

Spatial Release from Masking with and without Simulated Hearing Loss: Comparison of Different Spatial Audio Reproduction Methods

Thomas Deutsch¹, Luigi Falanga², Iring Koch², and Janina Fels¹

¹ Institute for Hearing Technology and Acoustics, RWTH Aachen University, 52074 Aachen,

² Institute of Psychology, RWTH Aachen University, 52056 Aachen,

E-Mail: thomas.deutsch@akustik.rwth-aachen.de

Introduction

The term *cocktail party effect* describes the ability of human listeners to focus on a single sound source amidst a multitude of different sound sources. Contributing to this crucial ability is the so-called *spatial release from masking* (SRM) effect: Trying to understand a target speaker in the presence of an additional masking sound source becomes significantly easier when the speaker and the other sound source are spatially separated, compared to a situation in which the two are co-located. Related studies have shown that this useful advantage is reduced for listeners with hearing loss [1][2][3]. In our ongoing research project, we are investigating the performance of young, normal-hearing listeners compared to elderly listeners with age-related hearing loss in the context of auditory selective attention. The question whether or to what extent performance differences can be attributed to sensory hearing deficits or rather cognitive decline has not been fully resolved. For a series of listening experiments, our idea is to intentionally impair the hearing capabilities of young, normal-hearing participants by means of software for the simulation of cochlear hearing loss. If the sensory hearing capabilities of this participant group can be reduced to a level comparable to the elderly, any remaining performance differences between these two groups in the context of selective listening must be attributed to other (i.e., cognitive) effects, such as *general slowing*. Among our goals for the ongoing experiment presented here are the integration of hearing loss simulation into an SRM setup as well as answering the following research question: How will different degrees of simulated hearing loss affect spatial release from masking of young, normal-hearing listeners? For research related to auditory selective attention, (spatial) audio reproduction methods represent important tools. As different methods exhibit individual advantages and disadvantages, an additional goal is to compare a selection of relevant reproduction methods in the context of our study.

Method

The listening experiment design includes three independent variables: *reproduction method*, *noise direction*, and *degree of hearing loss*. Four different levels for *reproduction method* were defined: playback via real, discrete loudspeakers, Higher Order Ambisonics (HOA, 7th order, using the AllRAD algorithm [4]), and two variants of static binaural synthesis using generic head-related transfer functions (HRTFs) [5], reproduced via Sennheiser HD 650 headphones. The differences between the two binaural synthesis variants are outlined in more detail in the next subsection. Loudspeaker-based reproduction was realized using the surrounding spherical cap loudspeaker array “SCaLAR” [6]. This system consists of 68 loudspeakers installed in the fully anechoic

chamber at the Institute for Hearing Technology and Acoustics. Four different noise directions were defined, with real and virtual sound sources being positioned in the horizontal plane of the loudspeaker array at 0° (in front of the listener) and to the right of the listener at azimuth angles of 18°, 42°, and 90°.

Three levels were defined for *degree of hearing loss*: *no hearing loss*, *mild hearing loss*, and *moderate hearing loss*. The target hearing thresholds for the *mild* and the *moderate* conditions to be simulated by the hearing loss simulation were chosen to be representative of 70- and 80-year-olds, respectively, based on data provided in ISO 7029 [7]. Figure 1 shows the corresponding audiograms.

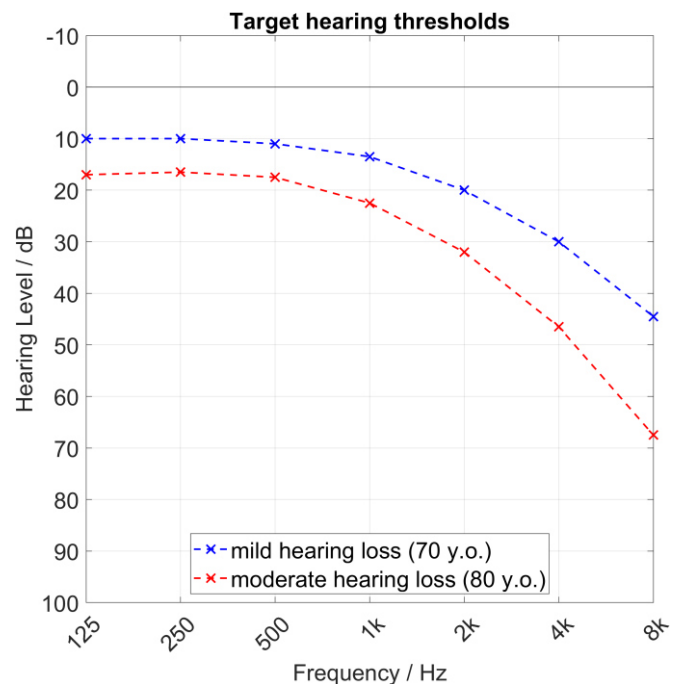


Figure 1: Audiogram for the target hearing threshold configuration of the *mild* and the *moderate* simulated hearing loss.

To quantify the SRM effect in a given context, the speech reception threshold (SRT) needs to be determined for both the co-located and the spatially separated case. The SRT is the specific signal-to-noise ratio (SNR) at which a listener can understand a predefined percentage (e.g., 50%) of presented speech material correctly. The same paradigm as the one presented in [8], which utilizes the digit triplet test as implemented by Viveros Munoz & Fels [9], is used to determine the SRT values. While the target speech is always presented from the front, the noise may be presented from either the front (co-located case) or one of the other directions (spatially separated cases). The level of the babble noise is fixed to 65 dB SPL (without hearing loss simulation) for all

source directions. Changes in SNR are realized by varying the level of the target speech across consecutive trials depending on participants' responses.

As dependent variables, the SRT is measured for each condition (to ultimately derive the SRM values). Furthermore, the response time is measured for each trial during testing.

Integration of hearing loss simulation

The signal processing software used for the simulation of cochlear hearing loss is the Wadai Hearing Impairment Simulator (WHIS) [10], which was selected based on the evaluation of hearing loss simulators presented in [11].

Figures 2 to 4 illustrate how the hearing loss simulator (HLS) is integrated into the signal chain for the different reproduction methods. Independent of the reproduction method used, a distinction needs to be made between the co-located case and the spatially separated cases: For the former, target and noise signals are superimposed prior to subsequent processing, while for the latter, the two signals are sent to separate instances of the following processing step. In the cases of the loudspeaker-based reproduction methods (see Figure 2), the HLS represents the first processing step. The (single-channel) output of each HLS instance is subsequently sent to a real loudspeaker (or a virtual loudspeaker when using HOA).

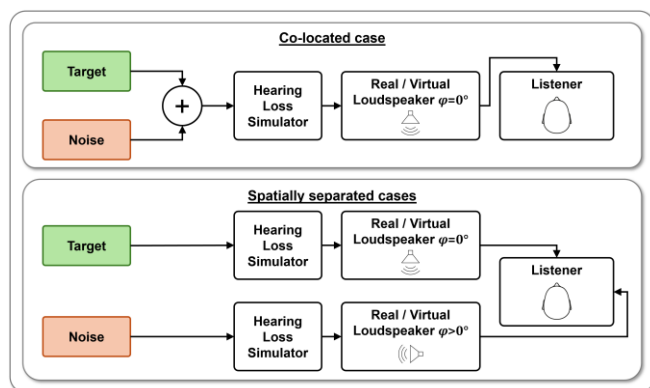


Figure 2: Signal flow diagram depicting the integration of hearing loss simulation for the loudspeaker-based reproduction methods (real loudspeaker and HOA).

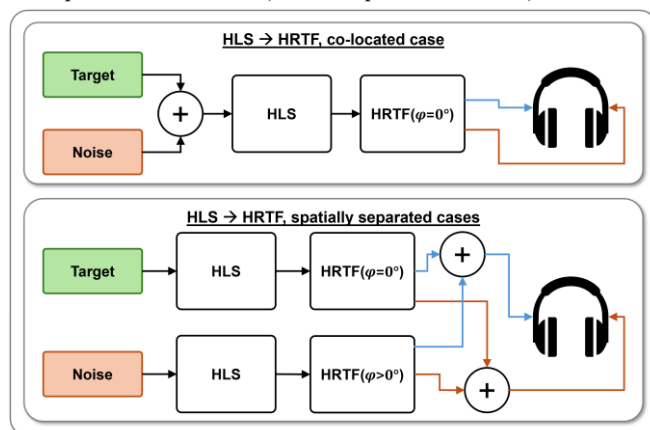


Figure 3: Signal flow diagram depicting the integration of hearing loss simulation for the first binaural synthesis variant (*HLS-HRTF*).

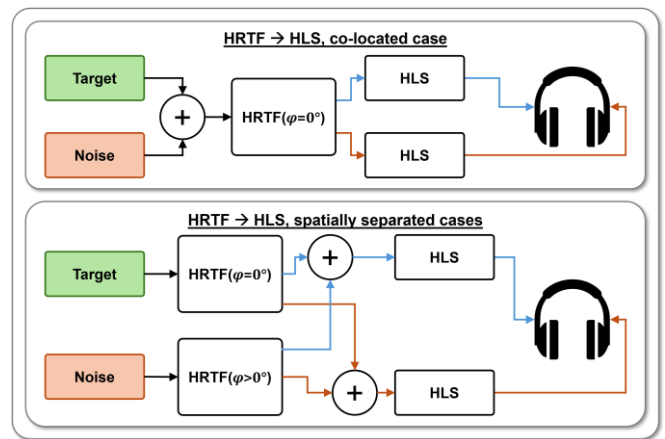


Figure 4: Signal flow diagram depicting the integration of hearing loss simulation for the second binaural synthesis variant (*HRTF-HLS*).

For the first binaural synthesis variant included in the experiment, *HLS-HRTF* (see Figure 3), the HLS also represents the first processing step. Spatialization is subsequently achieved by selecting and processing with the appropriate HRTFs for the desired sound source directions. In the second binaural synthesis variant, *HRTF-HLS* (see Figure 4), this order is reversed: The input signals are processed with the selected HRTFs first. This makes it possible to send the resulting left channel signal and right channel signal into two separate instances of HLS. For lateral sound source directions, the HRTF processing introduces interaural level differences (ILDs); in this variant, these ILDs can be considered by the non-linear (i.e., level-dependent) processing of the HLS appropriately. Thus, this method represents the most realistic one regarding the simulation of hearing loss among the included reproduction methods.

Participants and procedure

Participants are normal-hearing, young adults (hearing thresholds ≤ 20 dB HL in the range of 125 Hz to 8 kHz; age range 18 to 30 years). The initial pure-tone audiometry is followed by a headphone transfer function measurement used to create individual headphone equalization filters [12]. A calibration procedure for the optical tracking system is carried out to enable tracking of the head center and the interaural axis. Participants are familiarized with the task and the response interface during a short training phase. During the main part of the experiment, participants are seated on a chair in the center of the loudspeaker array. A headrest connected to the chair provides a reference point for the participant regarding head position and orientation. Conditions are presented in a counterbalanced order using a uniformly distributed random Latin square design. The duration of the experiment is approximately 2.5 hours, including preparations and breaks.

Preliminary results

Twelve participants (out of a determined total number of 24) have completed the experiment so far (8 female, 4 male, $M = 25.5$ y, $SD = 3.2$ y). For participant i and a particular condition j , the SRM benefit for noise direction φ is determined by

$$SRM(i, j, \varphi) = SRT(i, j, 0^\circ) - SRT(i, j, \varphi) \quad [\text{dB}] \quad (1)$$

Several trends can be identified from the preliminary descriptive statistics. Even for the smallest angular separation of 18° , a clear SRM benefit can be seen. When the angular separation is expanded to 42° , the SRM benefit increases considerably. An increase in the angular separation from 42° to 90° does not result in a clear increase in SRM values. No obvious pattern regarding reproduction method emerges – the preliminary results indicate an interaction between *reproduction method* and *degree of hearing loss*. Finally, a clear trend regarding the impact of *degree of hearing loss* can be observed: The preliminary results suggest that for all reproduction methods and all noise directions, the conditions with simulated hearing loss lead to lower SRM values compared to conditions without simulated hearing loss

Prior to response time analysis, differences in output latency between loudspeaker- and headphone-based reproduction are compensated. The preliminary results indicate slightly longer response times for trials of 0° conditions than for conditions of other noise directions. No clear tendencies for main effects or interactions of factors *reproduction method* and *degree of hearing loss* regarding response times can be observed thus far.

Discussion & Conclusion

Our preliminary results for the *no hearing loss* conditions closely match the results for the young, normal-hearing participant group in the SRM study by Duquesnoy [1]. The SRM benefit observed in [1] for the elderly participant group ($M=81.5$ y) with age-related hearing loss is even more reduced compared to the *moderate hearing loss* conditions of our study. This observation may be explained by the fact that for higher frequencies (i.e., 4 kHz and 8 kHz), the elderly participants in [1] exhibited hearing thresholds slightly higher (i.e., worse) than the thresholds defined for the *moderate hearing loss* condition of our study. With the decrease in SRM benefit observed for conditions with simulated hearing loss, our ongoing study already provides evidence that the performance deficits generally observed for elderly listeners can at least to a large part be explained by a sensory deficit (i.e., sensorineural hearing loss) – rather than a cognitive deficit.

Funding

This work is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) [project number 513033051].

References

- [1] Duquesnoy, A. J.: Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *JASA*, vol. 74, no. 3 (1983), 739-743
- [2] Gelfand, S. A., Ross, L., & Miller, S.: Sentence reception in noise from one versus two sources: effects of aging and hearing loss. *JASA*, vol. 83, no. 1 (1988), 248-256
- [3] Bronkhorst, A. W., & Plomp, R.: Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *JASA*, vol. 92, no. 6 (1992), 3132-3139
- [4] Zotter, F., & Frank, M.: All-round Ambisonic panning and decoding. *JASA*, vol. 60, no. 10 (2012)
- [5] Schmitz, A.: Ein neues digitales Kunstkopfmeßsystem. *Acta Acustica united with Acustica*, vol. 81, no. 4 (1995), 416-420
- [6] Pausch, F., Behler, G., & Fels, J.: SCaLAr – A surrounding spherical cap loudspeaker array for flexible generation and evaluation of virtual acoustic environments. *Acta Acustica*, vol. 4, no. 5 (2020)
- [7] ISO 7029:2017: Acoustics – Statistical distribution of hearing thresholds related to age and gender. ISO, (2017)
- [8] Deutsch, T., & Fels, J.: An assessment of Higher-Order Ambisonics reproduction in the context of spatial release from masking. *Proc. DAGA*, Stuttgart, Germany (2022), 985-988
- [9] Viveros Munoz, R., & Fels, J.: Digit triplet test. (2019)
- [10] Irino, T.: Hearing impairment simulator based on auditory excitation pattern playback: WHIS. *IEEE Access* (2023), 78419-78430
- [11] Deutsch, T., & Fels, J.: Evaluating models for the simulation of cochlear hearing loss. *Proc. DAGA*, Hamburg, Germany (2023), 200-203
- [12] Masiero, B., & Fels, J.: Perceptually robust headphone equalization for binaural reproduction. *AES 130th Convention*, London, UK (2011)