

# Comparing Binaural Pre-processing Strategies III: Speech Intelligibility of Normal-Hearing and Hearing-Impaired Listeners

Trends in Hearing  
2015, Vol. 19: 1–18  
© The Author(s) 2015  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/2331216515618609  
tia.sagepub.com  


Christoph Völker<sup>1,2</sup>, Anna Warzybok<sup>1,2</sup>, and Stephan M. A. Ernst<sup>1,2</sup>

## Abstract

A comprehensive evaluation of eight signal pre-processing strategies, including directional microphones, coherence filters, single-channel noise reduction, binaural beamformers, and their combinations, was undertaken with normal-hearing (NH) and hearing-impaired (HI) listeners. Speech reception thresholds (SRTs) were measured in three noise scenarios (multitalker babble, cafeteria noise, and single competing talker). Predictions of three common instrumental measures were compared with the general perceptual benefit caused by the algorithms. The individual SRTs measured without pre-processing and individual benefits were objectively estimated using the binaural speech intelligibility model. Ten listeners with NH and 12 HI listeners participated. The participants varied in age and pure-tone threshold levels. Although HI listeners required a better signal-to-noise ratio to obtain 50% intelligibility than listeners with NH, no differences in SRT benefit from the different algorithms were found between the two groups. With the exception of single-channel noise reduction, all algorithms showed an improvement in SRT of between 2.1 dB (in cafeteria noise) and 4.8 dB (in single competing talker condition). Model predictions with binaural speech intelligibility model explained 83% of the measured variance of the individual SRTs in the no pre-processing condition. Regarding the benefit from the algorithms, the instrumental measures were not able to predict the perceptual data in all tested noise conditions. The comparable benefit observed for both groups suggests a possible application of noise reduction schemes for listeners with different hearing status. Although the model can predict the individual SRTs without pre-processing, further development is necessary to predict the benefits obtained from the algorithms at an individual level.

## Keywords

noise reduction, binaural signal processing schemes, hearing aids, speech recognition, speech intelligibility index

## Introduction

The World Health Organization states that 360 million people worldwide have to deal with disabling hearing loss (World Health Organization, 2015). For adults, a disabling hearing loss refers to a hearing loss greater than 40 dB HL in the better ear. People with hearing loss suffer especially in complex acoustical situations. It is well known that they encounter great difficulties understanding speech in noisy and reverberant environments (e.g., Humes, 1991; Plomp, 1986). Modern digital hearing aids offer a number of approaches designed to solve this problem. The aim of hearing aids is to improve speech intelligibility while not degrading the signal quality by applying speech enhancement techniques, noise and feedback reduction schemes, or directional microphones.

In clinical practice, the benefit from hearing aids is usually measured by comparing the intelligibility scores of single words or sentences presented in quiet with and without hearing aids. However, an increasing number of studies have measured speech intelligibility in noise, which is closer to the acoustical environment that hearing-impaired (HI) listeners are faced with in their daily life. As there is no single unified way of assessing the

<sup>1</sup>Abteilung Medizinische Physik, Carl von Ossietzky Universität Oldenburg, Germany

<sup>2</sup>Cluster of Excellence “Hearing4all,” Oldenburg, Germany

## Corresponding author:

Christoph Völker, Medizinische Physik and Cluster of Excellence “Hearing4all,” Carl von Ossietzky Universität Oldenburg, Kükersweg 74, D-26129 Oldenburg, Germany.  
Email: christoph.voelker@uni-oldenburg.de



benefit from hearing aids in noise, some studies are restricted to the relatively simple condition involving speech recognition in stationary noise (e.g., Peeters, Kuk, Lau, & Keenan, 2009). Other studies have investigated the improvements caused by hearing aids within different more complex noise types, including babble or cafeteria noise (e.g., Cornelis, Moonen, & Wouters, 2012; Healy, Yoho, Wang, & Wang, 2013; Luts et al., 2010). Still, several studies have indicated that the benefit from hearing aids measured in such controlled acoustical conditions did not match the benefit reported by the users in the everyday listening conditions (Bentler, Niebuhr, Getta, & Anderson, 1993a, 1993b; Cord, Surr, Walden, & Dyrland, 2004).

In our view, a comprehensive evaluation incorporates (a) instrumental model-based measures determining the effectiveness of signal processing schemes and (b) subjective experiments accessing the efficacy in realistic test environments. Accordingly, the current article, being part of a collaborative research project to comprehensively evaluate state-of-the-art binaural signal pre-processing schemes, is aimed at further closing the gap between real-life performance and laboratory measures.

Eight advanced signal pre-processing strategies, including directional microphones, a coherence filter, single-channel noise reduction (SCNR), and binaural beamformers, as well as a no pre-processing (NoPre) condition serving as a reference, were implemented and subsequently evaluated. Three noise conditions, based on typical everyday listening situations, were designed to measure their potential benefit. These conditions involved speech recognition in multitalker babble noise, cafeteria ambient noise (CAN), and noise from an intelligible competing talker spatially separated from the target speaker.

This study tested the effects of the different strategies in normal-hearing (NH) and HI listeners. The results were compared with the outcomes from instrumental measures (Baumgärtel et al., 2015b) and results from bilateral cochlear implant (CI) users (Baumgärtel et al., 2015a).

## Methods

### Listeners

Ten NH listeners and 12 HI listeners participated in this study. The group of NH listeners consisted mostly of students and employees of our department with self-reported NH. Four female and six male listeners participated, ranging in age from 21 to 33 years, with an average age of 27.3 years.

The tested HI group included 8 male and 4 female listeners, ranging in age from 21 to 66 years (with a mean of 44.3 years). Table 1 summarizes the information

about the HI participants. The listeners are rank-ordered according to age, followed by their averaged thresholds over the frequencies 500, 1000, 2000, and 4000 Hz (4PTA) in their better ear. The first column ID is used to index the  $n$ th listener. The 4PTA for the better ear across HI listeners ranged from 20 dB HL to 55 dB HL, with a mean of 43.9 dB HL.

Figure 1 shows the audiometric data of the 12 HI listeners (left panel: right ear, right panel: left ear). The mean thresholds with corresponding standard deviation are displayed as thick lines; individual thresholds are shown as thinner solid lines. The HI listeners showed slight-to-moderate bilateral sensorineural hearing impairments. For all but one listener, the across-ear asymmetry in hearing thresholds was within 25 dB for any audiometric frequency between 125 Hz and 8 kHz. The participant with asymmetric hearing loss (ID 1, see Table 1) showed a mean asymmetry across ears of 17 dB, averaged across all audiometric frequencies. For the other listeners, the averaged asymmetry did not exceed 11 dB.

Except one listener (ID 9), all participants wore bilateral hearing aids on a daily basis. All but one participant with a longer distance to travel performed the tests on two different days. On the first test day, pure-tone thresholds and the speech reception thresholds (SRTs) for one noise condition were measured (2.5 hr). On the second day, the SRTs were measured for the other two noise conditions (2.5 hr). Sufficient breaks were provided between the measurements. The HI participants were remunerated for their participation and their travel expenses.

Approval for this experiment was obtained from the Carl von Ossietzky Universität Oldenburg ethics committee.

### Signal Processing Strategies

Eight signal pre-processing strategies were evaluated in this study. The algorithm conditions and corresponding references are listed in Table 2. The processing schemes were described in detail and evaluated instrumentally by Baumgärtel et al. (2015b). The adaptive differential microphone (ADM) scheme and the SCNR do not provide a binaural link between the input channels and can be regarded as being monaural. The remaining six schemes are binaural pre-processing strategies, consisting of the ADM in combination with a coherence-based noise reduction (ADM + coh), a fixed and an adaptive version of a binaural beamformer using minimum variance distortionless response (MVDR) technique, and three variants postfilters based on the binaural MVDR beamformer technique. Together with the NoPre condition serving as a reference, in total nine algorithm conditions were provided in the test protocol.

**Table 1.** ID, Gender (M/F), Age, and Hearing Thresholds of the HI Listeners.

| ID | M/F | Age | Audiometric thresholds (dB HL) |     |    |    |    |     |
|----|-----|-----|--------------------------------|-----|----|----|----|-----|
|    |     |     | 250                            | 500 | 1k | 2k | 4k | 8k  |
| 1  | M   | 21  | 5                              | 0   | 25 | 45 | 60 | 65  |
|    |     |     | 10                             | 30  | 50 | 65 | 60 | 70  |
| 2  | M   | 28  | 15                             | 25  | 45 | 50 | 50 | 50  |
|    |     |     | 20                             | 30  | 40 | 50 | 45 | 50  |
| 3  | M   | 28  | 5                              | 25  | 75 | 65 | 55 | 25  |
|    |     |     | 5                              | 20  | 75 | 70 | 60 | 75  |
| 4  | M   | 34  | 15                             | 25  | 50 | 55 | 65 | 60  |
|    |     |     | 20                             | 30  | 45 | 50 | 60 | 65  |
| 5  | F   | 40  | 35                             | 45  | 60 | 60 | 45 | 70  |
|    |     |     | 45                             | 50  | 65 | 60 | 50 | 80  |
| 6  | F   | 40  | 15                             | 45  | 50 | 55 | 60 | 75  |
|    |     |     | 25                             | 55  | 60 | 60 | 55 | 80  |
| 7  | M   | 43  | 15                             | 25  | 35 | 65 | 70 | 70  |
|    |     |     | 15                             | 20  | 35 | 55 | 60 | 70  |
| 8  | F   | 43  | 30                             | 40  | 50 | 55 | 55 | 60  |
|    |     |     | 30                             | 40  | 50 | 55 | 60 | 65  |
| 9  | M   | 61  | 5                              | 5   | 10 | 10 | 55 | 60  |
|    |     |     | 5                              | 10  | 15 | 30 | 55 | 65  |
| 10 | F   | 61  | 30                             | 45  | 45 | 45 | 65 | 85  |
|    |     |     | 40                             | 50  | 50 | 50 | 75 | 100 |
| 11 | M   | 66  | 15                             | 30  | 35 | 35 | 55 | 60  |
|    |     |     | 15                             | 20  | 20 | 35 | 50 | 80  |
| 12 | M   | 66  | 40                             | 45  | 45 | 50 | 75 | 105 |
|    |     |     | 40                             | 45  | 45 | 50 | 70 | 80  |

Note. The values in the upper lines for each listener specify thresholds of the left ear, and the values in the lower lines specify the right ear. HI = hearing-impaired; HL = hearing loss.

### Compensation of Hearing Loss

To compensate for the hearing loss of HI participants, a multiband compressor scheme was used. The compressor divides the left and right input signals into nine overlapping filter bands and measures the sound pressure levels (SPLs) in each band. Dependent on the actual input band levels and the individual hearing threshold levels of the HI listener, insertion gains are calculated using the nonlinear fitting procedure (NL1) by the National Acoustic Laboratories (NAL; Byrne, Dillon, Ching, Katsch, & Keidser, 2001). The NAL-NL1 prescription rule aims at maximizing speech intelligibility while constraining loudness to be normal or less. The obtained first fit of the procedure was used. The insertion gains are applied to the input signals by the multiband compressor. The time constants for attack and release were 20 ms and 100 ms, respectively.

When testing NH participants, the multiband compressor did not apply any gain. The dynamic compressor was—together with the evaluated noise reduction schemes—realized as part of the real-time capable Master Hearing Aid (MHA; Grimm, Herzke, Berg, & Hohmann, 2006).

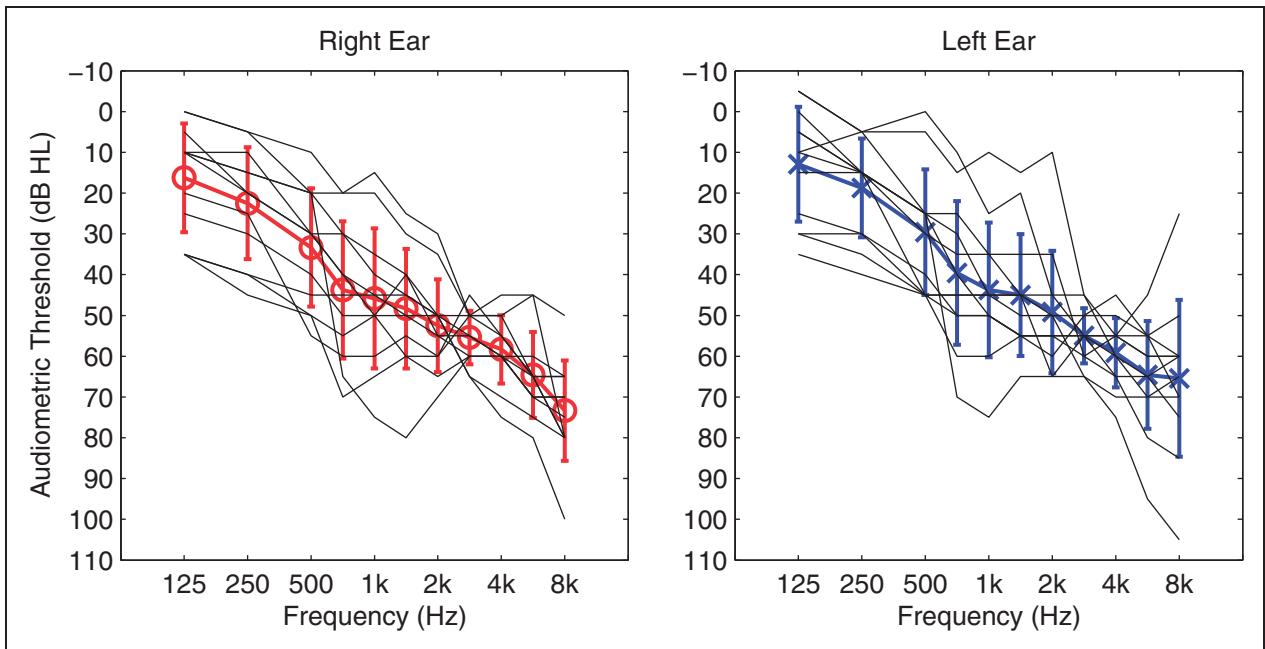
### Speech Reception Thresholds

In this study, the comparison of the binaural signal pre-processing schemes was done using adaptive speech intelligibility measurements in three noise conditions. The adaptive SRT measurement converged on the 50% intelligibility level. As speech material, the Oldenburg sentence test (OLSA; Wagener, Brand, & Kollmeier, 1999a, 1999b; Wagener, Kühnel, & Kollmeier, 1999c) was used, consisting of five-word semantically unpredictable sentences with the fixed grammatical structure, that is, *Name Verb Numeral Adjective and Noun*. Ten alternatives exist for each word category. One hundred twenty sentences are combined in 45 test lists, each containing 20 sentences. The test lists have a phonemic distribution similar to the German language.

The speech signals for actual measurements were generated by convolving the *dry* one-channel OLSA sentences with head-related impulse responses (HRIRs; Kayser et al., 2009). Here, the recorded multichannel impulse responses from behind-the-ear hearing aids with front and back microphone were used (four-channel). The HRIRs were recorded in a cafeteria, that is, the test subject is virtually seated at a table in a cafeteria, listening to the OLSA speaker directing toward him (or her) from 0° and a distance of 102 cm. The layout of the cafeteria is given in Figure 5 by Baumgärtel et al. (2015b), and the target OLSA speaker is placed at position A.

The evaluation of the algorithms was performed in (a) multitalker babble noise (20-talker babble [20T]), (b) cafeteria ambient noise (CAN), and (c) a single competing talker (SCT) located at an azimuth of 90°. The scenarios differ mainly in their spectro-temporal structure. The speech-shaped multitalker babble noise is stationary, the cafeteria noise has a typical quasistationary modulation of a cafeteria ambiance including noise of dishes and cutlery and snippets of conversations, and the SCT is speech-modulated. Each noise scenario has a duration of 600 s, which is a sufficient length for evaluating one pre-processing scheme with one test list containing 20 sentences. As the same signal material was used for the instrumental evaluation of the pre-processing schemes, further detailed information can be found in Baumgärtel et al. (2015b).

All three noise scenarios, including the unidirectional SCT condition, were scaled to a digital long-term root mean square level of −35 dB full scale averaged over all



**Figure 1.** Audiometric data of the 12 hearing-impaired listeners (left panel: right ear, right panel: left ear). The mean thresholds (marked for the right ear with circles and with crosses for the left ear)  $\pm$  standard deviation are displayed thick; individual thresholds are printed in thinner solid lines.

**Table 2.** List of Signal Pre-processing Strategies.

| # | Abbreviation        | Algorithm  |
|---|---------------------|--|
| 1 | NoPre               | No pre-processing  |
| 2 | ADM                 | Adaptive differential microphones (Elko & Nguyen Pong, 1995)   |
| 3 | ADM + coh           | Adaptive differential microphones in combination with coherence-based noise reduction (e.g., Grimm, Hohmann, & Kollmeier, 2009; Luts et al., 2010) |
| 4 | SCNR                | Single-channel noise reduction (e.g., Breithaupt, Gerkmann, & Martin, 2008; Gerkmann, Breithaupt, & Martin, 2008)                                  |
| 5 | fixed MVDR          | Fixed binaural MVDR beamformer (e.g., Doclo, Gannot, Moonen, & Spriet, 2010; Van Veen & Buckley, 1988)   |
| 6 | adapt MVDR          | Adaptive binaural MVDR beamformer (e.g., Griffiths & Jim, 1982)  |
| 7 | com PF (fixed MVDR) | Common postfilter based on fixed binaural MVDR beamformer (e.g., Simmer, Bitzer, & Marro, 2001)  |
| 8 | com PF (adapt MVDR) | Common postfilter based on adaptive binaural MVDR beamformer (e.g., Simmer et al., 2001)   |
| 9 | ind PF (adapt MVDR) | Individual postfilter based on adaptive binaural MVDR beamformer (e.g., Simmer et al., 2001)   |

Note. MVDR = minimum variance distortionless response.

four channels. The dry (one-channel) OLSA sentences intentionally fluctuate in level around a nominal value to provide the same intelligibility for each sentence. The convolved OLSA sentences (four-channel) were scaled to a new nominal level of  $-35$  dB full scale to match the noise scenarios and keep the internal level fluctuations intact. The sampling rate of sentence and noise files was 44.1 kHz. These four-channel signals mixed together adaptively at different signal-to-noise ratios (SNRs) formed the input for the processing by the algorithms.

The translation from digital full scale levels to SPLs was based on realistic pressure levels measured in a cafeteria ambiance by Kayser et al. (2009). The overall presentation level of the noise signals was 73.4 dB SPL averaged over both output channels.

The measurements of the SRTs were performed following the adaptive procedure by Brand and Kollmeier (2002). For the first stimuli presentation, speech and noise were mixed together at an SNR of 0 dB. Depending on the number of correctly understood

words, the speech level varies for the next presentation. The noise level is held constant during the measurement procedure, and the change of presentation level for the subsequent sentence follows

$$\Delta L = -\frac{f(i) \cdot (\text{previous} - \text{target})}{\text{slope}} \quad (1)$$

where the parameter *target* denotes the aimed value for intelligibility at which the procedure should converge. As we aimed at determining the SNR corresponding to 50% intelligibility, the parameter *target* is set to 0.5. Parameter *previous* denotes a value for the intelligibility of the previous sentence. The parameters *slope* and  $f(i)$ , which controls the rate of convergence, are set to standard values following the recommendation by Brand and Kollmeier (2002) to obtain a reliable bias-free SRT estimate. As all participants were able to use the test software and hardware by themselves, they performed the sentence tests in a closed-set response format, that is, during the test, the participants were able to see all 10 possible words for each of the five-word categories on a computer screen and were asked to select the understood words out of this closed  $10 \times 5$  matrix.

The processing schemes were evaluated successively in each noise condition. The order of the three noise conditions was pseudorandomized for each listener, having a balanced distribution over participants. The order of the eight processing schemes and the NoPre condition tested in a given noise condition was also randomized. Each processing scheme was evaluated with a random test list containing 20 sentences. Before the first SRT measurement on a test session, the participants were instructed and performed two random test lists without pre-processing in the upcoming noise condition for training. The sentence test was implemented inside the AFC Toolbox by Ewert (2013) for MATLAB. The necessary real-time processing was provided by the software SOUNDMEXPRO. The evaluated pre-processing schemes were realized inside the MHA developed by Grimm et al. (2006). This test setup runs on an Acer tablet Iconia W700 with external soundcard Maya 44 USB by ESI Audio.

To produce a realistic test scenario, the participants were equipped with behind-the-ear hearing aid dummies Acuris P by Siemens. Here, the hearing aid microphones were turned off, and the dummies only served as headphones. The two-channel output of the hearing aids was passed into the participants' ears by using ear plugs EARTONE 13A. The frequency responses of the hearing aid speakers were equalized in a calibration procedure with a 2 cc coupler (IEC 126). After equalization, correct output SPLs were ensured using the MHA output calibration routine with broadband noise. The experiments were performed in a soundproofed test room (ANSI/ASA S3.1-1999, 2008).

## Instrumental Evaluation

To compare the perceptually measured SRT benefits with the instrumental evaluation of the schemes (see Baumgärtel et al., 2015b), we calculated the Kendall rank correlation coefficient  $\tau$  between the subject data and the instrumental measures. These instrumental measures consist of (a) intelligibility-weighted SNR (iSNR), (b) short-time objective intelligibility (STOI), and (c) perceptual evaluation of speech quality (PESQ). For the instrumental evaluation, 120 OLSA sentences were mixed with the three noise scenarios 20T, CAN, and SCT at different long-term SNRs corresponding approximations of the averaged measured SRTs within the NoPre condition (*baselines*) in each scenario. These mixtures were processed by the schemes and evaluated by the three instrumental measures. The mixtures without any applied pre-processing strategies were also evaluated by the measures. The averaged better channel improvements regarding NoPre were used to determine the capability of each of these measures to predict the SRT benefits for NH and HI listeners.

## Binaural Speech Intelligibility Model

In addition to the instrumental measures described in Baumgärtel et al. (2015b), we evaluated the algorithms in this study by means of the binaural speech intelligibility model (BSIM; Beutelmann, Brand, & Kollmeier, 2010). In contrast to the instrumental measures, the BSIM was used to predict the individual SRTs of the NH and the HI listeners.

The first stage of the model applies a gammatone filter bank to analyze the binaural speech and noise signals. Internal noise is derived from the individual audiogram and is added to the external noise to account for hearing impairment. In each frequency band, an independent equalization-cancellation (EC) mechanism (Durlach, 1963) is applied to compute the maximally achievable SNR. It is achieved by eliminating the noise signal due the destructive interference by subtracting the two channels from each other. This maximal SNR is adapted to imperfect human binaural processing by applying binaural *processing errors*, which restrict the performance of the EC process by preventing the perfect cancellation of the noise signal. The speech and noise signals are processed separately for a reliable estimate of SNR. The resulting SNRs are then used as input for the speech intelligibility index (SII; ANSI 20S3.5-1997, 1997), which varies between 0 (completely unintelligible) and 1 (perfect intelligibility). The resulting SII is transformed into an intelligibility value using a nonlinear transform derived from a mapping function for sentence intelligibility (cf. Table 3, Figure 7, Fletcher & Galt, 1950). The SRT for a given condition is calculated by selecting a fixed

**Table 3.** Pairwise Comparison of Processing Schemes Regarding Speech Intelligibility Benefit Averaged Over Normal-Hearing and Hearing-Impaired Listeners Separated by Noise Conditions (a) 20 T, (b) CAN, and (c) SCT.

|                     | NoPre  | ADM     | ADM +<br>coh | SCNR   | fixed<br>MVDR | adapt<br>MVDR | com PF<br>(fixed MVDR) | com PF<br>(adapt MVDR) | ind PF<br>(adapt<br>MVDR) |
|---------------------|--------|---------|--------------|--------|---------------|---------------|------------------------|------------------------|---------------------------|
| Scheme in (a) 20T   |        |         |              |        |               |               |                        |                        |                           |
| NoPre               | –      |         |              |        |               |               |                        |                        |                           |
| ADM                 | 4.1*** | –       |              |        |               |               |                        |                        |                           |
| ADM + coh           | 3.7*** | –0.4    | –            |        |               |               |                        |                        |                           |
| SCNR                | –0.9   | –5.0*** | –4.6***      | –      |               |               |                        |                        |                           |
| fixed MVDR          | 4.3*** | 0.1     | 0.5          | 5.1*** | –             |               |                        |                        |                           |
| adapt MVDR          | 4.0*** | –0.1    | 0.3          | 4.8*** | –0.3          | –             |                        |                        |                           |
| com PF (fixed MVDR) | 3.6*** | –0.5    | –0.1         | 4.5*** | –0.6          | –0.3          | –                      |                        |                           |
| com PF (adapt MVDR) | 3.2*** | –0.9    | –0.5         | 4.1*** | –1.1**        | –0.8          | –0.5                   | –                      |                           |
| ind PF (adapt MVDR) | 3.4*** | –0.7    | –0.3         | 4.3*** | –0.8**        | –0.6          | –0.2                   | 0.2                    | –                         |
| Scheme in (b) CAN   |        |         |              |        |               |               |                        |                        |                           |
| NoPre               | –      |         |              |        |               |               |                        |                        |                           |
| ADM                 | 3.3*** | –       |              |        |               |               |                        |                        |                           |
| ADM + coh           | 3.7*** | 0.4     | –            |        |               |               |                        |                        |                           |
| SCNR                | –0.6   | –3.9*** | –4.3***      | –      |               |               |                        |                        |                           |
| fixed MVDR          | 4.0*** | 0.7     | 0.3          | 4.6*** | –             |               |                        |                        |                           |
| adapt MVDR          | 3.1*** | –0.2    | –0.6         | 3.7*** | –0.9          | –             |                        |                        |                           |
| com PF (fixed MVDR) | 3.5*** | 0.2     | –0.2         | 4.1*** | –0.5          | 0.4           | –                      |                        |                           |
| com PF (adapt MVDR) | 2.1*** | –1.2**  | –1.6***      | 2.6*** | –1.9***       | –1.0          | –1.4***                | –                      |                           |
| ind PF (adapt MVDR) | 2.6*** | –0.7    | –1.1**       | 3.2*** | –1.4***       | –0.5          | –0.9                   | 0.5                    | –                         |
| Scheme in (c) SCT   |        |         |              |        |               |               |                        |                        |                           |
| NoPre               | –      |         |              |        |               |               |                        |                        |                           |
| ADM                 | 2.6**  | –       |              |        |               |               |                        |                        |                           |
| ADM + coh           | 2.8*** | 0.3     | –            |        |               |               |                        |                        |                           |
| SCNR                | –1.4   | –4.0*** | –4.2***      | –      |               |               |                        |                        |                           |
| fixed MVDR          | 3.0*** | 0.5     | 0.2          | 4.4*** | –             |               |                        |                        |                           |
| adapt MVDR          | 4.8*** | 2.2     | 1.9          | 6.2*** | 1.8           | –             |                        |                        |                           |
| com PF (fixed MVDR) | 2.3*   | –0.2    | –0.5         | 3.7*** | –0.7          | –2.4          | –                      |                        |                           |
| com PF (adapt MVDR) | 3.4*** | 0.9     | 0.6          | 4.8*** | 0.4           | –1.4          | 1.1                    | –                      |                           |
| ind PF (adapt MVDR) | 4.0*** | 1.5     | 1.2          | 5.5*** | 1.0           | –0.7          | 1.7                    | 0.6                    | –                         |

Note. The numbers denote differences in dB for the averaged measured speech reception thresholds (schemes in rows minus schemes in columns), that is, positive values correspond to a better performance of the scheme in the respective row. Significant differences (criteria adjusted by the amount of comparisons) are marked with \* $p < .0014$ , \*\* $p < .00028$ , and \*\*\* $p < .00003$ . 20T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; NoPre = no pre-processing; ADM = adaptive differential microphone; ADM + coh = adaptive differential microphone in combination with a coherence-based noise reduction; SCNR = single-channel noise reduction; MVDR = minimum variance distortionless response.

reference SII value and varying the SNR until the SII equals this reference value.

BSIM was used to predict the SRTs in all noise conditions with signals for NH and HI listeners without any pre-processing. Because the effectiveness of the algorithms may depend on the input SNR, the processing schemes were evaluated in terms of SII for different input SNRs, ranging from  $-17.5$  dB to  $0$  dB with a step size of  $2.5$  dB. To validate the quality of the SII benefit predictions on an individual level, the measured SRT benefits are compared with the predicted SII benefits

for each HI listener separately. In line with the evaluation by the other instrumental measures, the empirical data and individual model predictions were compared by means of the Kendall rank correlation coefficient  $\tau$ .

The speech signal consisted of 20 concatenated OLSA sentences convolved with the same HRIRs as the signals in human experiments. For the model predictions, the signals from the front microphones were taken. The length of the speech stimuli corresponded to the ones that were presented to the listeners. For each noise condition, the extended version of the model was applied

(short-time BSIM; Beutelmann et al., 2010). This model version was proposed for speech intelligibility predictions in modulated or time variant interferers. The short-time BSIM analyses the signals in short time frames of 1,024 samples at 44100 Hz sampling rate and a frame shift of half the frame length. The effective frame length is about 12 ms. These parameters were set based on the previous findings of Rhebergen, Versfeld, and Dreschler (2006) and Beutelmann et al. (2010).

The measured SRTs of listeners with NH in the cafeteria noise and signals without pre-processing were used as the reference condition to normalize the SII. The reference SII was set to 0.09. This ensured that the predicted SRT was very close to the mean SRT measured in the reference condition. The reference SII was kept constant for other conditions. For speech intelligibility predictions of HI listeners, the individual audiograms were used to simulate hearing impairment. For NH listeners, the use of individual audiograms does not significantly influence the accuracy of speech intelligibility predictions in noise, and the variance observed in measured SRTs cannot be predicted by the model (Beutelmann et al., 2010). Therefore, for predictions of NH listeners, an average audiogram of 0 dB HL is assumed at all frequencies.

Because BSIM requires separate speech and noise signals at the input, the phase-inversion method (Hagerman & Olofsson, 2004) was used for separation of the processed signals.

### Statistical Analysis

Two separate analyses of variance (ANOVAs) were used to investigate the significance of within- (algorithm, noise condition) and between-subject factors (listener group: NH, HI) and their interactions on listeners performance. The first ANOVA analyzed the measured SRTs without pre-processing (baselines). For the second ANOVA, we calculated individual SRT differences (benefits) for each scheme and each noise condition with respect to the NoPre condition. Using the statistical software IBM SPSS, the data were fed into mixed-model ANOVAs with repeated measurements. Whenever necessary, violations of sphericity were adjusted using the Greenhouse–Geisser correction. To determine the sources of significant effects indicated by the ANOVA, post hoc tests for multiple comparisons were conducted and reported with adjusted criteria for significance by the amount of comparisons.

## Results

### Speech Reception Thresholds

The measured SRTs without pre-processing are shown in Figure 2. All data values are indicated separately for

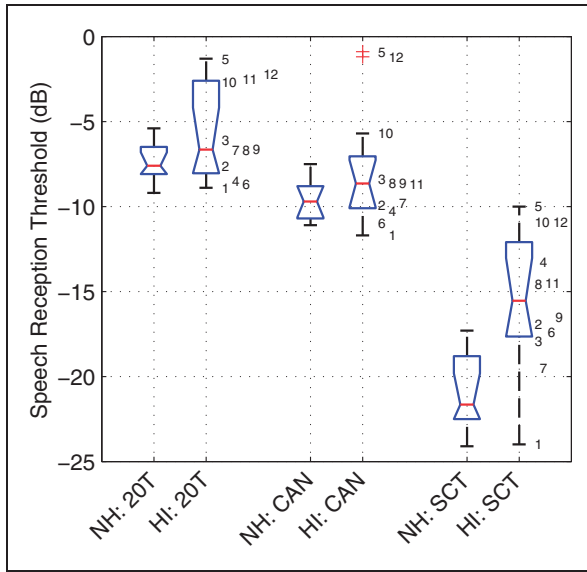
each of the three noise conditions and the two listener groups, that is, NH and HI listeners. The data from the HI listeners are also shown by index numbers, which are aligned with the information about the HI participants given in Table 1.

The ANOVA of the data supports that both within-subjects factor noise condition and listener group as between-subjects factor are statistically significant (noise condition:  $F(2,40)=438.3$ ,  $p < .001$ ; listener group:  $F(1,20)=411.7$ ,  $p < .001$ ). The interaction of factors noise condition and listener group is also significant,  $F(2,40)=12.7$ ,  $p < .001$ . Concerning the significant factor noise condition ( $p < .001$ ), the multitalker babble noise (20T) exhibits the highest thresholds, followed by the CAN and the SCT. All multiple comparisons of noise conditions differ significantly (all  $p < .0003$ ) for both groups, with normal and impaired hearing. Comparing the thresholds by subject group, mean thresholds for the NH listeners are lower than for the HI listeners. Multiple comparisons between NH and HI data only differ significantly in the SCT condition ( $p < .05$ ). In the 20T condition, the thresholds for the NH listeners are on average 1.8 dB lower ( $p > .05$ ) than for the HI listeners (mean SRTs of NH and HI group are  $-7.5$  dB and  $-5.7$  dB, respectively). The same difference of 1.8 dB ( $p > .05$ ) is observed in CAN (NH:  $-9.7$  dB, HI:  $-7.9$  dB). In the SCT condition, the thresholds of NH listeners are 5.5 dB significantly lower ( $p < .05$ ) with respect to the HI listeners (NH:  $-21.0$  dB, HI:  $-15.5$  dB). Furthermore, in all noise conditions, the variance between the HI listeners is considerably higher than in the group of NH listeners.

The individual thresholds in the NoPre condition, that is, without any active noise reduction algorithm, form the baselines for the further analysis of the signal processing schemes. For each scheme, the SRT difference with respect to the individual baseline is calculated (SRT benefit), that is, the benefit in terms of SRTs provided by each scheme is analyzed.

The mean individual SRT benefit averaged across NH participants is shown in Figure 3 for each of the eight processing schemes. The corresponding data for HI listeners are plotted in Figure 4. The error bars indicate intervals of  $\pm$  standard deviation from the mean value. The results are presented separately for each of the three noise conditions (20T, CAN, and SCT). Also, the averaged SRT values with corresponding standard deviation for the NoPre condition are shown for each noise condition in the figures legends.

The data from NH participants (Figure 3) reveal that all but the SCNR scheme lead to an improvement in SRTs. Excluding the SCNR scheme, the benefits provided by the schemes in the stationary 20T condition are quite similar across the algorithms, ranging from 3.1 dB (com PF based on adapt MVDR) to 4.4 dB (ADM). In the



**Figure 2.** Distribution of measured speech reception thresholds for the no pre-processing condition, pairwise for both listener groups (left: normal hearing, right: impaired hearing) and separated for noise conditions (from left to right: 20-talker babble [20T], cafeteria ambient noise [CAN], single competing talker [SCT]). The boxes have lines at the lower quartile, the median value, and the upper quartile. Whiskers extend from each end of the box to the adjacent values in the data. Outliers (+ sign) are data with values beyond the ends of the whiskers. Individual data from the HI listeners are shown by index numbers (cf. Table 1). NH = normal hearing; HI = hearing-impaired.

cafeteria condition also all but the SCNR scheme cause an improvement in speech recognition. Here, the benefits range from 2.2 dB (ind PF based on adapt MVDR) to 4.1 dB (fixed MVDR). In the SCT condition, the improvements range from 1.9 dB (com PF based on fixed MVDR) to 4.8 dB (adapt MVDR), again with the exception of the SCNR scheme. The variance of the individual improvements for each scheme is highest in the SCT condition, that is, the mean standard deviation of SRT benefit over all schemes is 2.5 dB (20T: 1.4 dB; CAN: 1.2 dB).

The results from HI listeners (see Figure 4) show similar trends as the data obtained from NH listeners. All but the SCNR scheme lead to improvements in speech recognition for all noise conditions. In 20T, the benefit ranges from 3.2 dB (ind PF based on adapt MVDR) to 4.3 dB (fixed MVDR). In CAN, the smallest improvement of 1.7 dB is observed for the com PF based on adapt MVDR scheme; the highest improvement of 3.6 dB is obtained for fixed MVDR. In SCT, the benefits range from 2.7 dB (ADM) to 4.7 dB (adapt MVDR). Again, the SCT condition shows the highest variance in SRT benefit for each scheme (mean standard deviations of SRT benefit are 20T: 1.2 dB; CAN: 1.3 dB; SCT: 2.2 dB).

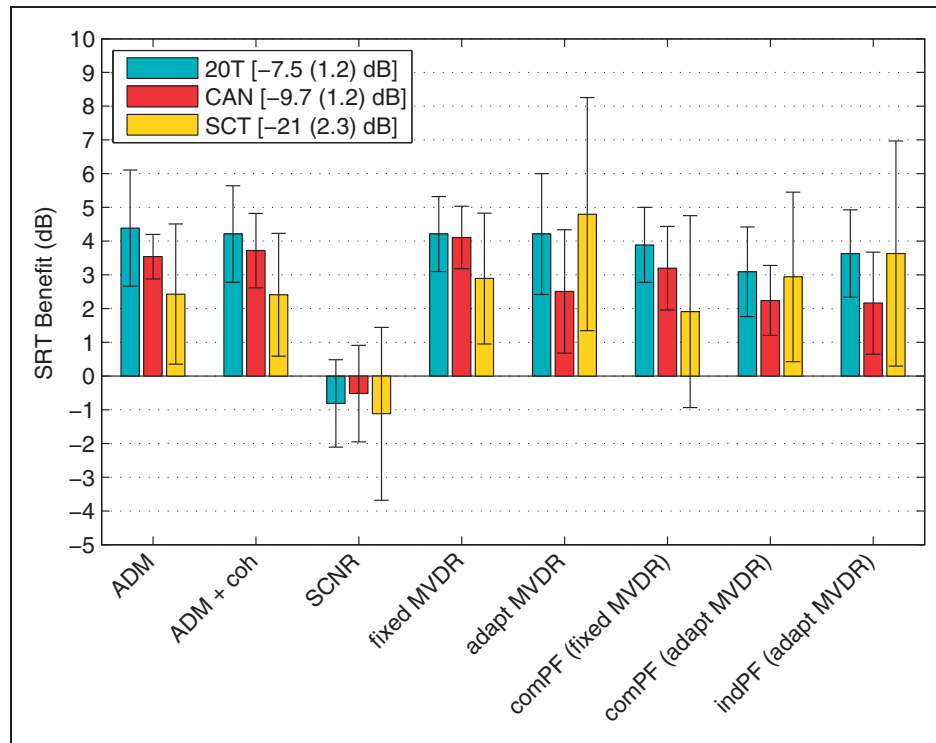
The between-subjects factor of listening group shows no significant effect on the measured SRT benefits,  $F(1,20)=0.024$ ,  $p>.8$ . This indicates that NH and HI listeners benefit equally from the different processing schemes. The within-subjects factor of algorithm has a significant effect on the SRT benefit,  $F(8,160)=96.5$ ,  $p<.001$ . The factor noise condition alone shows no effect,  $F(2,40)=1.6$ ,  $p>.2$ . The interaction of factors algorithm and noise condition is significant,  $F(4.4, 87.4)=4.3$ ,  $p=.002$ . We investigated this significant interaction further in a post hoc analysis by means of multiple comparisons of the algorithms in each noise condition separately.

The pairwise comparison of pre-processing schemes regarding speech intelligibility benefit averaged over NH and HI listeners separated by the three noise conditions is displayed in Table 3: The numbers denote differences in dB for the measured SRTs averaged across listeners (schemes in rows minus schemes in columns), that is, positive values correspond to a better performance of the scheme in the respective row. Significant differences (criteria adjusted by the amount of comparisons) are marked with \* $p<.0014$ , \*\* $p<.00028$ , and \*\*\* $p<.00003$ .

For the multitalker babble noise (20T), we find that with the exception of SCNR all intelligibility improvements caused by the schemes compared with NoPre are highly significant ( $p<.00003$ ). The largest improvement (4.3 dB) is produced by fixed MVDR. However, this benefit is not significantly higher than four other schemes, that is, ADM (4.1 dB), adapt MVDR (4.0 dB), ADM + coh (3.7 dB), and com PF based on fixed MVDR (3.6 dB). Fixed MVDR shows a significantly larger benefit than ind PF based on adapt MVDR ( $\Delta_{\text{benefit}}=0.8$  dB) and com PF based on adapt MVDR ( $\Delta_{\text{benefit}}=1.1$  dB), both with  $p<.00028$ . SCNR shows no significant difference to NoPre ( $p>.05$ ).

Identical to what was observed in the 20T conditions, all intelligibility improvements caused by the pre-processing schemes in the CAN scenario are highly significant with  $p<.00003$ . Again, the largest benefit is provided by fixed MVDR (4.0 dB), which is not significantly different from the four schemes ADM + coh (3.7 dB), com PF based on fixed MVDR (3.5 dB), ADM (3.3 dB), and adapt MVDR (3.1 dB). Regarding assertiveness, the fixed MVDR shows a significantly larger benefit than ind PF based on adapt MVDR ( $\Delta_{\text{benefit}}=1.4$  dB) and com PF based on adapt MVDR ( $\Delta_{\text{benefit}}=1.9$  dB), both with  $p<.00003$ . In contrast to the situation with 20T, the ADM + coh scheme prevails over ind PF based on adapt MVDR ( $\Delta_{\text{benefit}}=1.1$  dB,  $p<.00028$ ) and com PF based on adapt MVDR ( $\Delta_{\text{benefit}}=1.6$  dB,  $p<.00003$ ). As for the noise condition 20T, the SCNR scheme does not lead to significant changes in SRT with respect to NoPre ( $p>.05$ ).





**Figure 3.** Averaged individual SRT benefit caused by 8 signal processing schemes for 10 normal-hearing participants. The error bars denote an interval of  $\pm$  standard deviation from the mean value. The results are displayed separately for the three noise conditions: 20 T, CAN, and SCT. The legend shows the averaged SRT values measured without pre-processing including standard deviation in each noise condition.

SRT = speech reception threshold; 20 T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; ADM = adaptive differential microphone; ADM + coh = adaptive differential microphone in combination with a coherence-based noise reduction; SCNR = single-channel noise reduction; MVDR = minimum variance distortionless response.

In the SCT condition, all schemes but SCNR lead to significantly improved thresholds with respect to NoPre, whereas ADM brings a benefit of 2.6 dB ( $p < .00028$ ) and com PF (fixed MVDR) an improvement of 2.3 dB ( $p < .0014$ ). The benefits caused by the remaining schemes, ranging from 2.8 dB (ADM + coh) to 4.8 dB (adapt MVDR), are significant with  $p < .00003$ . No significant differences between the seven schemes leading to a prevailing strategy are found. Again, the SCNR scheme ( $-1.4$  dB) shows no difference to NoPre ( $p > .05$ ).

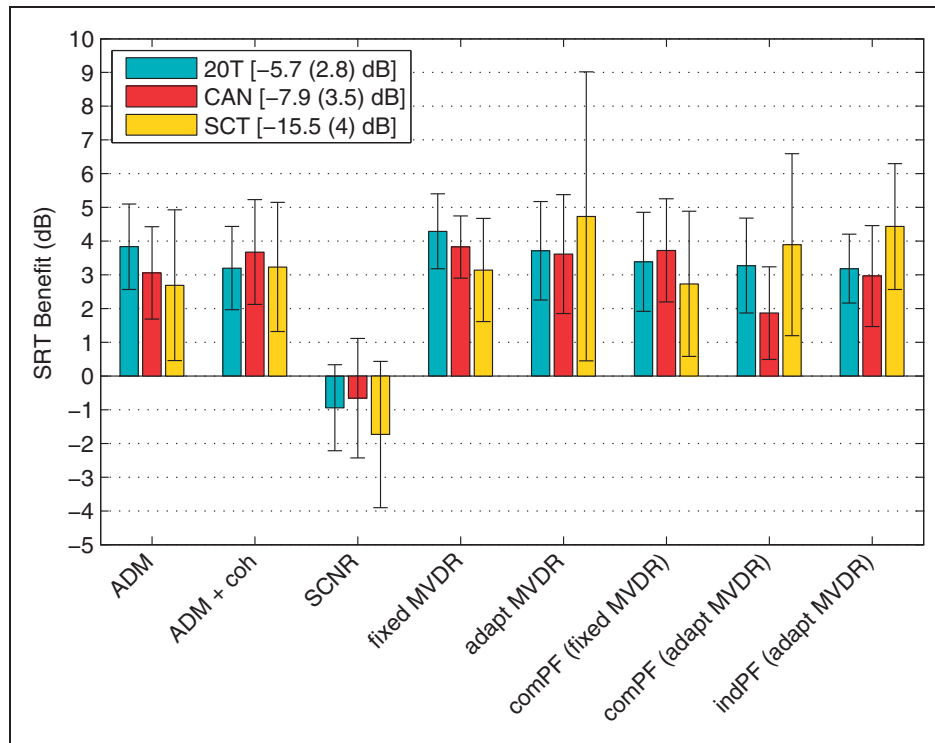
### Instrumental Evaluation

The correlation between the averaged individual SRT benefits of NH and HI listeners and the group-dependent improvements calculated within the instrumental evaluation with three different measures (Baumgärtel et al., 2015b) is displayed in Figure 5. The Kendall rank correlation coefficient data are shown separately for the instrumental measures, iSNR (left panel), STOI (middle panel), and PESQ (right panel). Within each panel, the data are displayed separately for each noise

condition; correlation coefficients  $\tau$  are given in the figure legend.

Averaging over all three noise scenarios, iSNR and STOI show a similar power to predict the subjective SRT benefits obtained with the tested pre-processing schemes ( $\tau_{\text{iSNR}} = .46$ ,  $p < .001$ ;  $\tau_{\text{STOI}} = .49$ ,  $p < .001$ ). Using PESQ, the overall correlation is  $\tau = .29$  ( $p < .01$ ). The predictions by all instrumental measures with respect to the different noise conditions have the highest statistical power ( $p < .001$ ) in the SCT scenario (iSNR:  $\tau = .69$ , STOI:  $\tau = .72$ , PESQ:  $\tau = .76$ ). Within CAN, the best performance is reached with STOI ( $\tau = .52$ ,  $p < .01$ ). None of the instrumental measures is able to produce reliable predictions of the algorithm rankings in the 20T condition ( $p > .05$ ).

Furthermore, the direct comparison of the subjective SRT benefit and the iSNR benefit predictions (Figure 5, left panel) showed a shift toward larger benefits derived from the iSNR measure, that is, all data points are above the dash-dotted line, which represents an idealized linear correlation between the instrumental results and subjective data. The median *gap* between iSNR improvements and SRT benefit, that is, the overestimation using iSNR



**Figure 4.** Averaged individual SRT benefit caused by 8 signal processing schemes for 12 hearing-impaired participants. The error bars denote an interval of  $\pm$  standard deviation from the mean value. The results are displayed separately for the three noise conditions: 20 T, CAN, and SCT. The legend shows the averaged SRT values measured without pre-processing including standard deviation in each noise condition.

SRT = speech reception threshold; 20 T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; ADM = adaptive differential microphone; ADM + coh = adaptive differential microphone in combination with a coherence-based noise reduction; SCNR = single-channel noise reduction; MVDR = minimum variance distortionless response.

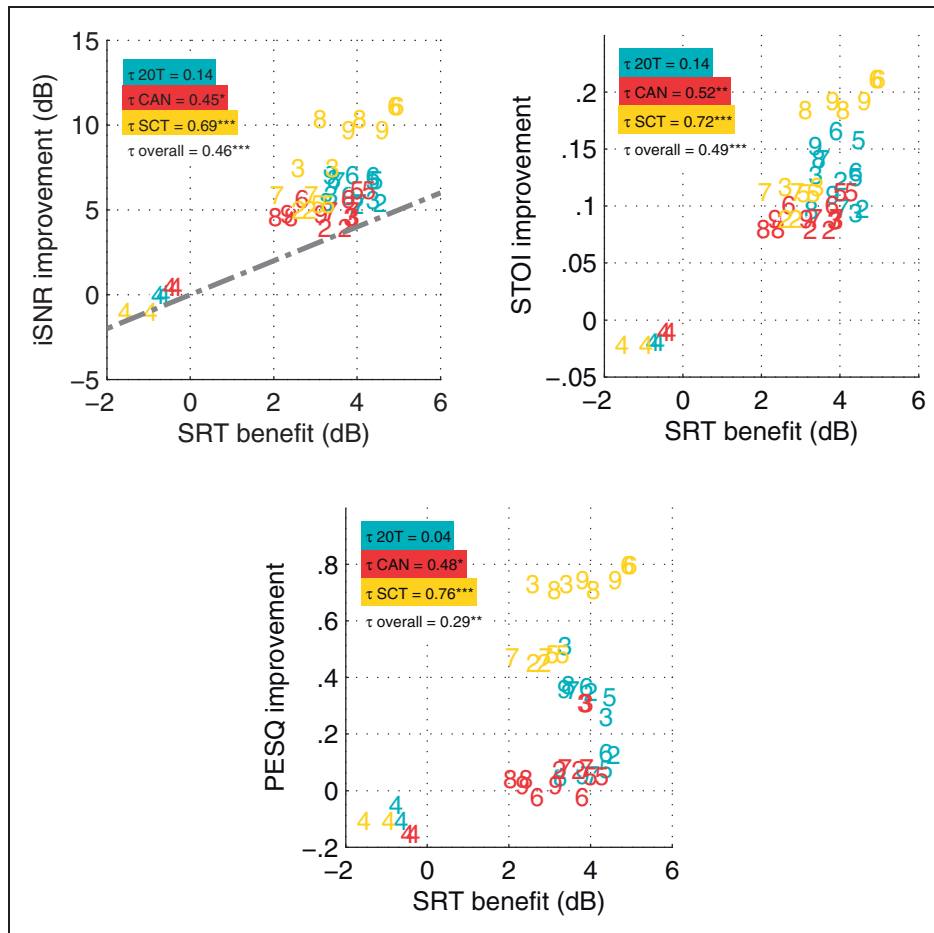
to predict the true SRT benefit, is 2.4 dB, which is in line with the earlier findings of, for example, Van den Bogaert, Doclo, Wouters, and Moonen (2009).

### Binaural Speech Intelligibility Model

Figure 6 shows the overall correlation between predicted and measured SRTs of NH and HI listeners for three noise conditions and signals without pre-processing. For listeners with NH, predicted SRT is compared with the mean measured SRT averaged across listeners. For HI listeners, the comparisons are done on individual data. The coefficient of determination ( $R^2$ ), the linear offset (bias) defined as the horizontal or vertical distance between the ideal mapping and the best fit with unity slope, and the root mean square prediction error ( $\text{rms}_e$ ) were calculated to evaluate the accuracy of model predictions. The coefficient of determination corresponds to fraction of the variance in the data, which can be explained by the model. The predictions of BSIM show a statistically significant correlation ( $p < .001$ ) to the measured data with the squared correlation coefficient of  $R^2 = .83$ , bias of 0.02, and  $\text{rms}_e = 3.1$  dB.

For each noise condition and each processing scheme, the SII benefit is shown as a function of input SNR in Figure 7. The SII benefit is calculated as a difference in SII between the respective processing schemes and the NoPre condition. The data are averaged across the predictions for all HI listeners. In all noise conditions, the SII benefit increases with increasing SNR for all the algorithms with exception of the SCNR scheme. The SCNR scheme does not show any benefit over the whole range of input SNRs, and in the SCT, noise shows even a slight deterioration in SII compared with NoPre. Three main groups of algorithms can be distinguished in the 20T and CAN noises. To the first one with the largest SII benefit belong all beamformers with post processing, ind PF (adapt MVDR), com PF (fixed MVDR), and com PF (adapt MVDR). The beamformers without post processing (adapt MVDR and fixed MVDR) and directional microphones (ADM and ADM + coh) create the second group with a moderate SII benefit. The smallest benefit was predicted for the SCNR scheme, which forms the third group.

In the SCT noise, the largest improvement is observed for the adapt MVDR. The SII benefit of adapt MVDR is

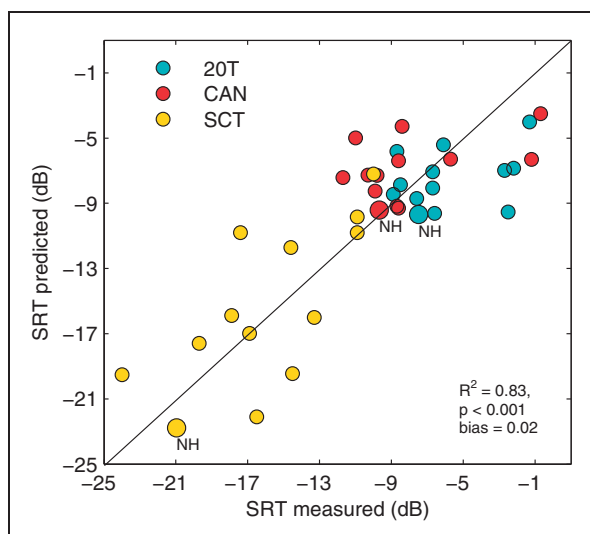


**Figure 5.** Correlation between the averaged individual SRT benefits for normal-hearing and hearing-impaired participants and the group-dependent improvements (better channel) calculated within the instrumental evaluation with three different measures (Baumgärtel et al. 2015b). The Kendall rank correlation coefficient data are shown separately for the instrumental measures, iSNR (left panel), STOI (middle panel), and PESQ (right panel). Within each panel, the data are displayed separately for each noise condition; correlation coefficients  $\tau$  are given in the figure legend. Significant correlation coefficients are marked with \* $p < .05$ , \*\* $p < .01$ , and \*\*\* $p < .001$ . The dash-dotted line in the left panel (iSNR) represents an idealized linear correlation between the instrumental results and subjective data. SRT = speech reception threshold; 20T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; iSNR = intelligibility-weighted signal-to-noise ratio; STOI = short-time objective intelligibility; PESQ = perceptual evaluation of speech quality.

larger in the SCT noise than in 20T and CAN noises. The lowest SII benefit is again observed for the SCNR scheme. All other processing schemes show moderate improvements in SII compared with NoPre. Similar to other two noise conditions, the adaptive beamformers with post processing demonstrate higher benefit than directional microphones (ADM and ADM + coh) and fixed beamformer without post processing (fixed MVDR).

The analysis of model predictions on individual level is shown in Table 4. The Kendall's rank correlation coefficients  $\tau$  are reported for each HI listener separately. Because the data shown in Figure 7 indicate that the predicted benefit depends on the noise condition, the accuracy of model predictions was calculated not only as average across all processing schemes and noise

conditions but also for each noise condition separately. The overall correlations indicate that the model is able to predict an individual benefit from different algorithms for half of the listeners (significant correlations are indicated in Table 4 with stars). For the SCT noise, the trends can be very well predicted for a few listeners with  $\tau$  up to .93. On the other side, in the same noise condition, the model fails to make reliable predictions of the individual SRT benefit for other listeners (e.g., ID 5, 6, 10, 12). Furthermore, the observed individual trends cannot be predicted in quasistationary noise conditions (20T and CAN). This suggests that the variance in the measured data for different processing schemes in each individual cannot be predicted well based on the individual audiograms. The possible reasons for the



**Figure 6.** Scatter plot of predicted SRTs against measured SRTs for normal-hearing and hearing-impaired listeners and three noise conditions (indicated by different colors). Solid line is the bisecting line. The coefficient of determination,  $R^2$ , was calculated based on individual data of HI listeners and mean data for NH listeners (marked as NH).

SRT = speech reception threshold; 20 T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; NH = normal hearing; HI = hearing-impaired.

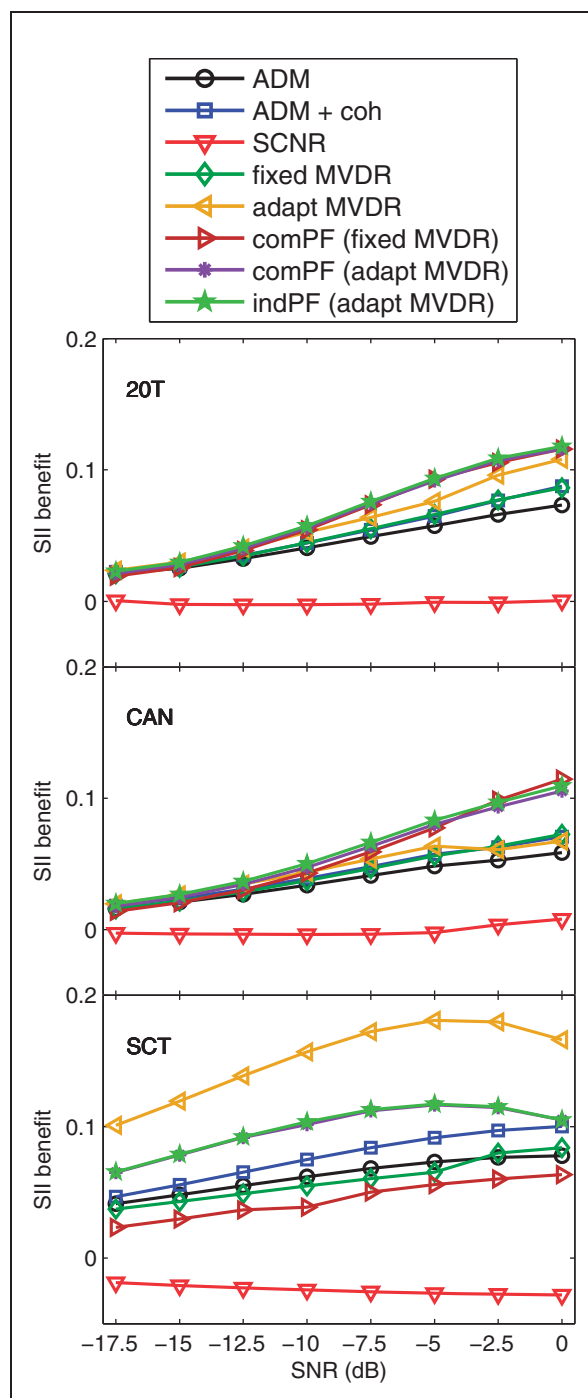
discrepancies between measured and predicted benefits are discussed in the next section.

## Discussion

### Speech Reception Thresholds

Comparing the measured thresholds without any signal pre-processing scheme (baselines) in the NH and HI groups, thresholds are significantly higher ( $p < .001$ ) for the HI group in every noise condition (see Figure 2). The HI listeners require on average 1.8 dB to 5.5 dB higher SNRs than NH listeners to achieve 50% intelligibility despite the individual compensation of hearing loss with a dynamic multiband compressor. This is a well-known finding in hearing research and can be explained by the limited processing capacities of the impaired auditory system, which leads to a limited benefit of hearing aids (Plomp, 1978). It should be stressed that a hearing aid is a supportive device but is not able to restore all the mechanisms affected by the hearing loss.

Both, the lowest thresholds and the highest variance measured in the SCT condition, can be explained by the spectro-temporal characteristics of the noise being a SCT. Due to silent gaps during speech pauses, this fluctuating noise exhibits the lowest masking of the target speaker and leads to low thresholds. The silent gaps offer the possibilities for the so-called listening in the dips



**Figure 7.** Mean SII benefit averaged across hearing-impaired listeners for each noise condition (20 T, CAN, and SCT) and each algorithm as a function of input SNR.

SII = speech intelligibility index; 20 T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker; ADM = adaptive differential microphone; ADM + coh = adaptive differential microphone in combination with a coherence-based noise reduction; SCNR = single-channel noise reduction; MVDR = minimum variance distortionless response; SNR = signal-to-noise ratio.

**Table 4.** Rank Correlation Coefficients (Kendall's  $\tau$ ) Between Individual Subjective SRT Benefits and Predicted Binaural SII Benefits Caused by the 8 Processing Schemes for the 12 HI Listeners.

| ID | Noise conditions |      |        | Overall |
|----|------------------|------|--------|---------|
|    | 20T              | CAN  | SCT    |         |
| 1  | .21              | .29  | .36    | .46**   |
| 2  | .00              | .47  | .86*** | .52***  |
| 3  | -.40             | -.11 | .93*** | .31*    |
| 4  | -.14             | .07  | .91*** | .04     |
| 5  | .62*             | .40  | -.07   | .17     |
| 6  | -.07             | -.04 | -.14   | -.03    |
| 7  | .18              | -.25 | .76**  | .46**   |
| 8  | .33              | -.07 | .52    | .14     |
| 9  | -.11             | .29  | .50    | .22     |
| 10 | .55              | .33  | .22    | .39**   |
| 11 | -.07             | -.25 | .43    | .38**   |
| 12 | .29              | -.07 | .00    | -.15    |

Note. Correlations are calculated separately for each noise condition, as well as the overall correlation. Significant correlation coefficients are marked with \* $p < .05$ , \*\* $p < .01$ , and \*\*\* $p < .001$ . Index numbers of the HI listeners (column "ID") correspond to audiometric data from Table 1 and measured thresholds from Figure 2. SRT = speech reception threshold; SII = speech intelligibility index; HI = hearing-impaired; 20T = 20-talker babble; CAN = cafeteria ambient noise; SCT = single competing talker.

(Peters, Moore, & Baer, 1998). Because the noise fragment presented with a respective sentence is not the same for each listener, the time points at which the dips are present also differ across listeners, possibly leading to a higher variance. Furthermore, the high variance is also due to the variable individual ability of listeners to attend information in the dips.

The statistical analysis revealed that the between-subjects factor listener group does not significantly affect the measured SRT benefits ( $p > .8$ ). Both listener groups, NH and HI, benefit equally from the investigated processing schemes. This finding gives more evidence for a reasonable use of noise reduction schemes also for NH persons.

The fact that between-subjects factor listener group was not significant allowed us to calculate benefits of the schemes averaged over both groups. From the pairwise comparisons of the schemes (cf. Table 3), we can calculate an estimate of the assertiveness of each scheme by counting the number of significant wins over the other processing strategies. Considering the restricted capacities of a digital hearing aid, it is most likely necessary to choose one pre-processing scheme over others. A sensible selection criterion would be to choose the one processing scheme with the most significant wins measured in this study. We find the fixed beamformer (fixed MVDR) belonging to the schemes with highest assertiveness in

all three noise scenarios. In the multitalker babble noise, the fixed MVDR was in the lead (significantly winning four times) and provides the highest benefit of 4.3 dB. In the CAN, the fixed MVDR (4.0 dB benefit) shares the same assertiveness as the ADM + coh scheme (3.7 dB benefit). Although we do not find an assertive scheme in the SCT condition, the fixed beamformer (3.0 dB benefit) belongs to the seven schemes that significantly improve the intelligibility in respect of the NoPre condition. However, although exhibiting the highest benefit in 20T and CAN, the benefit of the fixed MVDR scheme is not significantly higher than the benefits caused by four other schemes. Still, in terms of assertiveness over competing processing strategies generalized across the tested noise conditions, the fixed MVDR scheme emerges as being the best choice.

An alternative selection criterion for the pre-processing schemes could be to consider the implementation expense, that is, computing time and energy demand. The fixed MVDR is the best choice in terms of assertiveness. However, we found no significant difference between the binaural fixed MVDR and the monaural ADM scheme in any of the noise conditions. Thus, the ADM and the fixed MVDR provide comparable SRT benefits. In contrast to the fixed MVDR, the ADM does not require a binaural link across the ears to process the signals. Therefore, if the low costs are of great importance, the ADM can be considered as a reasonable alternative.

To estimate the implication of the SRT benefit values on the speech intelligibility in daily life, we assume that 1 dB increase in SNR produces a 20% increase in intelligibility for NH listeners and 5% for severely HI listeners (Brand & Kollmeier, 2002). Our tested group of HI showed slight-to-moderate hearing loss, that is, it is expected that 1 dB improvement in SNR will result in 5% to 20% (on average about 10%) increase in percentage correct scores. Thus, the measured benefit of 4.0 dB in the realistic CAN would lead to an increase of about 80% in intelligibility for NH and of about 40% for slight-to-moderate HI listeners. Relating this to our speech in noise test, these are four intelligible words out of five more for the NH and two intelligible words out of five more for the HI persons. Hence, for both listener groups the gained 3.0 dB to 4.3 dB benefit caused by fixed MVDR promises a considerable increase in terms of everyday life quality.

Due to the high variance in the measurements, the adaptive beamformer does not differ significantly from the fixed beamformer in the SCT noise condition. However, the adaptive beamformer scheme alone (4.8 dB) and the two postfilter variations based on the adaptive beamformer technique (4.0 dB and 3.4 dB) show a trend for a larger benefit than the fixed beamformer

(3.0 dB). This possible advantage of the adaptive beamformer schemes could be explained by the characteristic of the noise scenario. The competing talker coming from a single source ( $90^\circ$ ) allows the beamformer to adapt and to suppress this noise source. In the other more complex noise scenarios 20 T and CAN, the adaptation procedure does not operate equally well, that is, the adaptive beamformer does not outperform the fixed beamformer. This is in line with results from other evaluation studies with adaptive beamformer schemes (e.g., Maj, Royackers, Wouters, & Moonen, 2006).

The SRT measurements with the adaptive beamformer scheme in the SCT condition exhibit a noticeable large standard deviation (cf. Figures 3 and 4). We assume that due to the interactions of the speech signal and the competing talker, the adaptation process of the adaptive beamformer is unstable. This would lead to a high variance in the subjective measurement. The system's instability has consequences for separation of processed signals using the phase-inversion method (Hagerman & Olofsson, 2004). One of the assumptions of this method is stability of the system. Violation of this assumption can result in inaccurate separation of the speech and noise signals. The portions of the competing talker in the speech signal will lead to an overestimation of the SII benefit (see Figure 7).

Concerning the combination of schemes, we find the benefits caused by the additional coherence-based noise reduction serially coming after the ADM not significantly different from the ADM scheme alone. Still, in terms of assertiveness over competing processing schemes, the ADM+coh counts one more win than ADM in the realistic CAN condition. This can be explained considering the mostly incoherent nature of the cafeteria noise. Here, the noise reduction scheme can perform the separation of the coherent speech signal and unwanted incoherent noise signals very well. In the other two tested noise conditions, the additional coherence-based noise reduction does neither show an advantage nor disadvantage in respect of the ADM scheme alone.

In our evaluation, we found that in none of the tested noise scenarios any of the combinations of the two MVDR beamformer schemes and tested type of postfilter performed better than the corresponding beamformer alone. Furthermore, in the 20 T and the CAN condition, the adaptive MVDR scheme alone performs significantly better without the common postfilter ( $\Delta_{\text{benefit}} = 1.1$  dB in 20 T,  $p < .01$ ;  $\Delta_{\text{benefit}} = 1.9$  dB in CAN,  $p < .001$ ) or the individual postfilter ( $\Delta_{\text{benefit}} = 0.8$  dB in 20 T,  $p < .01$ ;  $\Delta_{\text{benefit}} = 1.4$  dB in CAN,  $p < .001$ ). This perceptual finding is not congruent with our objective calculations of the SII benefit (cf. Figure 7) or the instrumental evaluation of the processing schemes by Baumgärtel et al. (2015b) and gives evidence for another mismatch

between objective and perceptual evaluation. Possible explanations and solutions are discussed later.

A potential shortcoming in the measurements of the HI is the use of concatenated noise reduction and dynamic compression. When performing noise reduction before dynamic compression, the residual noise will receive more amplification compared with the speech. This might hamper the purpose of using noise reduction (Ngo, Spriet, Moonen, Wouters, & Jensen, 2012). Although the operation point of the algorithms was different for the NH and HI listeners, no significant differences were found in the SRT benefit for both listener groups. Therefore, for the algorithms tested, it may be assumed that as long as HI listeners with moderate hearing loss are considered, no degradation in efficiency due to the serial concatenation of noise reduction and dynamic compression was induced with respect to listeners with NH in the current study.

Having in mind that both listener groups of this study (NH and HI) benefit equally from the investigated processing schemes, we can now consider the third group—the bilateral CI users. The CI data are shown in Figure 2 in Baumgärtel et al. (2015a). Comparing the datasets, the benefits for the CI users are approximately 2 dB higher than for NH or HI listeners in the multitalker babble and the CAN. Also, the CI users profit up to 10 dB more from the adaptive beamformer scheme (with and without postfilter) than NH or HI listeners in the SCT condition.

In general, it may be concluded that the general trends observed for CI users are in principle comparable with the two investigated listener groups from this study. Therefore, it can be stated that the benefits caused by the signal processing schemes hold for listeners with a very different hearing status (NH, hearing aid users, and CI users).

### Instrumental Evaluation

The analysis considering the rank correlations between the instrumental measures presented in Baumgärtel et al. (2015b) and the subjective data from this study (cf. Figure 5) revealed significance in two of three noise conditions (SCT and CAN). However, the ranking correlations of the processing strategies tested here should be considered with caution. The empirical data of NH and HI listeners indicate that, for example, in the SCT noise, seven out of eight algorithms do not show significant differences in SRT benefit. Therefore, an attempt to make a reliable ranking of the processing strategies may be debatable. None of the instrumental measures described in the study of Baumgärtel et al. (2015b) can predict the trends in all noise conditions. The general discrepancy between instrumental and perceptual results, as also found in earlier studies (e.g., Luts et al., 2010),

underlines the importance of subjective evaluations with NH and HI listeners for a comprehensive hearing aid evaluation.

### *Binaural Speech Intelligibility Model*

The BSIM predictions can explain 83% of variance in the measured SRTs without any pre-processing. A relatively high and significant correlation between the measured and observed data as well as small bias indicates that the pure-tone threshold is a main factor for higher SRTs observed for HI listeners. In other words, the pure-tone audiometry is an efficient measure for describing the sensitivity loss, which is the main factor influencing speech perception. The effect of noise type is well predicted for the difference between stationary and single-talker noise. However, a considerable  $rms_e$  was observed between predicted and measured SRTs what indicates that despite accurate predictions of the general trends, the variance between listeners with similar audiograms cannot be predicted very well by the model. The accuracy of model predictions shown in this study is comparable with the previous finding of Beutelmann et al. (2010), who compared measured and predicted SRTs in stationary, babble, and single-talker noise conditions for NH and HI listeners. Beutelmann and colleagues used the same model as in the current study to predict speech intelligibility and reported an overall correlation between measured and predicted SRTs of .78. The bias and  $rms_e$  were  $-3.4$  dB and 3.0, respectively. A comparatively high bias in the study of Beutelmann et al. (2010) was greatest in the babble noise. The authors argued that it may be caused by the spectral difference between babble noise and speech that might be not properly handled by the SII, which frequency band importance function is linear and does not account for the correlations between adjacent or synergistic effects between the frequency bands.

The algorithms were further evaluated in terms of the SII, which was calculated for each HI listener and for each scheme over a broad range of SNRs. Based on that, the SII benefit was calculated corresponding to the difference in SII between signals with and without pre-processing. In agreement with the expectations, the SII benefit increased with increasing SNR. Also the differences between the algorithms were more prominent at high SNRs than at low SNRs. For the majority of the algorithms, the SII benefit was observed at the SNRs corresponding to measured SRTs. According to the model predictions, all postfilters based on binaural MVDR beamformer outperform other processing schemes in babble and cafeteria noise. These findings are not consistent with the measured data. The reason for the discrepancies might be the accuracy of the method used for separation of the processed signal into

speech and noise. It is known that the SCNR schemes are able to improve the SNR but at the same time may introduce distortions to the speech or noise signal. By separating the processed signals into speech and noise components using the phase-inversion technique (Hagerman & Olofsson, 2004), the distortions will be associated with the speech component. In the following analysis by the BSIM, the whole energy in the speech component (target speaker and also the introduced distortions) will be considered useful, leading to an overestimation of the speech recognition.

Although the individual SRTs can be well predicted in the NoPre condition (as described earlier), the model fails to make reliable estimates of the benefit from different noise reduction schemes for each individual. The analysis showed only overall (averaged across noise conditions) and in the SCT noise high and significant correlations for a subgroup of HI listeners. A number of studies give evidence for several reasons to be considered to explain the lack of the consistent correlations on individual level as well as the high observed variance. Beside the sensitivity loss, factors like age (Dubno, Horwitz, & Ahlstrom, 2002; Festen & Plomp, 1990), reduced sensitivity to temporal fine structure (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006), narrower frequency range, in which listeners are able to use the interaural phase/time differences (Neher, Laugesen, Jensen, & Kragelund, 2011; Warzybok, Rennies, Brand, & Kollmeier, 2014), or cognitive factors (Akeroyd, 2008) were shown to influence speech perception in noise. Thus, better individualization of model predictions for listeners with hearing impairment can be achieved when aspects other than pure sensitivity loss are accounted for. For example, Warzybok et al. (2014) showed that speech intelligibility in binaural conditions can be predicted more accurately when individual abilities in the detection of interaural phase differences are taken into account.

To overcome the shortcoming, a few solutions could be examined. The first one refers to the signal separation method. The phase-inversion method applied here could be replaced by the shadow filtering method (e.g., Fredelake, Holube, Schlueter, & Hansen, 2012). A second approach could consider a method suitable to estimate the distortions in the speech signal. Based on this estimate, a correction factor could be applied to the speech signal to account for the detrimental part of the speech energy.

Furthermore, another front end could be used in the model, like an EC stage, that does not require separate noise and speech signals as proposed by Hauth and Brand (2015). However, replacement of the front end alone will not solve the problem discussed here because SII requires knowledge about the SNRs in different frequency bands. After the EC processing stage, the SNRs will have to be estimated.

The most advanced solution could replace the EC stage as well as the back end to make the speech intelligibility predictions based on mixed signals. This could, for example, be achieved by combining the EC processing stage proposed by Hauth and Brand (2015) and the speech intelligibility model of Schädler, Warzybok, Hochmuth, and Kollmeier (2015). The model of Schädler et al. (2015) is based on an automatic speech recognition system and is able to make the predictions using mixed signals. In addition, in contrast to the SII, the predictions are done without any calibration of the model to the empirical data. This combination of the models could be a future step toward predictions that will not be influenced by the accuracy of the method for separating noise and speech signals after processing.

## Conclusion

In this study, NH and HI participants tested variations of noise reduction schemes with respect to the possible benefit for speech intelligibility. This subjective evaluation was expanded by objective individual model predictions using the BSIM (Beutelmann et al., 2010). The following conclusions can be drawn:

- Forced to choose one pre-processing scheme over the others, for example, in practical applications, the fixed MVDR beamformer scheme represents the best choice in our comparison study. In all noise scenarios, it was either the best placed or equal to other best placed schemes regarding assertiveness. Depending on the noise scenario, the fixed beamformer improved SRT from 3.0 dB to 4.3 dB. However, considering implementation expense, the monaural ADM scheme offers a reasonable alternative.
- Both tested listener groups (with normal and impaired hearing) benefit equally from the investigated processing schemes. It can be stated that the benefits caused by the signal processing schemes hold for subjects with different hearing status. Thus, the possible benefit of noise reduction schemes does not only apply to hearing aid users but calls for their promising use for NH persons as well.
- Model predictions using an individualized BSIM can explain up to 83% of the measured variance of the individual SRTs without any pre-processing, that is, the speech reception of NH as well as aided HI listeners without additional noise cancellation.
- At this stage of development, the individualized model scheme was able to estimate the possible benefits of the noise reduction algorithms for a subset of the participants. However, the model failed to give reliable predictions of signal processing benefits for each individual and all noise scenarios. Thus, further

developments are necessary that would include listener characteristics other than the audiogram.

## Acknowledgments

The authors would like to thank Regina Baumgärtel for providing the test software, as well as Simon Doclo, Timo Gerkmann, Martin Krawczyk-Becker, and Daniel Marquardt for answering questions about the algorithms. They also thank Tobias Herzke, Giso Grimm, and Graham Coleman for their support with the Master Hearing Aid (MHA), Felix Grossmann for help with the calibration, as well as Thomas Brand, Volker Hohmann, Mathias Dietz, and Birger Kollmeier for helpful advice and fruitful discussions.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Cluster of Excellence “Hearing4All” (EXZ1077) from the German Research Foundation DFG and the German Ministry of Education and Research (BMBF) project “Model-based Hearing Aids” (01EZ1127D).

## References

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(Suppl. 2), S53–S71.
- ANSI 20S3.5-1997. (1997) *Methods for the calculation of the speech intelligibility index* New York, NY: American National Standard Institute.
- ANSI/ASA S3.1-1999. (2008) *Maximum permissible ambient noise levels for audiometric test rooms (R2008 Ed.)*. Washington, DC: American National Standards Institute.
- Baumgärtel, R. M., Hu, H., Krawczyk-Becker, M., Marquardt, D., Herzke, T., Coleman, G., ... Dietz, M. (2015a). Comparing binaural pre-processing strategies II: Speech intelligibility of bilateral cochlear implant users. *Trends in Hearing*, 19, 1–18.
- Baumgärtel, R. M., Krawczyk-Becker, M., Marquardt, D., Völker, C., Hu, H., Herzke, T., ... Dietz, M. (2015b). Comparing binaural pre-processing strategies I: Instrumental evaluation. *Trends in Hearing*, 19, 1–16.
- Bentler, R. A., Niebuhr, D. P., Getta, J. P., & Anderson, C. V. (1993a). Longitudinal study of hearing aid effectiveness. I: Objective measures. *Journal of Speech and Hearing Research*, 36(4), 808–819.
- Bentler, R. A., Niebuhr, D. P., Getta, J. P., & Anderson, C. V. (1993b). Longitudinal study of hearing aid effectiveness. II: Subjective measures. *Journal of Speech and Hearing Research*, 36(4), 820–831.



- Beutelmann, R., Brand, T., & Kollmeier, B. (2010). Revision, extension, and evaluation of a binaural speech intelligibility model. *Journal of the Acoustical Society of America*, *127*(4), 2479–2497.
- Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility. *Journal of the Acoustical Society of America*, *111*(6), 2801–2810.
- Breithaupt, C., Gerkmann, T., & Martin, R. (2008). A novel a priori SNR estimation approach based on selective cepstrotemporal smoothing. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 4897–4900.
- Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology*, *12*, 37–51.
- Cord, M. T., Surr, R. K., Walden, B. E., & Dyrland, O. (2004). Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. *Journal of the American Academy of Audiology*, *15*, 353–364.
- Cornelis, B., Moonen, M., & Wouters, J. (2012). Speech intelligibility improvements with hearing aids using bilateral and binaural adaptive multichannel wiener filtering based noise reduction. *Journal of the Acoustical Society of America*, *131*(6), 4743–4755.
- Doclo, S., Gannot, S., Moonen, M., & Spriet, A. (2010). Acoustic beamforming for hearing aid applications. In S. Haykin & K. R. Liu (Eds.), *Handbook on array processing and sensor networks* (Chapter 9, pp. 269–302). Hoboken, NJ: Wiley.
- Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *Journal of the Acoustical Society of America*, *111*(6), 2897–2907.
- Durlach, N. I. (1963). Equalization and cancellation theory of binaural masking-level differences. *Journal of the Acoustical Society of America*, *35*(8), 1206–1218.
- Elko, G. W., & Nguyen Pong, A.-T. (1995). A simple adaptive first-order differential microphone. In IEEE (Ed.), *IEEE ASSP workshop on applications of signal processing to audio and acoustics* (pp. 169–172). New Paltz, NY: IEEE.
- Ewert, S. D. (2013). AFC – A modular framework for running psychoacoustic experiments and computational perception models. In: German Acoustical Society (DEGA) (Ed.), *Proceedings of the International Conference on Acoustics AIA-DAGA* (pp. 1326–1329). Merano, Italy: German Acoustical Society (DEGA).
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, *88*(4), 1725–1736.
- Fletcher, H., & Galt, R. H. (1950). The perception of speech and its relation to telephony. *Journal of the Acoustical Society of America*, *22*(2), 89–151.
- Fredelake, S., Holube, I., Schlueter, A., & Hansen, M. (2012). Measurement and prediction of the acceptable noise level for single-microphone noise reduction algorithms. *International Journal of Audiology*, *51*, 299–308.
- Gerkmann, T., Breithaupt, C., & Martin, R. (2008). Improved a posteriori speech presence probability estimation based on a likelihood ratio with fixed priors. *IEEE Transactions on Audio, Speech and Language Processing*, *16*(5), 910–919.
- Griffiths, L. J., & Jim, C. (1982). An alternative approach to linearly constrained adaptive beamforming. *IEEE Transactions on Antennas and Propagation*, *30*(1), 27–34.
- Grimm, G., Herzke, T., Berg, D., & Hohmann, V. (2006). The master hearing aid: A PC-based platform for algorithm development and evaluation. *Acta Acustica united with Acustica*, *92*, 618–628.
- Grimm, G., Hohmann, V., & Kollmeier, B. (2009). Increase and subjective evaluation of feedback stability in hearing aids by a binaural coherence-based noise reduction scheme. *IEEE Transactions on Audio, Speech and Language Processing*, *17*(7), 1408–1419.
- Hagerman, B., & Olofsson, A. (2004). A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. *Acta Acustica united with Acustica*, *90*(S), 356–361.
- Hauth, C., & Brand, T. (2015). Ein blindes Modell zur Vorhersage der binauralen Sprachverständlichkeit. In: German Acoustical Society (DEGA) (Ed.), *DAGA – 40. Jahrestagung für Akustik* (pp. 1035–1038). Oldenburg, Germany: German Acoustical Society (DEGA).
- Healy, E. W., Yoho, S. E., Wang, Y., & Wang, D. (2013). An algorithm to improve speech recognition in noise for hearing-impaired listeners. *Journal of the Acoustical Society of America*, *134*(4), 3029–3038.
- Humes, L. E. (1991). Understanding the speech-understanding problems of the hearing impaired. *Journal of the American Academy of Audiology*, *2*, 59–69.
- Kayser, H., Ewert, S. D., Anemüller, J., Rohdenburg, T., Hohmann, V., & Kollmeier, B. (2009). Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing*, *2009*, 1–10.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–18869.
- Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., Buechler, M., . . . , Spriet, A. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *Journal of the Acoustical Society of America*, *127*, 1491–1505.
- Maj, J.-B., Royackers, L., Wouters, J., & Moonen, M. (2006). Comparison of adaptive noise reduction algorithms in dual microphone hearing aids. *Speech Communication*, *48*(8), 957–970.
- Neher, T., Laugesen, S., Jensen, N. S., & Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *Journal of the Acoustical Society of America*, *130*(3), 1542–1558.
- Ngo, K., Spriet, A., Moonen, M., Wouters, J., & Jensen, S. H. (2012). A combined multi-channel wiener filter-based noise

- reduction and dynamic range compression in hearing aids. *Signal Processing*, 92(2), 417–426.
- Peeters, H., Kuk, F., Lau, C., & Keenan, D. (2009). Subjective and objective evaluation of noise management algorithms. *Journal of the American Academy of Audiology*, 20(2), 89–98.
- Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *Journal of the Acoustical Society of America*, 103(1), 577–587.
- Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *Journal of the Acoustical Society of America*, 63(2), 533–549.
- Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *Journal of Speech, Language, and Hearing Research*, 29, 146–154.
- Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2006). Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *Journal of the Acoustical Society of America*, 120(6), 3988–3997.
- Schädler, M. R., Warzybok, A., Hochmuth, S., & Kollmeier, B. (2015). Matrix sentence intelligibility prediction using an automatic speech recognition system. *International Journal of Audiology*, 54, 1–8.
- Simmer, K. U., Bitzer, J., & Marro, C. (2001). Post-filtering techniques. In: M. Brandstein & D. Ward (Eds), *Microphone arrays: Signal processing techniques and applications* (Chapter 3, pp. 39–60). Berlin, Germany: Springer-Verlag.
- Van den Bogaert, T., Doclo, S., Wouters, J., & Moonen, M. (2009). Speech enhancement with multichannel Wiener filter techniques in multimicrophone binaural hearing aids. *Journal of the Acoustical Society of America*, 125(1), 360–371.
- Van Veen, B., & Buckley, K. (1988). Beamforming: A versatile approach to spatial filtering. *ASSP Magazine, IEEE*, 5(2), 4–24.
- Wagener, K., Brand, T., & Kollmeier, B. (1999a). Entwicklung und Evaluation eines Satztests für die deutsche Sprache II: Optimierung des Oldenburger Satztests [Development and Evaluation of a set of tests for the German language II: Optimization of the Oldenburg sentence test]. *Zeitschrift für Audiologie*, 38(2), 44–56.
- Wagener, K., Brand, T., & Kollmeier, B. (1999b). Entwicklung und Evaluation eines Satztests für die deutsche Sprache III: Evaluation des Oldenburger Satztests [Development and Evaluation of a set of tests for the German language III: Evaluation of the Oldenburg sentence test]. *Zeitschrift für Audiologie*, 38(3), 86–95.
- Wagener, K., Kühnel, V., & Kollmeier, B. (1999c). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests [Development and Evaluation of a set of tests for the German language I: Design of the Oldenburg sentence test]. *Zeitschrift für Audiologie*, 38(1), 4–15.
- Warzybok, A., Rannies, J., Brand, T., & Kollmeier, B. (2014). Prediction of binaural speech intelligibility in normal-hearing and hearing-impaired listeners: A psychoacoustically motivated extension. In: German Acoustical Society (DEGA) (Ed.), *DAGA – 40. Jahrestagung für Akustik* (pp. 351–352). Oldenburg, Germany: German Acoustical Society (DEGA).
- World Health Organization. (2015). *Deafness and hearing loss—Fact sheet 300*. Retrieved from <http://www.who.int/mediacentre/factsheets/fs300/en/>