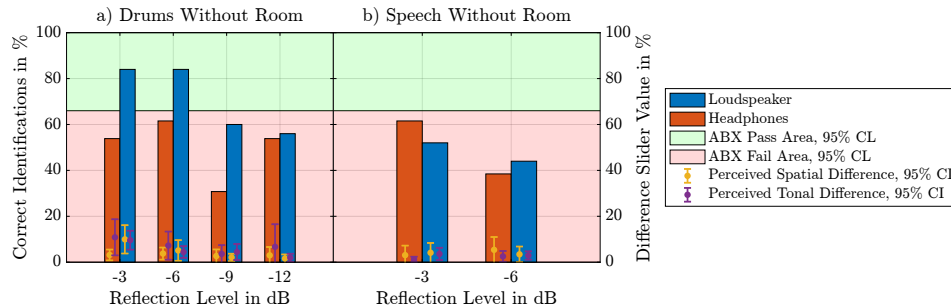

Fig. 5: Loudspeaker- and headphone-based listening test results for comparison between a reflection with an elevation angle of 10° and 0° at an azimuth angle of 0° with varying reflection levels and source stimuli without added room

loudspeaker-based experiment than in the headphone-based experiment.

The ABX responses for the drum signal with the *Listening Lab* SRIR indicate perceivable differences down to a ceiling reflection level of -9 dB for the loudspeaker-based experiment and -6 dB for the headphone-based experiment as shown in Figure 4 (d). Compared to the case with the *Big Hall* SRIR, the smaller *Listening Lab* with a shorter reverberation time hence seems to result in a lower reflection level detection threshold.

Comparing a reflection elevation angle of 10° to 0° with an azimuth angle of 0°, i.e. in the listener's look direction, as shown in Figure 5 leads to significantly reduced perceived differences compared to the previously described case of an elevation angle of 45°. The ABX responses of the headphone-based experiment in

combination with the reported tonal and spatial differences do not indicate significant audible differences for neither the drum signal without added room nor the speech signal without added room at all evaluated reflection levels. For the loudspeaker-based experiment on the other hand, the ABX responses for the evaluation of the drum signal without added room pass the 95 % confidence threshold both at -3 dB and -6 dB reflection levels.

Summarized, the results for the evaluated reflection elevation changes on the median plane indicate that:

- The amount of perceived differences decreases with reflection level.

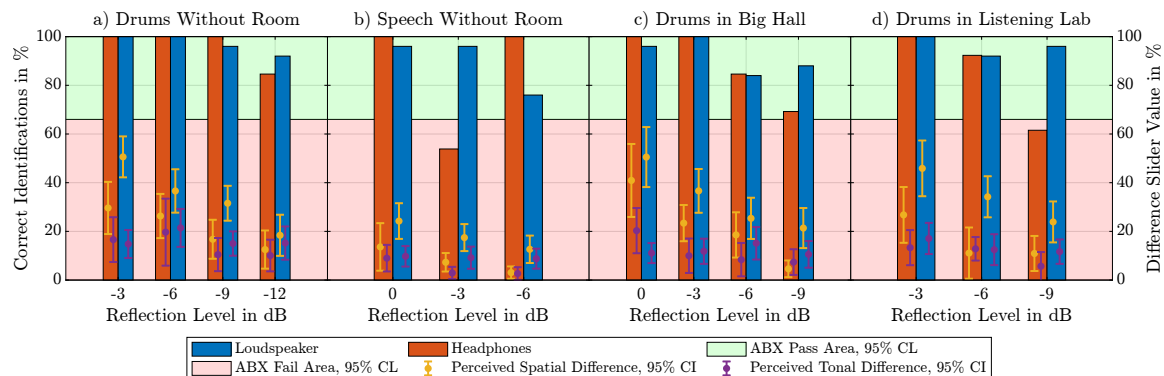- The amount of perceived differences increases with elevation angle.

Fig. 6: Loudspeaker- and headphone-based listening test results for comparison between a reflection with an elevation angle of 45° and 0° at an azimuth angle of 90° with varying reflection levels, source stimuli and room SRIRs

- The drum recording results in more audible differences than the speech signal.

- Comparisons of a single elevated reflection without added room reveal more differences than with added room.

- A room with a longer reverberation time results in less perceivable differences than a room with a shorter reverberation time.

- The loudspeaker-based reproduction method results in larger perceived differences than the headphone-based method with generic HRTFs.

3.2 Elevation Changes on the Frontal Plane

The results for the comparison between a reflection with an elevation angle of 45° and 0° on the frontal plane, i.e. with an azimuth angle of 90°, shown in Figure 6 confirm the previously described findings regarding the influence of reflection level, source signal, room and reproduction method for reflections on the median plane. However, comparing the overall detection thresholds between reflection elevation changes on the frontal plane and reflection elevation changes on the median plane indicates that a reflection arriving from a lateral direction results in significantly larger audible differences when changing its elevation angle than for reflections arriving from the front. While for the comparisons on the median plane both the perceived spatial and tonal differences were mostly reported to be in a similar order, the results for a reflection on the median

plane on the other hand show larger perceived spatial differences than perceived tonal differences, especially for the loudspeaker-based reproduction.

This can be explained by the fact that an elevation change on the median plane, i.e. with 0° azimuth angle in front of the listener, only causes difference in monaural localization cues, which are in general not as robust as interaural localization cues [12], especially when using generic HRTFs in a headphone-based reproduction. An elevation offset of a reflection arriving from the side on the other hand not only causes spectral changes in both ear signals but also changes in interaural cues such as the interaural time and level difference, which allow a more accurate source localization and hence result in larger perceived spatial differences.

3.3 Natural Rooms Without Added Reflection

While the previously described results always included an added strong ceiling reflection, the measured room SRIRs were also evaluated without added ceiling reflection by comparing the SRIRs spatially quantized to eight speakers on the horizontal plane and a single elevated speaker (Table 1, Speaker Nr. 1-9) to a SRIR quantization without added elevated speaker (Table 1, Speaker Nr. 1-8). This scenario might be considered as the most relevant for spatial audio applications since it only includes the natural reflections of the two measured acoustic spaces. Neither the ABX responses nor the reported perceived differences presented in Figure 7 indicate relevant audible differences for the *Big Hall* scenario for both reproduction methods which
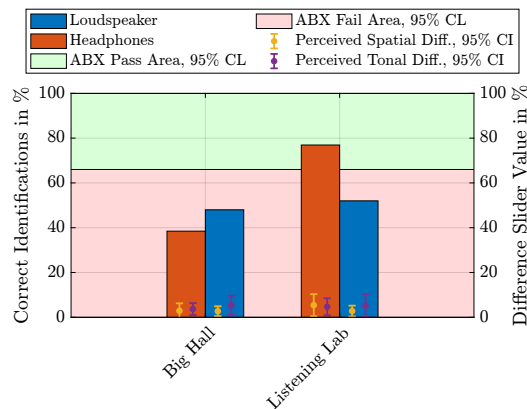
Fig. 7: Loudspeaker- and headphone-based listening test results for comparison between drum recording convolved with *Big Hall* and *Listening Lab* SRIRs, quantized with and without single elevated speaker

means that, for this specific room and this group of participants, including a single elevated speaker in the spatial reproduction of an otherwise horizontal-only setup does not affect the auditory perception.

The *Listening Lab* SRIR, which was measured in a smaller room with a lower reverberation time, resulted in sufficient correct ABX identifications to pass the 95 % confidence threshold for the headphone-based experiment. The reported perceived spatial and tonal differences are still relatively low which indicates that, if there are perceivable differences, they were barely noticeable for the majority of participants.

3.4 Consistency of ABX Responses and Perceived Differences

Figure 8 shows histograms and mean values of the overall reported spatial and tonal difference for both reproduction methods. Thereby, the data were divided into cases where the participants achieved a correct ABX identification and cases where the participants failed to identify the correct X stimulus. These plots allow to evaluate the consistency of the reported spatial and tonal difference values and the ABX responses.

One would expect that the participants reported 0 % perceived spatial and tonal differences whenever they were not able to identify the correct X stimulus. While this is true for the majority of the obtained responses, there are some outliers where participants reported high
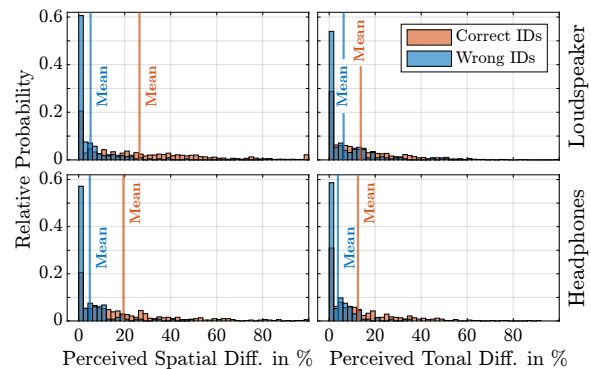

Fig. 8: Histograms and mean values of reported spatial and tonal differences for all comparisons of both reproduction methods, separated for correct and incorrect ABX identifications

perceived differences even though they failed the ABX test, which results in mean perceived differences of around 5 % for tonal and spatial differences in both experiments. Thereby, the extreme outliers can most likely be explained by user errors of participants accidentally selecting the wrong answer. The smaller inconsistencies could be caused by participants perceiving some barely audible differences that are not prominent enough for a correct ABX identification.

The mean perceived difference values for cases where participants achieved a correct ABX identification are, as expected, significantly higher than for the incorrect ABX identifications. Additionally, the mean values of the perceived spatial difference for correct ABX identifications indicate that the loudspeaker-based reproduction method results in larger perceived spatial differences than the headphone-based method.

3.5 Correlation Between Results

Figure 9 shows the Pearson correlation coefficients ρ obtained by correlating the number of correct answers as well as the averaged reported tonal and spatial differences for both the loudspeaker- and the headphone-based experiment. Additionally, the difference in the maximum normalized interaural cross correlation (IACC) between both stimuli of each comparison step was included in the correlation matrix. This IACC difference was obtained by first calculating the maximum normalized IACC for each stimulus by cross-correlating both ear signals of the binaural signals used for the headphone-based experiment as well as the ear
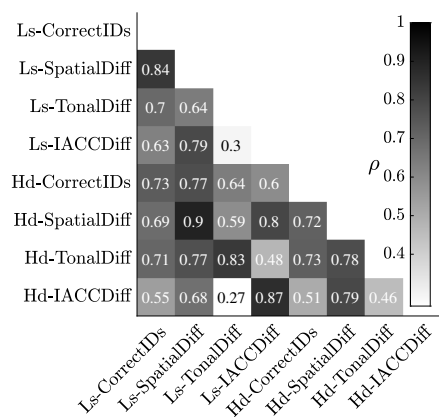
Fig. 9: Pearson correlation coefficients ρ of percentage of correct identifications, perceived spatial and tonal difference as well as measured IACC difference for both loudspeaker- (Ls) and headphone-based (Hd) experiment

signals of a dummy head recording of the loudspeaker reproduced stimuli and then subtracting the maximum normalized IACC values of each compared stimuli pair.

The correlation matrix shows that most of the compared parameters are highly correlated both within the data set of both experiments as well as when comparing the results of both reproduction methods to each other. Thereby, the correlation coefficients between loudspeaker- and headphone-based reproduction when comparing the perceived spatial difference ($\rho = 0.9$) and the perceived tonal difference ($\rho = 0.83$) are especially high which indicates that both reproduction methods result in similar trend of auditory differences. However, a high correlation can still mean an offset between both data sets such as the overall lower perceived spatial difference in the headphone-based experiment which was described before. The perceived spatial difference itself is highly correlated to the percentage of correct ABX identifications for both the loudspeaker experiment ($\rho = 0.84$) and the headphone experiment ($\rho = 0.72$). Additionally, it can be observed that the correlation between difference in interaural cross-correlation and perceived spatial differences is relatively high with $\rho = 0.79$ for both experiments while the correlation between IACC difference and perceived tonal difference is significantly lower ($\rho = 0.3$ and $\rho = 0.46$) which indicates that a change in IACC, i.e. a difference in interaural cues, is related more strongly to perceived spatial than to perceived tonal differences.

4 Discussion

The presented results show that, under certain conditions, the human auditory system is able to differentiate between early reflections with different elevation angles which justifies the need for elevated loudspeakers when reproducing acoustic spaces using methods like Auro-3D, even when the primary sound source itself is not elevated. Thereby, one key outcome is that elevation changes of lateral early reflections result in larger perceived differences than for reflections on the median plane which is a relevant finding for both spatial audio recording and reproduction techniques.

Some of the acoustic scenarios evaluated in the presented experiments can be considered as extreme cases with unnaturally strong early elevated reflections. Since the comparisons of different spatial quantizations for the measured SRIRs without added ceiling reflection mostly resulted in no perceivable differences, it can be assumed that an accurate elevation reproduction of elevated early reflections is not always necessary. Nevertheless, relevant acoustic scenarios with strong early elevated reflections such as some of the concert halls presented in [14] exist, even though the results of this study suggest that the long reverberation time of these halls might mask the perceivable differences due to elevation changes of early reflections.

The fundamental conceptual difference between the loudspeaker-based and the headphone-based reproduction in this study was the fact that the former employed individual HRTFs whereas the latter employed generic ones. The trends in the subjects' ABX responses were largely the same for both reproduction methods whereby the subjects typically reported less pronounced spatial differences for the headphone-based reproduction.

A limitation of this study is that the obtained results do not allow to conclude on universal detection thresholds for elevation changes in early reflections since only a limited amount of specific scenarios was evaluated. Conducting additional experiments with standardized source signals and a more comprehensive set of parameters of non-horizontal reflections would allow to determine more accurate detection thresholds.

One parameter that was not investigated is the influence of the delay of the reflection which, in the context of psychoacoustic mechanisms like the precedence effect, can be assumed to be a relevant factor for the amount of

perceived differences. Conducting a similar experiment and varying the reflection delay instead of the reflection level could further contribute to an understanding of the human perception of spatial reverberation.

5   Summary

This work used manipulated spatial room impulse responses to perform both loudspeaker- and headphone-based listening experiments regarding the effect of elevation changes in early room reflections. Thereby it was found that, under certain conditions, listeners could not identify a difference between elevated early reflections and non-elevated early reflections. Factors that affect the amount of perceivable differences are the reflection strength, the lateral reflection position, the acoustic environment and the source signal. Furthermore, it was found that a loudspeaker-based reproduction often results in larger perceived spatial differences than a headphone-based reproduction with generic HRTFs.

References

[1] Ahrens, J., Helmholz, H., Alon, D. L., and Amengual Garí, S. V., "Spherical Harmonic Decomposition of a Sound Field Based on Observations Along the Equator of a Rigid Spherical Scatterer," *J. Acoust. Soc. Am.*, 150(2), pp. 805–815, 2021, https://doi.org/10.1121/10.0005754.

[2] Rakerd, B. and Hartmann, W. M., "Localization of Sound in Rooms, II: The Effects of a Single Reflecting Surface," *J. Acoust. Soc. Am.*, 78(2), pp. 524–533, 1985, https://doi.org/10.1121/1.392474.

[3] Furuya, H., Fujimoto, K., Takeshima, Y., and Nakamura, H., "Effect of Early Reflections from Upside on Auditory Envelopment," *J. Acoust. Soc. Jpn. (E)*, 16(2), pp. 97–104, 1995, https://doi.org/10.1250/ast.16.97.

[4] Bech, S., "Perception of Reproduced Sound: Audibility of Individual Reflections in a Complete Sound Field, II," in *Audio Engineering Society Convention 99*, New York, US, 1995.

[5] Bech, S., "Spatial Aspects of Reproduced Sound in Small Rooms," *J. Acoust. Soc. Am.*, 103(1), pp. 434–445, 1998, https://doi.org/10.1121/1.421098.

[6] Robotham, T., Stephenson, M., and Lee, H., "The Effect of a Vertical Reflection on the Relationship Between Preference and Perceived Change in Timbre and Spatial Attributes," in *Audio Engineering Society Convention 140*, Paris, FR, 2016.

[7] Müller, L., *Perceptual Differences Caused by Altering the Elevation of Early Room Reflections*, Master's thesis, Chalmers University of Technology, Gothenburg, Sweden, 2021, https://doi.org/10.13140/RG.2.2.19874.35522.

[8] Ahrens, J., "Perceptual Evaluation of Binaural Auralization of Data Obtained from the Spatial Decomposition Method," pp. 65–69, IEEE, 2019, https://doi.org/10.1109/WASPAA.2019.8937247.

[9] Tervo, S., Pätynen, J., Kuusinen, A., and Lokki, T., "Spatial decomposition method for room impulse responses," *J. Audio Eng. Soc.*, 61, pp. 17–28, 2013.

[10] Gari, S. V. A., Arend, J., Calamia, P., and Robinson, P., "Optimizations of the Spatial Decomposition Method for Binaural Reproduction," *J. Audio Eng. Soc.*, 68(12), pp. 959 – 976, 2020, https://doi.org/10.17743/jaes.2020.0063.

[11] Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L., "Localization using nonindividualized head-related transfer functions," *J. Acoust. Soc. Am.*, 94(1), pp. 111–123, 1993, https://doi.org/10.1121/1.407089.

[12] Blauert, J., *Spatial Hearing: The Psychophysics of Human Sound Localization*, The MIT Press, 1996, https://doi.org/10.7551/mitpress/6391.001.0001.

[13] Boley, J. and Lester, M., "Statistical Analysis of ABX Results Using Signal Detection Theory," in *Audio Engineering Society Convention 127*, New York, US, 2009.

[14] Pätynen, J., Tervo, S., and Lokki, T., "Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses," *J. Acoust. Soc. Am.*, 133(2), pp. 842–857, 2013, https://doi.org/10.1121/1.4770260.