Prediction of binaural speech intelligibility in normal-hearing and hearing-impaired listeners: a psychoacoustically motivated extension

Anna Warzybok[1,2], Jan Rennies[2,3], Thomas Brand[1,2], Birger Kollmeier[1,2]

[1] *Medical Physics, Department of Medical Physics and Acoustics, University of Oldenburg, 26121 Oldenburg, E-Mail: anna.warzybok@uni-oldenburg.de, thomas.brand@uni-oldenburg.de, birger.kollmeier@uni-oldenburg.de*

[2] *Cluster of Excellence "Hearing4all", 26129 Oldenburg*

[3] *Fraunhofer IDMT Project Group Hearing, Speech and Audio Technology, 26129 Oldenburg, E-Mail: rns@idmt.fraunhofer.de*

Introduction

Understanding speech in complex environments is a challenge for most hearing-impaired (HI) listeners. For normal-hearing (NH) listeners, the auditory system takes advantage of early reflections in a room by integrating them with the direct sound and thereby increasing the effective speech level [1, 2]. The present study investigates the temporal and spatial processing of a single speech reflection in different noise conditions in normal-hearing and hearing-impaired listeners. The experimental data are compared to model predictions to investigate to what extent reductions in speech intelligibility and binaural processing can be explained by the pure tone audiogram [3]. A model extension is tested which accounts for the deficit in the detection of interaural phase differences in hearing-impaired listeners.

Method

EXPERIMENT: Speech reception thresholds (SRTs) were measured adaptively using the Oldenburg Sentence Test [4]. The binaural speech intelligibility tests were performed to examine the integration of a single, frontal speech reflection with frontal direct target sound with respect to reflection delay (0, 50, and 200 ms) and type of interferer (frontal, diffuse, lateral). To simulate different directions of the reflection and noise the clean speech and noise signals were convolved with binaural head-related impulse responses.

12 NH and 6 HI listeners with moderate, symmetrical sensorineural hearing loss participated in the experiments. Fig.1 shows the mean pure tone thresholds of the HI listeners.

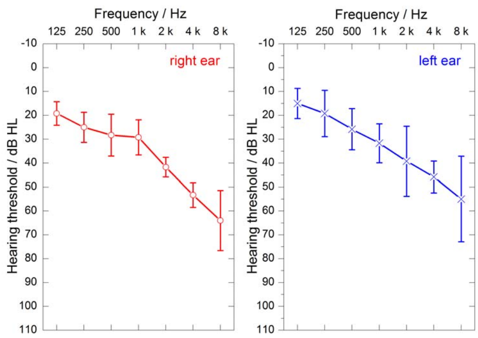

Figure 1: The mean audiograms for the right (red) and left (blue) ear of the hearing-impaired listeners.

To account for other deficits than loss of audibility psychoacoustic measurements were conducted including the upper frequency limit for detecting a 180° phase shift in a tone. The measurements were conducted using an AFC Toolbox [5].

MODEL: An extended model of binaural speech intelligibility accounting for the distinction between useful and detrimental reflections before the binaural stage [3] was used to predict the SRTs in experimental conditions for both groups of listeners. The first stage of the model applies a gammatone filter bank to analyse the binaural speech and noise signals. Internal noise is derived from the individual audiograms and added to the external noise to account for hearing impairment. An independent equalization-cancelation (EC) mechanism [6] in each frequency band is applied to take advantage of spatially separated speech and noise sources. This mechanism aims at optimizing the signal-to-noise ratio (SNR) by amplifying and delaying the left and right ear signals relative to each other followed by subtraction of the two channels from each other. The resulting SNRs are then used as input to the Speech Intelligibility Index (SII) from which SRTs are calculated.

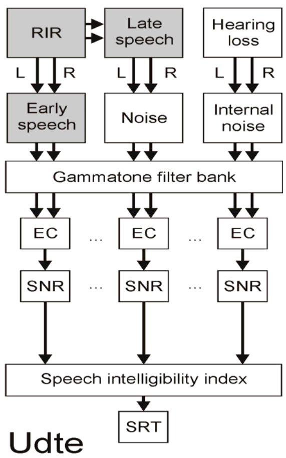

Figure 2: Schematic diagram of the binaural speech intelligibility model by Rennies et al. [3].

Results

EXPERIMENT: The mean SRTs averaged across listeners and corresponding standard deviations (SD) measured in conditions with frontal direct sound and frontal reflection are shown in Fig. 3 for NH (left) and HI (right) listeners as a function of reflection delay. The bottom panel of Fig. 3 shows the binaural intelligibility differences (BILDs).

The main findings were that the HI listeners (i) did not benefit as much as NH listeners from the spatial separation of the interferer from the direct sound and (ii) were more affected by a late reflection arriving from the same direction as the direct sound in all noise conditions.

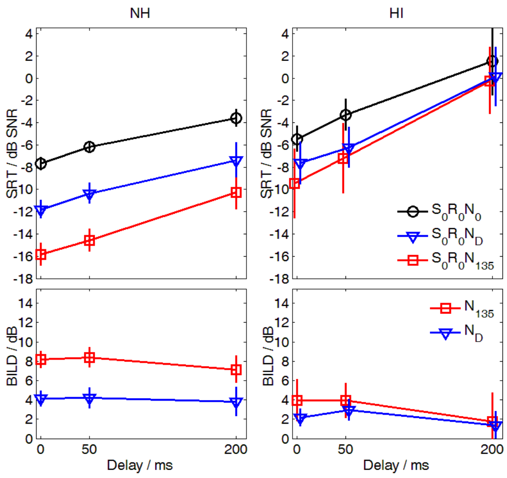

Figure 3: Mean SRTs with corresponding SD (top) and binaural intelligibility level differences (BILD, bottom) as a function of reflection delay in diotic (squares), diffuse (circles), and lateral noise (triangles) for NH (left) and HI listeners (right).

MODEL: The correlation of measured and predicted data for NH listeners is shown in Fig. 4. The prediction accuracy with an R^2 of 0.95 is close-to-optimum.

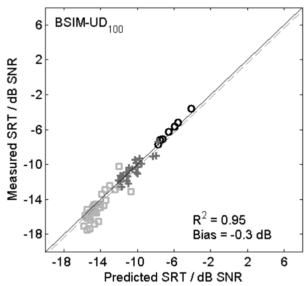

Figure 4: Measured vs. Predicted SRTs for NH listeners. Different gray scales represent data measured in diotic (black), diffuse (dark gray), and in lateral noise (light gray).

The model structure proposed in [3] was extended by accounting for suprathreshold deficits in HI listeners. The HI listeners were able to detect the 180° phase shift in a tone up to a carrier-frequency of 710 Hz on average, which is almost two times lower than in NH listeners. This disability was modeled by modifying the delay error which reduces the influence of the interaural cross correlation at high frequencies and can be interpreted as a very simple model of the decreasing phase coherence on the auditory nerve toward high frequencies. The comparison of the measured and predicted data using both versions of the model is shown in Fig. 5. As expected, the extended model did not change the predictions in diotic and diffuse noise conditions (Fig. 5, left and middle, respectively). However, an improvement was observed in the lateral noise condition.

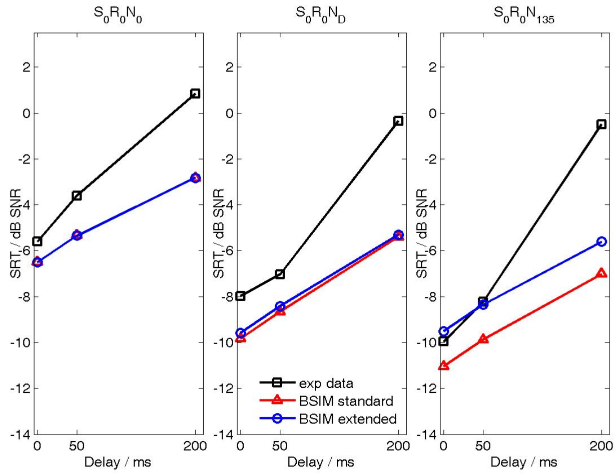

Figure 5: Comparison of measured data (black), predicted data using the model of Rennies et al. (3; red) and by the extended version accounting for suprathreshold deficit (blue) data as a function of reflection delay in diotic (left), diffuse (middle), and lateral noise condition (right).

Conclusions

In order to improve binaural speech intelligibility predictions in hearing-impaired listeners, an extended model was proposed accounting for deficit in the detection of interaural phase differences. The extended version improved the predictions of binaural effects in lateral noise. However, the large detrimental effect at long reflection delays cannot be modeled by accounting for binaural phase processing deficit.

References

[1] Lochner J. and Burger J.: The influence of reflections on auditorium acoustics. J. Sound Vib. 1(1964), 426-454.

[2] Warzybok A., Rennies J., Brand T., Doclo S. Kollmeier B.: Influence of a single, early reflection on speech intelligibility under different noise conditions. J. Acoust. Soc. Am. 133(1) (2013), 269-282.

[3] Rennies J., Warzybok A., Brand T., Kollmeier B.: Modeling the effects of a single reflection on binaural speech intelligibility. J. Acoust. Soc. Am. 135(3) (2014), 1556-67.

[4] Wagener K., Brand T., Kollmeier B., Entwicklung und Evaluation eines Satztests für die Deutsche Sprache III: Evaluation des Oldenburger Satztests. Z. Audiol. 38, 89-95.

[5] Ewert S.D.: AFC - A modular framework for running psychoacoustic experiments and computational perception models. In proceedings of AIA-DAGA, Merano, Italy (2013), pp. 1326-1329.

[6] Durlach N.I., Equalization and cancelation theory of binaural masking-level differences. J. Acoust. Soc. Am 35, 1206-1218.