Prediction of speech intelligibility based on a correlation metric in the envelope power spectrum domain

Relano-Iborra, Helia; May, Tobias; Zaar, Johannes; Scheidiger, Christoph; Dau, Torsten

Publication date:
2017

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

Link back to DTU Orbit

Citation (APA):
The sEPSMcorr model includes audibility thresholds, such that sensitivity loss can be predicted independently of hearing sensitivity, even when hearing sensitivity is similar. Therefore, the predictive power of the sEPSMcorr back-end was further investigated in combination with a more realistic model of the auditory processing of speech [4]. Here, the speech-based CASP (sCASP) was used to predict intelligibility based on a correlation metric in the envelope power spectrum domain.

The CASP model offers more flexibility to model hearing impairments, beyond the audiogram, due to the Dual Resonance Non-linear Filterbank (DRNL). [5]. The model has been shown to account for psychoacoustic data from individual HI subjects.

The models were evaluated in conditions with:
- Speech mixed with stationary or non-stationary interferers: Speech shaped noise (SSN), which was also used to fit the model; Amplitude modulated SSN (SAM) with f_{c,mod} = 8 Hz and modulation depth of 1. The speech level is non-stationary, but non-stationary international speech signal test signal (SI2) was used.
- Noisy speech in the presence of reverberation: T_{60} = 0, 0.4, 0.7, 1.3 and 2.3 s
- Noisy speech subject to different types of non-linear processing - Ideal Binary Mask processing (IBM) with four interferers.

The results were compared to human data as well as to fitting parameters for the CASP model. The fitting parameters were optimized for each condition, and the model was tested on different types of noise and interferers.

The sEPSMcorr model is specifically designed to capture the effects of hearing loss on speech intelligibility. It includes audibility thresholds that are independent of hearing sensitivity, allowing for more accurate predictions of intelligibility under various conditions. The model has been further enhanced by combining it with a more realistic auditory processing model, the sCASP, which accounts for the effects of non-linear processing such as reverberation and noise.

The results of the fitting process showed that the sEPSMcorr model can provide better predictions of speech intelligibility than the CASP model, especially in conditions with reverberation and non-linear processing. This makes the sEPSMcorr model a valuable tool for predicting speech intelligibility in real-world hearing loss scenarios.

To investigate the model's ability to account for individual hearing impairments, the parameters of the CASP model were optimized for each subject. This allowed for a more personalized prediction of speech intelligibility, which is crucial for improving speech communication in individual hearing loss conditions. The model parameters were determined for each subject, and the model was tested on different types of noise and interferers.

The results showed that the sEPSMcorr model can provide better predictions of speech intelligibility than the CASP model, especially in conditions with reverberation and non-linear processing. This makes the sEPSMcorr model a valuable tool for predicting speech intelligibility in real-world hearing loss scenarios.

The model can now serve as a foundation for the development of a HI model, since the DRNL-based framework allows for fitting to individual hearing loss conditions. This will provide more accurate predictions of speech intelligibility, which is crucial for improving speech communication in individual hearing loss conditions.

Summary of results

The sEPSMcorr model provides similar (and in some conditions better) results than the CASP model. This model can now serve as a foundation for the development of a HI model, since the DRNL-based framework allows for fitting to individual hearing loss conditions.

The model can now serve as a foundation for the development of a HI model, since the DRNL-based framework allows for fitting to individual hearing loss conditions. This will provide more accurate predictions of speech intelligibility, which is crucial for improving speech communication in individual hearing loss conditions.

Outlook

- Investigate the model's ability to account for individual hearing impairments using the parameters available in the CASP framework.
- Consider additional processing stages that could account for inner hair-cell loss and auditory nerve deafferentation [Sumner et al., 2002; Lopez-Poveda and Barron, 2013]
- Determine the conditions on which the HI model will be tested with special focus on supra-threshold distortions that might be challenging for HI subjects.