Predicting the benefit of binaural cue preservation in bilateral directional processing schemes for listeners with impaired hearing

THOMAS BRAND[1,*], CHRISTOPHER F. HAUTH[1], KIRSTEN C. WAGENER[2], AND TOBIAS NEHER[1,3]

[1] *Medizinische Physik and Cluster of Excellence "Hearing4All", Universität Oldenburg, Oldenburg, Germany*

[2] *Hörzentrum Oldenburg GmbH and Cluster of Excellence "Hearing4All", Oldenburg, Germany*

[3] *Institute of Clinical Research, University of Southern Denmark, Odense, Denmark*

Linked pairs of hearing aids offer various possibilities for directional processing providing adjustable trade-off between improving signal-to-noise ratio and preserving binaural listening. The benefit depends on the processing scheme, the acoustic scenario, and the listener's ability to exploit binaural cues. Neher *et al.* (2017) investigated candidacy for different bilateral processing schemes for 20 elderly listeners with symmetric and 19 age matched listeners with asymmetric hearing thresholds below 2 kHz. The acoustic scenarios consisted of a frontal target talker presented against two intelligible or unintelligible speech maskers from ±60° azimuth. In this study, the speech reception threshold (SRT) data were compared to predictions of the binaural speech intelligibility model (BSIM; Beutelmann *et al.*, 2010), which was used to model pure better-ear-glimpsing as well as additional binaural unmasking. The speech intelligibility index (SII), which served as backend of BSIM, was calibrated to an individual reference value at the SRT for each listener. This reference value mirrors the amount of acoustical information needed by the listener to achieve the SRT and correlated with the listeners' ability to process temporal fine structure. BSIM revealed a benefit due to binaural processing in well-performing listeners when processing provided low-frequency interaural timing cues.

INTRODUCTION

Due to wireless across-device links, bilateral processing schemes have become applicable in commercial hearing aids (HA). This allows improving the signal-to-noise-ratio (SNR) by exploiting interaural differences between target speech and interferers. This mimics the human binaural auditory system that is known to exploit interaural differences for binaural release from masking. Some bilateral processing schemes sacrifice interaural differences for the sake of SNR improvement. However,

*Corresponding author: thomas.brand@uol.de

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 6: Adaptive Processes in Hearing, August 2017, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C.-Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. The Danavox Jubilee Foundation, 2017.
 ISBN: 978-87-990013-6-1.

for the individual listener with impaired hearing it is unclear in advance if binaural release from masking provided externally by bilateral HA processing is beneficial or if the listener is able to achieve the same benefit using his or her own binaural processing. The latter has the advantage that binaural cues are preserved, enabling more natural listening including localisation and source separation. Neher *et al.* (2017) investigated the suitability of different bilateral directional processing schemes for listeners with hearing impairment and for both intelligible and unintelligible speech maskers. As the hearing loss of all listeners was compensated for by providing amplification in accordance with the NAL-R prescription rule, supra-threshold (e.g., binaural) processing played a major role.

Neher *et al.* used the binaural intelligibility level difference (BILD) for assessing the listeners' binaural processing abilities for speech in noise. The BILD is defined as the difference between two speech reception threshold (SRT) measurements obtained for a speech source located at 0° azimuth in presence of a noise source located at 90° (or −90°) azimuth. For the first SRT only the ear which benefits from the head shadow is used, for the second SRT both ears are used which enables binaural processing. Neher *et al.* found that listeners with BILDs larger than about 2 dB showed a larger benefit from preserved binaural cues at low frequencies compared to greater SNR improvement achieved by the beamforming algorithms used in their study. The opposite was true for listeners with smaller BILDs. Audiometric asymmetry reduced the influence of binaural hearing only slightly. Furthermore detection performance of an interaurally phase inverted 500-Hz sinusoid in interaurally coherent noise (N0Sπ) was an effective predictor of the benefit from preserved low-frequency binaural cues.

In this study, the data of Neher *et al.* (2017) were reanalysed using the binaural speech intelligibility model (BSIM) of Beutelmann *et al*. (2010), which combines the equalization-cancellation (EC) process (Durlach, 1963) as a model of the effective binaural processing with the speech intelligibility index (SII; ANSI, 1997) to predict binaural speech intelligibility in different acoustic scenarios. This model also considers the individual hearing status by taking the audiogram into consideration. BSIM can be used to analyse the relative contribution of binaural processing and better-ear-glimpsing (i.e., using in each frequency channel and each time frame the ear with the better SNR; Brungart and Iyer, 2013) on the predicted SRTs. Therefore, it can also be used to investigate whether or not individual listeners rely mainly on their better-ear to understand speech and to which extent they benefit from their own binaural processing. Furthermore it is evaluated if BSIM is able to predict the correlations to BILD and binaural masking level difference (BMLD) found by Neher *et al.* (2017).

The main research questions of this study were: (1) Is BSIM applicable to intelligible and unintelligible speech maskers? (2) Does binaural processing play a role in aided patients or does better-ear-glimpsing explain most of the benefit in spatial listening conditions? (3) Is BSIM able to separate the benefit due to processing of interaural differences of temporal fine structure from the benefit due to better-ear-glimpsing?

METHOD

Data

The data described in Neher *et al.* (2017) were used. Twenty elderly listeners (age: 63-80 years) with symmetric hearing thresholds (PTA4: 52 dB HL) and 19 elderly listeners (age: 62-80 years) with asymmetric hearing thresholds (PTA4: 53 dB HL) below 2 kHz took part in the experiment. Listeners were matched for age, hearing loss, and selective attention. Furthermore the listeners were split into a group with high BILD and a group with low BILD in this study.

The aided SRTs were measured using the Oldenburg sentence test (OLSA; Wagener *et al.*, 1999) with a frontal target talker and two speech maskers located at ±60° azimuth. To create the spatial arrangement of target and interfering signal, the signals were convolved with head related impulse responses (HRIRs; Kayser *et al.*, 2009), recorded with the microphones of two behind-the-ear (BTE) hearing aid dummies, which were equipped to a head-and-torso simulator.

The first masker with high informational masking (IM) consisted of Oldenburg sentences uttered by another male talker. The second masker with low IM was generated by transforming the unintelligible international speech test signal (ISTS; Holube *et al.*, 2010) to male pitch and vocal track length.

Bilateral directional processing simulating a linked pair of completely occluding BTE HAs was applied. The processing schemes differed in the trade-off between SNR improvement and binaural cue preservation. Scheme "pinna" simulated the directivity of the human pinna without any bilateral processing. Scheme "beamfull" simulated a bilateral beamformer steered towards the frontal direction which sacrificed all interaural cues to improve the SNR.

Binaural speech intelligibility model (BSIM)

The BSIM of Beutelmann *et al.* (2010) is shown in Fig. 1. It uses a gammatone filterbank with 30 frequency channels ranging from 143 to 8346 Hz to separate the input signals into different frequency bands. The individual hearing loss is considered after peripheral filtering by adding uncorrelated noises to the left and right ear to the interfering noise. In each band, an independent EC process according to Durlach (1963) is performed. Durlach's model assumes that the left and right ear signals are subtracted after an equalization of interaural level and time differences. As such, this approach implicitly requires an analysis of the temporal fine structure in the auditory system. Normally distributed internal processing errors are assumed, which limit the EC processing to be performed effectively at frequencies below 1500 Hz; At higher frequencies BSIM effectively performs better-ear-glimpsing.

The final step in the BSIM framework is the transformation of the SII value to an SRT value. This transformation belongs to the SII concept (ANSI, 1997) and defines the SII value representing the amount of acoustical information (expressed in the form of a frequency-weighted SNR) which is required to understand 50% of the speech, i.e., to reach the SRT. This SII reference is dependent on the speech material. The model's processing is performed in 23-ms time frames, and the final SRT is calculated by averaging the short-time SRTs across time, which takes into account that the used speech maskers fluctuate over time. To test the hypothesis that better-ear-glimpsing alone can explain the binaural benefit, BSIM was used both in “better-ear-glimpsing only” mode and in equalization-cancellation (EC) mode, respectively. In the first mode, in each 23-ms time frame and each auditory frequency band the SNR from the better ear was used for intelligibility prediction. In the EC mode, BSIM additionally incorporates interaural processing of the left and right ear signals according to the EC model.

An extension of BSIM incorporating individual internal processing errors derived from individual BMLD measurements was introduced by Hauth *et al.* (2017). Neher *et al.* (2017) measured BMLDs at 500 and 1000 Hz. The possible improvement of the individualized internal processing errors was evaluated by comparing the original BSIM (see Figs. 2 to 4) with the extended BSIM (see Fig. 5).

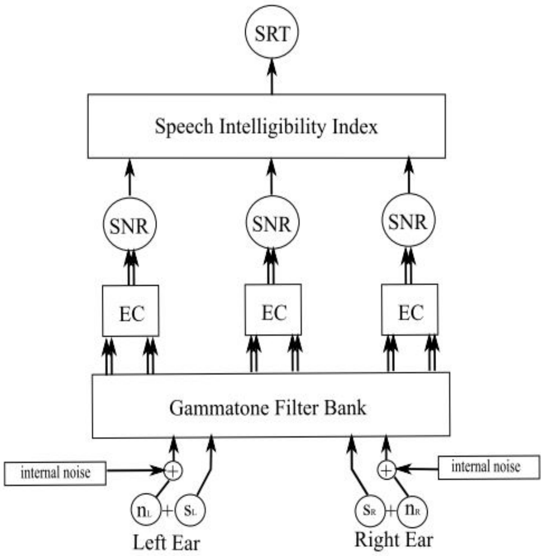

Fig. 1: Binaural Speech Intelligibility Model (BSIM) according to Beutelmann *et al.* (2010).

RESULTS AND DISCUSSION

A first analysis showed that the SII reference varied strongly across listeners. First, a fixed SII criterion was chosen to predict SRTs. This approach successfully predicted the effect of the audiogram on SRTs in hearing-impaired listeners (Beutelmann *et al.*, 2010). However, the stimuli analyzed in this study were amplified according to the audiogram (NAL-R) and audibility played a minor role. As a consequence, using a fixed SII criterion did not successfully predict the individual SRT in this study. Instead, it can be assumed that the observed variability of SRTs is mainly caused by supra-threshold and cognitive processing differences of the individual listeners. In order to test this hypothesis, one individual SII reference was fitted for each listener

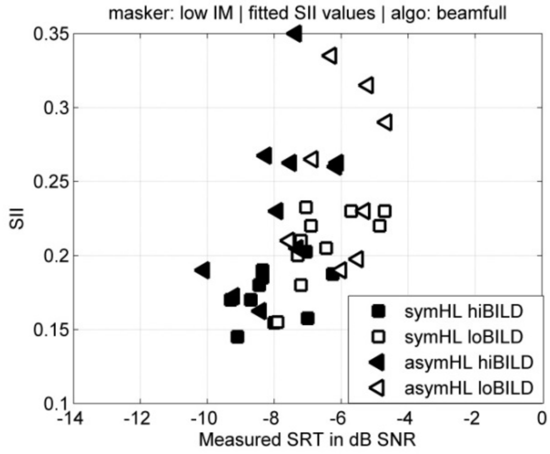

Fig. 2: Individual SII values (y-axis) for unintelligible speech maskers at the individual SRT (x-axis) for the diotic ("beamfull") algorithm. These SII values were used as individual SII references in the following predictions.

for the diotic situation ("beamfull" algorithm) and the unintelligible speech masker (low IM) and it was evaluated how BSIM using this individual SII reference is able to predict the SRTs for the other HA algorithms and for the unintelligible speech masker. Note that the algorithm used for fitting the SII reference does not provide any binaural cues and can thus be regarded as diotic (despite the compensation of hearing loss according to NAL-R which differed across ears). Consequently, the individual SII reference is independent of the listener's binaural processing. In Fig. 2, the resulting SII references for the different listeners are shown on the y-axis with the corresponding measured SRT on the x-axis.

Note that the individual SII reference value correlates with both the symmetry of the hearing loss and the the BILD. The lowest SII values were obtained for listeners with symmetric hearing loss and large BILDs. Slightly higher SII values were obtained for listeners with symmetric hearing losses and small BILDs. Listeners with asymmetric hearing losses showed higher SII values and a larger spread of these values across listeners compared to listeners with symmetric hearing losses. Note that the BILD is a binaural measure whereas the SRTs analyzed here were obtained with diotic stimuli. However, the BILD requires intact processing of temporal fine structure and was found to be correlated to monaural temporal fine structure sensitivity (Neher *et al.*, 2017).

Figure 3 shows BSIM predictions for the unintelligible speech masker (low IM) assuming only better-ear-glimpsing (left panel) or additional EC processing (right panel). The "pinna" algorithm was used here, meaning that binaural cues were available to the listeners. The diagonal line corresponds to perfect match between predicted and measured data; Points below the diagonal line represent underestimated SRTs and points above the diagonal overestimated SRTs. The left panel shows that using only better-ear-glimpsing results in an overestimation of the obtained SRTs. If the EC mechanism is applied, the predicted SRTs decrease by 1-2 dB, leading to a somewhat better aggreement with the perceptual data.

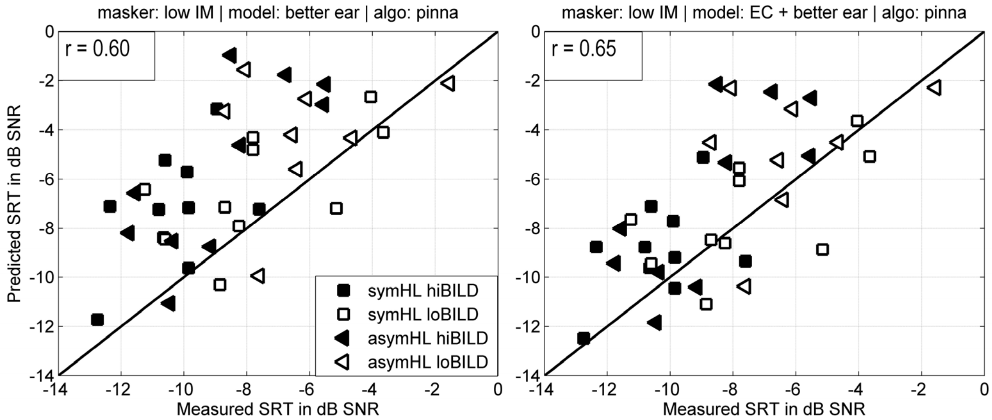

Fig. 3: The left panel shows predicted SRTs (y-axis) vs. measured SRTs (x-axis) for the *unintelligible* speech maskers (low IM) and the "pinna" condition (full binaural information available) for the case of better-ear-glimpsing only (no binaural interaction). The right panel shows corresponding predictions with EC processing included in the BSIM.

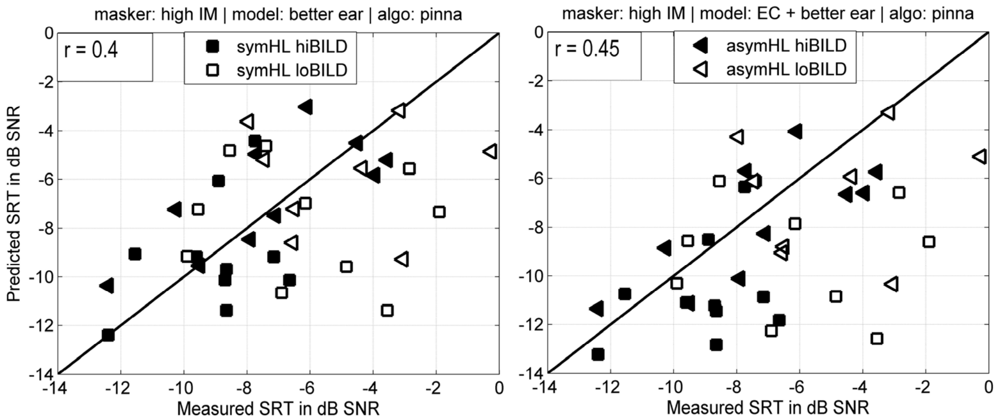

Fig. 4: Identical display as shown in Figu. 3 for the *intelligible* speech maskers (high IM).

Figure 4 shows BSIM predictions for the intelligible speech maskers (high IM). The left panel shows SRTs predicted using better-ear-glimpsing, the right panel shows SRTs predicted using both EC processing and better-ear-glimpsing. The predicted SRTs were mostly underestimated, which can be explained by the stronger IM of the intelligible speech maskers (which consisted of the same types of sentences as the target speech). This is not taken into account by BSIM as the SII was calibrated to the unintelligible speech maskers. Especially SRTs obtained at higher SNRs were underestimated for both better-ear-glimpsing only and better-ear-glimpsing plus EC.

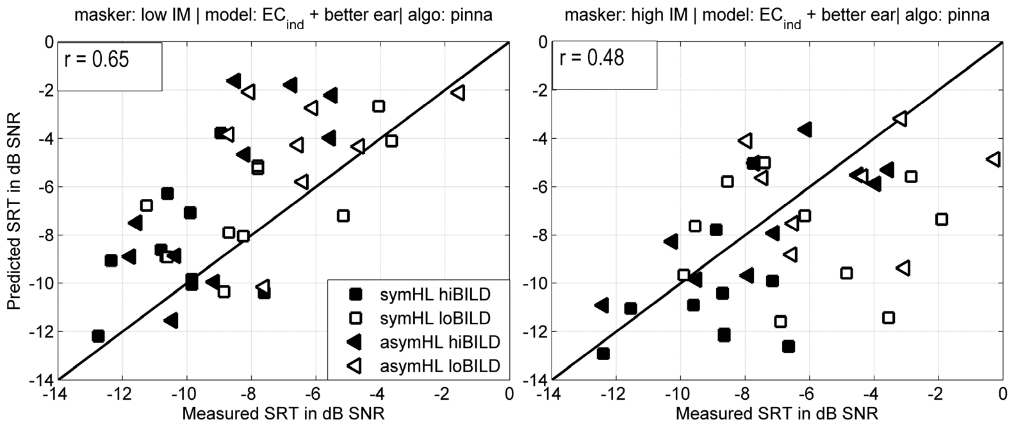

Fig. 5: Predictions obtained with extended BSIM incorporating individual binaural processing errors estimated from BMLDs measured at 500 Hz and 1000 Hz. The left panel shows results for the speech maskers with low IM and the right panel shows results for the speech maskers with high IM.

Figure 4 (high IM) shows that for several listeners with low BILD (open symbols) the observed SRT for the intelligible masker is much higher (worse) than predicted which is not the case for the unintelligible masker as shown in Fig. 3 (low IM). This suggests that the additional IM of the intelligible masker is more detrimental to these listeners than it is for listeners with large BILDs. Note that Neher et al. (2017) showed a significant correlation of BILDs with temporal fine structure sensitivity which might be relevant in scenarios with high IM, where target and interferer can better be separated, for instance, using pitch information.

Figure 5 shows predictions of the extended BSIM (Hauth *et al.*, 2017) incorporating individualized binaural processing errors estimated from BMLDs measured at 500 and 1000 Hz for characterizing supra-threshold binaural processing deficits. In general, larger processing errors were found compared to normal-hearing data. As a consequence, the predicted benefit from binaural processing is reduced, with better-ear-glimpsing defining the upper bound.

The additional individualization of binaural processing errors led to only slight improvement of the predictions. But based on the findings from Neher *et al*. (2017) it can be assumed that both SII and EC individualization mirror supra-thresholds processing deficits in temporal fine structure and are, therefore, highly correlated.

CONCLUSIONS

BSIM is applicable to speech maskers. This was achieved by calibrating the SII back-end of BSIM to a diotic condition comprising an unintelligible speech masker individually for each listener. This took into account that different listeners required different amounts of acoustical information (as quantified by the SII) to reach the SRT in a diotic condition. Using this "monaural" individualization, significant correlations between predictions and observations were achieved for an intelligible masker as well as for bilateral processing schemes that also included binaural processing by the listeners.

BSIM predicted a benefit due to 'true' binaural processing for aided listeners with impaired hearing.

Virtually no improvement of prediction accuracy was achieved, when additionally to monaural individualization the binaural processing errors in BSIM were individualized based on BMLDs measured at 500 and 1000 Hz. This might be due to the fact that the BMLDs were also correlated to the diotic SRTs, so that the binaural individualization did not add information.

ACKNOWLEDGMENTS

This study was supported by the DFG Cluster of Excellence EXC 1077/1 "Hearing4All". The research reported in (Neher *et al.*, 2017) was additionally supported by Sonova AG, Switzerland.

REFERENCES

ANSI (**1997**). *Methods for Calculation of the Speech Intelligibility Index (S3.5-1997).* American National Standard Institute, New York, NY

Beutelmann, R., Brand, T., and Kollmeier, B. (**2010**). "Revision, extension, and evaluation of a binaural speech intelligibility model," J. Acoust. Soc. Am., **127**, 2479-2497. doi: 10.1121/1.3295575

Brungart D.S., and Iyer N. (**2012**) "Better-ear glimpsing efficiency with symmetrically-placed interfering talkers," J. Acoust. Soc. Am., **132**, 2545-2556.

Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am. **35**(8), 1206–1218. doi: 10.1121/1.1918675.

Hauth, C.F., Brand, T., and Kollmeier B. (**2017**). "Modelling the frequency dependency of binaural masking level difference and its role for binaural unmasking of speech in normal hearing and hearing impaired listeners," J. Acoust. Soc. Am. **141**, 3638. doi: 10.1121/1.4987846

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (**2010**). "Development and analysis of an international speech test signal (ISTS)," Int. J. Audiol., **49**, 891-903. doi: 10.3109/14992027.2010.506889

Kayser, H., Ewert, S.D., Anemüller, J., Rohdenburg, T., Hohmann, V., and Kollmeier, B. (**2009**). "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," EURASIP J. Adv. Signal Process., 1-10.

Neher, T., Wagener, K.C., and Latzel, M. (**2017**). "Speech reception with different bilateral directional processing schemes: Influence of binaural hearing, audiometric asymmetry, and acoustic scenario," Hear. Res. **353**, 36-48.

Wagener, K., Brand, T., and Kollmeier, B. (**1999**). "Development and evaluation of a sentence test for the German language. I-III: Design, optimization and evaluation of the Oldenburg sentence test," Z. für Audiol. Audiol. Acoust., **38**, 86e95, 4-15, 44-56.