ISRA 2019 AMSTERDAM

PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

Perceptual differences of position dependent room acoustics in a small conference room

Christian SCHNEIDERWIND[1]; Annika NEIDHARDT[1,2]

[1] Technische Universität Ilmenau, Germany

ABSTRACT

In virtual and augmented reality acoustical illusions are usually perceived as more vivid if they endure an interactive change of the listening position by the user. The room acoustics play an essential role in the localization and the naturalness of a sound source. However, not every room acoustical detail can be perceived by most listeners. This suggests a certain potential to improve the efficiency of a position dependent reproduction. However, the exact perceptual tolerances remain to be determined. The study presented in this paper investigates perceptual differences between different local room acoustical conditions with a small conference room with standard furniture. Various listening positions are compared for the case of similar direct sound but strong variations in the early reflection pattern. Two different test scenarios, one with strong direct sound, the other with low direct sound, are considered. A psychoacoustic paired comparison experiment was conducted to identify and quantify the perceptual differences. The results contribute to the determination of the perceptual requirements for a position dependent auralization of virtual sound source in virtual or real rooms.

Keywords: Augmented/Virtual acoustics, Acoustic of small rooms, Binaural synthesis, Perception

1. INTRODUCTION

In virtual and augmented reality acoustical illusions are usually perceived as more vivid, if they endure an interactive change of the listening position by the user. Room acoustics play an essential role in the localization and the impression of naturalness of a sound source. However, not every room acoustical detail can be perceived by most listeners. The determination of the exact perceptual tolerances will allow a psychoacoustic optimization of reproduction algorithms.

In the past many experiments were conducted to achieve a detailed understanding regarding the perception of distinct early acoustical reflections (1, 2, 3, 4). However, most of these investigations are based on test conditions that represent only a small part of the acoustical scenarios occurring in the real world. This includes e.g. the assumption of point sources, dirac impulses as reflections (2) or a strong direct sound caused by a receiver position within the critical distance of a sound source (1,4). The consideration of typical properties of reflections in test scenarios is important (5).

The present work investigates the perceptual and physical differences of the local room acoustics at five different listening positions in a small conference room. Previously, two studies (4, 6) were conducted to find out, whether subjects could not only perceive differences, but whether these differences were distinct enough to allow the recognition of the listening position in the room. With some training few participants were able to assign the positions correctly.

This study intends to provide an analysis of the available physical and perceptual differences between the chosen positions. For the physical analysis classic room acoustic parameters such as Direct-to-Reverberant Ratio (DRR), Early Decay Time (EDT) and Reverberation Time (T_{60}) were considered. For the identification of perceptual differences, a paired comparison experiment with respect to five different attributes was conducted.

[1] christian.schneiderwind@tu-ilmenau.de
[2] annika.neidhardt@tu-ilmenau.de

Peer reviewed

2. Room acoustic measurements

The room acoustic measurements were conducted in a conference room of the university in Ilmenau. The conference room has a volume of 185.2 m^3 (10.3 m x 5.8 m x 3.1 m). The same arrangement of source and receiver was placed at five different positions in the room as shown in Figure 1. The positions were chosen with the goal to get a representative variation of the early reflection pattern that can be expected for a listener moving through that room. For all listening positions, the distance between source and receiver was set to 2.5 m in order to keep the direct sound as similar as possible. Consequently, the main differences between the impulse responses of the various positions arise in the structure of the early reflection patterns. In addition to the common measurement where the source faces the receiver position, a second case was introduced with the source turned around facing away from the receiver. Therefore, with the second scenario the experiment considers the directivity of the sound source which has an overall impact on the measured physical quantities as well as the perception.

The impulse responses were measured with the swept-sine method generating a sine sweep which exponentially rises in frequency from 40 Hz to 20 kHz. A Genelec 1030A studio monitor was used as the directional sound source. Two different types of receivers were employed:

- Omnidirectional Microphone: Microtech Gefell Capsule Mk221 + Amplifier MV203
- Dummy head: KEMAR 45BA for BRIR measurements (5° azimuthal resolution)

The omnidirectional RIRs were captured to determine the classic room acoustic parameters DRR, EDT and T_{60}. In order to reproduce the different positions via dynamic binaural synthesis the KEMAR 45BA dummy head was used for measuring the BRIRs with an azimuthal resolution of 5° allowing for head rotation. All measurements are part of a data set documented in (7).

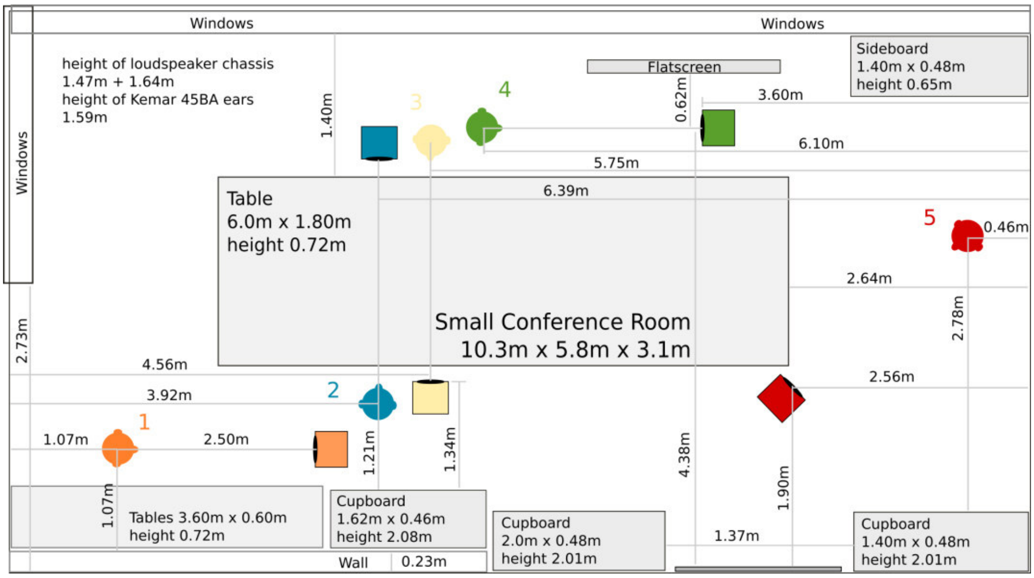

Figure 1: Measurements of the room, relevant furniture, the chosen listening and source positions.

3. Physical analysis

Different room acoustic parameters were investigated in order to explain potential differences in the perception. This chapter gives a brief overview about the physical measurements. A more detailed analysis can be found in (7).

3.1 Early Decay Time (EDT)

The EDT was inspected as it relates better to the perception of reverberance than the T_{60} values (8). The EDT is defined as the time interval in which the sound pressure level drops 10 dB below the initial level. Table 1 shows the EDTs of all positions for the direct and indirect irradiation scenario. The EDTs are averaged from the values of the 500 Hz and 1 kHz bands. A rise in the EDT is expected to induce a perception of an increased reverberance.

Table 1: EDT values (in ms) from the RIRs of all five positions

Position	1	2	3	4	5
EDT for LS towards HATS	262	504	409	432	331
EDT for LS turned by 180°	505	561	420	430	498

3.2 Direct-to-Reverberant Ratio (DRR)

The ratio of the direct sound energy and the reverberant sound energy is described by the Direct-to-Reverberant Ratio (DRR). It corresponds to the perceived distance of sound source in a room (9). Table 2 denotes the DRR values for the RIRs and BRIRs of the frontal direction of the five positions for both irradiation scenarios. For the calculation of these DRR-values the time window (rectangle) for the direct sound was set to ±1 ms from the direct sound peak.

Table 2: DRR values (in dB) of all positions for RIRs and BRIRs for the direct (0°) and indirect (180°) irradiation scenarios

Position	1	2	3	4	5
DRR – 0°, RIR	-2.7	-0.63	-0.59	0.69	-2.2
DRR – 180°, RIR	-13	-15.4	-15.2	-14.1	-14.3
DRR – 0°, BRIR (L|R)	1.7 | -3.7	1.8 | 2.0	2.7 | 1.9	2.4 | 4.3	-1.4 | 2.1
DRR – 180°, BRIR (L|R)	-25.3 | -25.7	-29.0 | -27.9	-26.6 | -27.3	-26.0 | -26.0	-25.8 | -26.7

3.3 Interaural Cross-Correlation (IACC)

The Interaural Cross-Correlation is a parameter that is typically associated with the apparent source width (ASW). It measures the similarity between the left and right ear signals and therefore considers differences in time and level as well as timbre. The IACC values range between -1 and 1. Table 3 lists the values of the IACCs for all positions and both irradiation scenarios.

Table 3: IACC values for the frontal direction all positions for the direct (0°) and indirect (180°) irradiation scenarios

Position	1	2	3	4	5
IACC – 0°	0.54	0.74	0.74	0.74	0.51
IACC – 180°	0.16	0.41	0.30	0.27	0.07

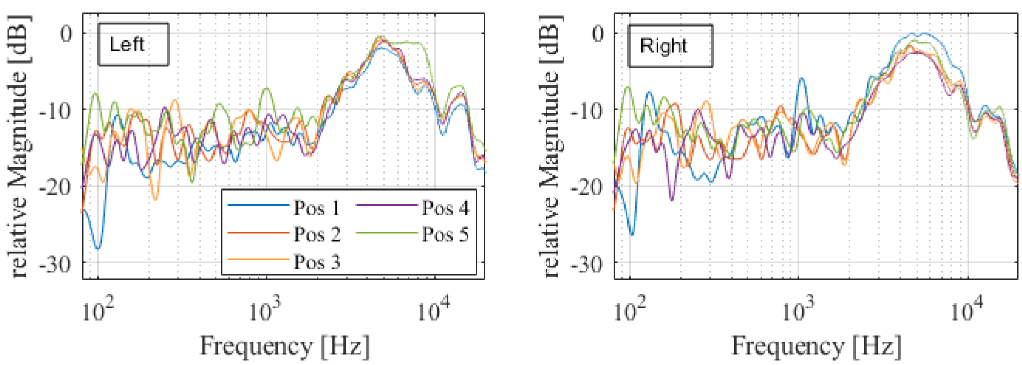

Figure 2: Comparison of the spectra of the BRIRs of all five positions for the left and right ear in the case of direct sound irradiation.

3.4 Spectrum

The smoothed spectra (1/6-octave) of the BRIRs (0° azimuth) of all positions for the direct and indirect irradiation scenario are shown in Figure 2 and Figure 3 respectively. The curves in these plots are normalized to the maximum value of all curves.

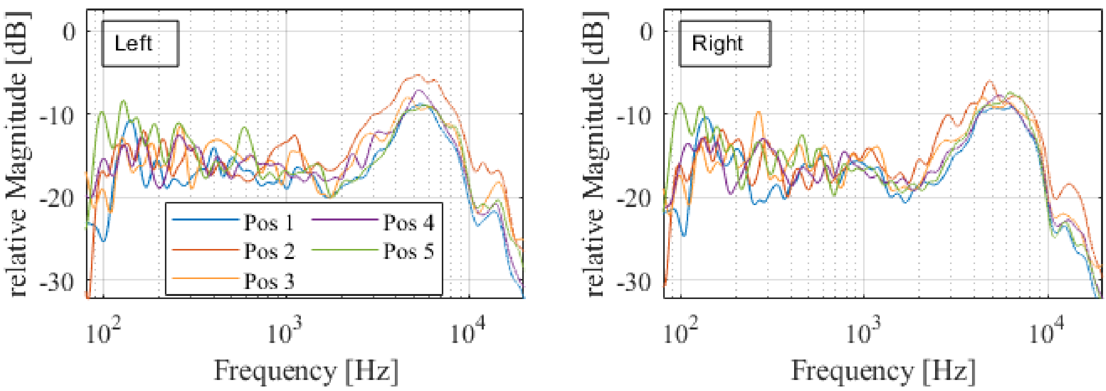

Figure 3 – Comparison of the spectra of the BRIRs of all five positions for the left and right ear in the case of indirect sound irradiation.

4. Listening Experiment

In order to evaluate the perceptual differences, the BRIRs were used to auralize the different scenarios. The following section documents the realization of the psychoacoustic experiment.

4.1 Description & Setup

For the psychoacoustic evaluation a paired-comparison experiment was conducted. 15 participants had to compare the five positions in pairs regarding five attributes including localization (left-right), reverberance (less-more), apparent source width (ASW, more narrow - wider), perceived distance (closer – further away) and timbre (darker – brighter). In a pretest a group of experts selected the attributes to cover the main audible differences. The participants had to rate the differences between a position and a reference position on an interval-scale ranging from -5 to +5. No perceived difference was associated with a rating of 0. All pair-combinations of the positions had to be rated twice with both items being the reference once. Overall, 20 pairs had to be rated in five categories. To avoid listener fatigue, the experiment was split into two sessions. In the first session only differences in localization and reverberance were rated, the remaining three attributes in the second session. An eight-minute-long dry male speech sample was chosen as a stimulus. All ratings were given orally by the participants to supervisor. The subjects took a short training session introducing all audio scenes in order to get familiar with the test items and make better use of the rating scale.

During the test the participants had to wear a head-mounted display (HTC VIVE) basically for tracking purposes. It showed a neutral grid scene providing no visual information about the reproduced audio scene. The different listening perspectives were presented via dynamic binaural synthesis over STAX SR202 headphones using the pybinsim-Toolbox (10) and the measured BRIRs. In the paired comparison the items could be switched arbitrarily by giving a signal to the supervisor.

Two of the 15 participants were female, the others male. Table 4 contains additional information about the test subjects.

Table 4: Details about test subjects

Age	24 – 36 (avg. 28.8y)
Hearing aid/ Hearing impairment	0 / 15
Interested in music/ acoustics	15 / 15
Experience in acoustics/ audio processing	14 / 15
Experience with listening tests employing binaural synthesis	14 / 15

4.2 Results & Discussion

As all position pairs were rated twice it can be seen if a subject has given similar ratings for the same pair. The audible differences between the items are very small. For this reason, the ratings often varied slightly in the repeated evaluation of the same pair. In the following the rating will be considered as consistent, if the repeated rating exhibits the same sign (same side of the scale).

In the case of the direct irradiation scenario the localization ratings yielded the highest consistency. For about 79.3% (mean) of the repeated evaluations the participants rated the deviations in direction consistently. A lower consistency across the items was observed for the attributes reverberance and ASW with only 43.3% and 46.7% of the repetitions respectively. The comparison of position 2 and 3 showed a notably high consistency for all attributes compared to the other position pairs. This particular pair was rated with “0” most of the time indicating that the acoustic impression for both positions sounds rather alike. Several participants did not even notice the switching between the positions.

For the indirect scenario, the overall consistency is higher suggesting that audible differences are more distinct in the case of lower direct sound energy. Especially attributes like reverberance and ASW exhibit a higher consistency (78.6% and 66.3% respectively) in comparison to the direct irradiation scenario.

The following figures show the boxplots of the rated position pairs grouped by the attributes. The labels denote the reference position (first number) and the position to be rated (second number). A Wilcoxon-Signed-Rank test was conducted to test whether differences were statistically significant. The red boxes depict position pairs where no significant deviations from a zero median were observed.

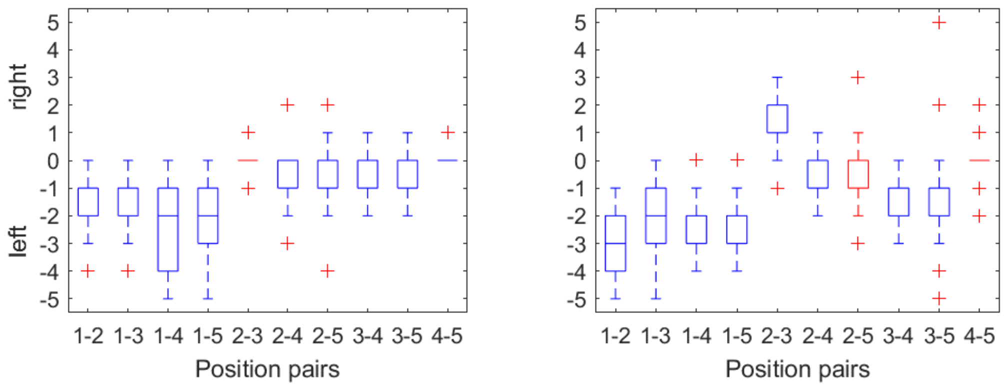

Figure 4 – Results for the perceived difference in localization of the second position in comparison to the position given first. - LEFT: Direct irradiation - RIGHT: Indirect radiation.

Figure 4 shows the results for the ratings of the localization differences for both irradiation scenarios. At position 1 the source was localized more on the right side than for all other positions. Position 1 exhibits a first strong reflection from the nearby wall on the right side. This particular reflection arriving 2.2 ms after the direct sound seems to cause a shift in perceived direction of arrival. For position 4 and 5 the source was perceived more on the left than for positions 2 and 3. At position 4 the reflection from the TV screen may be the cause and for position 5 the nearby wall. In the case of the indirect sound irradiation scenario a significant difference for pairing 2-3 arises. An explanation could be that for both positions the localization shifts to different directions due to the closer lateral wall. At position 5 the influence of the nearby wall decreases for the case of the indirect irradiation.

Figure 5 depicts the ratings for the reverberance. No distinct perceptual differences were observed at any of the position pairs in the case of the direct irradiation scenario. This indicates that it was challenging for participants to interpret perceived sound differences regarding reverberation. In comparison, the scores for the turned speaker case show significant differences at every position pair. The overall sound impression at every position is dominated by the early reflection patterns as the direct sound energy is rather low. Additionally, participants rated positions as less reverberant if the position had more pronounced early reflections. Such pronounced reflections seem to be interpreted as a direct sound in terms of perception. Looking at the DRR and EDT values in tables 1 and 2 no correlation between the ratings for reverberance and the physical parameters was observed. Instead it may be of interest to study the relation between the first dominant wave front and the late reverberation. Furthermore, a correspondence between reverberance and distance perception (Figure 7) can be observed.

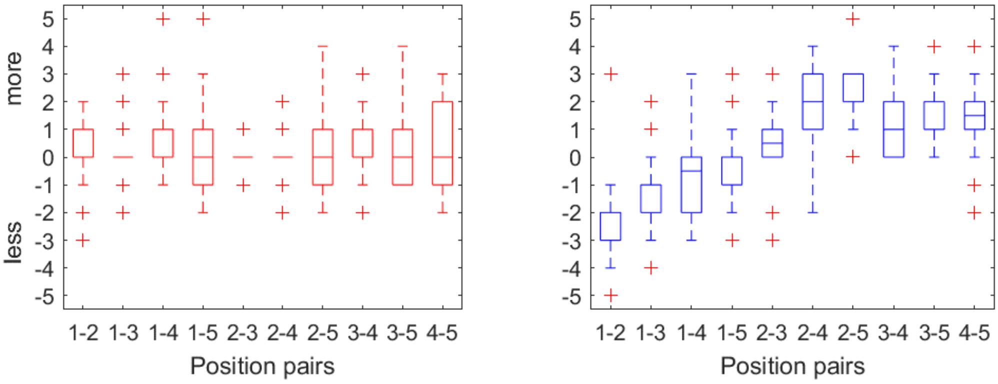

Figure 5 – Results for the perceived difference in reverberance. LEFT: Direct irradiation

RIGHT: Indirect radiation.

The results for ASW are shown in Figure 6. For position 5 the source image position was perceived as wider compared to the positions 1, 3 and 4. The early strong lateral reflection from the nearby wall may be an explanation. In comparison to the strong lateral reflection from position 1 (about 45°) which did not cause a significant increase in the ASW the reflection at position 5 arrives slightly from the back. Although both positions exhibit a similar IACC, the reflections at position 5 induce an increased ASW. The ratings for the remaining positions pairs did not yield any significant perceivable differences. ASW, like reverberance, seems to be an attribute which was hard to evaluate for the participants. However, the indirect irradiation case exhibits significant differences for most position pairs. In that special scenario the impression of an apparent source is dominated by the spatio-temporal structure of the first reflections instead of the direct sound. Only for the position pairs 2-3 and 2-4 no differences in ASW were observed. At those positions the main beam of the rotated speaker radiates onto a perpendicular surface generating a strong reflection that arrives at the listener's position from the frontal direction. In this case the apparent source shifts towards the first mirror image source. The ASW depends mainly on the distance to reflecting surface and the directivity of the sound source.

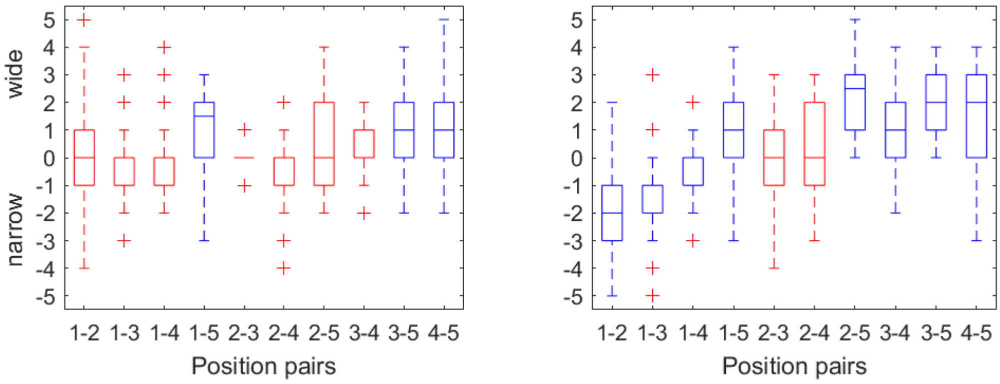

Figure 6 – Results for the perceived difference in ASW. LEFT: Direct irradiation

RIGHT: Indirect radiation.

Figure 7 displays the scores for the differences in perceived distance of the sound source. For the direct irradiation scenario significant differences are observed for position pairs including position 1. The DRR values of the BRIRs exhibit differences close to the smallest JNDs found in (11). Therefore, it is uncertain to what degree the DRR contributes to a change in distance perception. Small differences regarding the overall level and timbre can have an effect as well. For the JNDs of DRR in BRIRs Larsen et al. conclude that DRR changes "are primarily discriminated based on spectral cues" (11). Then again, the indirect sound irradiation scenario exhibits more distinct differences in distance perception. As for reverberance and ASW the first dominant reflection characterizes the apparent sound source. Consequently, the DRR values from table 2 do not correspond to the perceived distance. Instead the relation of the first strong reflection and the remaining reverberation may be of interest. Generally, many participants reported, that especially for position 5 the sound source was hardly localizable. This may be an explanation for the insignificant differences for 3-5 and 4-5.

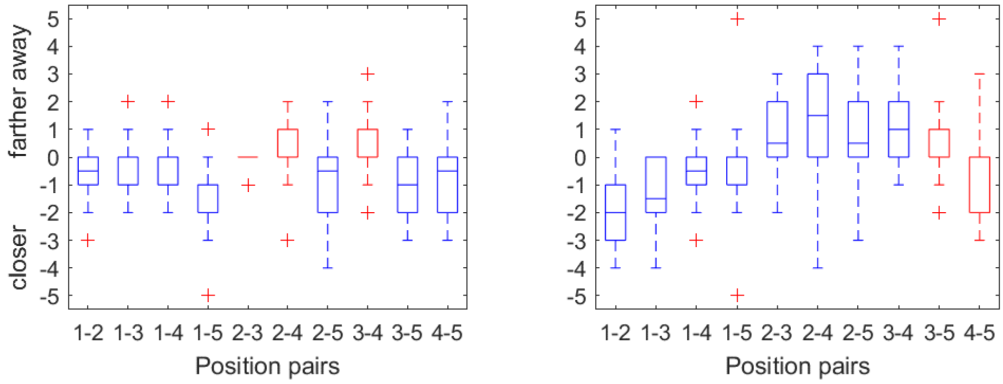

Figure 7 – Results for the perceived difference in distance. LEFT: Direct irradiation

RIGHT: Indirect radiation.

Timbre is an attribute that can be described by a lot of different terms. For practical reasons the participants were asked to rate the difference in coloration on a scale from darker to brighter sounding. Figure 8 shows the scores for this attribute. In the case of direct sound irradiation positions 5 were rated as significantly darker compared to the other positions. Already in previous studies people described position 5 as more muffled/darker (4, 6). The pairings 1-4, 2-3, 2-4 and 3-4 did not yield significant differences. The whiskers of all boxplots range from the positive to the negative part of the scale which indicates that differences in timbre are either not distinct or the provided attribute construct (darker/brighter) was not suitable enough to describe the differences. The indirect sound irradiation scenario shows that position 5, even with the turned speaker, has the most energy in the low frequency range which may be caused by the nearby wall. Position 2 was perceived as brighter than the other positions which is confirmed by the spectrum shown in Figure 3. Due to the short distance to the window the high frequencies of the first strong reflection are preserved.

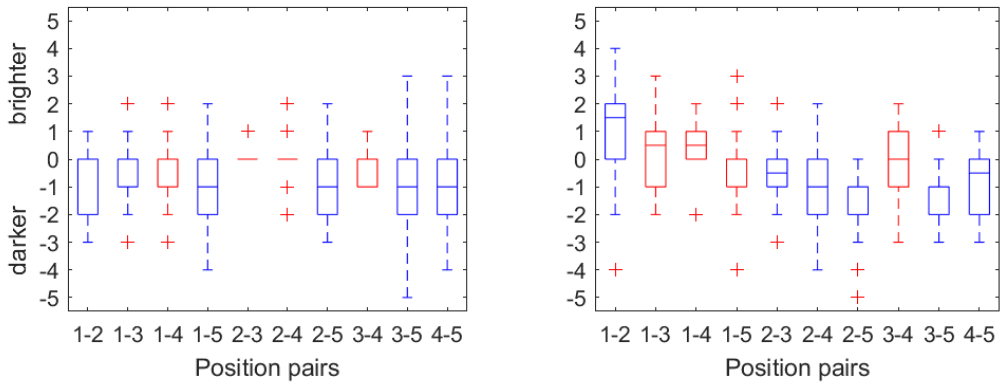

Figure 8 – Results for the perceived difference in timbre. LEFT: Direct irradiation

RIGHT: Indirect radiation.

5. SUMMARY AND CONCLUSIONS

This paper presents a case study with the goal to investigate the physical and perceptual differences between the early reflection patterns at different positions inside a small room. Understanding the perception of the reverberation in small rooms is particularly interesting for auditory augmented reality, because people spend more time of their everyday lives in small rooms than in concert halls. The results of the listening test indicate that people are able to perceive differences depending on the deviation in the structure of the early reflections. In the case of strong lateral reflections, a shift in localization was observed when it was compared with a position lacking this type of reflection. A difference in localization was observed although initial time delay gaps were greater than 2.2 ms where the occurrence of the precedence effect is expected.

Furthermore, it was found that the directivity of the sound source has a great influence on the impact of the early reflections on the overall perceived sound. In many common studies the strong direct sound masks differences in the reverberation. Considering low direct sound energy is of relevance for perceptual studies regarding room acoustics. The audible deviations caused by the early reflections are more distinct. The sound source directivity and its impact on the perception as well as the physical parameters have to be investigated in further studies.

To be able to explain the perceived differences classic room acoustic parameters were calculated from omnidirectional RIRs and the BRIRs which were measured at the desired positions. It was found that most parameters like EDT or DRR are not corresponding to the perception of reverberance and source distance as expected. It is therefore not possible to explain certain perceptual differences with these types of parameters. The parameters and JND values given in part 1 of ISO-3382 (12) have been determined for large rooms like concert halls and omnidirectional sound sources. Part 2 only discusses the determination of the reverberation time for 'usual' rooms. A model for the correspondence between the perception of room acoustics in small rooms and suitable physical parameters remains to be developed. The presented study can only deliver a first impression, which parameters may be of interest in such a model. A systematic investigation considering different room sizes and properties as well as all potential source and listening positions and the source directivity is required.

ACKNOWLEDGEMENTS

This research was funded by the German Research Council (DFG BR 1333/18-1). Thanks to all participants of our psychoacoustic experiment.

REFERENCES

1. Shinn-Cunningham B, Ram S. Identifying where you are in the room: Sensitivity to room acoustics. In Proc. of Int. Conf. on Auditory Display, Boston, MA, USA, July 6-9, 2003.
2. Freyman RL, Clifton RK, Kitovsky, RY. Dynamic processes in the precedence effect. J. Acoust. Soc. Am. 90 (2), Pt. 1, August 1991.
3. Seeber BU, Müller M, Menzer F. Does learning a room's reflections aid spatial hearing? In 22nd Int. Congress on Acoustics, Buenos Aires, Argentina, September, 2016.
4. Neidhardt A. Perception of the reverberation captured in a real room, depending on position and direction. In 22nd Int. Congress on Acoustics (ICA), Buenos Aires, Argentina, 2016.
5. Wendt F, Höldrich R. Reflected properties influencing the precedence effect, 48th Annual Conf. on Acoustics, Germany, 2018.
6. Klein F, Neidhardt A, Seipel M, Sporer T. Training on the acoustical identification of the listening position in a virtual environment. 143rd Int. AES Convention New York, NY, USA, 2017
7. Schneiderwind C, Neidhardt A, Klein F, Fichna S. Data set: Eigenmike-DRIRs, KEMAR 45BA-BRIRs, RIRs and 360° pictures captured at five positions of a small conference room. 45th Annual Conference on Acoustics (DAGA), Rostock, Germany 2019.
8. Barron, M. Using the standard on objective measures for concert auditoria, ISO 3382, to give reliable results. Acoustical Science and Technology, (26)2:162-196, 2005.
9. Zahorik P, Brungart DS, Bronkhorst AW. Auditory Distance Perception in Humans: A Summary of Past and Present Research, Acta Acustica united with Acustica, Volume 91, Number 3, May/June 2005, pp. 409-420(12).
10. Neidhardt A, Klein F, Knoop N, Köllmer T. Flexible python tool for dynamic binaural synthesis applications. Proc. 142nd International AES Convention, Berlin, Germany, 2017.
11. Larsen E, Iyer N, Lansing CR, Feng AS. On the minimum audible difference in direct-to-reverberant energy ratio. In: Journal of the Acoustical Society of America, 124, p. 450-461, 2008.
12. Akustik – Messung von Parametern der Raumakustik – Teil 1: Aufführungsräume. Standard DIN EN ISO 3382-1:2009, DIN Deutsches Institut für Normung e.V., Berlin, D, Oktober 2009.