Sound localization with and without hearing aids

T. Van den Bogaert[1], E. Carette[1], J. Wouters[1]

[1] *ExpORL, K.U.Leuven, Herestraat 49 bus 721, B-3000 Leuven, Belgium, Email: tim.vandenbogaert@med.kuleuven.be*

Introduction

Building a realistic spatial representation of our surrounding environment is an important task of our auditory system. This allows us to localize sounds and to understand speech in adverse scenarios [4],[5]. An inaccurate spatial map will degrade localization and speech perception and will lead to an unnatural perception of sounds. The psycho-acoustical mechanisms used to localize sound sources are fairly well known [6]. The human auditory system mainly uses interaural cues such as interaural level differences (ILDs) and interaural time differences (ITDs) [7, 8] to localize sounds in the horizontal hemisphere. However, sound sources positioned in the front and the back hemisphere generate almost identical interaural properties when they are positioned on the so called cone of confusion. Therefore the auditory system relies on the spectral filtering of the head, torso and especially pinna to resolve the front-back confusions [9, 10]. Each human listener possesses two highly personalized pinnae. The cavities of the pinna reflect and diffract a sound wave dependent on the angle of arrival of the signal. This can be used to resolve front-back confusions and to localize sounds in the vertical plane.

The question whether hearing aids preserve or transmit enough binaural cues to the hearing aid user and whether hearing aid users are still capable of using these cues has been gaining a lot of interest recently [11] [12]. This is partly due to recent technology which enabled the design of binaural hearing aids in which binaural cues may be controlled and preserved. However, the question whether monaural spectral information is preserved by different hearing aids has been studied to a lesser extent. Still some attention has been given to the question whether a natural microphone positioning, such as used in in-the-canal (ITC) hearing aids, has a benefit in terms of preserving monaural and binaural cues. Results of these studies, done more than a decade ago, have been conflicting. Studies such as [2] and [3] found no significant benefit of having a microphone in the ear compared to behind the ear, while [1] did claim a benefit of microphone position. In [3], these differences were explained by difference in test setup since all three studies investigated different dimensions of sound localization simultaneously leading to results which are hard to compare.

Three different dimensions of sound localization are discussed in this manuscript: localization in the frontal horizontal hemisphere, localization in the frontal vertical plane and the ability to resolve front-back confusions. The localization experiment in the frontal horizontal hemisphere is dominated by the processing of binaural cues. The two other dimensions are known to be dominated by the processing of monaural spectral information. Several hearing aids were evaluated: behind-the-ear (BTE), in-the-ear (ITE), but also a newly designed in-the-pinna (ITP) hearing aid. This new hearing aid, i.e. the Be hearing aid, released by GN Resound, uses a flexible positioning mechanism which positions its microphone in the anthelix of the pinna of the hearing impaired subjects. This should give benefits in terms of wind noise reduction and naturalness of sounds. This has renewed the interest in studying the effect of microphone positioning on the monaural spectral cues and on sound localization performance.

Evaluation

Methods

Tests were carried out in a room with a reverberation time of $T_{60} = 0.61s$. For left-right localization, test subjects were placed inside a horizontal array of 13 loudspeakers. The array is illustrated in Figure 1. Speakers were located from -90°(at the left side of the subject) to 90°(at the right side of the subject), with a spacing of 15°and were labelled 1 to 13. For the front-back localization task, listeners were turned side-on to the array, with their right ear facing the array of speakers. A resolution of 30°was used during this experiment by using only the 7 odd numbered loudspeakers. Subjects were seated in a vertical array of loudspeakers during vertical plane localization. Ten speakers were placed at 1m distance of the subject from -45°to 90°with a spacing of 15°. The speaker placed at 0°was located in the horizontal plane while the speaker at 90°was located directly above the subject. During all tests a roving level of 6dB was used to avoid the use of loudness cues. All experiments were done in a test-retest configuration. At present, no significant difference between test and retest has been found for all tests.

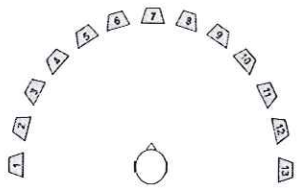

Figure 1: The speaker array used for left-right and front-back localization.

7 hearing impaired subjects were evaluated using different hearing aids. These subjects were used to wearing behind-the-ear (BTE) hearing aids in daily live. The hearing impaired subjects were fitted with GN ReSound

in-the-canal Canta7 or Azure (ITC), GN Resound Be in-the-pinna (ITP), GN Resound Air BTE (BTE-A) and Oticon XW Epoq receiver in the ear, BTE hearing aids (BTE-E). The left-right communication link of the BTE-E was switched on and all optional noise reduction schemes were switched off. An unaided condition, in which the stimuli were compensated for audibility was added to the protocol as a reference. The latter condition was currently only completed for the left-right localization task. The audibility correction was done by using the half gain rule to amplify the stimulus per octave band.

Left-right

First, left-right localization in the frontal horizontal plane was studied. Three stimuli were presented to the subjects: a broadband steady state noise with an average spectrum of a dutch male talker (VU-noise, [13]), a 1/3rd octave noise band centered around 500Hz and a 1/3rd octave noise band centered around 3150Hz. A summary of the data is presented in Figure 2. Since [11] studied localization with normal hearing (NH) subject and hearing impaired (HI) subjects with bilateral hearing aids in very similar conditions, these data were added to the figure (last 3 bars).

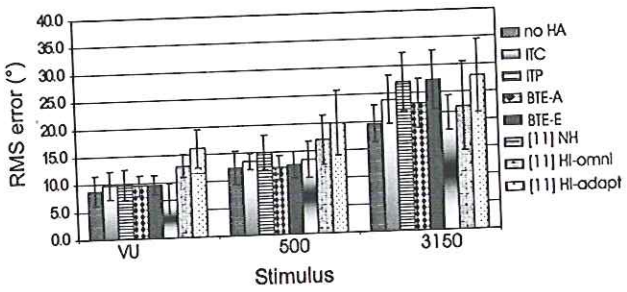

Figure 2: Average RMS errors and group standard deviations from the left-right localization experiment. The data obtained in [11] in similar conditions are also added to the figure.

First, it was observed that broadband stimuli were much easier localized than narrowband stimuli. Using only low-frequency cues led to a better localization performance when compared to using only high-frequency information. It can be assumed that the first condition (500Hz) is dominated by binaural processing of ITDs while the second condition (3150Hz) is dominated by the processing of ILD information. When presenting both ITD and ILD information, localization accuracy improved significantly (VU-stimulus). Second, and probably most importantly is the fact that when comparing the unaided with the aided conditions, averaged over all stimuli, only the BTE-A condition shows no difference with the unaided performance. All the other hearing aids introduced a small but significant degradation in left-right localization. The observed differences are especially prominent when localizing the high-frequency sound. No significant differences were found between hearing aids although a general trend is found that localization with the BTE-A hearing aids was more accurately than with the other hearing aid systems.

These conclusions are very similar to the work presented in [11] and in [12]. In these studies it was found that hearing aids introduce a small but significant degradation in sound localization in the horizontal plane. This degradation becomes significantly larger when switching on additional algorithms such as an adaptive directional microphone.

Front-back

Since, broadband information is needed to distinguish the front from the back hemisphere, front-back experiments were only performed with the broadband vu stimulus. 6 repetitions were used per loudspeaker leading to 42 presentations per test. The accumulated results of the hearing aid users are presented in Figure 3.

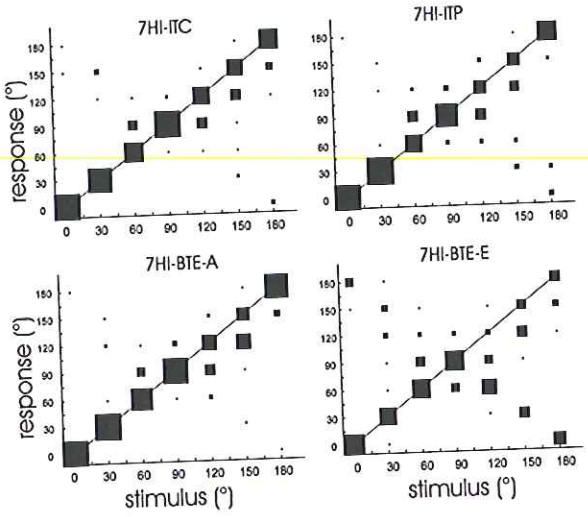

Figure 3: Accumulation of responses (y-axis) given by the group of hearing impaired subjects wearing different types of hearing aids (N=7).

As shown in Figure 3, the results obtained with the different hearing aid systems were surprisingly good except for the BTE-E condition. An error score of 8%±9, 11%±14, 4%±7 and 30%±8 was obtained when using ITC, ITP, BTE-A and BTE-E hearing aids respectively.

To identify the amount of spectral cues, generated by each hearing aid, objective measurements were performed. First, a CORTEX MK2 manikin was fitted with the different hearing aids. Afterwards the stimulus was played from all the different loudspeakers and the sound arriving at the eardrums of the manikin were recorded. Then the spectra of these recordings were calculated and normalized with respect to its overall RMS level. This way, loudness cues were removed from the recordings. To visualize the information which could be used to distinguish the front from the back (in this case 30°from 150°), the spectra of these angles were subtracted and the absolute value was taken. This shows the spectral information which could be used by a listener to discriminate 30°from 150°or the front from the back. The result of this procedure is shown in Figure 4.

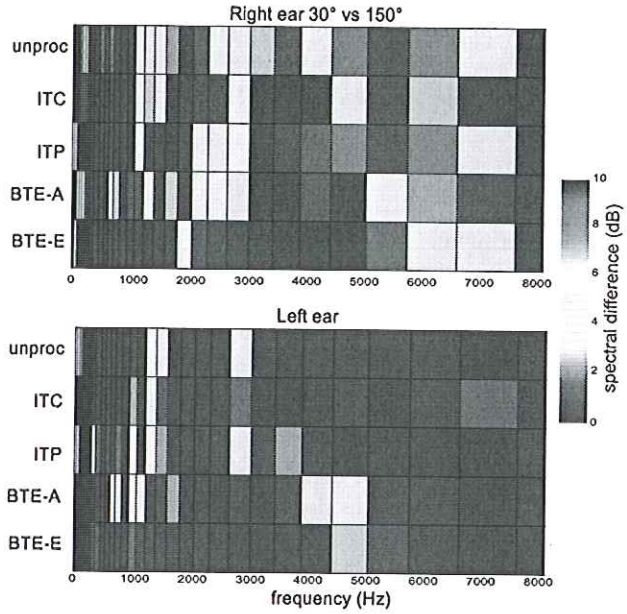

Figure 4: The spectral information present in each hearing aid system to distinguish 30°from 150°.

In this figure, it is observed that especially the right ear, which is the ear facing the loudspeakers, receives spectral information to distinguish 30°from 150°. When comparing the ITC and ITP hearing aids with the unprocessed condition, it is observed that these device almost preserve all spectral information. Both BTE hearing aids however, do not preserve the spectral information. This can be explained by the fact that the microphone of a BTE hearing aid is positioned on top of the pinna instead of close to the eardrum. Still, it is observed that the BTE-A hearing aid preserves some spectral information (between 2 and 3kHZ) while the BTE-E hearing aid removes almost all information. The perceptual evaluation shows that the little information preserved by the BTE-A seemed to be sufficient for the hearing impaired subjects to distinguish the front from the back while performance dropped significantly when using the BTE-E. It should be noted that Figure 4 illustrates the potential present in the different hearing aids. The hearing aid should produce enough amplification in all frequency regions such that the spectral information can be used by the hearing impaired user. Combining the good perceptual performance obtained with the BTE-A hearing aid and the spectral measurements suggests that although positioning a microphone on top of the pinna reduces spectral information, it is not necessarily the main limiting factor with respect to front-back localization.

Elevation

The same group of hearing aid users were evaluated on their capability to localize sounds in the frontal vertical plane. Again the broadband noise was used. An accumulation of the results is given in Figure 5.

Figure 5 illustrates that the differences between the different groups are relatively small. Averaged over all subjects, an RMS error score of 23.2±4.4, 27.5±7.0,

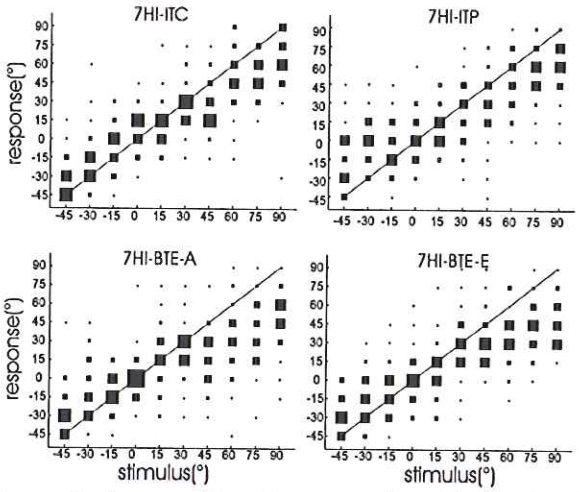

Figure 5: Accumulation of responses (y-axis) given by the group of hearing aid users (N=7) wearing different types of hearing aids on the elevation task.

30.1±9.0 and 27.8±5.7 was obtained by the hearing aid users with ITC, ITP, BTE-A and BTE-E hearing aids respectively.

To quantify the amount of spectral cues preserved or generated by each hearing aid, spectra were measured using a CORTEX MK2 manikin in a similar way as discussed during the front-back experiment. Each spectrum was normalized with respect to its own average RMS level and with respect to 90°. This way, overall loudness cues were removed from the spectrum and a figure is created which illustrates the spectral information which could be used to distinguish an angle x from 90°in elevation. The spectral information generated by the different hearing aid configurations for an angle of 0°are shown in Figure 6. Due to the symmetric test setup, measurements performed with the left and the right hearing aid were similar. Therefore, only the measurements at the right ear are shown.

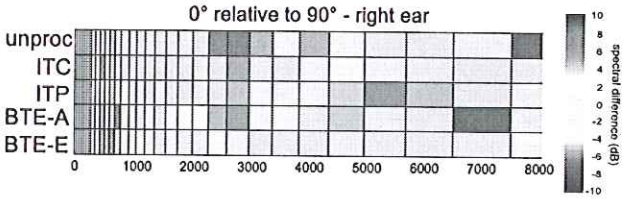

Figure 6: The spectral information present in each hearing aid system to identify 0°of elevation. The spectra are normalized with respect to their own average RMS to remove overall loudness cues and with respect to 90°.

Figure 6 shows that although large variations are present between the recorded spectra, all hearing aids are capable of generating useful elevation cues. This corresponds to the perceptual data in which no large differences were found between hearing aids. The measured spectra show that especially the region between 2 and 3kHz is well preserved by all systems. This is the same region which was well preserved by most hearing aids when analyzing the front-back spectra (Figure 4). In [10], it was shown that this region is dominated by the spectral filtering

of head and torso. Moreover, they could demonstrate that normal hearing subjects could be trained to localize sounds in the vertical plane by only using this kind of filtering. It can be assumed that hearing impaired subjects, used to wearing a BTE hearing aid and used to listening to low frequency dominated sounds (since a sloping hearing loss is the most common type of hearing loss) have been highly trained on using this kind filtering.

Conclusion

This manuscript presents a study on sound localization by hearing impaired subjects in three different dimensions: left-right localization in the frontal horizontal plane, front-back discrimination in the horizontal plane and elevation. The first dimension is dominated by the processing of binaural cues while the other two are dominated by the processing of monaural spectral cues. This document shows that hearing aids can affect localization performance. Evidence was found of a reduced left-right accuracy and a reduced ability to resolve front-back confusions when using some of the hearing aids. Although microphone position has a very high impact on the spectral cues crucial to separate the front from the back hemisphere it is not necessarily the main limiting factor. This is probably due to the fact that hearing impaired subjects are highly trained on low frequency cues from head and torso to resolve front-back confusions.

Acknowledgements

We would like to thank all subjects for their tremendous listening efforts. This study was partially funded by GN Resound and FWO VIRTAK research project G.0334.06N.

References

[1] Westermann, S. et al. "Comparing BTEs and ITEs for localizing speech". Hearing Instruments, **36**(2):20–24, 1985.

[2] Leeuw, A.R. et al. "Speech understanding and directional hearing for hearing-impaired subjects with in-the-ear and behind-the-ear hearing aids". Scand. Audiol., **16**(1):31–36, 1987.

[3] Noble, W. et al. "A comparison of different binaural hearing aid systems for sound localization in the horizontal and vertical planes". British Journal of Audiology, **24**:335–346, 1990.

[4] Bronkhorst, A.W. et al. "The effect of head-induced interaural time and level differences on speech intelligibility in noise". J. Acoust. Soc. Am., **83**(4):1508–1516, 1988.

[5] Bronkhorst, A.W. et al. "Binaural speech intelligibility in noise for hearing impaired listeners". J. Acoust. Soc. Am., **86**(4):1374–1383, 1989.

[6] Hartmann, W.M. "How we localize sound". Physics today, pages 24–29, 1999.

[7] Kuhn, G.F. "Model for the interaural time differences in the azimuthal plane". J. Acoust. Soc. Am., **62**(1):157–167, 1977.

[8] Wightman, F.L. et al. "The dominant role of low-frequency interaural time differences in sound localization". J. Acoust. Soc. Am., **91**(3):1648–1661, 1992.

[9] Langendijk, E.H.A. "Contribution of spectral cues to human sound localization". J. Acoust. Soc. Am., **112**(4):1583–1596, 2002.

[10] Algazi, V.R. "Elevation localization and head-related transfer function analysis at low frequencies". J. Acoust. Soc. Am., **109**(3):1110– 1122, 2001.

[11] Van den Bogaert, T. et al. "Localization with bilateral hearing aids: without is better than with". J. Acoust. Soc. Am., **119**(1):515–526, 2006.

[12] Keidser, G. et al. "The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers". International Journal of Audiology, **45**(10):563–579, 2006.

[13] Versfeld, N.J. "Method for the selection of sentence materials for efficient measurement of the speech reception threshold". J. Acoust. Soc. Am., **107**(3):1671–1684, 2000.

Preserving binaural hearing of hearing impaired subjects with binaural noise reduction systems for hearing aids

T. Van den Bogaert[1], S. Doclo[2,3], J. Wouters[1], M. Moonen[2]

[1] *ExpORL, K.U.Leuven, Herestraat 49 bus 721, B-3000 Leuven, Belgium, Email: tim.vandenbogaert@med.kuleuven.be*

[2] *ESAT-SCD, K.U.Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium;*

[3] *Presently at NXP Semiconductors, Interleuvenlaan 80 (C034), 3001 Leuven, Belgium;*

Introduction

Noise reduction algorithms in hearing aids are important for hearing impaired persons to improve speech intelligibility in background noise. Multi-microphone systems are able to exploit spatial in addition to spectral information and are hence preferred to single-microphone systems [1], [2]. However, hearing aid users often localize sounds better when switching off the noise reduction in their hearing aids [3], [4]. This is not suprising, since noise reduction algorithms currently used in hearing aids, typically bilateral adaptive noise reduction systems, are designed to optimize signal to noise ratio (SNR) in a monaural way, and not to preserve binaural cues. Binaural cues are especially important for sound source localization and for speech segregation in noisy environments (a.k.a. 'the cocktail party effect').

The progress made in wireless communication between hearing aids is slowly enabling the design of a full binaural hearing aid. Changing from a bilateral to a binaural hearing aid design, i.e. generating an output signal for both ears using all microphone signals, could enhance the amount of noise reduction and could increase the capability to control the adaptive processes to preserve the binaural cues between left and right hearing aid. An important limitation of most noise reduction array systems studied thus far is that most of them are designed to have a single or monaural output. Extending these to a binaural output is not always trivial.

In this manuscript we present an overview of work done on two different binaural adaptive multi-microphone noise reduction algorithms: the multichannel Wiener filter (MWF) and the MWF with partial noise estimation (MWF-N). A classic bilateral adaptive directional microphone (ADM) was taken as a reference.

Algorithms

In general, the goal of a single channel or multichannel Wiener filter is to filter out noise corrupting a desired signal given all the available microphone signals. A Wiener filter in hearing aids is typically calculated by first estimating and then using the second-order statistical properties of the desired signal and the noise components. It generates an output signal which approaches the target signal as closely as possible in a mean-square error (MSE) sense. See [9] for an overview on Wiener filtering.

Doing noise reduction in a binaural framework leads to the system illustrated in Figure 1. A target signal and one or multiple noise sources are present in an acoustic environment. A mixture of both components, $\mathbf{Y}(\omega)$ is recorded by the microphones. $\mathbf{Y}(\omega)$ is a complex vector with size 1xM with M the number of microphones present in the system. We define the M-dimensional signal vector $\mathbf{Y}(\omega)$, with $M = M_{Left} + M_{Right}$, as

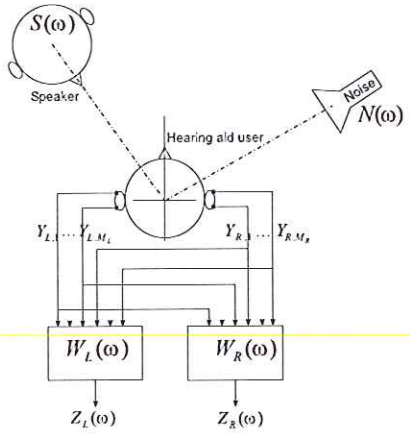

Figure 1: Noise reduction in a binaural framework. Calculate the filters $\mathbf{W_L}$ and $\mathbf{W_R}$ to remove the noisy part of the signal and which preserve the binaural information.

$$\mathbf{Y}(\omega) = [Y_{L,1}(\omega) \dots Y_{L,M_{Left}}(\omega) Y_{R,1}(\omega) \dots Y_{R,M_{Right}}(\omega)]^T.$$

$\mathbf{Y}(\omega)$ equals the sum of the recorded target component $\mathbf{X}(\omega)$ and noise component $\mathbf{V}(\omega)$ which are defined similarly as $\mathbf{Y}(\omega)$. For conciseness, the frequency domain variable ω is now omitted from the manuscript.

MWF

In the proposed MWF approach, an estimate of the unknown target signal picked up at the front microphone of the left ($X_{L,1}$) and the right ($X_{R,1}$) hearing aid is produced for the left and the right ear respectively. This is done by taking all microphone signals into account. Mathematically the problem can be formulated by minimizing the following cost function:

$$J_{MSE}(\mathbf{W}) = \mathcal{E}\left\{ \left\| \begin{bmatrix} X_{L,1} - \mathbf{W}_L^H \mathbf{Y} \\ X_{R,1} - \mathbf{W}_R^H \mathbf{Y} \end{bmatrix} \right\|^2 \right\}. \quad (1)$$

By rewriting the vector $\mathbf{Y}$ as a sum of the target component $\mathbf{X}$ and the noise component $\mathbf{V}$ and by introducing a trade-off parameter μ (see [7]), this cost

function can be rewritten as

$$J_{MWF}(\mathbf{W}) = \mathcal{E}\left\{\left\|\begin{bmatrix} X_{L,1} - \mathbf{W}_L^H \mathbf{X_L} \\ X_{R,1} - \mathbf{W}_R^H \mathbf{X_R} \end{bmatrix}\right\|^2 + \mu \left\|\begin{bmatrix} \mathbf{W}_L^H \mathbf{V_L} \\ \mathbf{W}_R^H \mathbf{V_R} \end{bmatrix}\right\|^2\right\} \quad (2)$$

The parameter μ was introduced in monaural hearing aid research to increase the amount of noise reduction by allowing the system to introduce some speech distortion. In eq. 2, it can be seen that a large μ emphasis the amount of noise reduction (the second term of the equation) while allowing some speech distortion (the first term of the equation). The solution to this minimization problem is known from Wiener filter theory and equals

$$\begin{bmatrix} \mathbf{W}_L \\ \mathbf{W}_R \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{x,L} + \mu \mathbf{R}_{v,L} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{R}_{x,R} + \mu \mathbf{R}_{v,R} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{R}_{x,L} e_L \\ \mathbf{R}_{x,R} e_R \end{bmatrix} \quad (3)$$

with e_L and e_R being vectors with one element equal to 1 and the other elements equal to zero, defining the reference microphones used at both hearing aids, i.e., in case of the front omnidirectional microphone $e_L(1) = 1$ and $e_R(1) = 1$. $\mathbf{R}_x$ and $\mathbf{R}_v$, which are at present still unknown, are defined as the $M \times M$-dimensional speech and noise correlation matrices, containing the auto-correlations and cross-correlations (or the statistical information) of, respectively, the speech and noise components $\mathbf{X}$ and $\mathbf{V}$ over the different input channels, e.g., $\mathbf{R}_{x,L} = \mathcal{E}\{\mathbf{X_L}\mathbf{X_L^H}\}$. To find the filters $\mathbf{W}$ using eq. (8), a voice activity detector (VAD) is used to discriminate between 'speech and noise periods' and 'noise only periods'. The noise correlation matrix $\mathbf{R}_v$ can be calculated during the 'noise only periods'. By assuming a sufficient stationary noise signal, the speech correlation matrix $\mathbf{R}_x$ can be estimated during speech and noise periods by subtracting $\mathbf{R}_v$ from the correlation matrix $\mathbf{R}_y$ for the noisy signal Y. By using these correlation matrices, the filters $\mathbf{W}$ can be found.

Using a MWF in a binaural reference framework has the advantage that the binaural cues of the target component are always preserved by the algorithm. This was mathematically proven in [8] and can be easily understood by the fact that a binaural MWF estimates the target signal present at the left and right front microphone as well as possible for respectively the left and right ear. However, in [8] it was also proven that the binaural cues of the remaining noise component are changed into those of the original target component.

MWF-N

A binaural Wiener filter with partial noise estimation (MWF-N) is designed to remove only part (1-η) of the noisy signal. The remaining part (η) of the noise component is preserved and presented to the hearing aid user. The reasoning for this approach is the assumption that the small but significant amount of unprocessed sound will result in a more natural perception of the surrounding sound environment and that the binaural cues of the unprocessed sound will enable the user to localize both the target component and the noise component(s). Increasing η off course degrades the obtained improvement in SNR by the MWF algorithm. The minimization criterion of eq. 1 now changes into

$$J_{MSE\eta}(\mathbf{W}) = \mathcal{E}\left\{\left\|\begin{bmatrix} X_{L,1} + \eta V_{L,1} - \mathbf{W}_L^H \mathbf{Y_L} \\ X_{R,1} + \eta V_{R,1} - \mathbf{W}_R^H \mathbf{Y_R} \end{bmatrix}\right\|^2\right\}. \quad (4)$$

The solution to this problem can be written similar as eq.3. However, there is also a simple relationship between the filter output of the MWF and the MWF-N:

$$Z_{MWF_\eta L}(\eta, \mu) = \eta Y_{L,1} + (1 - \eta) Z_{MWF,L}(\mu) \quad (5)$$

$$Z_{MWF_\eta R}(\eta, \mu) = \eta Y_{R,1} + (1 - \eta) Z_{MWF,R}(\mu) \quad (6)$$

During the evaluation of the algorithms we have chosen to put η to $\eta = 0.2$, thereby removing 80% of the noise component. Algorithms were always evaluated using a perfect VAD.

Reference

Two different reference conditions were evaluated: a condition without noise reduction and a condition in which a bilateral ADM was used. In the latter condition a single tap ADM was running independently for each ear which used both microphone inputs of the left and right dual microphone hearing aid respectively.

Localization

A CORTEX MK2 manikin was placed in an array of 13 loudspeakers. The speakers were located from -90°(at the left side) to +90°(at the right side) of the manikin with a spacing of 15°between them. All the speakers were located at a distance of 1m from the manikin. Then, impulseresponses between the loudspeakers and the microphones of two dual microphone behind-the-ear hearing aids placed on the manikin were measured. Three different spatial scenarios were generated with these impulseresponses: S_0N_{60}, $S_{45}N_{-45}$, $S_{90}N_{-90}$. The spatial scenarios are denoted as S_xN_y with x the location of the target component and y the location of a single noise component. A steady state noise, weighted by the average spectrum of a dutch male talker, was used as target component (S) and a multitalker babble (Auditec of St. Louis) was used as noise source (N). To evaluate localization performance in an unprocessed reference condition, the speech and noise signal were convolved with the impulseresponses from angles x and y, measured with the front left and the front right microphone. These stimuli were then presented to the left and the right ear of the listener using headphones. The listeners were sitting in the loudspeaker array and had to localize the stimulus. An accumulation of these results is given in the top left figure of Figure 2. The x-axis presents were the stimulus was played (e.g. N-45 denotes that in this test a multitalker babble was played arriving from -45°of the subject) while the y-axis shows were the sound was localized. Figure 2 shows that localization is almost perfect in the unprocessed condition with a slight degradation at the sides of the subject.

When evaluating the noise reduction algorithms, the mixture of the speech and the noise component of S_xN_y

were first processed by the noise reduction algorithms. This way, the adaptive filters had the opportunity to adapt and convergence to that specific condition. Afterwards, the filters were fixed and the speech and noise source were filtered separately through the filters. The filtered speech and noise source were then presented to the listener who had to localize the sound. Target and noise source were presented separately to the subject since this way interaction effects such as masking, localizing two sound sources is different from localizing a single source, etc., were avoided. This enabled a clear interpretation on what the filters were doing to the target and the noise component. The accumulated responses for all three noise reduction algorithms are given in Figure 2.

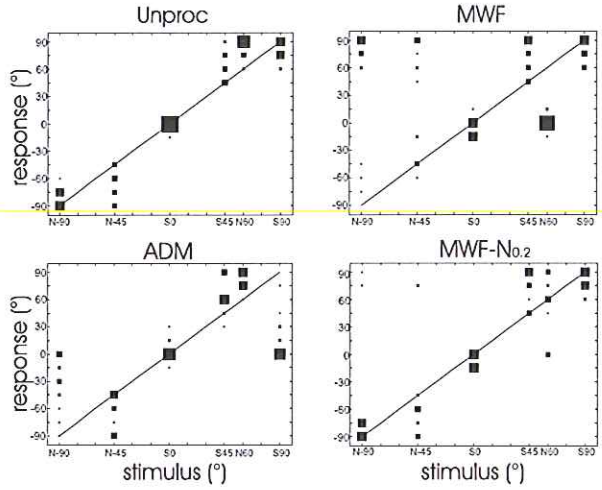

Figure 2: Accumulation of the responses of 5 normal hearing subjects on the localization task in the frontal horizontal hemisphere. Three different spatial scenarios (S_0N_{60}, $S_{45}N_{-45}$, $S_{90}N_{-90}$) were processed by three different algorithms. The x-axis shows were the sound was presented. The size of the square is proportional to the number of responses given at this location.

The top right of Figure 2 shows that all target signals (S) were localized correctly when using the MWF algorithm. The noise components (N) however were localized at the angle of where the speech source was played. E.g. both S and N are localized around +90°in the spatial condition $S_{90}N_{-90}$. This is in accordance with the theory of [8]. When using the MWF-N algorithm (bottom right of Figure 2), it was observed that both the speech and noise components were localized correctly. This proofs that preserving 20% of the original noise component was sufficient to preserve a correct sound source localization. The bottom left figure of Figure 2 shows the results obtained when using a bilateral ADM. This figure shows that all sounds between -60°and +60°were localized correctly when using the ADM. This because an ADM used in hearing aids is constrained to preserve the sound signal arriving from the front of the subject. Hence the signal and its binaural cues were not distorted. If a sound source, target or noise component, is arriving from the sides of the head, then localization deteriorated. For these angles of arrival, subjects localized the sound around 0°. However, when asked, subjects identified the sound signals as being diffuse with no directional information present. This can be explained by the fact that the ADM tries to remove sound signals arriving from these angles. Hence, a diffuse sound component remains.

A final remark should be that in [5] it was shown that when filtering a noise component with the MWF algorithm, typically a "dual sound" is perceived. Due to the fact that a very small part (typically the frequency regions with a very low SNR) of the noise component are wrongly classified as target component (e.g. due to errors in the VAD algorithm), the binaural information of this part is preserved. On the other hand, the binaural cues of the part of the noise component which was correctly classified, are changed into those of the target component and a "dual sound" is heard. If speech and noise are presented simultaneously to the subject, masking effects occur. The combination of misclassification and masking effects can result in a correct localization of the noise signal even when using the MWF algorithm. The misclassification of the noise component off course reduces the noise reduction performance of the algorithm.

Noise reduction

Speech reception thresholds (SRTs) in multitalker babble of ten normal hearing subjects were measured using an adaptive test procedure and dutch sentence material. The obtained gain in SRT, ΔSRT_{algo}, is given in Table 1 for each algorithm. Three different spatial scenarios were generated by using the impulsresponses measured with the manikin (see previous section). Stimuli were processed off-line. First, the mixture of the speech and the noise component (S_xN_y) was processed by the different adaptive algorithms. After the filters reached convergence, the filters were fixed and stored. This was done for a range of different SNRs between +10 and -30dB. Afterwards, an adaptive test procedure used the converged filters to produce the stimuli which were presented to the subjects. Three different spatial scenarios were evaluated: two single noise source scenarios, i.e. S_0N_{60} and $S_{90}N_{-90}$, and one scenario with three noise sources, $S_0N_{90/180/270}$. Similar to the localization study, three algorithms were evaluated: the ADM, the MWF and the MWF-N with 20% noise preservation, i.e. MWF-$N_{0.2}$. In contrast with the previous study, the number of microphones which was used by the MWF algorithms was varied. All MWF algorithms made use of the two microphones present on the ipsilateral hearing aid. In some conditions microphone signals from the contralateral hearing aid were added (M_C). The algorithms are denoted as MWF_{2+M_C}. For example, MWF_{2+1} shown in Table 1 illustrates that this MWF algorithm used one additional microphone from the contralateral hearing aid.

Objective measurements, fully described in [6], were also performed to analyze the performance of the noise reduction algorithms. In these measurements, a speech intelligibility weighted improvement in SNR (Δ SNR_{SI}) was calculated for the different algorithms using different microphone combinations. SNR improvements were

calculated for both the left and the right hearing aid. The measured Δ SNR$_{SI}$ are also added to Table 1.

Bilat/bin SRT [dB]	S_0N_{60}			$S_{60}N_{270}$			$S_0N_{60,(180,270)}$		
	perceptual	Left	Right	perceptual	Left	Right	perceptual	Left	Right
ADM	2,1 ± 1,9	2,7	2,8	-1,3 ± 1,3*	1,3	-3,2	1,3 ± 1,1	6,0	5,9
MWF$_{4+4}$	4,3 ± 1,5*	4,9	9,6	0,7 ± 1,1	10,0	2,5	4,6 ± 0,8*	7,1	7,2
MWF$_{4+1}$	3,8 ± 1,6*	4,0	6,2	0,3 ± 2,0	9,6	2,1	4,0 ± 1,5*	6,6	6,0
MWF$_{4+0}$	1,0 ± 0,7*	1,9	3,3	-1,2 ± 1,6	3,8	1,0	2,8 ± 1,3*	5,1	4,9
MWF$_{4+0}$-N$_{0,2}$	3,6 ± 1,1*	3,3	5,1	2,0 ± 1,1*	4,3	1,9	3,2 ± 0,8*	4,1	4,2
MWF$_{4+1}$-N$_{0,2}$	2,7 ± 1,3*	2,6	3,0	1,5 ± 1,6	3,9	1,6	3,1 ± 0,8*	3,7	3,3
MWF$_{4+0}$-N$_{0,2}$	1,0 ± 2,1	1,1	0,9	0,0 ± 1,5	1,0	0,7	2,3 ± 1,4*	2,8	2,6
SNR-unproc [dB]	-6,2±1,8			-9,1±1,7			-7,2±1,6		

Table 1: The gain in SRT, ΔSRT_{algo}, averaged over ten normal hearing subjects. The bottom row show the SNR at which the unprocessed reference SRT was measured. A '*' depicts a significant noise reduction performance compared to the unprocessed condition. ΔSNR_{SI}, calculated for the left and right hearing aid in the objective evaluation, is also added to the table.

Several conclusions can be drawn from Table 1. First it is observed that the objective performance measure, Δ SNR$_{SI}$, gives a good indication of the perceptual outcome of the algorithm, ΔSRT_{algo}. However, Δ SNR$_{SI}$ is a performance measure originating from monaural hearing aid research and therefore some of the effects of listening with two ears have to be taken into account. First, one has to take into account the best ear benefit. In an assymetric spatial scenario, Δ SNR$_{SI}$ will be different for both ears. ΔSRT_{algo} correlates only with the Δ SNR$_{SI}$ measured at the ear which has the best input SNR. Typically, this is the ear with the lowest gain in SNR. E.g. in condition S_0N_{60}, the perceptual results correlate best with Δ SNR$_{SI}$ obtained at the left ear which is the ear with the best input SNR. Second, Δ SNR$_{SI}$ does not take into account spatial release from masking effects which explains some of the small differences between Δ SNR$_{SI}$ and ΔSRT.

Second, it is observed that a two microphone MWF and ADM have approximitely the same performance, except when the target signal is not arriving from the front direction, i.e. $S_{90}N_{90}$. Since the ADM is designed to preserve signals arriving from the front, it will start to remove both the noise and the target component in this spatial scenario.

Third, it is observed that feeding contralateral microphone signals into the noise reduction schemes significantly improves noise reduction performance. The largest gain is observed when the first contralateral microphone is added to the system. A lower gain is observed when adding the second contralateral microphone to the signal. This can be intuitively explained by the fact that the first contralateral microphone signal adds more new information to the noise reduction scheme in comparison with the second contralateral microphone signal.

Fourth, it is observed that the decrease in SRT performance, when replacing the MWF into an MWF-N algorithm is smaller than predicted by the objective SNR measurements. In condition $S_{90}N_{-90}$, the MWF-N even outperforms the MWF algorithm. This may be due to the fact that the listener is able to localize both the target and the noise component correctly when using the MWF-N algorithm. This leads to an improved speech perception due to release from masking effects.

Conclusions

This study presented objective and perceptual measurements done with a binaural MWF and MWF-N noise reduction scheme. An unprocessed condition and a bilateral ADM were taken as a reference. First, it was shown that in contrast with an ADM, the MWF preserved the ability to localize the target signal independent of its angle of arrival. When using the MWF-N algorithm, both the target and the noise component were accurately localized. However, this comes at the cost of SNR improvement. Later it was shown that this decrease not necessarily implies a loss in speech perception due to spatial release from masking. Further it was shown that adding microphone signals of the contralateral hearing aid to a noise reduction scheme running on an ipsilateral hearing aid improves noise reduction performance.

References

[1] Welker, D.P. et al. "Microphone-array hearing aids with binaural output-part II: A two-microphone adaptive system," IEEE Trans. Speech and Audio Processing **5** (1997), 543-551

[2] Lotter, T., "Single and multimicrophone speech enhancement for hearing aids," Ph.D. thesis, RWTH Aachen (2004)

[3] Keidser, G. et al."The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers," Int. J. Audiol. **45**, 563–579. (2006)

[4] Van den Bogaert, T. et al. "Localization with bilateral hearing aids: without is better than with," J. Acoust. Soc. Am. **119**, 515–526. (2006)

[5] Van den Bogaert, T. et al. "The effect of multimicrophone noise reduction systems on sound source localization in binaural hearing aids", Journal of the Acoustical Society of America, 124(1), pp. 484-497, (2008)

[6] Van den Bogaert, T. et al. "Speech enhancement with multichannel Wiener filter techniques in binaural hearing aids", Journal of the Acoustical Society of America, 125(1), pp. 360-371. (2009)

[7] Doclo, S. et al. "GSVD-based optimal filtering for single and multi-microphone speech enhancement," IEEE Trans. Signal Process. **50**, 2230–2244 (2002).

[8] Doclo, S. et al. "Theoretical analysis of binaural cue preservation using multichannel Wiener filtering and interaural transfer functions," in Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC), Paris, France, pp. 1–4. (2006).

[9] Haykin, S., Adaptive Filter Theory 4th ed., chapter 2 pp94-135. Prentice Hall, New Jersey, 2002