DOI: 10.61782/fa.2023.1118

VIRTUAL ACOUSTICS AND AUDIOLOGY: SPEECH INTELLIGIBILITY IN STANDARD SPATIAL CONFIGURATIONS AND IN A LIVING ROOM

Julia Schütze[1*] **Stephan D. Ewert**[1] **Christoph Kirsch**[1] **Birger Kollmeier**[1]
[1] Medizinische Physik and Cluster of Excellence Hearing4All, Universität Oldenburg, Germany

ABSTRACT

A mismatch between the outcomes of "classical" audiological tests and the perceived benefits of aided hearing impaired patients leads to an interest in more ecologically valid testing methods, more closely reflecting the patients' performance in real-world situations.

In daily-life, the living room constitutes a highly relevant indoor environment. Here, listening and speech communication typically involve different target or interferer sources, such as a TV set and/or talkers located on a chair or sofa in an adjacent room connected with an open door.

This study compares such an ecologically relevant condition to anechoic audiological test conditions by testing speech intelligibility (German matrix sentence test) in two standard spatial configurations with frontal target and either collocated or 90-degree-separated masker (S0N0 and S0N90). Dummy head recordings of a controllable laboratory environment resembling an average German living room with an adjacent kitchen were used. Target and interferer positions were permuted over four different positions, including an acoustically challenging target speaker in the adjacent kitchen with obstructed line of sight. Speech recognition thresholds were measured with normal hearing listeners. Speech intelligibility in the living room with a target in the coupled room is more difficult than the standard collocated S0N0 condition.

Keywords: *virtual acoustics, speech intelligibility, coupled room, room acoustics, ecological validity.*

1. INTRODUCTION

Hearing aids are designed to improve a person's ability to hear and communicate in a variety of settings, including noisy environments and social situations. However, if the fitting process does not involve ecologically valid testing conditions, the hearing aid might not perform as expected in real-word situations. Hearing aid wearers might experience a mismatch between the proposed benefit of hearing aids and the benefit in real-life situations [1]. Therefore, the integration of acoustically challenging daily-life situations in laboratory-based testing can extend the ecological validity in hearing research [2] and hearing aid fitting, leading to an overall improved hearing aid benefit in everyday situations and environments.

Virtual acoustic environments (VAE) offer the potential of more ecological valid testing in hearing research and audiology. Reproductions of real-world environments enable controllable and consistent ecologically valid measurements (e.g., [3]). VAEs have been used to increase the ecological validity in laboratory-based measurements in a number of studies [3 - 11]. Using VAEs, speech intelligibility has been measured in several simulated real-world conditions: in an office scenario [12], in a simulated cafeteria [3], in an underground station [13] and in a reverberant loft apartment [14], including the comparison to the comparable anechoic condition (represented by the visual rendering of a snowy outdoor environment).

Reverberation and masking by interfering sound sources are the two components that may affect speech intelligibility in challenging acoustic conditions. In real rooms, a complex mixture of direct sound, early reflections from different directions and late reverberation arrives at a listeners' ear.

*__Corresponding author__: julia.schuetze@uni-oldenburg.de.

When target and noise sources are spatially separated, the listener can use binaural cues like interaural time delay (ITD) and interaural level difference (ILD) to selectively attend to the target sound and suppress the noise. Spatial release from masking (SRM) is based on ITD and ILD cues and describes how much the masking effect is when the target and the noise source(s) are separated in space in comparison to co-located target and noise. SRM has important implications for many everyday situations, such as having a conversation in a crowded restaurant.

One highly relevant, typical acoustic communication environment for hearing impaired and normal hearing people is the living room [15-17]. Here, speech understanding might be interfered by other talkers or, e.g., the TV set, while acoustically challenging conditions might occur if the target talker is in an adjacent room with obstructed line of sight. Schulte et al. [17] assessed different challenging listening scenarios and found that the highest combined importance and occurrence ratings were given to following news broadcasts on TV, followed by situations where someone speaks to the listener from an adjacent room.

In the current study, we focus on a direct comparison of speech intelligibility measures with simulated acoustic reproductions of a real living room environment, established as "living room lab" in the university building [18, 19] and audiological standard spatial configurations, to investigate the effect of room acoustics on speech intelligibility.

2. METHODS

2.1 Listeners

Twelve normal hearing native German speakers (eight females, four males; aged 18–27 year) participated in the measurement. All participants were naïve to the test material. They had hearing thresholds for pure tones better than 20 dB HL at all audiometric frequencies between 125 and 8000 Hz. An hourly compensation was paid to the listeners for their participation in the measurement.

2.2 Acoustic environment and conditions

A laboratory environment, resembling a typical German living room, was built at the University of Oldenburg. This living room lab (LRL; see also [18], [19]) consists of the living room (18.8 m²) and an adjacent smaller kitchen (10m²), both connected by a door. In the living room is a seating group, a sofa table, a carpet, a TV and a TV board among other furniture. A floorplan of the lab environment is shown in Fig. 1.

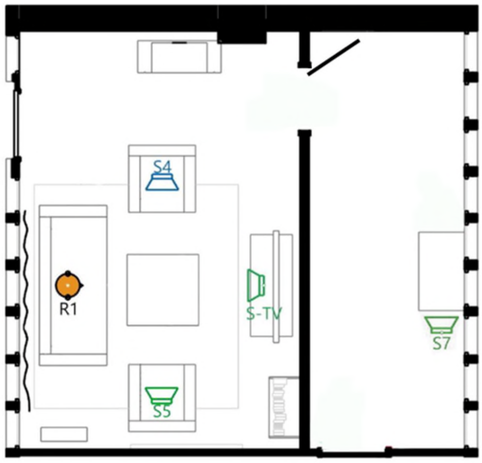

Figure 1. Floorplan of the living room lab at the University of Oldenburg. The receiver is positioned on the sofa, indicated by the head symbol R1, the masker is at position S4 (shown in blue), and the target is positioned either at STV, S5, or S7 (green).

The receiver was positioned at R1, seated on the couch. The target was presented from three different positions: from directly in front of the receiver (S-TV), from the right of the receiver on a chair (S5), or from the coupled room with obstructed line of sight (S7). An interfering sound source (masker) was presented from a fixed position on the sofa chair at S4.

Two standard audiological, anechoic spatial configurations with frontal target and either collocated or separated masker at a side (S0N0 and S0N90) were compared to the three more ecologically valid conditions in the living room.

Acoustical reproductions, based on measured binaural room impulse responses (recorded using a G.R.A.S KEMAR type 45BM head and torso simulator and Genelec 8030C loudspeakers; sampling rate 44.1 kHz) of the LRL were used in the measurement. The reverberation time of the living room lab with open door is T_{30} = 0.56 s and the early decay time is EDT = 0.46 s.

Table 1. Binaural levels for left and right ear (see left and right column, respectively) at receiver position R1 with the measured BRIRs convolved with the speech shaped OLnoise from the source positions STV, S4, S5 and S7.

Position	Level (dB)	
STV	65.02	64.97
S4	66.04	64.19

S5	64.60	65.76
S7	55.99	54.10

For the standard spatial configurations (S0N0 and S0N90), the MK2-database was used, which was obtained with the same measurement set-up as presented in [20].

2.3 Apparatus and procedure

Speech intelligibility (50% speech recognition threshold, SRT) was measured using the AFC framework [21] with the OLSA matrix sentence test [22] in the presence of a stationary speech-shaped noise (OLnoise) [23]. For the standard spatial configurations, the masker level was fixed at 65 dB SPL and the level of the target was varied depending on the number of the correctly identified words. For the living room, the STV source was calibrated to 65 dB at the receiver position (averaged across both ears; see Table 1). For the other source positions, the levels varied (see Table 1) according to the room acoustics. For the SRTs, signal-to-noise ratios (SNR) at the ears are reported.
The stimuli were presented via Sennheiser HD 650 headphones. The measurements were carried out in a double-walled, sound attenuating hearing booth. No head-tracking was used in the binaural rendering, therefore head movements were not considered.
Participants used a Matlab interface to mark the words (ten options for each word), which they had understood from the sentence. The participants started with a training condition, which was not used in the main listening test.

3. RESULTS

Results for the three conditions in the LRL and the two standard spatial audiological test conditions are represented in Fig. 2. The mean SRTs and the corresponding standard deviations, averaged across all 12 participants, are shown.
The lowest SRT score was obtained for S0N90. The highest SRT value was obtained for S7, the target position in the coupled room. A one-way, repeated-measures ANOVA (analysis of variance) revealed statistically significant differences between all measurement conditions, except for STV and S0N0.

4. DISCUSSION

With co-located target and interfering source, speech intelligibility is lower compared to conditions, where target and masker are spatially separated. This SRM has been earlier reported by several studies (e.g., [24-26]) and is also clearly visible in the current anechoic, audiological test conditions: S0N0 with collocated masker shows considerably higher SRTs compared to S0N90 indicating a profound SRM of about 9 dB.
Regarding the LRL conditions, the target position S5 has the lowest SRT of the three conditions, in line with the largest angular separation and distance between masker and target of the living room conditions. The target position STV is a bit more challenging in comparison to S5. This is mainly caused by the smaller spatial separation between STV and the masker position S4. The highest SRT, and therefore the most challenging condition, is S7. In this case no direct sound reaches the receiver position from the connected kitchen room, and only reflections and reverberated diffracted sound reach the receiver due to the occluded position of the source. Additionally, the angular separation between the door and the masker S4 masker position is quite small, observed from the receiver position. This likely reduces SRM for the sound components coming from the door opening. Moreover, S7 is the condition with the largest distance between receiver and target. A considerable decrease in speech intelligibility for room-in-room conditions has been reported [27], compared to single rooms with the same T_{30} and same total distance.
The audiological test condition S0N90 resulted in the lowest SRTs, indicating a large SRM and the anechoic presentation.

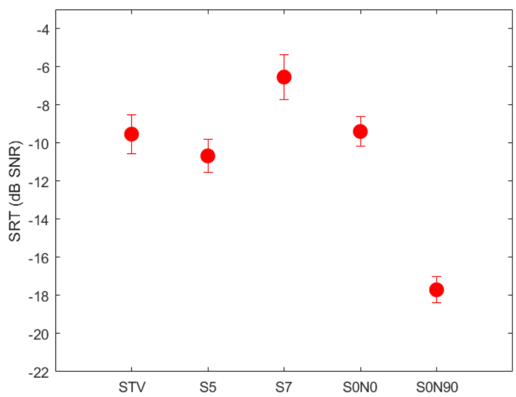

Figure 2. Average speech recognition thresholds and standard deviations across participants for the five different measurement conditions STV, S5, S7, S0N0 and S0N90.

5. CONCLUSION

In this study we compared SRTs from standard audiological test conditions in an anechoic environment with spatially collocated or separated target and masker to a more ecologically valid set-up with different spatial configurations in a living room. The reverberation of the (coupled) room increased the SRT scores. For the target in the adjacent kitchen room, with an obstructed line of sight, higher SRTs than for the unnatural anechoic colocated conditions were observed. The living room lab thus provides natural, challenging conditions requiring a relatively high signal-to-noise ratio at threshold. Such ecologically valid conditions are particularly relevant for assessing hearing aid performance.

6. ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) –Projektnummer 352015383 –SFB 1330 C4 and C5.

7. REFERENCES

[1] M. Cord, D. Baskent, S. Kalluri and B. C. Moore, "Disparity Between Clinical Assessment and Real-World Performance of Hearing Aids," Hear. Rev., vol. 14, pp. 2226 (2007 Jun.).

[2] G. Keidser *et al.*, "The Quest for Ecological Validity in Hearing Science: What It Is, Why It Matters, and How to Advance It," *Ear & Hearing*, vol. 41, no. Supplement 1, pp. 5S-19S, Nov. 2020, doi: 10.1097/AUD.0000000000000944.

[3] V. Best, G. Keidser, J. M. Buchholz, and K. Freeston, "An examination of speech reception thresholds measured in a simulated reverberant cafeteria environment," *International Journal of Audiology*, vol. 54, no. 10, pp. 682–690, Oct. 2015, doi: 10.3109/14992027.2015.1028656.

[4] Cubick J, Dau T. Validation of a Virtual Sound Environment System for Testing Hearing Aids. Acta Acustica united with Acustica. 2016. 102(3):547–57.

[5] Seeber, B.U., Kerber, S., Hafter, E.R. A System to Simulate and Reproduce Audio-Visual Environments for Spatial Hearing Research. Hearing Research 260(1-2), 2010.

[6] G. Grimm, S. Ewert, and V. Hohmann, "Evaluation of Spatial Audio Reproduction Schemes for Application in Hearing Aid Research," *Acta Acustica united with Acustica*, vol. 101, no. 4, pp. 842–854, Jul. 2015, doi: 10.3813/AAA.918878

[7] T. Weller, V. Best, J. M. Buchholz, and T. Young, "A Method for Assessing Auditory Spatial Analysis in Reverberant Multitalker Environments," *J AM ACAD AUDIOL*, vol. 27, no. 7, pp. 601–611, Jul. 2016, doi: 10.3766/jaaa.15109.

[8] A. Ahrens, M. Marschall, and T. Dau, "Measuring speech intelligibility with speech and noise interferers in a loudspeaker-based virtual sound environment," *The Journal of the Acoustical Society of America*, vol. 141, no. 5_Supplement, p. 3510, May 2017, doi: 10.1121/1.4987360.

[9] A. Westermann and J. M. Buchholz, "The effect of nearby maskers on speech intelligibility in reverberant, multi-talker environments," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 2214–2223, Mar. 2017, doi: 10.1121/1.4979000.

[10] V. Hohmann, R. Paluch, M. Krueger, M. Meis, and G. Grimm, "The Virtual Reality Lab: Realization and Application of Virtual Sound Environments," *Ear & Hearing*, vol. 41, no. Supplement 1, pp. 31S-38S, Nov. 2020, doi: 10.1097/AUD.0000000000000945.

[11] J. F. Culling, "Speech intelligibility in virtual restaurants," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2418–2426, Oct. 2016, doi: 10.1121/1.4964401.

[12] N. Mansour, M. Marschall, T. May, A. Westermann, and T. Dau, "Speech intelligibility in a realistic virtual sound environment," *The Journal of the Acoustical Society of America*, vol. 149, no. 4, pp. 2791–2801, Apr. 2021, doi: 10.1121/10.0004779.

[13] Ľ. Hládek, S. D. Ewert, and B. U. Seeber, "Communication Conditions in Virtual Acoustic Scenes in an Underground Station," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, Sep. 2021, pp. 1–8. doi: 10.1109/I3DA48870.2021.9610843.

[14] S. Fichna, T. Biberger, B. U. Seeber, and S. D. Ewert, "Effect of Acoustic Scene Complexity and Visual Scene Representation on Auditory Perception in Virtual Audio-Visual Environments," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, Sep. 2021, pp. 1–9. doi: 10.1109/I3DA48870.2021.9610916.

[15] F. Wolters, K. Smeds, E. Schmidt, E. K. Christensen, and C. Norup, "Common Sound

Scenarios: A Context-Driven Categorization of Everyday Sound Environments for Application in Hearing-Device Research," *J Am Acad Audiol*, vol. 27, no. 07, pp. 527–540, Jul. 2016, doi: 10.3766/jaaa.15105.

[16] A. Weisser *et al.*, "The Ambisonic Recordings of Typical Environments (ARTE) Database," *Acta Acustica united with Acustica*, vol. 105, no. 4, pp. 695–713, Jul. 2019, doi: 10.3813/AAA.919349.

[17] M. Schulte, M. Vormann, M. Meis, K. Wagener, B. Kollmeier, "Vergleich der Höranstrengung im Alltag und im Labor," *16. Jahrestagung der Deutsche Gesellschaft für Audiologie*, Rostock (2013).

[18] S. van de Par *et al.*, "Auditory-visual scenes for hearing research," *Acta Acust.*, vol. 6, p. 55, 2022, doi: 10.1051/aacus/2022032.

[19] J. Schütze, C. Kirsch, K. C. Wagener, B. Kollmeier, and S. D. Ewert, "Living room environment." Zenodo, Sep. 28, 2021. doi: 10.5281/zenodo.5747753.

[20] F. Brinkmann *et al.*, "A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations," *J. Audio Eng. Soc.*, vol. 65, no. 10, pp. 841–848, Oct. 2017, doi: 10.17743/jaes.2017.0033.

[21] S. D. Ewert, "AFC—A modular framework for running psychoacoustic experiments and computational perception models," in *Proceedings of the international conference on acoustics AIA-DAGA*, 2013, pp. 1326–1329.

[22] K. Wagener, T. Brand, and B. Kollmeier, "Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: Evaluation des Oldenburger Satztests," Zeitschrift für Audiologie. 1999. 38:86–95.

[23] K. C. Wagener and T. Brand, "Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters," *International Journal of Audiology*, vol. 44, no. 3, pp. 144–156, Mar. 2005, doi: 10.1080/14992020500057517.

[24] R. Plomp and A. M. Mimpen, "Improving the Reliability of Testing the Speech Reception Threshold for Sentences," *Int J Audiol*, vol. 18, no. 1, pp. 43–52, Jan. 1979, doi: 10.3109/00206097909072618.

[25] J. Peissig and B. Kollmeier, "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *The Journal of the Acoustical Society of America*, vol. 101, no. 3, pp. 1660–1670, Mar. 1997, doi: 10.1121/1.418150.

[26] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta Acustica united with Acustica*, vol. 86, no. 1, pp. 117–128, 2000.

[27] A. Haeussler and S. van de Par, "Crispness, speech intelligibility, and coloration of reverberant recordings played back in another reverberant room (Room-In-Room)," *The Journal of the Acoustical Society of America*, vol. 145, no. 2, pp. 931–944, Feb. 2019, doi: 10.1121/1.5090103.