CARL VON OSSIETZKY universität OLDENBURG

SFB 1330
Hearing Acoustics

Spatial Speech Intelligibility Map Rendering
for Hearing Device Users with TASCAR, openMHA, and FADE

Marc René Schädler
Medizinische Physik and Cluster of Excellence Hearing4all
University of Oldenburg

CONCLUSIONS

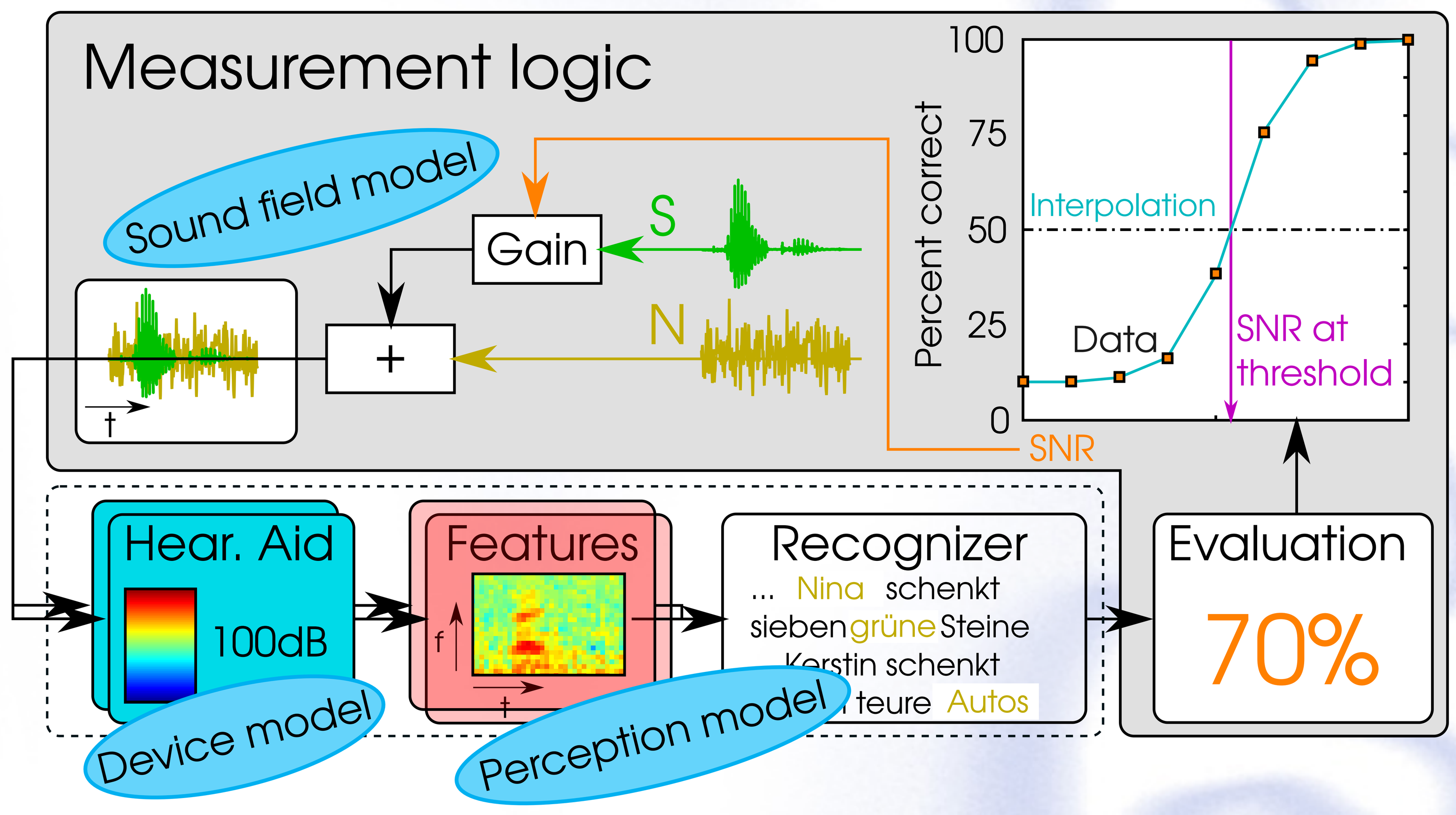
Position-dependent speech recognition thresholds (SRTs) for hearing device users can be plausibly simulated by linking simulation models of A) acoustic scenes, B) hearing devices, and C) human speech recognition.

Binaural interaction in human speech perception should be considered, and was sucessfully modelled in the feature extraction stage of the speech recognition model.

The spatial representations can be intuitively interpreted but should be related to the speech levels that normal-hearing listeners would actually employ.

The framework might be used for hearing device fitting.

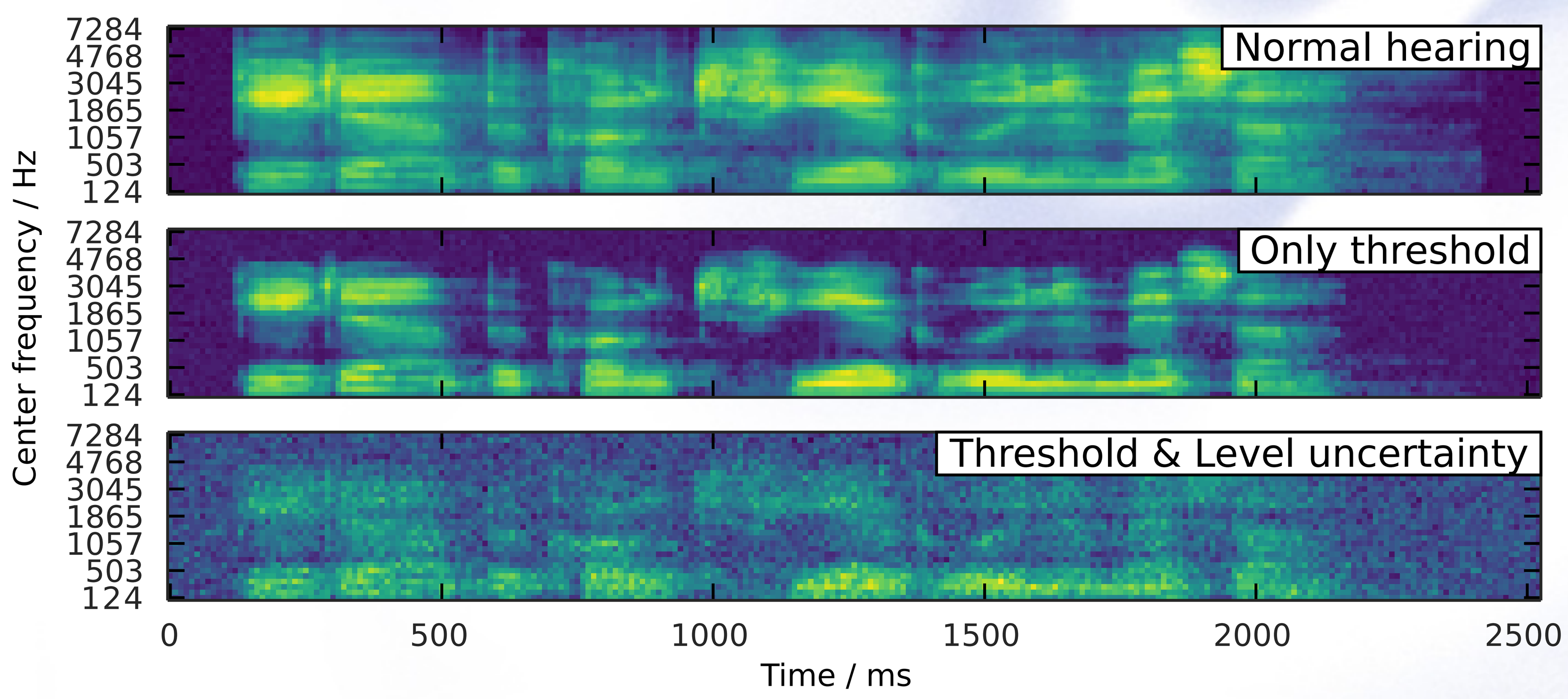
METHODS



Sound field model:
Toolbox for Acoustic Scene Creation and Rendering (TASCAR) [1] with KEMAR head-related impulse responses from the OlHeaD-HRTF database [2].

Hearing device model:
Open Master Hearing Aid (openMHA) [3] with adaptive differential microphones and multi-band dynamic compressor.

Perception model:
Framework for auditory discrimination experments (FADE) [4] with a modified feature extraction stage.



DISCUSSION

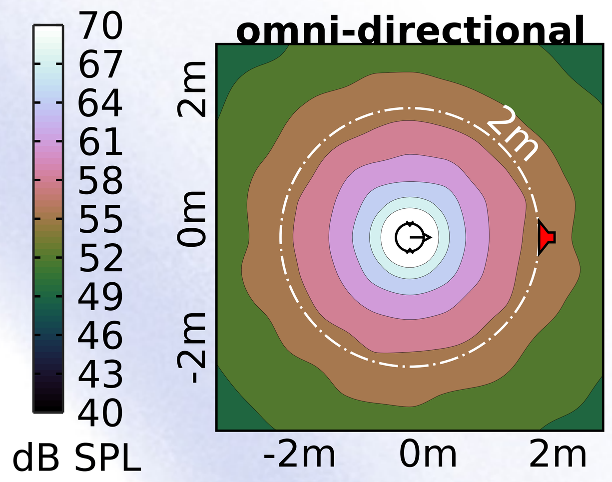
A major challenge when linking the models was to correctly implement the physical links between them, e.g., from the sources to the eardrums or device microphones, and from the device loudspeakers to the eardrums, where the hearing threshold was applied.

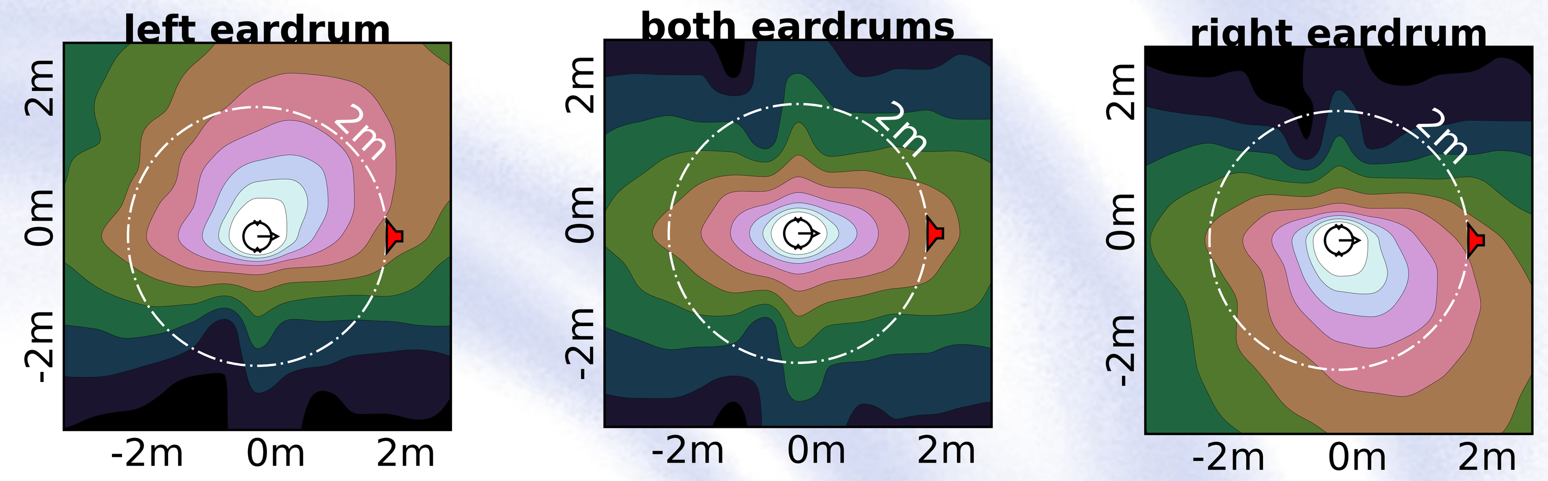
The resulting SRTs in dB SPL need to be interpreted in terms of the speech level that normal hearing listeners would use intuitively in that condition.

More scenes need to be implemented and the model needs to be validated against the benefits of listeners.

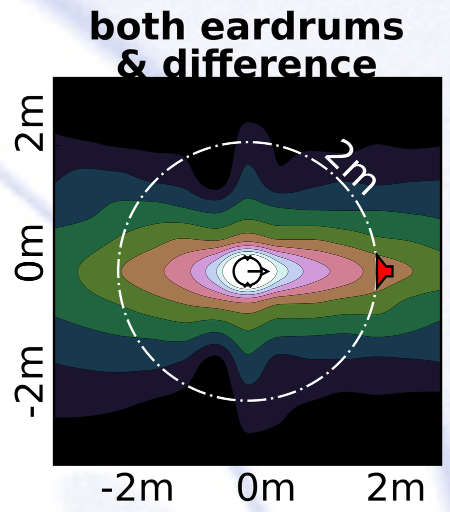
RESULTS: binaural anechoic condition

The unaided normal hearing listener in the center looks to the right. The depicted levels indicate the required speech level from the loudspeaker position to achieve 50% correct speech recognition when a stationary noise source (ICRA1 [5]) with 65dB SPL is located at that position.
An omni-directional microphone yields an isotropic pattern, with a required speech level of about 57dB SPL, i.e., -8 dB SNR, at the loudspeaker distance.

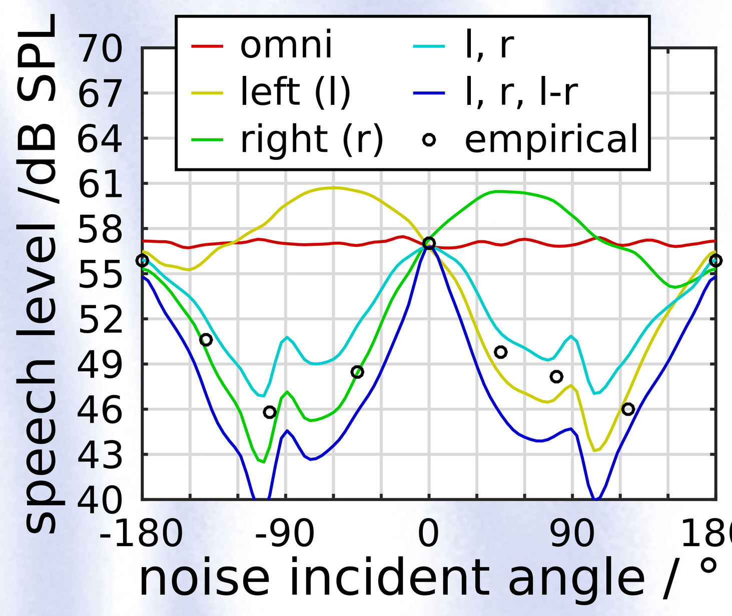




Using dummy-head(KEMAR)-related impulse responses from [2] to simulate the signals at the ear drums, the head shadow becomes visible. The speech recognition model can integrate both feature vectors almost to better-ear performance.

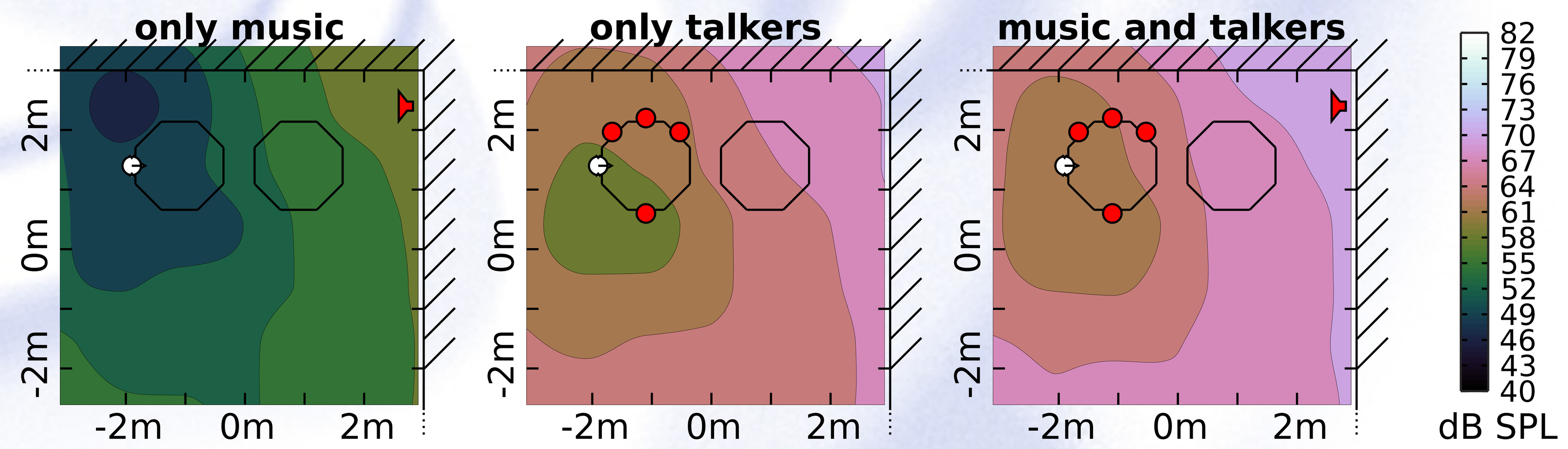


If the difference of the left and right feature vectors is used as additional feature, recognition performance can be increased to reach empirical performance from [6].

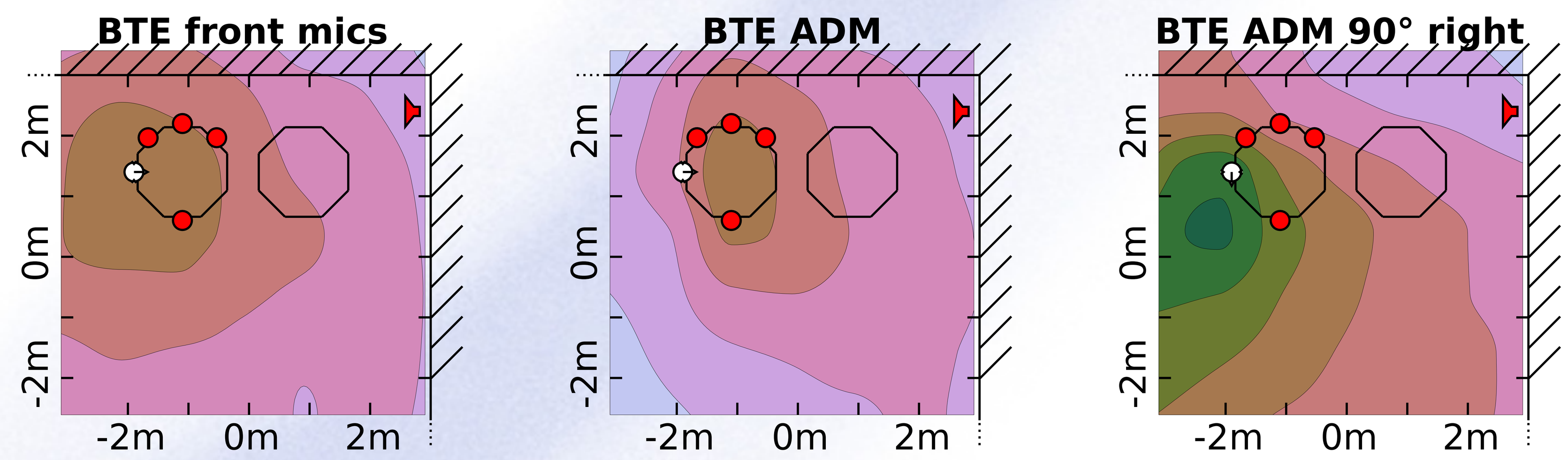


RESULTS: complex acoustic conditions

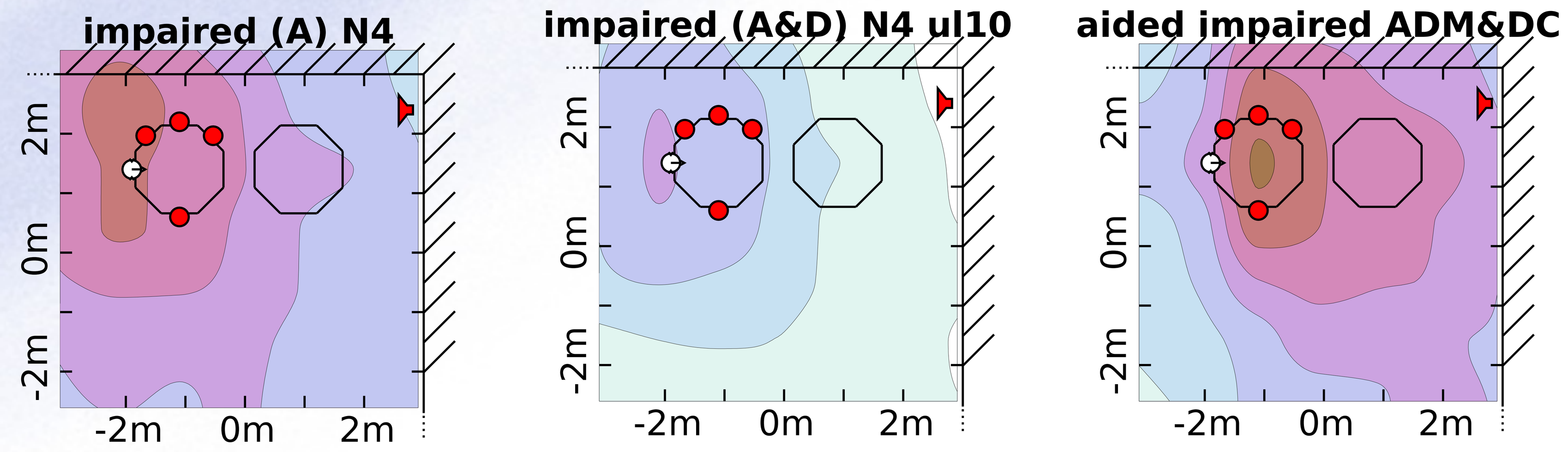
It is more intuitive and practical to interpret a scene with fixed interferers and a variable target speaker position. The depicted level is the required speech level from that position to achieve 50% correct speech recognition.



Here, the music (saxophone) and concurrent talkers (different languages) add up to a scene where the required speech level is about 61 dB SPL.



Adaptive differential microphones (ADM) show only a small frontal effect, however, head movements can help to improve speech recognition.



Here, a combined hearing impairment with elevated hearing thresholds according to Bisgaard profile N4 [7] and a supra-threshold component of hearing loss with a level uncertainty of 10 dB can be aided with ADM and dynamic compression coording to NAL-NL2, according to the simulation.

References
[1] G. Grimm, J. Luberadzka, T. Herzke, V. Hohmann (2015): Toolbox for acoustic scene creation and rendering (TASCAR) - Render methods and research applications. In: Proceedings of the Linux Audio Conference, Mainz, 2015
[2] F. Denk S.M.A Ernst, S.D. Ewert, B. Kollmeier (2018): Adapting hearing devices to the individual ear acoustics: Database and target response correction functions for various device styles. Trends in Hearing, vol 22, pp. 1-19
[3] T. Herzke, H. Kayser, F. Loshaj, G. Grimm, V. Hohmann (2017): Open signal processing software platform for hearing aid research (openMHA). Proceedings of the Linux Audio Conference. Université Jean Monnet, Saint-Étienne, pp. 35-42
[4] M. R. Schädler, A. Warzybok, S.D. Ewert, B. Kollmeier (2016): A simulation framework for auditory discrimination experiments: Revealing the importance of across-frequency processing in speech perception. JASA, Volume 139, Issue 5, pp. 2708–2723
[5] W.A. Dreschler, H. Verschuure, C. Ludvigsen, S. Westermann (2001): ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. Int. J. Audiol. 40(3), pp. 148–157
[6] R. Beutelmann, T. Brand (2006): Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. JASA, Volume 120(1), pp. 331-342.
[7] N. Bisgaard, M.S. Vlaming, M. Dahlquist (2010): Standard audiograms for the IEC 60118-15 measurement procedure. Trends in amplification, 14(2), pp. 113-120
[8] G. Keidser, H. Dillon, M. Flax, T. Ching, S. Brewer (2011). The NAL-NL2 prescription procedure. Audiology research, 1(1)

https://github.com/m-r-s

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 352015383 – SFB 1330 A 3
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 352015383 – SFB 1330 A 3.