The Relative Influence of Reverberation at the Contralateral versus Ipsilateral Ear on Perceived Externalization of a Lateral Sound Source

Song Li[1], Roman Schlieper[1], Stephan Preihs[1], Jürgen Peissig[1]
[1] *Leibniz Universität Hannover, Institut für Kommunikationstechnik*
Appelstr. 9A, 30167 Hannover, Email: {song.li, schlieper, peissig}@ikt.uni-hannover.de

Abstract

Reverberation, especially the early reflection part, is important to perceived externalization of headphone-reproduced virtual sound sources. The result of our previous experiment demonstrated that the degree of externalization decreased substantially by reducing the amount of reverberation at the contralateral ear. In contrast, the degree of externalization changed slightly by removing the reverberation part in the ipsilateral ear signal. In that study, a one-second long white noise was used as the stimulus and the simulated virtual sound source was located at a distance of 1.7 m for an azimuth angle of 45°. However, it is still unknown whether or not this effect depends on the type of stimuli and source-listener distances, and how this effect changes as a source moves from lateral to frontal incidence angles. For these purposes, the non-individual binaural room impulse responses (BRIRs) were measured for different azimuth angles and distances in a listening room. The amount of reverberation was reduced in BRIRs of either the contralateral or the ipsilateral ear. Such modified BRIRs were convolved with different stimuli, and subjects rated the degree of externalization of the rendered binaural signals presented over headphones. The results will be discussed in the present study.

Introduction

Headphone-based three-dimensional (3D) audio reproduction technology has become more and more important due to the ever-increasing market of mobile devices, virtual-, augmented- and mixed-reality (VR/AR/MR) applications. Perceived externalization, i.e., out of the head [1], is an essential indicator for building up immersive acoustic environments.

In free-field conditions, head-related transfer functions (HRTFs) are responsible for perceived externalization, which depend strongly on individual anatomies. However, non-individual HRTFs are often used in commercial binaural rendering systems because the measurement of high-quality individual HRTFs is impracticable in consumer scenarios. The degree of perceived externalization might be reduced by applying non-individual HRTFs for binaural rendering systems [1, 2].

In reverberant conditions, reverberation is essential to perceived externalization of headphone-reproduced virtual sound sources [3]. Several studies [4, 5, 6, 7] showed that the reverberation between 20 ms and 80 ms (early reflection parts) had a substantial influence on perceived externalization. In our previous study [7], we have investigated the relative influence of reverberation at the contralateral versus the ipsilateral ear on externalization of a lateral (45°) sound source. A one-second long white noise was used as a stimulus in the listening experiment. The results showed that the reverberation at the contralateral ear had more influence on externalization than that at the ipsilateral ear; the perceived externalization remained nearly unchanged by removing the reverberation only in the ipsilateral ear signal. Furthermore, Li *et al.* [8] investigated this effect for different azimuth angles. The results revealed that the contribution of the reverberation at the contralateral ear to perceived externalization increased as the source moved laterally. For azimuth angles larger than 30°, the reverberation at the ipsilateral ear did not have noticeable contributions to externalize sound images. However, visual cues (loudspeakers) were provided in the previous studies, which might influence the externalization results. In addition, it is still unknown whether or not this effect depends on the type of stimuli and source-listener distances.

In the present study, non-individual binaural room impulse responses (BRIRs) were measured for various azimuth angles and distances in a listening room. The amount of reverberation was reduced in BRIRs of either the left or the right ear. Such modified BRIRs were convolved with speech and pop music stimuli, and subjects rated the degree of perceived externalization of the binaurally rendered virtual sound sources presented over headphones without visual cues.

Methods

In total, 12 pairs of non-individual BRIRs were measured with a low-noise dummy head KEMAR 45BC-12 at azimuth angles of 0°, 30°, 60°, and 90° with distances of 0.5 m, 1.5 m and 4 m in a listening room, which was designed under the ITU-R BS.1116 standard and has a reverberation time of about 260 ms [9]. A five-second long exponential sweep [10] from 20 Hz to 20 kHz was used as a measurement signal, and the BRIR measurement for each azimuth angle and distance was repeated five times. The BRIR measured was truncated to 260 ms.

The reverberation was removed in (i) BRIRs of both ears (condition “RB”), (ii) BRIRs of the left ear (condition “RL”), (iii) BRIRs of the right ear (condition “RR”). The removal of reverberant parts in BRIRs was achieved by applying a 2.5 ms long time window with a 0.5 ms long half raised-cosine fall time (only the direct parts of BRIRs remained). All truncated BRIRs were zero-padded to a length of 260 ms. The non-processed BRIRs were used as references (condition "RN"). Note that in the case of lateral (30°, 60° and 90°) sound sources, the left ear is the contralateral ("opposite-side") ear, and the right ear is the ipsilateral ("same-side") ear. For a frontal (0°) sound source, both ears are facing the loudspeaker, and they are neither contralateral nor ipsilateral ears.

Experimental Setup

Five subjects (one female and four males, aged between 26 and 29) listened to the test stimuli presented over a pair of Sennheiser HD800 headphones. The effect of headphones was compensated by measuring the headphone transfer function (HpTF) on the KEMAR 45BC-12 [11]. As shown in Table 1, a subjective rating scale with four possible externalization degrees from 0 to 3 was used to evaluate the degree of perceived externalization. Subjects could rate each stimulus by using a slider with a step-size of 0.1 between 0 and 3. During the listening experiment, listeners were allowed to repeat every sequence.

Degree	Meaning of the degree
3	The sound is externalized and far away from me.
2	The sound is externalized but within my arm's reach.
1	The sound is externalized but very close to me.
0	The sound is in my head.

Table 1: A subjective rating scale to rate the degree of perceived externalization.

The stimuli used in the listening experiment were truncated speech sentence (Track 50) and pop music (Track 70) with a length of 2.5 s taken from the European Broadcasting Union (EBU) Sound Quality Assessment Material (SQAM) [12]. For the direction of 60°, a low-pass and a high-pass condition were additionally considered, where the stimuli (speech and music) were filtered at 1 kHz (cut-off frequency of the low-pass filter) and 2 kHz (cut-off frequency of the high-pass filter), respectively. The binaurally-rendered signals with the non-processed BRIR was presented at a level of 64 dBA over the headphones.

Results

Figure 1 shows the externalization results of music and speech stimuli for "RN", "RB", "RL", and "RR" conditions in different azimuth angles and distances. For non-processed BRIRs ("RN" condition), the mean externalization ratings increased with the increasing azimuth angle and source-listener distance. The degree of perceived externalization reduced noticeably by removing the reverberant parts in BRIRs of both ears ("RB" condition) for all azimuth angles, especially for distances larger than 1.5 m.

For a 0° sound source, the externalization ratings were almost the same for the "RL" and "RR" conditions. For lateral sound sources, similar to the "RN" condition, the externalization ratings for the "RR" condition increased with the increasing azimuth angle and source-listener distance, but slightly lower than those for the the "RN" condition. The externalization ratings for the "RL" and "RB" conditions increased as the sound source moved

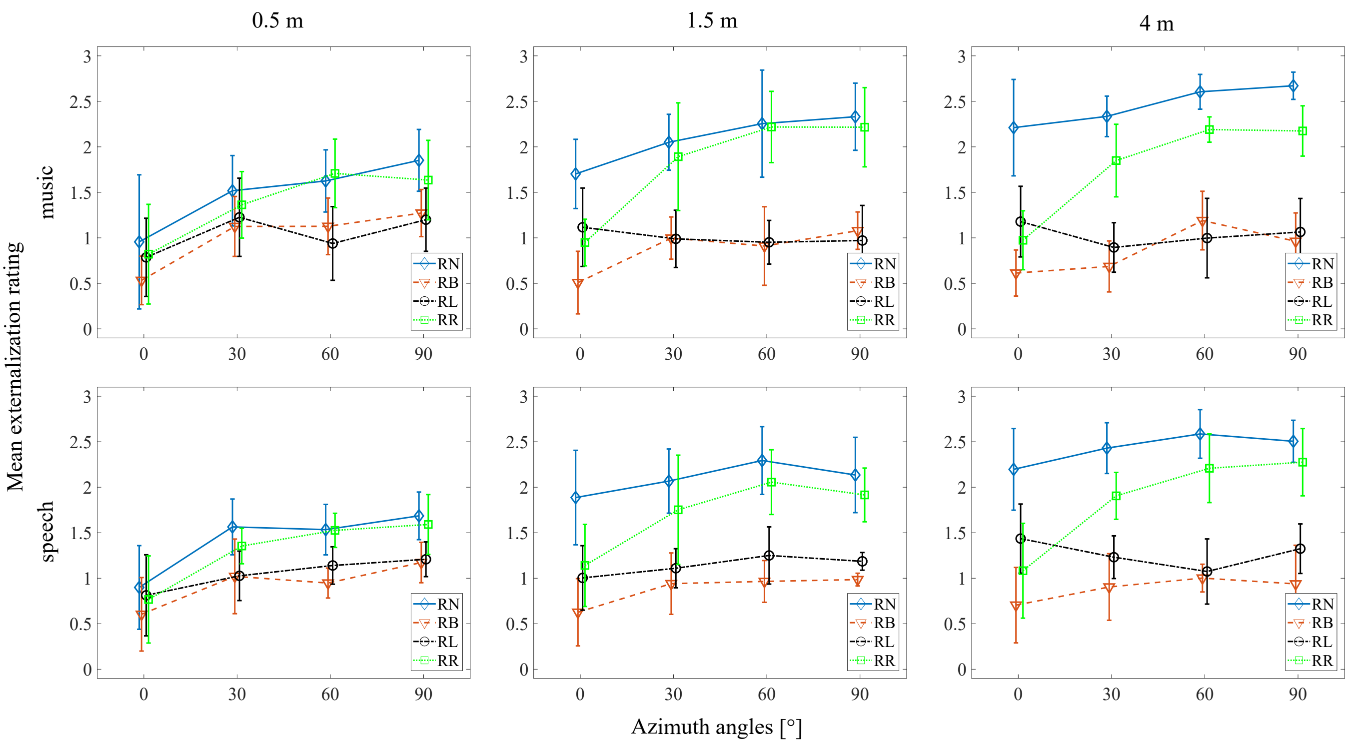

Figure 1: Mean externalization ratings of test stimuli across listeners for "RN" (diamonds), "RB" (triangles), "RL" (circles) and "RR" (squares) conditions in different azimuth angles and distances. The error bars indicate the standard deviations of the mean values.

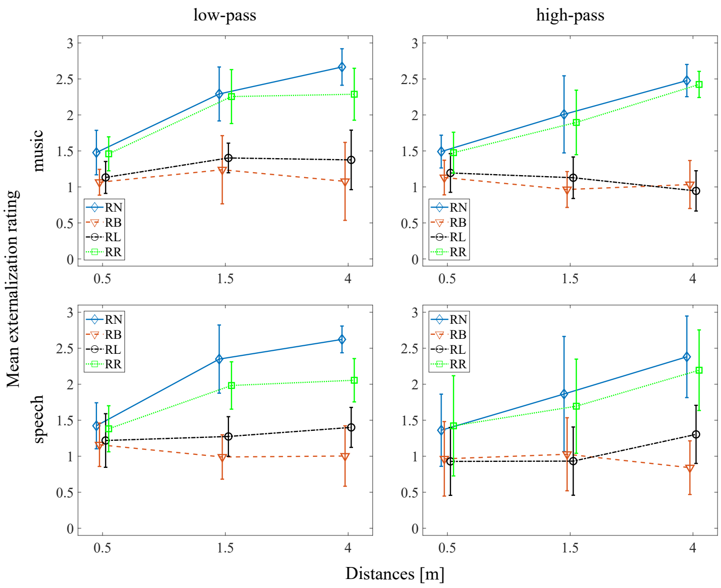

Figure 2: Mean externalization ratings of low- and high-pass filtered stimuli across listeners for "RN" (diamonds), "RB" (triangles), "RL" (circles) and "RR" (squares) conditions in different distances. The error bars indicate the standard deviations of the mean values.

from frontal to lateral incidence angles, but they did not change noticeably by increasing the source-listener distance. The ratings were overall lower or around 1.0, corresponding to sound sources being internalized or very close to subjects' heads. The differences between "RL" and "RR" conditions regarding the degree of perceived externalization were obvious by distances larger than 1.5 m and azimuth angles larger than 30°. In addition, the differences of externalization ratings between "RN" and "RR" conditions were more pronounced by the distance of 4 m than 0.5 m and 1.5 m.

For each azimuth angle, distance and stimulus, an analysis of variance (ANOVA) with post-hoc-test (5 % significance level with Bonferroni adjustment) was performed to show the significant differences of externalization ratings between "RL" and "RR" conditions. For the distance of 0.5 m, there was overall no significant difference between the results for these two conditions for both stimuli, apart from one exception at the azimuth angle of 60°, where externalization was significantly higher for the "RR" condition than the "RL" condition (p=0.01). For distances of 1.5 m and 4 m, the results for the "RL" condition differed significantly from the results for the "RR" condition for azimuth angles of 30°, 60° and 90°, apart from one exception for the speech stimulus at the azimuth angle of 30° with the distance of 1.5 m, where there was no significant differences between the results for the two conditions (p=0.06).

Figure 2 demonstrates the externalization ratings of low- and high-pass filtered stimuli for "RN", "RB", "RL", and "RR" conditions at 60° with different distances. Similar to the results showed in Figure 1, for both low- and high-pass filtered stimuli, the externalization ratings for "RN" and "RR" conditions increased with increasing the distances, while the externalization ratings for "RL" and "RB" conditions were almost constant and lower than those for "RN" and "RR" conditions. It can be seen that the difference of externalization ratings between "RN" and "RR" conditions for high-pass filtered stimuli was not as pronounced as that for low-pass filtered stimuli for distances larger than 1.5 m.

ANOVA shows no significant difference between the externalization results for "RL" and "RR" conditions in the case of the distance of 0.5 m. For distances of 1.5 m and 4 m, the results for the "RL" condition differed significantly from the results for the "RR" condition, apart from one exception for the high-pass filtered speech stimulus at the distance of 1.5 m, where there was no significant differences between the results for the two conditions (p=0.06).

Conclusions

The present study investigated the relative influence of reverberation at the contralateral versus ipsilateral ear on perceived externalization of virtual sound sources located at different azimuth angles with different distances without visual cues. The result of the listening experiment revealed that the reverberation at the contralateral ear had more influence on perceived externalization than that at the ipsilateral ear, especially for azimuth angles larger than 30° and source-listener distances larger than 1.5 m in the present study. This statement was also valid for the low- and high-pass filtered signals. In addition, this effect was more pronounced for high-pass filtered stimuli than for low-pass filtered stimuli for distances larger than 1.5 m.

References

[1] W. M. Hartmann and A. Wittenberg, "On the externalization of sound images," *The Journal of the Acoustical Society of America*, vol. 99, no. 6, pp. 3678–3688, 1996.

[2] S.-M. Kim and W. Choi, "On the externalization of virtual sound images in headphone reproduction: A wiener filter approach," *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3657–3665, 2005.

[3] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 904–916, 2001.

[4] F. Völk, "Externalization in data-based binaural synthesis: effects of impulse response length," in *NAG/Tagungsband Fortschritte der Akustik-DAGA 2009*, 2009.

[5] R. Crawford-Emery and H. Lee, "The subjective effect of brir length on perceived headphone sound externalization and tonal coloration," in *Audio Engineering Society Convention 136*. Audio Engineering Society, 2014.

[6] J. Catic, S. Santurette, and T. Dau, "The role of reverberation-related binaural cues in the externalization of speech," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 1154–1167, 2015.

[7] S. Li, R. Schlieper, and J. Peissig, "The effect of variation of reverberation parameters in contralateral versus ipsilateral ear signals on perceived externalization of a lateral sound source in a listening room," *The Journal of the Acoustical Society of America*, vol. 144, no. 2, pp. 966–980, 2018.

[8] S. Li, R. Schlieper, and J. Peissig, "The role of reverberation and magnitude spectra of direct parts in contralateral and ipsilateral ear signals on perceived externalization," *Applied Sciences*, vol. 9, no. 3, pp. 460, 2019.

[9] R. Hupke, M. Nophut, S. Li, R. Schlieper, S. Preihs, and J. Peissig, "The immersive media laboratory: Installation of a novel multichannel audio laboratory for immersive media applications," in *Audio Engineering Society Convention 144*. Audio Engineering Society, 2018.

[10] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.

[11] Z. Schärer and A. Lindau, "Evaluation of equalization methods for binaural signals," in *Audio Engineering Society Convention 126*. Audio Engineering Society, 2009.

[12] EBU, "3253: Sound quality assessment material recordings for subjective tests," 2008.