Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures

Ahrens, Axel; Marschall, Marton; Dau, Torsten

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Ahrens, A., Marschall, M., & Dau, T. (2016). Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures. Poster session presented at 5th Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan, Honolulu, Hawaii, United States.

3pAAb2

Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures

Axel Ahrens, Marton Marschall, Torsten Dau
Hearing Systems group, Technical University of Denmark, Lyngby, Denmark
aahr@elektro.dtu.dk

HEARING SYSTEMS

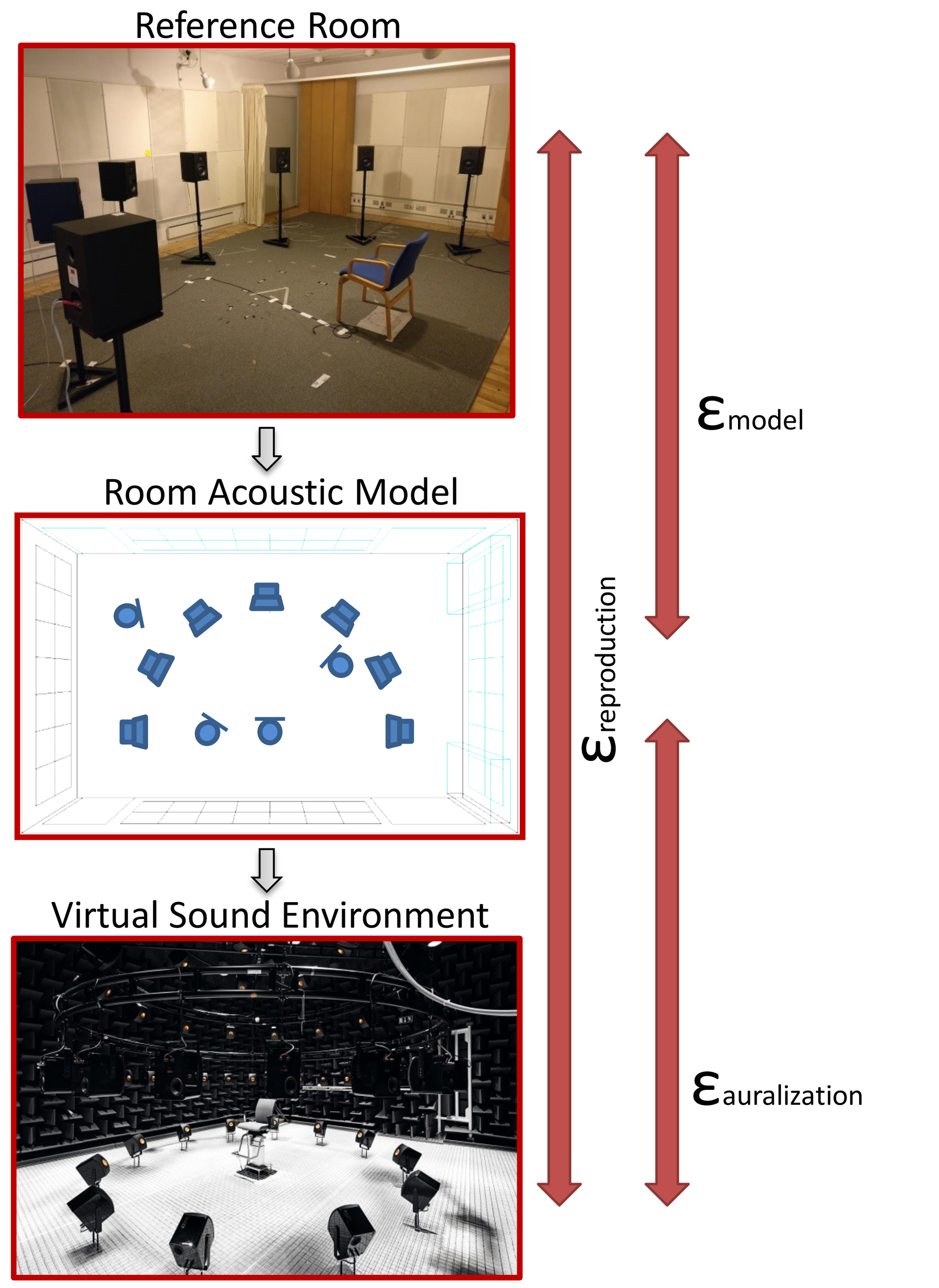
Introduction

To study human **auditory perception** in **realistic environments**, loudspeaker-based reproduction techniques have recently become state-of-the-art. To **evaluate the accuracy** of a simulation-based room auralization of a small room, **objective measures** were evaluated. In particular:

- early-decay time (**EDT**) & reverberation time (**T20, T30**)
- clarity (**C7, C50, C80**)
- interaural cross-correlation (**IACC**)
- speech transmission index (**STI**)
- direct-to-reverberant ratio (**DRR**)

Impulse responses (**IRs**) were measured in an IEC listening room. The room was then modeled in the room acoustics software ODEON, and the same objective measures were evaluated for auralized versions of the room. The auralizations were realized using higher-order ambisonics (**HOA**), mixed-order ambisonics (**MOA**), and a nearest-loudspeaker method (**NL**) and reproduced in a virtual sound environment.

Method



Reproduction techniques

- Nearest loudspeaker (**NL**; *Favrot&Buchholz, 2010*)
- Higher-order ambisonics (**HOA**, 5th order)
- Mixed-order ambisonics (**MOA**, 7th/5th order; *Daniel, 2000*)

Modeling

- ODEON v13.04 *(Rindel&Naylor, 1991)* model of IEC listening room (7.5*5.75*2.8m)
- Material properties optimized using ODEON's genetic material optimizer *(Christensen et al., 2014)*

IR recording

- 7 source positions (Dynaudio BM6)
- 4 receiver positions (B&K **4192** and B&K **HATS Type 4100**)
- Processing and analysis using ITA-toolbox and Two!Ears framework

Room Acoustic Measures

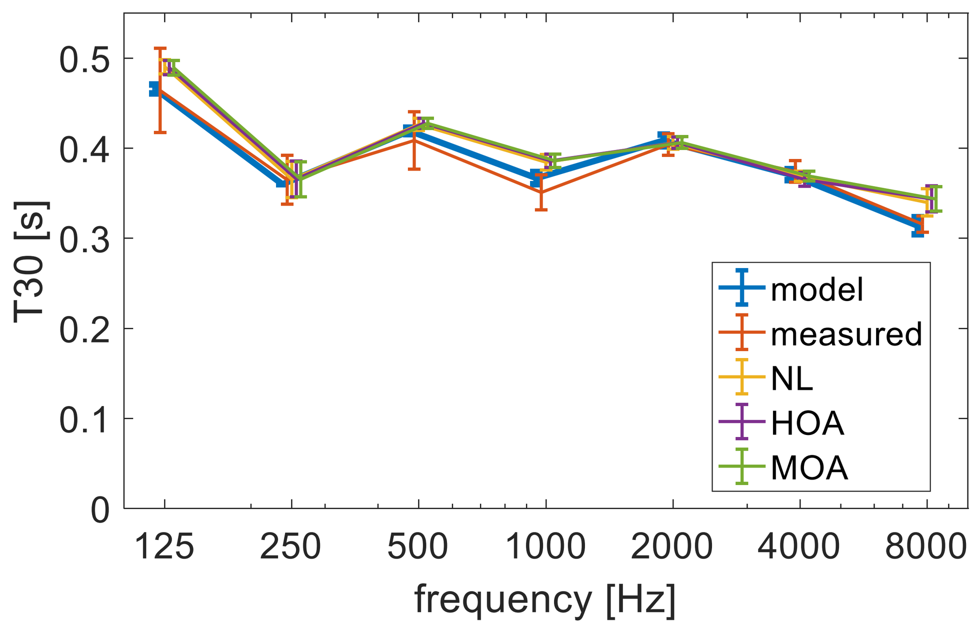


Figure 1: Reverberation time (T30; mean±standard deviation) over octave bands, measured at 7 source and 4 receiver positions. The blue and red curves indicate the ODEON model and the reference room, respectively. The remaining curves are auralized versions of the room.

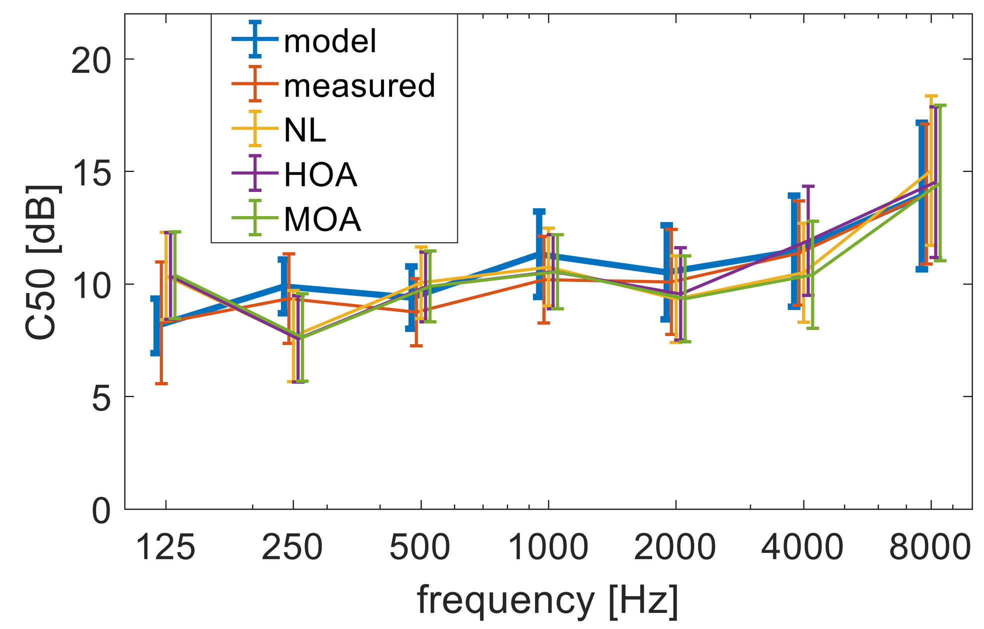


Figure 2: Clarity (C80; mean±standard deviation) over octave bands, measured at 7 source and 4 receiver positions. The blue and red curves indicate the ODEON model and the reference room, respectively. The remaining curves are auralized versions of the room.

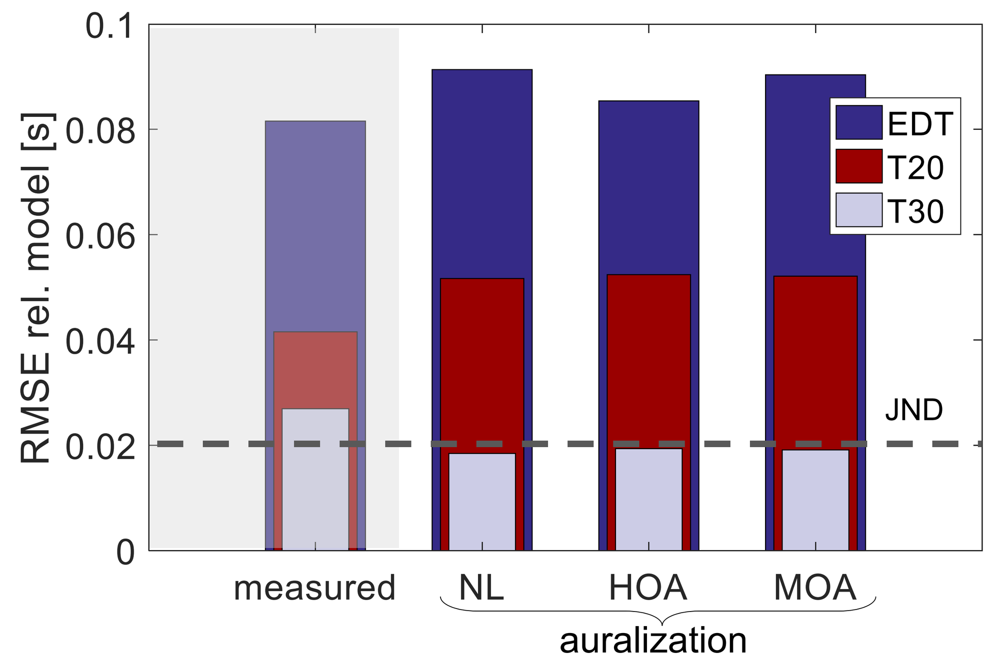


Figure 3: Root-mean-square error (RMSE) of the energy decay measures relative to the ODEON model. The dashed line indicates the perceptual just-noticable-difference for reverberation time and EDT (5%; Álvarez-Morales et al. (2016)).

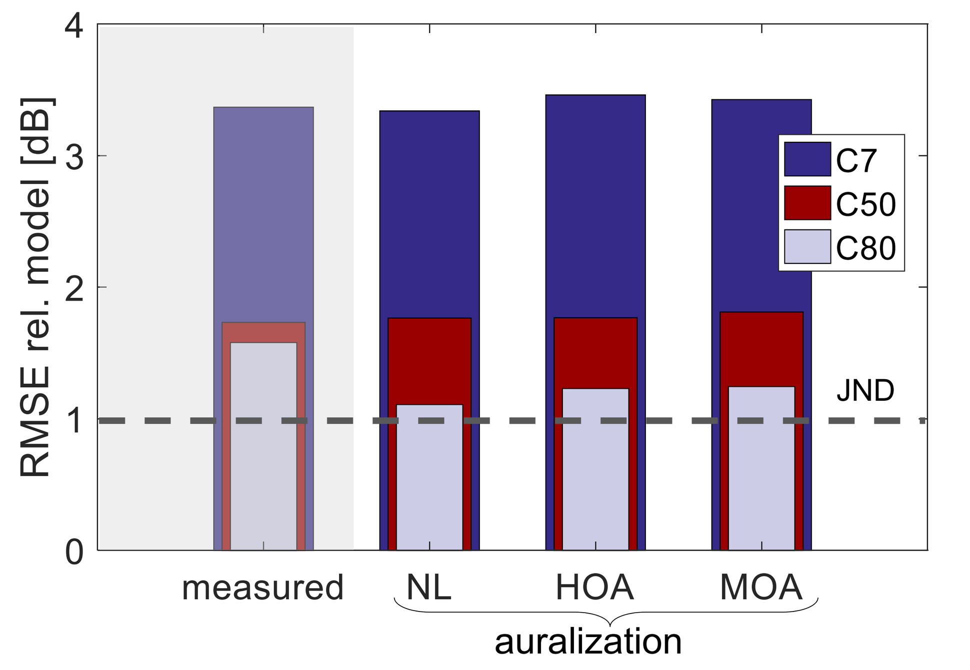


Figure 4: Root-mean-square error (RMSE) of the clarity measures relative to the ODEON model. The dashed line indicates the perceptual just-noticable-difference for clarity (1 dB; Álvarez-Morales et al. (2016)).

Binaural Measures

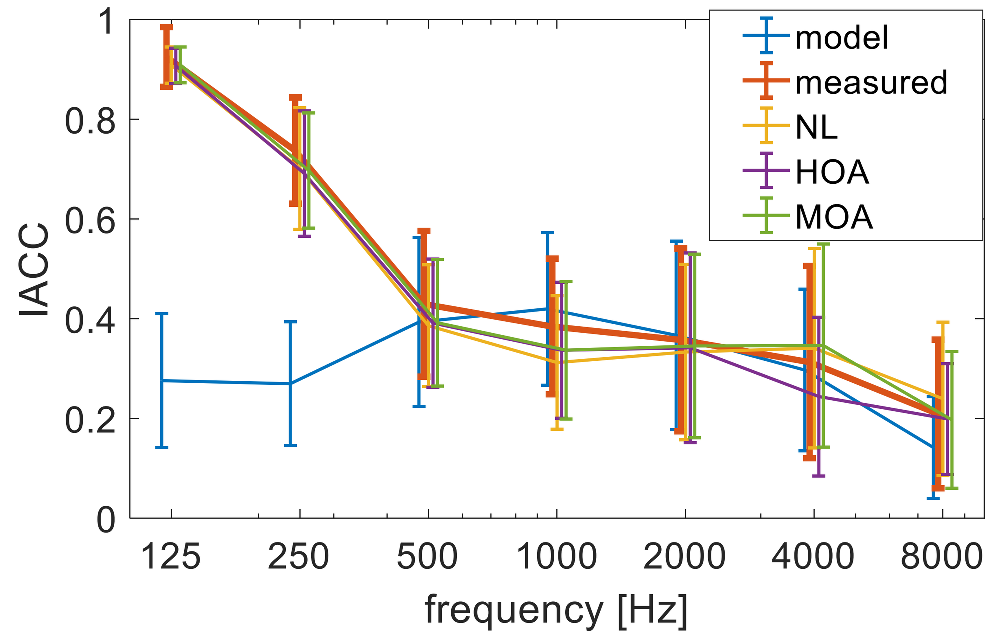


Figure 5: Interaural cross-correlation (IACC; mean±standard deviation) computed from the entire impulse response in octave bands, measured at 7 source and 4 receiver positions. The blue and red curves indicate the ODEON model and the reference room, respectively. The remaining curves are auralized versions of the room.

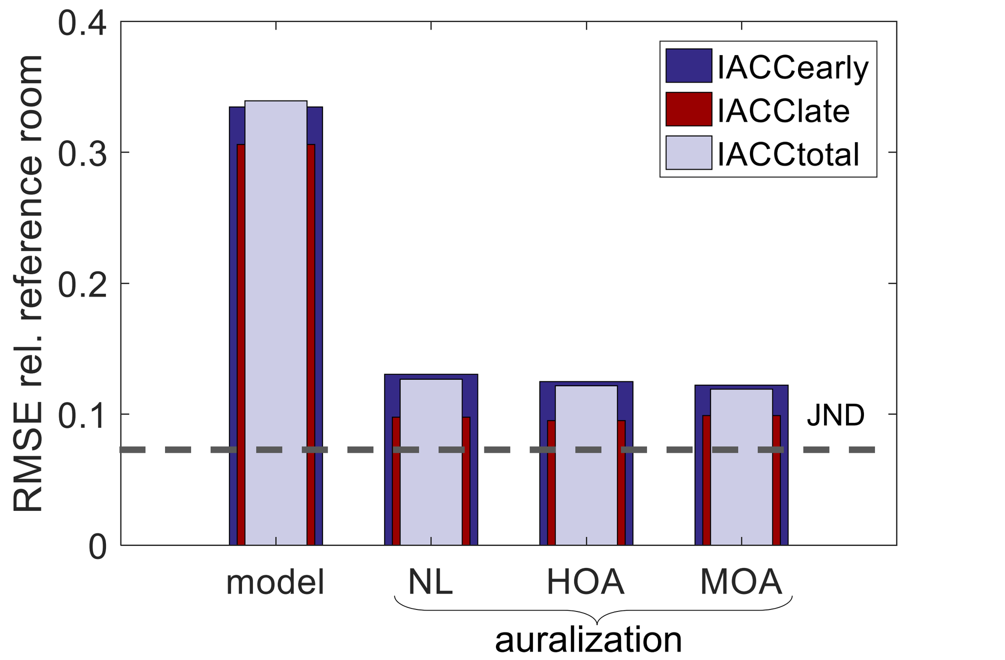


Figure 6: Root-mean-square error (RMSE) of the IACC measures relative to the reference room. The early IACC is calculated over the first 80ms of the impulse response. The late IACC from 80ms onwards. The dashed line indicates the perceptual just-noticable-difference for IACC (0.075; Álvarez-Morales et al. (2016)).

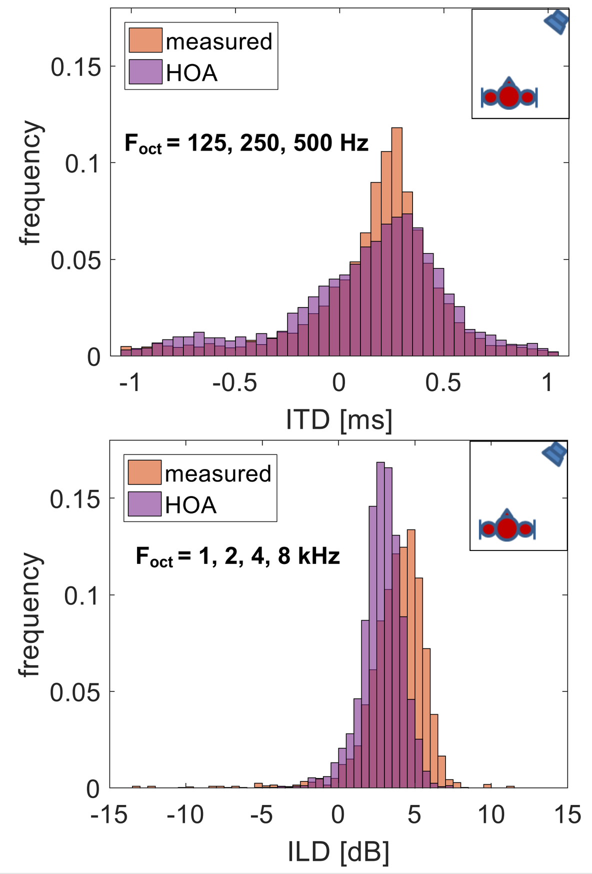


Figure 7: Distribution of short-term ITDs (top) and ILDs (bottom) calculated for a source 30° right of the HATS (see pictogram). Interaural differences were analyzed in 20ms windows with 50% overlap over a 10s long pink noise sample. The distributions were calculated over the indicated octave bands.

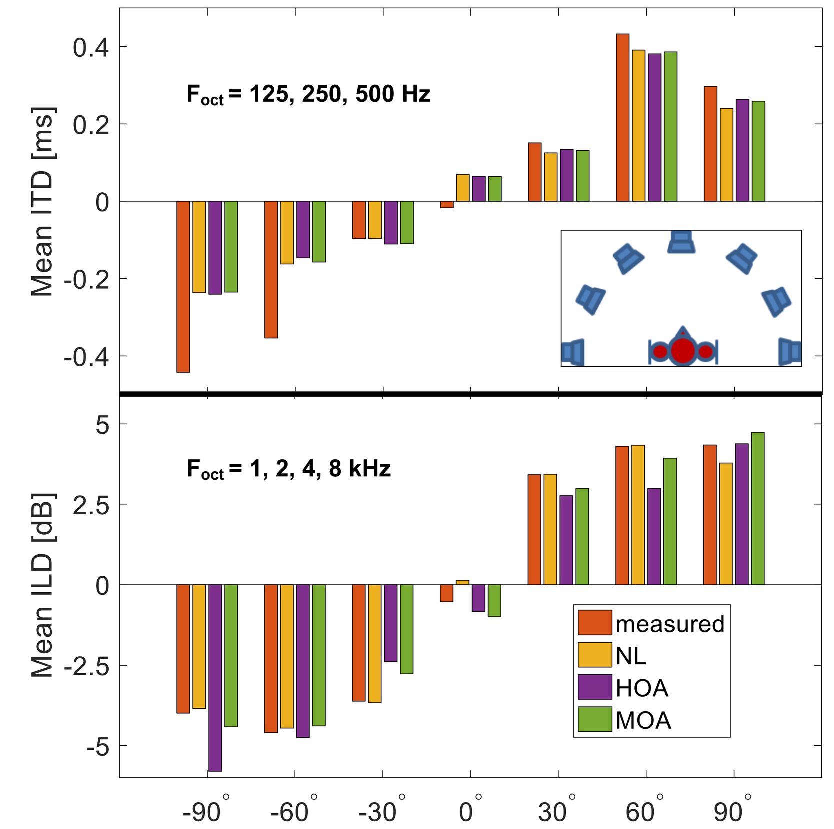


Figure 8: Mean ITDs and ILDs calculated from distributions as shown in Fig. 7 as a function of loudspeaker position and auralization technique. The same frequency ranges were applied. The pictogram depicts the source/receiver setup.

Binaural Direct-to-Reverberant Ratio

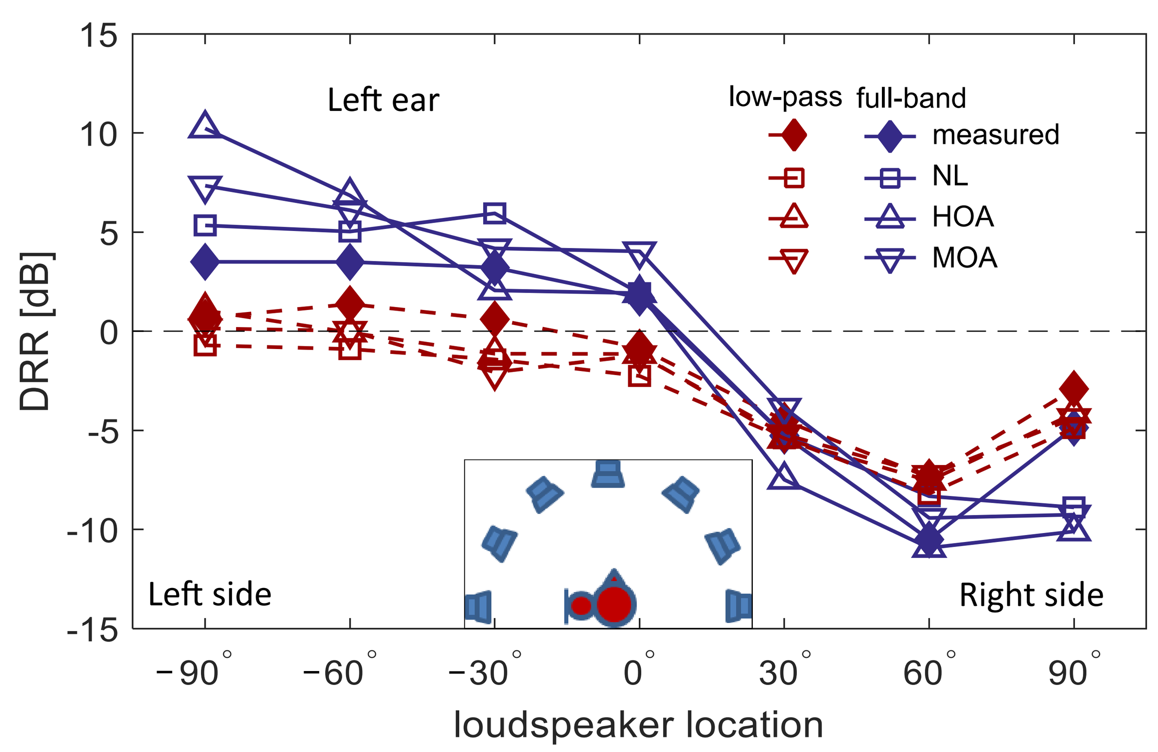


Figure 9: Direct-to-reverberant ratio (DRR) for the depicted source/receiver combinations, recorded with the left ear of a head-and-torso-simulator as a function of loudspeaker position. The blue markers depict the full-band DRR, the red markers the low-passed (cut-off frequency 2.73kHz) DRR. The closed and open symbols represent the DRR measurement from the room and the DRR of the auralization, respectively.

Speech Intelligibility and STI

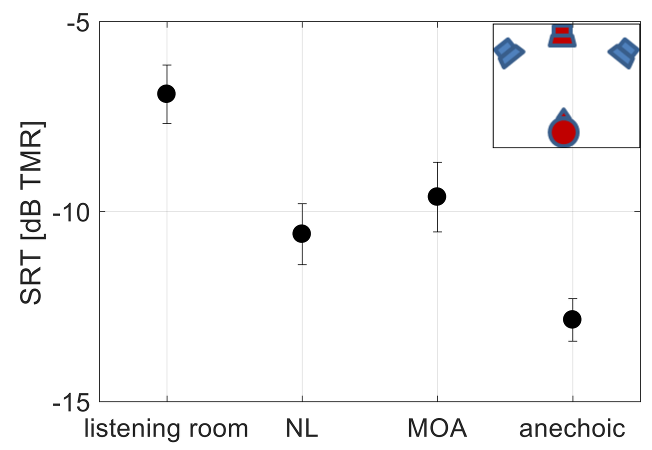


Figure 10: Speech reception thresholds (SRT) as target-to-masker-ratio (TMR) measured in the reference room, in the auralizations using NL/MOA and in an anechoic condition. The target speech was presented from 0° and 2 interfering talkers from ±30° (see pictogram). The material was the matrix sentence test Dantale II.

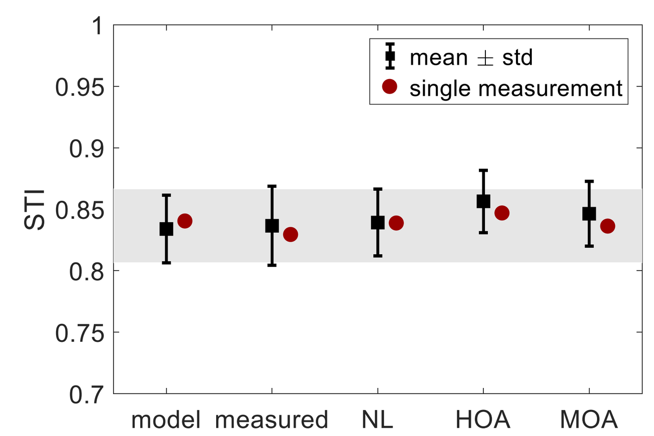


Figure 11: Speech transmission index (STI; mean±standard deviation) measured at 7 source and 4 receiver positions (black markers). The red markers indicate the STI between the listener and target speaker location as depicted in Fig. 10. The grey bar shows the perceptual just-noticable-difference around the measured STI (0.03; Álvarez-Morales et al. (2016)).

Conclusions

- **Long-term**, averaged measures are reproduced in the range of ~1 JND (T20/30, C50/80, STI, IACC)
- **Short-term** features of the impulse response are more difficult to capture leading to higher errors in e.g. EDT and C7
- Similar performances were obtained across **reproduction techniques**
- **Auralization errors** (auralization vs. model) are in the range of **modeling errors** (model vs reference)
- Dynamic **binaural cues** appear to be well captured
- **Perceptual differences** (e.g. speech intelligibility) occur, but not reflected in shown objective measures

  ⇒ Further investigations needed to link **perceptual differences** to **objective** measures

Literature

S. Favrot & J. M. Buchholz (2010), LoRA: A loudspeaker-based room auralization system, Acta Acustica united with Acustica, 96 (2).
J. Daniel (2000), Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia. PhD thesis. Université Paris 6.
J.H. Rindel & G.M. Naylor (1991), Odeon - A Hybrid Computer Model for Room Acoustic Modelling. 4th Western Pacific Regional Acoustics Conference, Brisbane. Proceedings.
C.L. Christensen, G. Koutsouris & J.H. Rindel (2014), Estimating absorption of materials to match room model against existing room using a genetic algorithm. Forum Acusticum, Poland.
L. Álvarez-Morales, M. Galindo, S. Girón, T. Zamarreño & R.M. Cibrián (2016), Acoustic Characterisation by Using Different Room Acoustics Software Tools: A Comparative Study, Acta Acustica united with Acustica, 102 (3).

Acknowledgements

The authors would like to thank the developers of the ITA-toolbox from RWTH Aachen as well as the developers of the Two!Ears Framework. Additionally, thanks to Oticon Research Centre Eriksholm, for providing the speech material of the Dantale II.
The project was partly funded by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS).