# Benefit from different beamforming schemes in bilateral hearing aid users: Do binaural hearing abilities matter?

MATTHIAS LATZEL[1,*], KIRSTEN C. WAGENER[2], MATTHIAS VORMANN[2] AND TOBIAS NEHER[3]

[1] *Sonova AG, CH-8712 Stäfa, Switzerland*

[2] *Hörzentrum Oldenburg, D-26129 Oldenburg, Germany*

[3] *Institute of Clinical Research, University of Southern Denmark, DK-5230 Odense, Denmark*

Using a hearing aid simulator and virtual acoustics, Neher *et al.* (2017) recently showed that binaural hearing abilities influence speech-in-noise reception through different bilateral directional processing schemes. The current study aimed to extend this finding to real acoustic environments and commercial devices. Three beamforming schemes were tested – they differed in signal-to-noise ratio (SNR) improvement and binaural cue preservation. The participants were 38 elderly experienced hearing aid users. Speech understanding and localisation performance were measured. Binaural hearing abilities were assessed using the binaural intelligibility level difference (BILD). The analyses revealed a clear effect of the BILD on speech understanding in noise, but no interaction with the beamformer conditions. Greater SNR improvement was generally beneficial. In contrast, localisation of static and dynamic stimuli was more accurate when low-frequency binaural cues were preserved. Furthermore, the interaction with the BILD was marginally significant for dynamic stimuli ($p = 0.054$). Altogether, these results suggest that when selecting directional processing schemes in bilateral hearing aid fittings both speech understanding and aspects of spatial awareness perception should be considered.

## INTRODUCTION

Almost all hearing aids (HAs) comprise directional microphones as directionality is the only feature that improves speech intelligibility (Dillon, 2012). Although considerable effort is dedicated to parameterizing these systems to provide the optimal benefit to the intended target population, the acceptance and benefit of the so-called "FirstFit" vary remarkably across the individual users (Gatehouse *et al.,* 2003; Lunner, 2003). Thus, it is of interest to investigate how HA settings can be better tailored to the individual needs and medical/audiological parameters of the user. There have been a number of investigations looking into several factors and their ability to explain individual differences in HA outcomes, but recommendations for translating this information into a meaningful prescribed fitting are rather rare.

---

*Corresponding author: matthias.latzel@sonova.com

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ISSN: 2596-5522.

In a recent study, Neher *et al.* (2017) focused on individual binaural hearing abilities determined through the binaural intelligibility level difference (BILD; Kollmeier, 1996). They investigated how individual binaural hearing abilities play a role to understand speech-in-noise and how this information correlates with the setting of directional microphone systems in HAs. They found that the BILD was correlated with speech perception in situations with lateral interferers, while in spatially diffuse situations the speech perception was driven by the signal-to-noise ratio (SNR) improvement (or directivity index) provided by the different beamformer technologies independent of the BILD values. Together, these findings provide a base for adapting directional processing to the HA user, its binaural hearing abilities, and the acoustic scenario.

The current study investigated the extent to which the results of Neher *et al.* (2017) can be transferred to directional processing strategies used in commercially available hearing aids. Thereby, real acoustical coupling, real acoustical scenes and the possibility of head movements are of relevance. In addition to speech reception in noise, it was investigated how binaural cue preservation in the different directional processing schemes affects aspects related to spatial awareness.

## METHODS

### Participants

In the current study, 38 experienced HA users (16 women) with an average age of 74.7 yrs (range: 63-82 yrs) and moderate-to-severe bilateral hearing losses participated. All of them had participated in the Neher *et al.* (2017) study. The participants were divided into two groups according to their binaural hearing abilities as assessed using the BILD measure: BILD < 2.5 dB (‘BILD–’; $N = 18$) and BILD ≥ 2.5 (‘BILD+’; $N = 20$). The individual BILD values were equally distributed between the minimum value of -0.4 dB and maximum value of 5,2 dB. The two groups were balanced in terms of four-frequency pure-tone average hearing loss (55 dB HL and 51 dB HL, respectively). All 38 participants completed a set of speech-in-noise measurements, similar to those performed by Neher *et al.* (2017). A subset of 26 paricipants (9 women) with the same mean age, hearing loss and distribution of BILD values completed a set of additional spatial awareness measurements (see below).

### HA conditions

The participants were fitted with Phonak Audéo V90-312 devices using the xP-receiver with closed sShells and a flat real-ear-to-coupler difference to maximize the acoustic differences between the different beamformer settings. For gain prescription, a modified version of the NAL-NL1 fitting rule (Dillon, 1999) was used with linear amplification based on the gain prescribed for 65 dB input level of NAL-NL1. To ensure adequate audibility and comparability with the results of Neher *et al.* (2017), a minimum gain of 6 dB was defined in the frequency range from 250 Hz to 500 Hz. This was verified using real-ear measurements. The beamformer settings were all steered in the 0° direction and set at non-adaptive. Five settings were tested (see polar patterns in Figure 1):

(1) Real Ear Sound (RES): A commercially available beamformer setting simulating the pinna effect (Latzel, 2013) of the outer ear with a small degree of directivity (mean directivity index, DI: -1,0 dB) > 1 kHz. Output: Dichotic stimulus with binaural cues preservation over the entire frequency range.
(2) UltraZoom (UZ): A commercially available unilateral beamformer setting (Latzel, 2013) providing SNR improvement over the whole frequency range (mean DI: 2,3 dB). Output: Dichotic stimulus with binaural cue preservation over the entire frequency range.
(3) StereoZoom (SZ): A commercially available bilateral beamformer setting (Latzel, 2013) providing SNR improvement at frequencies < 2 kHz (mean DI: 4.7 dB). Output: predominantly diotic stimulus < 2 kHz and dichotic stimulus above.
(4) StereoZoom INV (SZ-inv): An experimental beamformer setting based on SZ that provides SNR improvement > 800 Hz (mean DI: 4.2 dB). Output: Dichotic stimulus < 800 Hz and diotic stimulus above.
(5) FullBeam (FB): An experimental beamformer setting based on SZ that provides SNR improvement over the whole frequency range (mean DI: 4.9 dB). Output: Diotic stimulus over the entire frequency range (no binaural cue preservation).

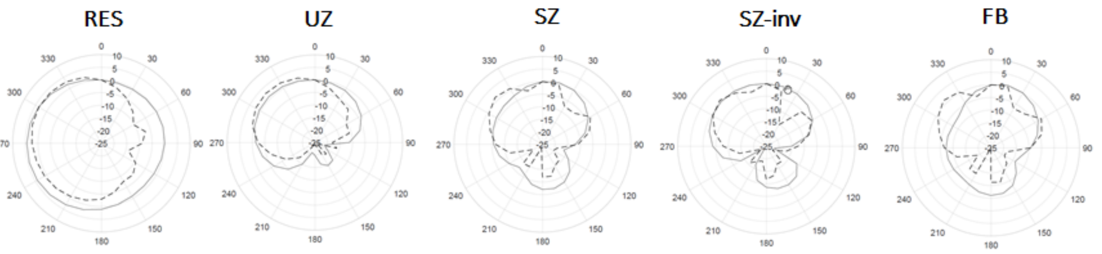

**Fig. 1:** Polar patterns of the five beamformer settings (left ear) calculated in octave bands with centre frequencies of 517 (low frequencies, solid line) and 1981 (high frequencies, dashed line). The azimuth is in degrees and the gain in dB.

### Acoustic scenarios and speech-in-noise measurements

The different beamformer conditions were tested in two different acoustic scenarios. In both cases, the Oldenburg sentence test (OLSA; Wagener *et al.*, 1999) was performed with the target speech presented from 0° and 1-m distance. The participants' task was to repeat as many of the five words per sentence as possible. For the background noise, two different masker scenarios were implemented:

(1) *Lateral interferers*: 10 sentences of a male speaker of an alternatively recorded OLSA (Hochmuth *et al.*, 2015) were concatenated without any pauses and presented from two loudspeakers placed at ±60°. To ensure that different sentences were played from both speakers, an offset of about 9s between the speakers was applied.

(2) *Diffuse interferer:* Recording made in a large cafeteria ($T_{60}$ = 1.25 s) during a busy lunch hour (Kayser *et al.*, 2009) presented through 11 loudspeakers placed around the participant in 30° steps (excluding 0°).

For each combination of acoustic scenario and beamformer condition, two speech reception threshold (SRT) measurements were determined per participant. A correlation analysis revealed high test-retest reliability ($r$ = .81, $p$ < .0001). For the statistical analyses, the average SRT per condition was used.

### Spatial awareness measurements

In addition to the speech-in-noise measurements, sound localization was assessed using both static and dynamic stimuli. A traffic-junction scene was simulated using TASCAR (Toolbox for Acoustic Scene Creation and Rendering) (Grimm *et al.*, 2015). The overall level of the noise scenario was 60.2 dB SPL at the listening position. The total length of the traffic-junction scene was 360 s. Within this scene, different target stimuli were presented at random time intervals:

(1) *Static localization*: Barking dog 1 (length: ~3.7 s; overall level: 68.4 dB SPL) or barking dog 2 (length: ~3.0 s; overall level: 68.5 dB SPL), placed at different angles (0°, ±45°, ±60°, ±75° or ±90°), presented twice per angle.

(2) *Dynamic localization*: Ambulance (length: ~4.5 s; angular velocity: ~13°/s; overall level: 68.3 dB SPL) or car horn (length: ~4.4 s; angular velocity: ~13°/s; overall level: 68.5 dB SPL), moving on a circle around the subject (-90°➔-30°; -30°➔+30°; +30°➔+90°; +90°➔+30°; +30°➔-30°; -30°➔-90°), with constant velocity. Three measurements per target movement were conducted.

The task of the participant was to pay attention to the different signals by turning the head into the direction of the source (static signals) or following the sources by moving the head synchronously (moving signals). The trajectories were recorded using a head tracker. The order of the presentation angles/trajectories was varied randomly. The order of the two localization tasks/stimuli was randomized across participants.

## RESULTS

### Speech reception in noise

Do binaural hearing abilities and/or the masker scenarios correlate with speech intelligibility in noise? To answer this question the data were grouped either according to BILD+ or BILD- or to the masker scenario. Figure 2 shows the average SRTs for all beamformer conditions either pooled according to *masker scenario* (Figure 2 right) or according to the *BILD* (Figure 2 left). Figure 2 left shows a clear difference in speech intelligibility between BILD+ and BILD-. Figure 2 right shows a different pattern of speech intelligibility performance depending on the different beamformer conditions. When the noise scenario was diffuse the binaural beamformer conditions were better than for the noise scenario with lateral interferers: the better the DI the better the speech intelligibility in a diffuse noise scenario. A three factor ANOVA revealed a significant main effect of the *BILD* ($p$ < .000), *masker scenario* ($p$ < .005) and the *beamformer conditions* ($p < 0.05$). A posthoc test (Bonferroni corrected) showed that SZ is statistically significantly better than all other beamformer conditions regardless of the binaural hearing abilities (SZ⇔ RES ($p < .01$), SZ⇔ UZ($p < .01$), SZ⇔ SZ-inv ($p < .05$) SZ⇔ FB ($p < .05$)). In Figure 3 the individual SRT data are visualized according to the BILD values with a regression line plotted for each beamformer condition. For the *diffuse interferer* scenario, the regression lines are spread for small BILD values but much narrower for large BILD values. This suggests that for participants with good binaural hearing abilities the selection of the beamformer is not relevant for speech intelligibility in noise and should be individually selected based on other parameters (see section environmental awareness test). For participants with poor binaural hearing abilities, the beamformer providing the highest DI values should be selected as it allows for better speech intelligibility performance. Almost the opposite trend can be observed for the *lateral interferer* scenario: the choice of beamformer is not relevant for participants with poor binaural hearing abilities, as the listeners do not benefit from binaural cues anyway. Therefore, other parameters are potentially more relevant to select the most effective individual beamformer condition. For participants with good binaural hearing abilities, the beamformer condition that preserves most binaural cues showed the best speech intelligibility.

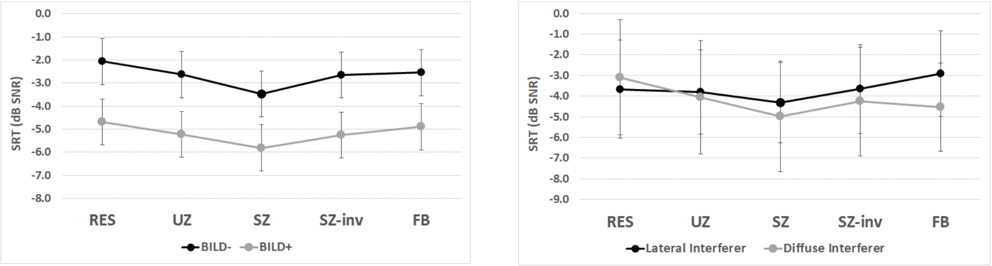

**Fig. 2:**Left panel: Average (standard deviation) speech reception thresholds (SRT) in noise for beamformer conditions grouped according to binaural hearing abilities: BILD < 2,5 dB (BILD-, black), BILD ≥ 2,5 dB (BILD+, grey). Right panel: Average (standard deviation) speech reception thresholds (SRT) in noise for beamformer conditions grouped according to masker scenarios: Lateral interferer (black), diffuse interferer (grey).

## Spatial awareness measurements

*Static localization*: The localization ratings for both static signals were averaged for analyses since no statistically significant differences was detected between the ratings for a given presentation angle. To condense the results, the localization errors were analysed depending on the angle deviation from the front, independent of hemisphere since no statistically significant differences were detected for the same angle deflections in either left or right direction from the front. Figure 4 (left) shows the distribution of localization errors for all beamformer conditions after this data condensation.

*Static localization*: The localization error was analysed by a repeated measures two factor ANOVA with the factors *beamformer condition* and *presentation angle*. Analyses revealed significant main effects of *beamformer condition* ($p$ < .001) and *presentation angle* ($p$ < .001). In addition, a significant interaction of *beamformer condition* and *presentation angle* was detected ($p$ < .001). A post hoc test (Bonferroni corrected) revealed that all *beamformer conditions* were significantly different from SZ and FB where SZ also performed significantly different from FB (all $p$ < .05). RES, UZ, and SZ-inv were not significantly different. The post hoc analysis for *presentation angle* revealed that 0° and 45° both significantly differed from 75° and 90°, as well as that 60° significantly differed from 90°. All other data were not significantly different. The between-subject factor *BILD* was not statistically significant.

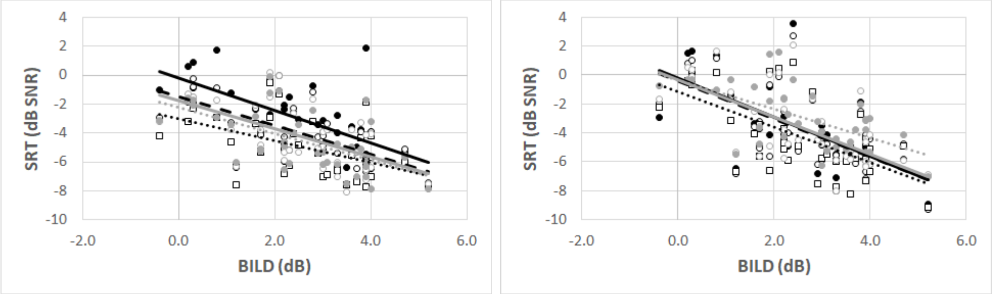

**Fig. 3:** Scatterplot of the BILD and SRT data. Left panel: diffuse interferer, right panel: lateral interferer. Least square regression lines corresponding to RES ( black solid line, filled black circles), UZ (dashed black line, unfilled black circles), SZ (dotted black line, unfilled black squares), SZ-inv (solid grey line, unfilled grey circles) and FB (dotted grey line, filled grey circles). SRT is SNR in dB

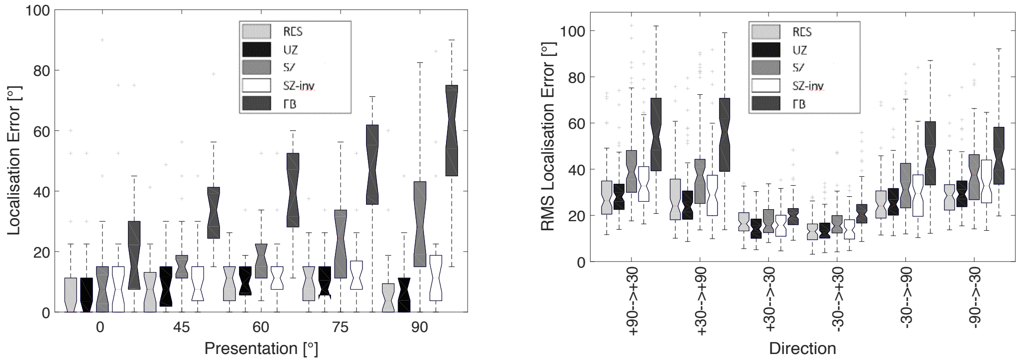

**Fig. 4:** Left panel: Box plots of the static localization errors in % for beamformer conditions and presentation angles. Right panel: Box plots of the RMS dynamic localization errors in % for beamformer conditions and trajectories.

*Dynamic localization*: The head movements of the participants were recorded as individual trajectories per participant and per measurement condition. In case of any discontinuities or missing data these were replaced by continuous completion. In addition, the individual trajectories were resampled (125Hz). From these trajectories the differences between the perceived angles and target angles were calculated resulting in the outcome measure RMS localization error (see Figure 4 right). For further analysis, the three measurement repetitions per target movement were averaged and analysed by a repeated measures two factor ANOVA with the factors *beamformer condition* and *direction* and individual *BILD* as covariate. Statistically significant main effects were revealed for *beamformer condition* ($p < .001$) and *target movement* ($p < .001$). In addition, an interaction trend was detected between *beamformer condition* and *BILD* ($p = 0.054$). A post hoc test for *beamformer condition* revealed that RES and UZ did not differ in RMS localization error but both showed significantly lower RMS localization errors than SZ-inv. SZ-inv yielded significantly lower RMS localization errors than SZ, that yield significantly lower RMS localization errors than FB (all $p < .05$). The post hoc analysis for *direction* revealed that the internal (-30°➔+30, +30°➔-30°) and external (all other directions) target movements significantly differed in terms of RMS localization error.

In summary the data showed that beamformer conditions with preservation of binaural cues at low frequencies provided better localisation performance for static and dynamic objects.

## SUMMARY AND CONCLUSIONS

We investigated the influence of binaural hearing abilities and acoustic scenarios on speech intelligibility in noise and on spatial awareness for five different beamformer approaches that are available in commercial hearing devices. The main differences between the beamformer approaches were the preservation of binaural cues, especially at low frequencies, and the DI, the ability to emphasize the target source from the front.

The analyses revealed that speech intelligibility in noise depends on binaural hearing abilities, masker scenario and beamformer conditions. Listeners with poor binaural hearing abilities have worse speech perception in noise compared with listeners with good binaural hearing abilities. An interaction effect between masker scenario and beamformer was demonstrated as well, but there was no interaction effect between binaural hearing abilities and beamformer condition. A post hoc analysis revealed that the commercially available beamformer SZ outperformed all other beamformers, independent of masking scenario and binaural hearing abilities. This means that we could only partly replicate the results of Neher *et al.* (2017), a study that included the same participants. This may be due to some differences in the set-up of the study, such as allowing for real head movements, real acoustic scenarios, and real acoustic couplings. Additionally, the algorithms differed slightly from those in the Neher *et al.* study as the systems used here were already fine-tuned to be effective under real life conditions. Therefore, we could not emanate from a clear distinction between diotic and dichotic output of the hearing devices which was the case in the Neher *et al.*

study's experimental setup. This can explain the higher variances in this study. The additional measurements of environmental awareness revealed a clear advantage of algorithms that preserve the binaural cues at low frequencies when localizing static or moving sound sources in a noisy environment.

Together, these findings of the study provide a basis for adapting beamformer (settings) in commercial hearing devices to the individual binaural hearing abilities and the noise situation meaning that both speech understanding and aspects of spatial awareness perception should be considered.

## REFERENCES

Dillon, H. (**1999**) "NAL-NL1: A prescriptive fitting procedure for non-linear hearing aids," The Hear J. **58**(10): 10-16.

Dillon, H. (**2012**). "Hearing Aids," 2nd ed., Boomerang Press, Sydney, Australia.

Gatehouse, S., Naylor, G., and Elberling, C. (**2003**). "Benefits from hearings aids in relation to the interaction between the user and the environment,' Int. J. Audiol. **42**(suppl.1). 77-85.

Grimm, G., Luberadzka, J., Herzke, T., and Hohmann, V. (**2015**). "Toolbox for acoustic scene creation and rendering (tascar): Render methods and research applications," In F. Neumann, editor, Proceedings of the Linux Audio Conference, Mainz, Germany, 2015. Johannes-Gutenberg Universitat Mainz.

Hochmuth, S., Jürgens, T., Brand, T., and Kollmeier, B. (**2015**). " Talker- and language-specific effects on speech intelligibility in noise assessed with bilingual talkers: Which language is more robust against noise and reverberation?," Int. J. Audiol. **54**:sup2, 23-34.

Kayser, H., Ewert, S.D., Annemüller, J., Rohdenburg, T., Hohmann, V., and Kollmeier, B. (**2009**). "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse response," EURASIP J. Adv. Signal Process., **298605**. 1-10.

Kollmeier, B. (**1996**). "Computer-controlled speech audiometric techniques for the assessment of hearing loss and the evaluation of hearing aids," In; Kollmeier, B. (Ed.) Psychoacoustics, Speech and Hearings Aids. World Scientific, Singapore, pp. 57-68.

Latzel, M. (**2013**). "StereoZoom and auto StereoZoom," Phonak compendium.

Lunner, T. (**2003**). "Cognitive function in relation to hearing aid use,' Int. J. Audiol. **42**(suppl.1), 49-58.

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., Kühnel, V., 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. Audiol. (Audiol. Acoustics). **38**. 4-15, 44-56, 86-95.