Influence of multi-microphone signal enhancement algorithms on auditory movement detection in acoustically complex situations

MICHA LUNDBECK[1,2], LAURA HARTOG[1,2], GISO GRIMM[1,2], VOLKER HOHMANN[1,2], LARS BRAMSLØW[3], AND TOBIAS NEHER[1,4]

[1] *Medizinische Physik and Cluster of Excellence Hearing4all, Oldenburg University, Oldenburg, Germany*

[2] *HörTech gGmbH, Oldenburg, Germany*

[3] *Eriksholm Research Centre, Snekkersten, Denmark*

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

The influence of hearing aid (HA) signal processing on the perception of spatially dynamic sounds has not been systematically investigated so far. Previously, we observed that interfering sounds impaired the detectability of left-right source movements and reverberation that of near-far source movements for elderly hearing-impaired (EHI) listeners (Lundbeck *et al.*, 2017). Here, we explored potential ways of improving these deficits with HAs. To that end, we carried out acoustic analyses to examine the impact of two beamforming algorithms and a binaural coherence-based noise reduction scheme on the cues underlying movement perception. While binaural cues remained mostly unchanged, there were greater monaural spectral changes and increases in signal-to-noise ratio and direct-to-reverberant sound ratio as a result of the applied processing. Based on these findings, we conducted a listening test with 20 EHI listeners. That is, we performed aided measurements of movement detectability in two acoustic scenarios. For both movement dimensions, we found that the applied processing could partly restore source movement detection in the presence of reverberation and interfering sounds.

INTRODUCTION

Listeners with sensorineural hearing loss exhibit considerable difficulties in complex acoustic environments. Hearing aids (HAs) can help by restoring audibility and by improving the signal-to-noise ratio (SNR). This can improve speech reception in noise, but it may also compromise spatial hearing abilities including movement perception. To start addressing this possibility, we recently conducted a study where we observed that for elderly hearing-impaired (EHI) listeners interfering sounds impaired the detectability of left-right source movements, and reverberation that of

*Corresponding author: micha.lundbeck@uni-oldenburg.de

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

near-far source movements (Lundbeck *et al.*, 2017). These results raise the question of how to compensate these deficits with HAs. In the current study, we therefore investigated the influence of different multi-microphone signal enhancement algorithms on source movement detection in acoustically complex situations. To that end, we used a higher-order Ambisonics-based system for simulating complex sound scenes together with a computer simulation of bilateral multi-microphone HAs. To start with, we investigated the influence of different multi-microphone signal enhancement algorithms on acoustic measures that are presumed to be related to movement perception. We then evaluated the most promising HA settings in a listening test to explore the potential of improving source movement detection with HAs. In summary, the current study had the following aims:

1. To identify HA settings that can enhance acoustic cues that are presumed to underlie left-right (L-R) and near-far (N-F) source movement detection;
2. To evaluate the most promising HA settings for improving L-R and N-F source movement detection with a group of EHI listeners.

METHODS

Experimental setup

We simulated a complex acoustic environment using a toolbox for creating dynamic virtual environments (TASCARpro version 0.128; Grimm *et al.*, 2015). We configured our setup such that it produced 48 virtual loudspeaker signals in the horizontal plane with a spatial resolution of 7.5°. The virtual listener was seated at the center of the loudspeaker array. As the aim of this study was to include different HA algorithms, we generated multi-microphone signals by convolving the loudspeaker signals with binaural room impulse responses from the database of Thiemann and van de Par (2015) for the corresponding directions.

Stimuli

We made use of five different environmental sounds. For the target, we used a broadband noise-like fountain signal (S1; at 0° azimuth and 1 m distance re. the listener in the reference position). As interfering sounds, we used recordings of ringing bells, bleating goats, pouring water and humming bees (S2-S5: at ±45° and ±90° and 1 m distance each). We presented the target sound (S1) at 65 dB SPL (nominal) and the other sounds (S2-S5) at 62 dB SPL (nominal) each, as measured under reverberant conditions at the position of the virtual listener. The duration of each sound was 3.1 s.

HA signal processing

We used the Master Hearing Aid (MHA) research platform (Grimm *et al.*, 2006) for simulating five HA settings: *unproc*, *dir*, *coh*, *dircoh*, and *beam*. The *unproc* condition corresponded to a pair of omnidirectional microphones that we simulated using the front microphones of two behind-the-ear (BTE) devices without any additional processing. The *dir* condition corresponded to a pair of static forward-facing cardioid microphones (e.g., Dillon, 2012), which were realized based on the front and rear microphone signals of each BTE device. We then spectrally equalized the output signals to ensure that the frontal target signals sounded highly similar across the unproc and dir settings. The *coh* condition corresponded to a binaural noise reduction scheme for attenuating incoherent signal segments (Grimm *et al.*, 2009). The gains applied to the left and right channels were always the same, so that interaural level and time differences (ILDs and ITDs) were unaffected, while incoherent sounds (as caused by early reflections and late reverberation, for example) were attenuated. The *dircoh* condition consisted of the serial combination of the dir and coh settings. The *beam* setting corresponded to a bilateral beamforming algorithm with a post-filter for binaural cue preservation (Rohdenburg *et al.*, 2007). For the current study, we used six input signals (three per side) and the front BTE microphone signals as reference signals for the binaural post-filter. We then also spectrally equalized the output signal so that the frontal target signals sounded similar across the unproc and beam settings. In the following, we will concentrate on the dircoh and beam settings, as they showed the clearest effects relative to the unproc setting.

Technical measurements

General setup and procedure

For the technical measurements, we generated stimuli based on the median L-R and N-F detection thresholds of the EHI listeners tested previously (Lundbeck *et al.*, 2017). Specifically, we generated stimuli where the target signal moved 28° in the L-R direction or 1.5 m in the N-F direction re. the reference position. The signal processing chain used for the acoustical analyses is shown in Fig. 1.

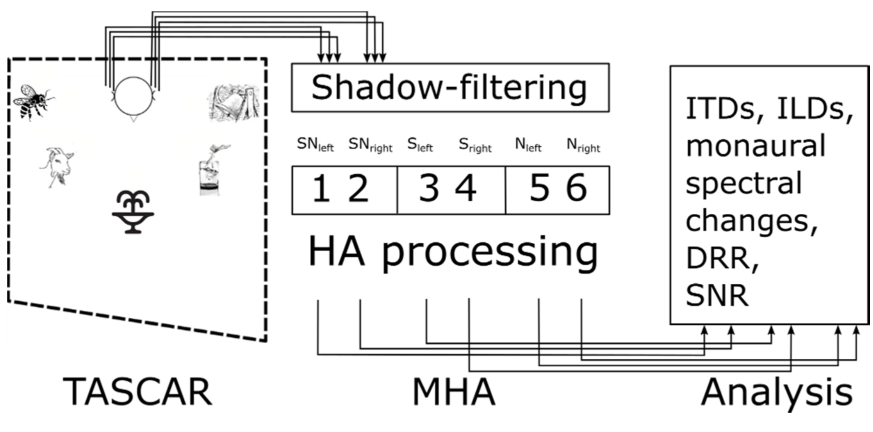

Fig. 1: Signal processing chain used for the acoustical analyses. Following the generation of the stimuli using TASCAR (left) and the shadow-filtering in the MHA (middle), different output channels (1-6) were analyzed using different measures (right). SN_{left}, SN_{right} = Left and right channels of the signal mixture; S_{left}, S_{right} = Left and right channels of the target signal; N_{left}, N_{right} = Left and right channels of the interfering signals.

We equipped the virtual listener with two BTE devices with up to three microphones each. We then processed the microphone signals with the MHA. We used the so-called shadow-filtering method to apply the processing computed for the signal mixture separately to the target and interferers. Depending on the measure of interest (see below), we then analyzed different output signals. To reveal short-time changes in the chosen measures, we used a 100-ms analysis window with 50% overlap.

Monaural spectral changes

To analyze the influence of the different HA settings on monaural spectral cues, we applied a spectral coloration measure of Moore and Tan (2004). We always analyzed the stimulus channel ipsilateral to the movement direction and referenced it to the stationary equivalent of the same stimulus. In this way, we measured relative monaural spectral changes due to the source movement and the HA settings.

Signal-to-noise ratio (SNR) changes

For estimating the SNR, we used the separate target and interferer signals (see middle panel of Fig. 1, channels 3+4 and 5+6, respectively). We then calculated the short-term level ratio between the target and the interferers at either the ipsilateral side (L-R dimension) or averaged across the two sides (N-F dimension).

Direct-to-reverberant sound ratio (DRR) changes

For the stimuli moving along the N-F dimension, we estimated short-term changes in the DRR. To that end, we created two stimuli per condition: one with and one without reverberation. We then subtracted the anechoic stimulus (comprising the direct sound only) from the reverberant stimulus and fed the direct and reverberant sound separately into the MHA. By comparing the DRR at the input and output of the MHA, we could estimate DRR changes due to the applied HA processing.

Perceptual measurements

Participants

For the perceptual measurements, we used 20 EHI listeners aged 63-80 yr (mean: 72.4 yr). Fifteen of them had bilateral HA experience of at least 2 yr. All participants had symmetric, sloping mild-to-moderate sensorineural hearing losses. We divided the participants into two groups according to their performance on a target detection task (see below). The mean pure-tone average hearing loss calculated across 0.5, 1, 2 and 4 kHz and both ears (PTA4) differed significantly across the two groups (group 1: 58 dB HL; group 2: 47 dB HL; $p < 0.001$), whereas age did not (group 1: 75 yr; group 2: 70 yr; $p > 0.1$).

General setup and procedure

To investigate the perceptual consequences of the tested HA settings, we carried out a listening test with 20 EHI listeners. Initially, we assessed each listener’s ability to detect the target signal in the presence of the four interferers. For the participants who could not consistently detect the target signal in the presence of the interferers (group 1; $N = 9$), we performed the movement detection threshold measurements without the interferers. For the other participants (group 2; $N = 11$), we performed the measurements with all five signals.

The listening test was carried out under reverberant conditions ($T_{60} \approx 0.8$ sec). Stimulus presentation was via a 24-bit RME (Haimhausen, Germany) Hammerfall DSP 9632 soundcard, a Tucker-Davis Technologies (Alachua, USA) HB7 headphone buffer and a pair of Sennheiser (Wennebostel, Germany) HDA200 headphones. For the psychoacoustic measurements, we used the "psylab" toolbox (Hansen, 2006). To ensure adequate audibility for each participant, we spectrally shaped all stimuli in accordance with the "National Acoustics Laboratories–Revised-Profound" (NAL-RP) fitting rule (Dillon, 2012).

Source movement detection thresholds

We presented stimuli with moving target sounds on half of the trials and stimuli with static target sounds (at the reference position) on the other trials. For the angular measurements, we randomized the direction of movement (towards the left or right). For the radial measurements, we always simulated a withdrawing (N-F) movement. In this way, the starting position of the target sound source was the same in all conditions (0°, 1 m re. the listener). To control the extent of the movement, we varied the velocity (in °/s or m/s) in the adaptive procedure. For the adaptive procedure, we used the single-interval-adjustment-matrix procedure of Kaernbach (1990) to ensure unbiased measurements. A run was terminated after 12 reversals, and the first four reversals were discarded from the analyses. Before the actual measurements, each participant completed two training runs (one with *unproc* and one with *beam*).

We estimated the detection thresholds by taking the arithmetic mean of the last eight reversal points. In this manner, we quantified the smallest displacement (in ° or m) of the target source that the participants could perceive within the 2.3 s over which the movements occurred. In the following, we will refer to these thresholds as the minimum audible movement angle (MAMA) or distance (MAMD) thresholds. We performed the L-R and N-F measurements in separate blocks. Within each block, we tested the various conditions in randomized order. After 1-2 weeks, we conducted retest measurements. In total, we measured six detection thresholds per movement dimension (L-R and N-F) and listener (and thus 240 thresholds in total). According to Kolmogorov-Smirnov's test, all datasets fulfilled the requirement for normality (all $p > 0.05$). We therefore used parametric statistics to analyze our data. Whenever appropriate, we corrected for violations of sphericity using the Greenhouse-Geisser correction.

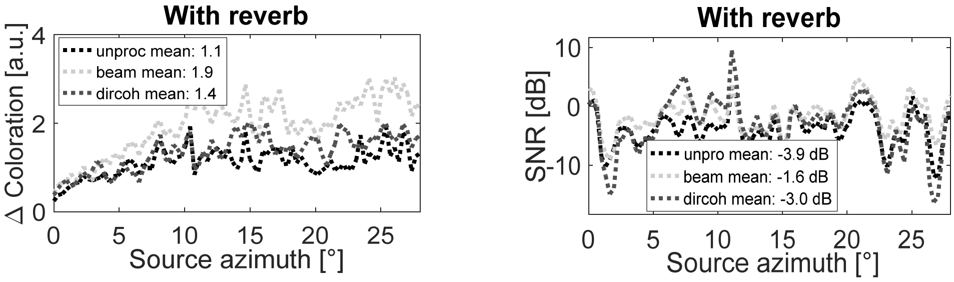

Fig. 2: Left panel: Monaural spectral coloration re. a static stimulus subjected to the same processing for unproc (black), beam (light gray) and dircoh (dark gray) as a function of source azimuth. Right panel: SNR for unproc (black), beam (light gray) and dircoh (dark gray) as a function of source azimuth. Legends show mean values for the different HA settings calculated over the whole stimulus duration.

RESULTS

Technical measurements

L-R dimension

Concerning the L-R dimension, the changes in the measures of interest that we observed were generally as expected. Regarding the monaural spectral changes, our analyses revealed that the beam and dircoh settings both increased this measure, suggesting that they are suited for improving source movement detectability. The left panel of Fig. 2 shows the resultant spectral coloration relative to the static condition in the presence of the four interferers and reverberation.

The right panel depicts the SNR caused by the three HA settings over the course of the target source movement in the presence of the four interferers. It is noticeable that the SNR varied substantially over the course of the source movement. This was because of the spectro-temporal fluctuations inherent to the environmental sounds that we used. Concerning the influence of dircoh and beam, beam increased the SNR more relative to unproc.

N-F dimension

Concerning the N-F dimension, the changes in the chosen measures were generally as expected (data not shown). The DRR generally decreased with increasing source distance, irrespective of the HA setting. Furthermore, the beam and especially the dircoh setting increased the DRR. The same was essentially true for the monaural spectral coloration, suggesting that monaural spectral cues may provide salient information about source movements. Regarding the SNR improvement relative to unproc, beam and especially dircoh led to clear increases.

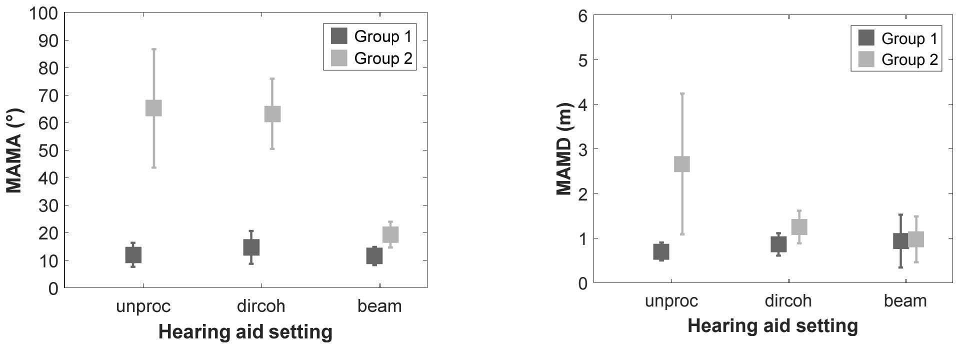

Fig. 3: Means and standard deviations of the MAMA (left) and MAMD (right) thresholds for the two groups and three HA settings.

Perceptual measurements

L-R dimension

Figure 3 (left panel) shows means and standard deviations of the MAMA thresholds for the two groups and three HA settings. For group 1, the thresholds varied little across HA settings and listeners. For group 2, the thresholds were much higher with unproc and dircoh than with beam. Furthermore, unproc was characterized by the largest spread and beam by the smallest spread.

To test for statistical differences among the three HA settings, we conducted two analyses of variance (ANOVA), that is, one per group with the within-subject factor HA setting (unproc, dircoh, beam). For group 1, we found no effect of HA setting [$F(2,16) = 2.5$, $p = 0.14$]. For group 2, the effect of HA setting was highly significant [$F(2,20) = 38.1$, $p < 0.0001$]. A series of planned contrasts showed that the beam setting differed significantly from both unproc and dircoh (both $p < 0.001$).

N-F dimension

Figure 3 (right panel) shows means and standard deviations of the MAMD thresholds for the two groups and three HA settings. As can be seen, group 1 obtained thresholds of around 1 m or lower in all conditions. In other words, the different HA settings did not appear to affect their performance. In contrast, for group 2 there was a clear influence of HA setting on movement detectability. To test for statistical differences among the three HA settings, we conducted an ANOVA per group with HA setting (unproc, dircoh, beam) as within-subject factor. For group 1, the effect of HA setting was not significant [$F(2,14) = 1.8$, $p = 0.2$], while for group 2 it was strongly significant [$F2,18 = 13.6$, $p < 0.001$]. A series of planned contrasts showed that the beam and dircoh settings differed significantly from unproc (both $p < 0.05$) and also from each other ($p < 0.01$).

Summary

The current study, which we conducted based on a setup for simulating complex virtual environments, showed that selected multi-microphone signal enhancement algorithms can enhance acoustical cues presumed to underlie source movement perception. Furthermore, the subsequent listening test showed substantial improvements in source movement detectability for a group of EHI listeners in complex scenarios with reverberation and interfering signals. In view of the fact that our study focused on one particular spatial dimension (i.e., source movement detection), it is of interest to extend this research to other aspects of spatial awareness perception and to head-worn HAs in future studies.

ACKNOWLEDGEMENTS

This research was funded by the Oticon Foundation and the DFG Cluster of Excellence EXC 1077/1 “Hearing4all”.

REFERENCES

Dillon, H. (**2012**). *Hearing aids / Harvey Dillon* (Boomerang Press; Thieme, Sydney, New York).

Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (**2006**). “The master hearing aid: a PC-based platform for algorithm development and evaluation,” Acta Acust United Ac., **92**, 618-628.

Grimm, G., Hohmann, V., and Kollmeier, B. (**2009**). “Increase and subjective evaluation of feedback stability in hearing aids by a binaural coherence-based noise reduction scheme,” IEEE T. Audio Speech, **17**, 1408-1419.

Grimm, G., Luberadzka, J., Herzke, T., and Hohmann, V. (**2015**). “Toolbox for acoustic scene creation and rendering (TASCAR)-Render methods and research applications,” Proceedings of the Linux Audio Conference, Mainz.

Hansen, M. (**2006**). “Lehre und Ausbildung in Psychoakustik mit psylab: Freie Software fur psychoakustische Experimente,” Fortschritte der Akustik, **32**, 591.

Kaernbach, C. (**1990**). “A single-interval adjustment-matrix (SIAM) procedure for unbiased adaptive testing,” J. Acoust. Soc. Am., **88**, 2645-2655.

Lundbeck, M., Grimm, G., Hohmann, V., Laugesen, S., and Neher, T. (**2017**). “Sensitivity to angular and radial source movements as a function of acoustic complexity in normal and impaired hearing,” Trends Hear., **21**, 2331216517717152.

Moore, B.C., and Tan, C.-T. (**2004**). “Development and validation of a method for predicting the perceived naturalness of sounds subjected to spectral distortion,” J. Audio Eng. Soc., **52**, 900-914.

Rohdenburg, T., Hohmann, V., and Kollmeier, B. (**2007**). “Robustness analysis of binaural hearing aid beamformer algorithms by means of objective perceptual quality measures,” in *Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop on* (IEEE), pp. 315-318.

Thiemann, J., and van de Par, S. (**2015**). “Multiple model high-spatial resolution HRTF measurements,” Proc. DAGA 2015.