KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Afdeling ESAT-SCD
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)
In samenwerking met:
FACULTEIT GENEESKUNDE
DEPARTEMENT NEUROWETENSCHAPPEN
Afdeling ExpORL
Herestraat 49, 3000 Leuven

PRESERVING BINAURAL CUES IN NOISE REDUCTION ALGORITHMS FOR HEARING AIDS

Promotoren:
Prof. dr. J. Wouters
Prof. dr. ir. M. Moonen

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen

door

Tim Van den Bogaert

June 2008

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Afdeling ESAT-SCD
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)
In samenwerking met:
FACULTEIT GENEESKUNDE
DEPARTEMENT NEUROWETENSCHAPPEN
Afdeling ExpORL
Herestraat 49, 3000 Leuven

PRESERVING BINAURAL CUES IN NOISE REDUCTION ALGORITHMS FOR HEARING AIDS

Jury:
Prof. dr. A. Haegemans, voorzitter
Prof. dr. J. Wouters, promotor
Prof. dr. ir. M. Moonen, promotor
Prof. dr. P. Joris
Prof. dr. ir. D. Van Compernolle
Prof. dr. ir. J. Vandewalle
Prof. dr. sc.techn. N. Dillier - Universität Zürich
Prof. dr. dr. B. Kollmeier - Universität Oldenburg

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen

door

Tim Van den Bogaert

U.D.C. 534.7-621.3

June 2008

D/2008/7515/39

ISBN 978-90-5682-930-8

Voorwoord

Voilà, het zit erop. Vele jaren zwoegen en zweten gebundeld en samengevat, klaar voor publicatie. Tenslotte resten mij enkel nog deze pagina's om even terug te blikken en een aantal mensen te bedanken.

Vooreerst zijn er natuurlijk Jan en Marc, de twee promotoren van dit werk. Twee promotoren, twee invalshoeken op hetzelfde probleem, twee werelden, één doel. Ik wil graag zowel Jan als Marc uitvoerig bedanken, niet omdat dit zo voorgeschreven staat in de doctoraatsreglementen, maar ik wil hen oprecht bedanken omdat zij dit multi-disciplinair onderzoek met zeer veel enthousiasme en steun mogelijk hebben gemaakt. Multi-disciplinariteit is niet enkel een trendy woord, het vereist basiskennis, inzet en kritische massa verspreid over de verschillende disciplines. Met Jan en Marc was deze kritische massa ruimschoots aanwezig om dit doctoraatsproject tot een goed einde te brengen.

Daarnaast had ik Jan ook nog graag bedankt voor de mooie tijd aan ExpORL. Wanneer ik weer eens droevig naar mijn data zat te staren kon Jan altijd iets uit zijn hoed toveren waardoor het plots leek dat dit de resultaten waren die ik altijd al wilde hebben. Dit gevoel ebde typisch weer weg na een week, waardoor een tweede en soms derde bezoek aan Jan's kantoor aan de orde waren. Business as usual. Niet enkel heb ik kunnen genieten van een prettige wetenschappelijke samenwerking, maar het toffe was dat 'de Jan' altijd te vinden was voor een grap of één of andere idioterij. Persoonlijk ben ik niet zeker of dit werk voor u zou gelegen hebben mocht dit niet zo zijn geweest. Bedankt chef.

Daarnaast zijn er zoveel mensen te bedanken. Vooreerst wil ik benadrukken dat dit werk geen één-mans verhaal is. Naast Jan en Marc waren er gedurende dit project vruchtbare samenwerkingen met Monika, T.J. en Simon. Simon wil ik extra bedanken voor zijn geduldige uitleg en zijn grote inspanningen op het vlak van ruisonderdrukking. Ook de proefpersonen moet ik bedanken. Doordat zij urenlang telefoontjes lokaliseerden is dit werk voltooid geraakt. Ik vermoed dat sommigen er een telefoon-trauma aan hebben overgehouden, waarvoor mijn excuses. Ik wil hier ook mijn financiële partners bedanken. IWT-Vlaanderen heeft me gedurende 4 jaar financieel ondersteund waarvoor een woord van dank terecht is. Tenslotte wil ik ook de jury bedanken voor hun tijd en inspanningen om dit werk na te lezen en te becommentariëren waar nodig, met extra dank aan de externe leden van de jury, Prof. Kollmeier en Prof. Dillier, die speciaal naar Leuven zijn gekomen voor de publieke doctoraatsverdediging.

Naast de wetenschappelijke medewerkers mag ik zeker de vrienden van het labo niet vergeten. Met Heleen, Johan, JB en het charmante audio-team, Ann D., Katleen en Lies zette ik mijn eerste pasjes in de wereld van de psycho-akoestiek. Johan had doorslaggevende argumenten om in ExpORL te komen werken: flexibele uren, nabijheid van videotheek en SPAR. Intussen heb ik geleerd dat flexibele uren ook lange uren kan betekenen en zijn ook videotheek en SPAR verdwenen, maar no hard feelings hoor. Later heb ik het geluk gehad om de multitaskers te leren kennen. Jane, Lot en vooral Brammeke. Dat dit soms gepaard ging met een exponentiële toename van het aantal decibels was toeval en volledig te wijten aan de slechte akoestische isolatie van het ventilatiesysteem. Bedankt hé, jullie zijn schatjes. Mijn buddy Bram en zijn vrouwtje Ilse wil ik nog eens extra bedanken voor het ontdekken van Mogwaii en alle andere fijne momenten in alle uithoeken van de wereld en op alle mogelijke en onmogelijke uren van de dag.

En dan zijn er nog al die andere mensen van het labo waar je steeds op kon rekenen binnen en buiten de uren: Nathalie, Dirk en kleine Ben-piraat bedankt voor de steun, nog een marathon die ten einde is. En ook de sportieve logo's met al hun aanhangsels: Ellen, Catherine, Joke en Tinne, dienen vermeld te worden. Laat het duidelijk zijn dat het een leuke tijd is geweest. Verder zijn er nog onze goedlachse Astrid, onze lieftallige secretaresse Frieda, the French guy who moved to England, crazy Chinese guy, Ellen B., Ann S., Koen, Tom, Matthias, Michael, Eric, Wivine, Inge en manusjes van alles Dennis en Wilfried.

Verder wil ik nog een aantal mensen bedanken voor de zalige momenten buiten de uren. Zo waren er Evi-poes; een bende eikels: Bram, Cedric, Ann Verstr., Hans, Filip en alle anderen van de A en de B lijst (protest!); de indies: Pedro, Marjoleine en Dimi, Geert en Kathleen, Ann Versch. en Tom; Han; TTK Gierle en BASF; de läufmeisters van Dirk en co., enfin te veel mensen om op te noemen, die allemaal voor zeer veel leven in de brouwerij zorgden.

Last but not least, wil ik ook mijn zeer enthousiast thuisfront bedanken. Evelyne, die van wel zéér dichtbij de laatste hectische maanden mee heeft getrotseerd, mijn ouders, mijn broer Sven, Bomma en René en de rest van de familie die van kinds-af-aan altijd klaar stonden in geval van nood, pa, hoe maak ik een draagbare testopstelling van 2m diameter, of wanneer de was weer eens dringend gedaan moest worden, ma, moet ik dit nu zelf doen?

Bedankt allemaal en veel plezier tijdens het lezen van de rest van dit werk (Jan is de moordenaar met de kandelaar in de bibliotheek).

Tim

Abstract

Hearing aid users experience great difficulty in understanding speech in noisy environments. This has led to the introduction of noise reduction algorithms in hearing aids. The development of these algorithms is typically done monaurally. However, the human auditory system is a binaural system, which compares and combines the signals received by both ears to perceive a sound source as a single entity in space. Providing two monaural, independently operating, noise reduction systems, i.e. a bilateral configuration, to the hearing aid user may disrupt binaural information, needed to localize sound sources correctly and to improve speech perception in noise.

In this research project, we first examined the influence of commercially available, bilateral, noise reduction algorithms on binaural hearing. Extensive objective and perceptual evaluations showed that the bilateral adaptive directional microphone (ADM) and the bilateral fixed directional microphone, two of the most commonly used noise reduction algorithms in hearing aids, can significantly distort the binaural properties of the sound signals. These distortions are well within the range used by the human auditory system. In what follows, three binaural algorithms, based on a multichannel Wiener filter (MWF) approach, were developed and evaluated. These algorithms assume a communication link between both hearing aids. It was observed that a binaural hearing aid design significantly increased noise reduction performance. Moreover, the binaural MWF, the binaural MWF with partial noise estimation (MWF-N) and the binaural MWF with interaural transfer function extension (MWF-ITF) provided a better combination of noise reduction performance and preservation of binaural cues compared to the bilateral ADM algorithm.

Korte Inhoud

Hoorapparaatgebruikers ervaren vaak grote moeilijkheden om spraak te verstaan in lawaaierige omstandigheden. Om aan dit probleem tegemoet te komen wordt er gebruik gemaakt van ruisonderdrukkingsalgoritmen. De ontwikkeling hiervan gebeurt vaak monauraal. Het auditief systeem is echter een binauraal systeem, dit wil zeggen dat beide oren samenwerken om een geluidsbron waar te nemen als één enkele entiteit in de ruimte. Een bilaterale aanpassing, bestaande uit twee onafhankelijke monaurale systemen, neemt de binaurale werking van het menselijk auditief systeem niet noodzakelijk in rekening en verstoort daarom mogelijk de binaurale informatie nodig voor het correct lokaliseren van geluidsbronnen en voor een verbeterd spraakverstaan in lawaaierige omstandigheden.

In dit project werd eerst de invloed van hedendaagse, bilaterale, ruisonderdrukkingssystemen op het binauraal horen onderzocht. Theoretische, objectieve en perceptuele evaluaties tonen aan dat de twee meest gebruikte commerciële ruisonderdrukkingsalgoritmen, namelijk een bilaterale directionele en een bilaterale adaptief directionele (ADM) microfoonsconfiguratie, de binaurale informatie significant kunnen verstoren. Deze algoritmen bieden typisch geen mogelijkheid om ruisonderdrukking te combineren met het bewaren van alle binaurale informatie. Nadien werden drie nieuwe, binaurale, algoritmen ontworpen en geëvalueerd. Deze zijn gebaseerd op de werking van een meerkanaals Wiener filter (MWF) en veronderstellen de aanwezigheid van een communicatiekanaal tussen beide hoorapparaten. Er werd aangetoond dat de binaurale link tussen de hoorapparaten een significante winst in ruisonderdrukking oplevert. De binaurale MWF, de binaurale MWF met partiële ruisschatting en de binaurale MWF met interaurale transferfunctie zorgen bovendien voor een betere combinatie van ruisonderdrukking met het bewaren van de binaurale informatie in vergelijking met de bilaterale ADM.

Glossary

Mathematical Notation

$\sim$	Is proportional to
$\| \ \|$	Absolute value
$\\| \ \\|$	Vector norm
$\mathbf{0}_M$	MxM matrix with all elements=0
$\mathbf{1}_M$	MxM matrix with all elements=1
$\mathbf{I}_M$	MxM unity matrix
$\mathcal{E}\{\ \}$	Expected value operator
$\mathbf{\Gamma}_x(\omega)$	Coherence matrix of vector $\mathbf{X}(\omega)$
a	Scalar a
$\mathbf{a}$	Vector $\mathbf{a}$
$\mathbf{A}$	Matrix $\mathbf{A}$
$A(\omega)$	Discrete time Fourrier transform of a[k]
$\mathbf{A}(\omega)$	Vector of discrete time Fourrier transformed elements $A(\omega)$
$A_a(\omega)$	The a-th element of $\mathbf{A}(\omega)$
$\mathbf{A}^{-1}$	Inverse of matrix $\mathbf{A}$
$\mathbf{A}^T$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^*$	Complex conjugate of matrix $\mathbf{A}$
$\mathbf{A}^H = (\mathbf{A}^*)^T$	Hermitian transpose of matrix $\mathbf{A}$
$\mathbf{R}_{yx} = \mathcal{E}\{\mathbf{Y}\mathbf{X}^H\}$	Cross correlation matrix of vectors $\mathbf{X}(\omega)$ and $\mathbf{Y}(\omega)$
$\mathbf{R}_{yy} = \mathcal{E}\{\mathbf{Y}\mathbf{Y}^H\}$	Correlation matrix of vector $\mathbf{Y}(\omega)$

Fixed Symbols

ΔSNR_L	SNR improvement at the left hearing aid
ΔSNR_R	SNR improvement at the right hearing aid
η	Trade-off parameter MWF-N between binaural cue preservation and noise reduction
$\Phi_{algo}(\omega, \theta)$	Phase transfer function of a noise reduction algorithm

μ	Trade-off parameter SDW-MWF between speech distortion and noise reduction
$\omega = 2\pi f$	Pulsation
θ	Angle of arrival of the signal
$\tau(\omega)$	Internal delay of a directional microphone
$A_{ILD}(\omega)$	Frequency dependent weight, used when calculating the ILD error
$A_{ITD}(\omega)$	Frequency dependent weight, used when calculating the ITD error
$\mathbf{A}(\omega)$	The acoustic transfer functions between the speech source and all microphones
$B(\omega, \tau)$	Amplitude transfer function of a noise reduction system
$c_v^{in/out}(\omega)$	Cross-correlation of the noise component at the input/output of the algorithm
$c_x^{in/out}(\omega)$	Cross-correlation of the speech component at the input/output of the algorithm
$\mathbf{e}_r$	Vector defining the reference microphone, the r-th element of $\mathbf{e} = 1$
f	Frequency
f_c	Cut-off frequency
f_s	Sampling frequency
$G_x(\omega)$	Power transfer function of the speech component
$G_v(\omega)$	Power transfer function of the noise component
$H(\omega, \tau)$	Transfer function of a noise reduction system
$I(\omega_i)$	The importance of the i-th third octave band for speech intelligibility
$ITF_v^{in/out}(\omega)$	Interaural transfer function of the noise component at the input/output of the algorithm
$ITF_x^{in/out}(\omega)$	Interaural transfer function of the speech component at the input/output of the algorithm
J_{MSE}	MSE cost function
$L_v^{in/out}(\omega)$	The estimated ILD of the noise component at the input/output of the algorithm
$L_x^{in/out}(\omega)$	The estimated ILD of the speech component at the input/output of the algorithm
M	Total amount of microphone signals used in each hearing aid
M_C	Number of microphone signals received from the contralateral hearing aid
M_L	Number of microphones of the left hearing aid
M_R	Number of microphones of the right hearing aid
$N(\omega)$	Noise signal
$P_{v,m}$	Power spectral density of the noise component of the m-th microphone signal
$P_{x,m}$	Power spectral density of the speech component of

	the m-th microphone signal
$P_{y,m}$	Power spectral density of the m-th microphone signal
r_L	The reference microphone used at the left hearing aid
r_R	The reference microphone used at the right hearing aid
$S(\omega)$	Target speech signal
S_xN_y	Sound scenario with a speech source at x° and a noise source at y°
T_{60}	Reverberation time
$\mathbf{V}(\omega)$	Noise component input vector of both the left and the right hearing aid
$V_{L,m}(\omega)$	Noise component of $Y_{L,m}(\omega)$
$V_{R,m}(\omega)$	Noise component of $Y_{R,m}(\omega)$
$\mathbf{W}(\omega)$	Computed Wiener filters for both the left and the right hearing aid
$\mathbf{W}_L(\omega)$	Computed Wiener filters at the left hearing aid
$\mathbf{W}_R(\omega)$	Computed Wiener filters at the right hearing aid
$\mathbf{X}(\omega)$	Speech component input vector of both the left and the right hearing aid
$X_{L,m}(\omega)$	Speech component of $Y_{L,m}(\omega)$
$X_{R,m}(\omega)$	Speech component of $Y_{R,m}(\omega)$
$\mathbf{Y}(\omega)$	Signal input vector of both the left and the right hearing aid
$\mathbf{Y}_L(\omega)$	Signal input vector of the left hearing aid
$\mathbf{Y}_R(\omega)$	Signal input vector of the right hearing aid
$Y_{L,m}(\omega)$	m-th microphone signal of the left hearing aid
$Y_{R,m}(\omega)$	m-th microphone signal of the right hearing aid
$Z_L(\omega)$	Output signal of the left hearing aid
$Z_R(\omega)$	Output signal of the right hearing aid
$Z_{vL}(\omega)$	Noise component at the output of the left hearing aid
$Z_{vR}(\omega)$	Noise component at the output of the right hearing aid
$Z_{xL}(\omega)$	Speech component at the output of the left hearing aid
$Z_{xR}(\omega)$	Speech component at the output of the right hearing aid

Acronyms and Abbreviations

ACVN	Anteroventral cochleus nuclei
ADM	Adaptive directional microphone
a.k.a.	also known as
ALP	Advanced localization procedure
ANOVA	Analysis of variance
ANC	Adaptive noise canceller
AR1	Anechoic room, loudspeakers are at 1m distance
ASA	Auditory scene analysis

BILD	Binaural intelligibility level difference
BMLD	Binaural masking level difference
BRIR	Binaural room impulse response
BSS	Blind source separation
BTE	Behind the ear
CASA	Computational auditory scene analysis
dB A	A-weighted decibels
dB HL	Decibel hearing level
dB SPL	Decibel sound pressure level
DI	Directivity index
DSP	Digital signal processor
EE	Excitation-excitation
e.g.	`exempli gratia` : for example
etc.	`etcetera` : and so on
ExpORL	Experimental oto-rhino-laryngology
FDM	Fixed directional microphone
FFT	Fast Fourrier transformation
GSC	General sidelobe canceller
HPM	Headphones, manikin measured impulse responses
HPO	Headphones, ODEON generated impulse responses
HRTF	Head related transfer function
IC	Inferior colliculus
i.e.	`id est` : that is
ILD	Interaural level difference
IPD	Interaural phase difference
ISM	Image source method
ITC	In the canal
ITD	Interaural time difference
ITE	In the ear
ITF	Interaural transfer function
LSO	Lateral superior olive
MAA	Minimal audible angle
MAE	Mean average error
MNTB	Medial nucleus of the trapezoid body
MSE	Mean-square-error
MSO	Medial superior olive
MWF	Multichannel Wiener filter
MWF-db	Binaural MWF with distributed processing
MWF-front	Binaural MWF with a front contralateral microphone
MWF-ITF	Binaural MWF with ITF extension
MWF-N	MWF with partial noise estimation
nme	No main effect
OE	Stimuli presented with loudspeakers, own ears condition
PSD	Power spectral density
RIR	Room impulse response

RMS	Root mean square
RR1	Reverberant room, loudspeakers are at 1m distance
RR2	Reverberant room, loudspeakers are at 2.4m distance
RTM	Ray tracing method
SDW	Speech distortion weighted
SDW-MWF	Speech distortion weighted multichannel Wiener filter
SI	Speech intelligibility weighted
SISTA	Signals, Identification, System Theory and Automation
SNR	Signal to noise ratio
SPL	Sound pressure level
SRT	Speech reception threshold
SSQ	Speech and spatial quality questionnaire
VAD	Voice activity detector
VIRTAC	Virtual acoustics
vs.	versus

Contents

Het bewaren van binaurale cues bij ruisonderdrukking in hoorapparaten

Motivatie

Slechthorendheid is één van de grootste gezondheidsproblemen van de westerse wereld. De WHO (World Health Organisation) schat dat tegen 2010 10 à 15 % van de bevolking een gehoorprobleem zal hebben. Om de nadelige gevolgen van slechthorendheid te compenseren wordt vaak gebruik gemaakt van één of twee hoorappara(a)t(en). Hoewel hoorapparaten reeds een zeer grote evolutie hebben doorgemaakt, blijft het gebrek aan spraakverstaanbaarheid in lawaaierige omstandigheden één van de grootste oorzaken van ontevredenheid bij hoorapparaatgebruikers. Dit heeft geleid tot de ontwikkeling en implementatie van ruisonderdrukkingssystemen in hoorapparaten.

Bij het ontwikkelen van ruisonderdrukkingsalgoritmen worden deze doorgaans geoptimaliseerd voor één enkel oor. Bij een bilateraal gehoorverlies, i.e. een gehoorverlies aan beide oren, wat de meest voorkomende vorm van gehoorverlies is, worden twee zulke apparaten aangepast. Men spreekt dan van bilaterale hoortoestellen. Het menselijk auditief systeem is echter geen bilateraal, bestaande uit twee onafhankelijk werkende receptoren, maar een binauraal systeem bestaande uit twee samenwerkende receptoren. Zo worden de signalen van het linker- en het rechteroor met elkaar vergeleken en gecombineerd om één enkele auditieve waarneming te bekomen, gelokaliseerd in tijd en ruimte. Bovendien helpt de binaurale informatie om verschillende geluidsstromen van elkaar te onderscheiden wat leidt tot een verbeterd spraakverstaan in lawaaierige omstandigheden, i.e. het zogenaamde 'cocktail-party effect'.

De invloed van hoorapparaatalgoritmen op binaurale informatie, i.e. de informatie die vrijkomt bij het vergelijken van de signalen ontvangen aan het linker- en rechtertrommelvlies, is, door de monaurale ontwikkeling, lange tijd genegeerd geweest. De laatste tiental jaren is de interesse naar binauraal horen echter fel toegenomen. De analyse van grote data-sets verkregen door het ondervragen van monaurale en bilaterale hoorapparaatgebruikers toont het belang van binauraal horen aan. Zo werd de 'speech and spatial quality questionaire' ontwikkeld welke, in tegenstelling tot klassieke vragenlijsten, zeer duidelijke vragen omtrent binauraal horen bevat (Gatehouse and Noble, 2004; Noble, 2006). Door de komst van een draadloze link tussen beide hoorapparaten is nu ook de commerciële interesse naar binauraal horen en naar binaurale ruisonderdrukking fel toegenomen. De evolutie naar binaurale hoortoestellen lijkt dan ook een logische voortzetting van de evolutie van monaurale naar bilaterale hoortoestellen die zich in de jaren '90 voltrok.

Dit onderzoek gaat na wat de invloed van ruisonderdrukking is op binaurale informatie voor twee van de meest gebruikte ruisonderdrukkingstechnieken in hoorapparaten zijnde een directionele en een adaptief directionele microfoon (**hoofdstuk 2**). Nadien worden nieuwe algoritmen voorgesteld (**hoofdstuk 4**) en geëvalueerd (**hoofdstuk 5** en **hoofdstuk 6**) die een groter potentieel bieden op het gebied van het combineren van ruisonderdrukking met het bewaren van binaurale cues. Aangezien het evalueren van de nieuwe algoritmen bij voorkeur gebeurt door middel van geluidsaanbiedingen via hoofdtelefoon is een kleine tussenstap vereist die deze methodologie voor lokalisatie-experimenten valideert (**hoofdstuk 3**).

Hoofdstuk 1: Inleiding

Gehoorverlies (**paragraaf 1.2**) zorgt ervoor dat het detecteren van geluiden en het verstaan van spraak gedeeltelijk of volledig wegvalt. Een gehoorverlies zorgt niet enkel voor verzwakking maar ook voor distortie van geluiden. Distortie treed op onafhankelijk van de geluidsintensiteit en zorgt ervoor dat de slechthorende steeds een 5 à 10dB hogere signaal-ruis verhouding (SNR) nodig heeft dan een normaalhorende om dezelfde hoeveelheid spraak te verstaan. Dit heeft geleid tot de ontwikkeling van ruisonderdrukkingsalgoritmen voor hoorapparaten.

Verschillende ruisonderdrukkingstechnieken (**paragraaf 1.3**) zijn reeds onderzocht naar hun toepasbaarheid in hoorapparaten. Hoorapparaten hebben dan ook zeer specifieke eisen: weinig tot geen voorkennis over de opgevangen signalen, een zeer kleine afstand tussen de microfoons, laag vermogen, lage complexiteit, etc.. De beste ruisonderdrukking wordt typisch behaald door gebruik te maken van adaptieve meerkanaalssystemen. Door de adaptiviteit passen deze zich aan aan de luistersituatie van de gebruiker en door het combineren van meerdere microfoons wordt de ruimtelijke scheiding tussen geluidsbronnen gebruikt om de SNR te verbeteren. In dit manuscript wordt er vooral aandacht besteed aan de meerkanaals Wiener filter (MWF) en aan de fixed directionele en adaptief directionele microfoon (FDM en ADM). Deze laatste algoritmen zijn zeer eenvoudige, maar in hoorapparaten de meest toegepaste, voorbeelden van respectievelijk vaste en adaptieve meerkanaals beamforming technieken. De MWF heeft hierbij het voordeel ten opzichte van beamformingtechnieken dat er geen a priori assumpties over de invalsrichting van het spraaksignaal en over de microfoonkarakteristieken nodig zijn voor een goede werking van het algoritme. Het nadeel van een MWF is de hoge complexiteit die lange tijd de toepasbaarheid in hoorapparaten heeft verhinderd. Door de implementatie van subband stochastische gradiënt oplossingen door Spriet et al. (2004) en Spriet et al. (2005) is hier echter verandering in gekomen.

Om de invloed van hoorapparaten op binaurale informatie en ruimtelijke gewaarwording perceptueel op te meten wordt er in dit werk gebruik gemaakt van lokalisatie-experimenten (**paragraaf 1.4**). Lokalisatie van een geluidsbron in het frontale horizontale vlak is dan ook een taak gedomineerd door binaurale informatie. Rayleigh (1907) stelde in 1907 reeds een theorie voor waarbij twee binaurale mechanismen het lokaliseren van geluidsbronnen verklaarden. Dit is de zogenaamde 'duplex-theorie'. De belangrijkste component hierin is het verschil in aankomsttijd van een geluid aan de beide oren. Door de eindige geluidssnelheid zal een signaal namelijk sneller het ene dan het andere oor bereiken. Hierdoor ontstaat er een richtingsafhankelijk verschil in aankomsttijd. De tweede component van de 'duplex-theorie' is het verschil in luidheid. Doordat het hoofd een akoestische schaduw creëert ontstaan er hoekafhankelijke intensiteitsverschillen tussen de signalen aan beide trommelvliezen.
Buiten interaurale tijds- en intensiteitsverschillen zijn er nog andere informatiedragers, zogenaamde cues, die ervoor zorgen dat een geluidsbron gelokaliseerd kan worden. Zo zijn er nog spectrale cues, visuele cues, monaurale luidheidscues en hoofdbewegingen die elk bijdragen aan de ruimtelijke waarneming. Ongeacht het feit dat deze cues een minder grote rol spelen, dienen ze gecontroleerd te worden in elk lokalisatie-experiment.

Correcte binaurale informatie, i.e. interaurale tijds- en intensiteitsverschillen zijn niet enkel cruciaal voor een correcte lokalisatie van geluidsbronnen maar zorgen ook voor een verbeterde detectie en herkenning van geluidssignalen (**paragraaf 1.5**). Verschillen in binaurale informatie van spatieel gescheiden geluidsbronnen zorgen er immers voor dat het menselijk auditief systeem de verschillende geluidsstromen beter kan onderscheiden wat leidt tot een verbeterd spraakverstaan in ruis. Distortie van binaurale informatie kan dus mogelijk leiden tot een verminderde lokalisatieperformantie en een verminderd spraakverstaan in lawaaierige omstandigheden.

Een groot verschil tussen theoretische evaluaties enerzijds en objectieve en perceptuele evaluaties anderzijds zijn de akoestische parameters van de testruimte en imperfecties die deel uitmaken van het hoorapparaat, zoals de microfoonskarakteristieken (**paragraaf 1.6**). Het is algemeen geweten dat reverberatie een negatieve impact heeft op het spraakverstaan en op ruisonderdrukkingssystemen. Hoe minder reflecties, hoe dichter de performantie van een ruisonderdrukkingssysteem zal aanleunen bij de theoretische evaluatie.
Een ander aspect van belang zijn de microfoon-karakteristieken. Deze hebben voornamelijk een grote invloed op de performantie van meerkanaals ruisonder-

drukkingssystemen. Vaak wordt er tijdens het ontwerp, vooral bij beamforming, ervan uit gegaan dat de karakteristiek van elke microfoon identiek is. In realiteit zullen de microfoons van dit gedrag afwijken wat grote gevolgen kan hebben op de performantie van het algoritme en op de binaurale cues (zie ook hoofdstuk 2).

Hoofdstuk 2: De impact van commerciële ruisonderdrukkingssystemen op de binaurale cues.

In hoofdstuk 2 wordt de impact van commerciële ruisonderdrukkingssystemen op de binaurale cues en op de lokalisatie van geluidsbronnen besproken. Verschillende publicaties behandelden reeds het lokaliseren van geluidsbronnen met hoorapparaten (**paragraaf 2.1**). Een algemene conclusie kan echter moeilijk getrokken worden aangezien deze werken vaak moeilijk vergelijkbaar zijn door een verschil in methodologie (o.a. een verschil in performantiematen en resolutie van de testopstelling). Toch zijn er aanwijzingen dat de signaalverwerking in hoorapparaten een invloed kan hebben op de lokalisatieperformantie.

In **paragraaf 2.2** worden twee van de meest gebruikte ruisonderdrukkingssystemen in hoorapparaten theoretisch geëvalueerd, namelijk de bilaterale directionele microfoon (FDM) en de bilaterale adaptieve directionele microfoon (ADM) waarbij de term bilateraal duidt op het feit dat elk oor gebruik maakt van een onafhankelijk ruisonderdrukkingssysteem. Er kan worden aangetoond dat vanuit theoretisch oogpunt een ideale FDM de binaurale cues niet beïnvloedt aangezien de vertraging en de verzwakking gegenereerd door ideale FDM's identiek is voor beide hoorapparaten. Indien er echter realistische imperfecties in het model worden ingevoerd, zoals niet-identieke microfoon-karakteristieken, dan worden zowel de interaurale tijds- als de intensiteitsverschillen bëinvloed, wat kan leiden tot een verkeerde lokalisatie van geluidsbronnen.
Een bilaterale ADM heeft de eigenschap om, voor elk oor, zich aan te passen aan de luistersituatie om de meest dominante ruisbron te onderdrukken. Aangezien deze verschillend kan zijn voor beide hoorapparaten kan een ideale ADM reeds invloed uitoefenen op interaurale intensiteitsverschillen. Bij het toevoegen van realistische imperfecties zoals verschillen in microfoon-karakteristiek wordt er, net zoals bij de FDM, een distortie van interaurale tijdsinformatie geobserveerd. Beide systemen introduceren de grootste interaurale distorties rond de invalshoeken met de meeste ruisonderdrukking.

In **paragraaf 2.3** worden hoorapparaatgebruikers geëvalueerd met en zonder hoorapparaten in een lokalisatie-experiment. Bij het dragen van hoorapparaten wordt er gebruik gemaakt van een bilaterale omnidirectionele instelling, i.e. geen ruisonderdrukking aanwezig, en een bilaterale ADM. Een groep normaalhorenden zijn geëvalueerd als referentie. Vier verschillende stimuli worden aangeboden: lage frequenties (lokalisatie is gebaseerd op interaurale tijdsverschillen), hoge frequenties (lokalisatie is gebaseerd op interaurale intensiteitsverschillen), een breedband stimulus (lokalisatie is gebaseerd op tijds- en intensiteitsverschillen) en een breedband stimulus met ruisbronnen aan beide zijden van het hoofd. Een eerste observatie is dat voor alle groepen de breedband stimulus het best lokaliseerbaar is. De slechtste resultaten worden behaald bij het testen met hoge frequenties. Een tweede vaststelling is dat de groep slechthorenden, wanneer zij niet gebruik maken van hun hoorapparaten, iets minder goed lokaliseren dan de groep normaalhorenden. Deze groepen zijn echter niet gematched in leeftijd wat dit verschil zou kunnen verklaren. De belangrijkste bevinding is echter dat de groep slechthorenden wel degelijk nog een relatief goede lokalisatieperformantie behalen wat verder onderzoek naar hoorapparaten en binaurale cues ondersteunt. Verder is er de vaststelling dat de slechthorenden de beste lokalisatieperformantie behalen indien zij niet gebruik maken van hun hoorapparaten. De slechtste performantie wordt behaald bij het gebruik van de bilaterale ADM wat vooral te wijten is aan fouten gemaakt bij het lokaliseren van geluiden aan de zijkanten van het hoofd. Hieruit kan worden geconcludeerd dat hoorapparaten wel degelijk een negatieve invloed uitoefenen op de lokalisatieperformantie en dat een bilaterale ADM configuratie de lokalisatie verder bemoeilijkt (**paragraaf 2.4**).

Hoofdstuk 3: het gebruik van virtuele akoestiek bij het evalueren van lokalisatieperformantie

In dit manuscript worden een aantal nieuwe ruisonderdrukkingsalgoritmen voorgesteld en geëvalueerd (hoofdstuk 4 tot 6). Om nieuwe algoritmen te evalueren, wordt er vaak gebruik gemaakt van off-line bewerkte signalen die vervolgens aan luisteraars worden gepresenteerd door middel van een hoofdtelefoon. Hierdoor wordt het ontwikkelingsproces sneller en eenvoudiger.

In dit werk wordt ondermeer de invloed van ruisonderdrukkingsalgoritmen op de lokalisatie-performantie in het horizontale vlak geëvalueerd. Het gebruik van hoofdtelefoon experimenten voor lokalisatie doeleinden is echter niet vanzelfsprekend. Meer nog, om zeer tijds-intensieve, gepersonaliseerde in-de-oor metingen te vermijden, wordt er bij voorkeur gebruik gemaakt van metingen met een kunsthoofd. Dit heeft echter ook een invloed op de lokalisatieperformantie en is afhankelijk van het gebruikte kunsthoofd (Møller et al., 1999). Deze factoren hebben ervoor gezorgd dat een evaluatie is uitgevoerd om de nauwkeurigheid van lokalisatie-experimenten met behulp van hoofdtelefoon en kunsthoofd op te meten (**paragraaf 3.2**). Door een samenwerking tussen ExpORL, SISTA-SCD en de groep Akoestiek en Thermische Fysica werd dit onderzoek uitgebreid met de vraag of geavanceerde virtuele akoestische modellen (**paragraaf 3.1**) kunnen gebruikt worden bij het evalueren van hoorapparaatalgoritmen. Dit zou ervoor zorgen dat de nood aan de fysisch beschikbaarheid van verschillende akoestische omgevingen tijdens de evaluatie van algoritmen wordt opgelost.

De data (**paragraaf 3.3**) van 7 normaalhorenden toont aan dat het gebruik van een kunsthoofd (CORTEX MK2) slechts een kleine, maar significante daling van de lokalisatieperformantie introduceert bij het lokaliseren van smalband hoog-frequente stimuli (**paragraaf 3.4**). Bij het lokaliseren van breedband of laag-frequente stimuli wordt er geen significante invloed geconstateerd. Het gebruik van akoestische modellen heeft ook enkel invloed bij het lokaliseren van hoog-frequente stimuli, en dit enkel in experimenten waarbij de originele lokalisatieperformantie hoog is.

Dit hoofdstuk toont aan dat er significante verschillen kunnen optreden tussen natuurlijke lokalisatie en lokalisatie met hoofdtelefoonaanbieding, vooral bij het lokaliseren van hoog-frequente stimuli. Aangezien deze verschillen echter klein zijn, zeker bij gebruik van breedband stimuli, kan en zal deze techniek worden toegepast om de invloed van ruisonderdrukkingsalgoritmen op lokalisatie in het horizontale vlak te onderzoeken.

Hoofdstuk 4: Het bewaren van binaurale cues d.m.v. de meerkanaals Wiener filter: MWF, MWF-N, MWF-ITF

Zoals vermeld in hoofdstuk 2, bieden de bilaterale FDM en ADM geen optimale combinatie van ruisonderdrukking en binauraal horen. De komst van een binaurale link tussen beide hoorapparaten biedt echter de mogelijkheid om binaurale ruisonderdrukkingssystemen te ontwerpen. Deze systemen hebben toegang tot alle microfoons van beide hoorapparaten. De toename van het aantal microfoons verhoogt de potentiële ruisonderdrukking terwijl met de binaurale link de binaurale cues beter gecontroleerd kunnen worden. Hierdoor kan interferentie van ruisonderdrukking met ruimtelijk horen vermeden worden. Dit hoofdstuk stelt een aantal MWF-gebaseerde binaurale algoritmen voor, ontworpen om ruisonderdrukking te combineren met binauraal horen.

In **paragraaf 4.2** wordt de context waarin de binaurale ruisonderdrukkingsalgoritmen ontworpen worden gedefinieerd. Zo worden de binaurale signalen en filters mathematisch beschreven samen met een aantal theoretische maten zoals ruisonderdrukking en interaurale tijds- en intensiteitsverschillen. Deze maten zullen gebruikt worden om de performantie van de systemen te beschrijven en hun invloed op de binaurale cues te voorspellen.

Paragraaf 4.3 stelt een binaurale MWF voor. Dit is een uitbreiding van de monaurale MWF, geïntroduceerd in het werk van Doclo and Moonen (2002). De binaurale MWF is een systeem dat inherent de binaurale cues van de spraakcomponent bewaart. De cues van de ruiscomponent worden echter gewijzigd in die van de spraakcomponent. Om de ruimtelijke gewaarwording van de slechthorende te bewaren en om het "cocktail-party effect"te kunnen benutten moeten echter de binaurale cues van zowel de spraak- als de ruiscomponent bewaard worden. Daarom worden er twee nieuwe uitbreidingen van een MWF voorgesteld: de MWF-N en de MWF-ITF.

De MWF met gedeeltelijke ruisschatting (MWF-N), besproken in **paragraaf**

4.4, is ontworpen om niet de volledige, maar enkel een gedeelte van de ruiscomponent van het signaal te verwijderen. Het resterende deel zorgt dan voor een correcte lokalisatie van de ruiscomponent. Vanzelfsprekend leidt dit tot een verlies in ruisonderdrukking. De parameter η is een trade-off parameter die de hoeveelheid onverwerkte ruis bepaalt. Bij $\eta = 0$ herleidt de MWF-N zich tot de standaard MWF met maximale ruisonderdrukking . Indien $\eta = 1$ worden de binaurale cues van de spraak en de ruis perfect bewaard maar is er geen ruisonderdrukking. De binaurale MWF en MWF-N worden verder geëvalueerd en vergeleken met een bilaterale ADM in hoofdstukken 5 en 6.

De binaurale MWF met interaurale transfer functie (MWF-ITF), besproken in **paragraaf 4.5**, voegt een term toe aan de kostfunctie van de binaurale MWF. Deze term beperkt de oplossingsruimte van de kostfunctie tot filters die, in zekere mate (afhankelijk van het gewicht β), voldoen aan het bewaren van de binaurale cues van de ruiscomponent. Indien β te groot wordt gekozen, veranderen de binaurale cues van de spraakcomponent echter in deze van de ruiscomponent. Uitgebreid onderzoek naar dit algoritme is nog volop aan de gang. De gerapporteerde pilootexperimenten tonen echter reeds de mogelijkheden van de MWF-ITF.

Om het overzicht van de ontwikkelde binaurale MWF algoritmen te vervolledigen, beschrijft **paragraaf 4.6** onderzoek naar algoritmen met gereduceerde bandbreedte. Aangezien de binaurale link vanuit commercieel standpunt bij voorkeur een draadloze link is, vraagt het oversturen van microfoonsignalen tussen beide hoorapparaten een grote investering van het beperkte vermogen. Door het combineren van microfoonsignalen vooraleer ze worden doorgestuurd naar het ipsilaterale hoorapparaat kan de benodigde bandbreedte nodig om een maximale performantie te bereiken worden verminderd. De onderzochte mogelijkheden zijn: een binaurale MWF die gebruik maakt van slechts één contralateraal microfoonsignaal, een binaurale MWF die gebruik maakt van een contralaterale superdirectieve beamformer, een binaurale MWF die gebruik maakt van een monaurale contralaterale MWF en een binaurale MWF die gebruik maakt van een gedistribueerde processing. Deze oplossingen behalen een performantie tussen die van de bilaterale en de volledig binaurale MWF in, waarbij de gedistribueerde MWF het resultaat van een volledig binaurale MWF benadert.

Hoofdstuk 5: Ruisonderdrukking van de binaurale MWF en MWF-N t.o.v. de bilaterale ADM.

In dit hoofdstuk wordt nagegaan wat de realistische ruisonderdrukking is bij gebruik van een binaurale MWF en MWF-N in verschillende akoestische omgevingen en in verschillende ruimtelijke condities. Een bilaterale ADM, op dit moment de meest gebruikte commerciële adaptieve ruisonderdrukkingstechniek, wordt gebruikt als referentie (**paragraaf 5.1**). Aangezien het doorsturen van signalen van het contralaterale naar het ipsilaterale hoorapparaat een dure investering is van het beschikbare vermogen worden verschillende microfooncombinaties onderzocht. Hierbij wordt een bilaterale MWF en MWF-N uitgebreid met respectievelijk geen, één en twee contralaterale microfoonsignalen.

De evaluatie gebeurt aan de hand van objectieve en perceptuele evaluaties (**paragraaf 5.2**). Als objectieve maat wordt er gebruik gemaakt van de spraakgewogen verbetering in SNR, gedefinieerd door Greenberg et al. (1993), berekend op de in- en uitgangssignalen van het algoritme. Bij de perceptuele evaluatie wordt er door middel van een adaptieve procedure de SNR bepaald bij dewelke 50% van de spraak wordt verstaan, de zogenaamde spraakverstaanbaarheidsdrempel (SRT). Hoe lager deze waarde hoe beter de performantie van het algoritme.

De objectieve evaluatie (**paragraaf 5.3.1**) toont aan dat de ruisonderdrukking van zowel de MWF, de MWF-N als de ADM sterk wordt beïnvloed door de aanwezige reverberatie. In een omgeving met reverberatietijd $T_{60} = 0.21s$ worden waarden tot 23dB genoteerd in aanwezigheid van één enkele ruisbron, bij een $T_{60} = 0.61s$ daalt deze waarde naar 12dB. Het blijkt ook dat het toevoegen van contralaterale microfoons wel degelijk de hoeveelheid ruisonderdrukking vergroot. Deze winst is echter sterk afhankelijk van de plaatsing van de spraak- en ruisbronnen (Figuur 5.1 en Figuur 5.2). De performantie van de bilaterale ADM en de bilaterale MWF, beide 2-microfoonssystemen, blijken gelijklopend te zijn behalve wanneer de spraakbron niet frontaal gepositioneerd is. In dat geval daalt de performantie van de ADM en wordt deze sterk overtroffen door de performantie van de MWF. Dit is logisch aangezien de ADM, in tegenstelling tot de MWF, veronderstelt dat de spraakbron zich recht voor de luisteraar bevindt. Bij het vergelijken van de resultaten van de MWF-N en de MWF wordt er geobserveerd dat de ruisonderdrukking, zoals verwacht uit hoofdstuk 4, significant daalt bij het verhogen van de parameter η.

De perceptuele evaluaties (**paragraaf 5.3.2**) tonen dezelfde trends als de objectieve evaluaties, waarbij de 2-microfoons ADM en MWF ongeveer dezelfde ruisonderdrukkingsperformantie vertonen, behalve als de spraakbron niet rechtvoor geplaatst wordt en waarbij het toevoegen van vooral één contralaterale microfoon aan de MWF een duidelijke verbetering in spraakverstaan biedt. De MWF-N heeft typisch een lagere performantie dan de MWF, behalve wanneer spraak en ruis zeer ver van elkaar gescheiden zijn. Dit kan mogelijk verklaard worden door een verbeterd "cocktail-party effect" bij het gebruik van de MWF-N, doordat deze, in tegenstelling tot de MWF, de cues van zowel de spraak- als de ruiscomponent bewaart.

Hoofdstuk 6: Het lokaliseren van geluidsbronnen met de binaurale MWF en MWF-N t.o.v. de bilaterale ADM.

Het doel van de binaurale MWF en MWF-N is uiteindelijk om het spraakverstaan in lawaaierige omstandigheden te verbeteren en daarbij de binaurale cues te bewaren. Dit hoofdstuk bespreekt een perceptuele evaluatie van het binauraal horen bij gebruik van de MWF en MWF-N (**paragraaf 6.1**). Hiervoor wordt gebruik gemaakt van een lokalisatie-experiment in het frontale horizontale vlak. Een bilaterale ADM en een conditie zonder ruisonderdrukking worden gebruikt als referentie condities.

Om de invloed van de MWF te begrijpen worden de spraak- en ruiscomponent door de geconvergeerde MWF filters (geconvergeerd op het totaalsignaal bestaande uit spraak- én ruiscomponent) gefilterd en daarna afzonderlijk gepresenteerd. Hierdoor worden maskeringseffecten tijdens het lokalisatie-experiment vermeden. In een tweede fase worden de spraak- en ruiscomponent samen aangeboden (**paragraaf 6.2**). De taak bestaat er telkens in om de spraak- én de ruiscomponent te lokaliseren. Als stimulus wordt er telkens gebruik gemaakt van breedbandige signalen. Drie scenario's worden onderzocht met verschillende invalshoeken voor de spraak- en ruiscomponent.

De data gepresenteerd in **paragraaf 6.3** toont aan dat de lokalisatie bij gebruik van een bilaterale ADM zeer slecht is indien de te lokaliseren stimulus afkomstig is van de linker- of de rechterzijde van het hoofd. Deze signalen worden vaak gelokaliseerd als komende van rechtvoor in plaats van links of rechts van het hoofd. Een kwaliteitsanalyse toont aan dat deze signalen vaak gepercipieerd worden als diffuse signalen zonder enige richtingsinformatie. Doordat de lokalisatietaak een antwoord vereist, wordt de neutrale richting, 0°, vaak aangeduid als positie waar de geluidsbron zich bevindt. De diffuusheid van het signaal kan verklaard worden door het feit dat een ADM de correlatie tussen microfoonsignalen gebruikt om de ruissignalen, i.e. de signalen niet komende van rechtvoor, weg te filteren.

De MWF gedraagt zich grotendeels zoals verwacht uit de theoretische analyse. Dit wil zeggen dat de lokatie van de spraakcomponent correct wordt gepercipieerd maar dat de ruiscomponent ook wordt gepercipieerd op de lokatie van de spraakcomponent. De kwaliteitsanalyse toont echter aan dat, indien geen maskeringseffecten aanwezig zijn, de ruiscomponent vaak gepercipieerd wordt als komende van twee richtingen. Eén richting is de richting van de spraakcomponent en de andere richting is de richting van de oorspronkelijke ruiscomponent. Dit kan verklaard worden aan de hand van de schatting van de correlatiematrices in het MWF algoritme (Figuur 6.2). De kwaliteit van deze schatting is sterk afhankelijk van de SNR in elke frequentieband. Indien de schatting van de spraakcorrelatiematrix correct verloopt is er veel ruisonderdrukking maar wordt de ruiscomponent gepercipieerd op de plaats van de spraakcomponent. Verloopt de schatting van de spraakcorrelatiematrix slecht (in frequentiebanden met lage SNR) is er weinig ruisonderdrukking maar blijven de binaurale cues van de ruiscomponent bewaard. Bij het simultaan aanbieden van beide componenten, waardoor maskering optreedt, kunnen de fouten in de schatting van de correlatiematrices er dan ook voor zorgen dat de ruiscomponent correct wordt gelokaliseerd (**paragraaf 6.4**).

Bij het gebruik van een MWF-N met $\eta = 0.2$ blijft de lokalisatie van de spraak- en de ruiscomponent bewaard. De MWF-N benadert de lokalisatieperformantie van de conditie zonder ruisonderdrukking in alle geteste scenario's.

Besluit en suggesties voor verder onderzoek

Dit werk stelt de invloed van ruisonderdrukkingsalgoritmen op binaurale cues in vraag (**paragraaf 7.1**). Deze cues zijn belangrijk voor de ruimtelijke gewaarwording van de hoorapparaatgebruiker en voor een verbeterd spraakverstaan in lawaaierige omstandigheden door het zogenaamde "cocktail-party effect".

Het effect van de twee meest geïmplementeerde ruisonderdrukkingsalgoritmen op de binaurale cues werd geanalyseerd door middel van theoretische en perceptuele evaluaties. De theoretische evaluatie toont aan dat zowel de bilaterale FDM als de bilaterale ADM een negatieve invloed kunnen uitoefenen op de lokalisatie van geluidsbronnen. Perceptuele evaluaties, enkel uitgevoerd voor de bilaterale ADM, bevestigen dit. De evaluaties in hoofdstuk 2, 5 en 6 wijzen erop dat de bilaterale ADM de binaurale cues bewaart van alle signalen komende uit de meest frontale richtingen. Signalen uit andere richtingen worden onderdrukt wat echter ook aanleiding geeft tot de distortie van de binaurale cues. Meer nog, signalen worden dan vaak waargenomen als zijnde diffuus zonder enige richtingsinformatie.

De binaurale MWF is een uitbreiding van de bilaterale MWF en leidt tot een verbeterde ruisonderdrukking t.o.v. de bilaterale MWF en ADM door gebruik te maken van contralaterale microfoonsignalen. Bij het gebruik van de MWF worden de binaurale cues van de spraakcomponent inherent bewaard, ongeacht de invalshoek van het signaal. Dit in tegenstelling tot de ADM. Uit een theoretische analyse blijkt echter dat de cues van de ruiscomponent worden gewijzigd in die van de spraakcomponent. Luistertesten tonen aan dat dit effect afhankelijk is van de kwaliteit van de geschatte correlatie-matrices en de SNR van de desbetreffende luistersituatie. Hierdoor zijn de lokalisatieresultaten soms beter dan verwacht. Voor het verbeteren van de lokalisatie van de ruiscomponent werden twee varianten van de binaurale MWF voorgesteld: de binaurale MWF-ITF en de binaurale MWF-N.

De binaurale MWF-ITF is gebaseerd op het toevoegen van een extra term in de kostfunctie van de MWF. Deze term beperkt de oplossingsruimte van de kostfunctie tot filters die in zekere mate voldoen aan het bewaren van de binaurale cues van de ruiscomponent. Indien de nadruk op deze term wordt opgedreven, d.m.v. de parameter β, veranderen de cues van de spraakcomponent echter in deze van de ruiscomponent. Uitgebreid onderzoek naar dit algoritme is nog steeds aan de gang. Toch werd reeds aangetoond dat, in scenarios met één enkele ruisbron, de MWF-ITF de gemiddelde lokalisatieperformantie verbeterd in vergelijking met de MWF.

De binaurale MWF-N is gebaseerd op een gedeeltelijke ruisschatting. Door de ruis slechts gedeeltelijk te verwijderen kan het overgebleven signaal gebruikt worden voor een correcte lokalizatie van de ruiscomponent. Dit leidt tot een verbeterde lokalisatieperformantie maar logischerwijze ook tot een verminderde ruisonderdrukking. Dit laatste kan, indien spraak en ruisbron voldoende ruimtelijk gescheiden zijn, gecompenseerd worden door het "cocktail-party effect"dat optreedt wanneer voldoende binaurale cues van zowel de spraak- als de ruiscomponent bewaard zijn gebleven.

Gedurende dit werk werd het potentieel van de binaurale MWF, MWT-ITF en de MWF-N aangetoond. In het geval van de MWF-ITF werden enkel pilootexperimenten gerapporteerd. Verder onderzoek omvat de nood aan een grondige evaluatie van de MWF-ITF in aanwezigheid van meerdere ruisbronnen en het gebruik van verschillende microfooncombinaties in verschillende akoestische omstandigheden.

Ook zijn er mogelijkheden tot uitbreiding van de MWF-N (**paragraaf 7.2**). Zo is op dit moment de proportie onverwerkt signaal, gebruikt om de overblijvende ruiscomponent te maskeren, identiek voor alle frequentiebanden. Door een frequentie-specifieke weging, gerelateerd aan SNR of het belang van die specifieke frequentiebanden voor de lokalisatie van geluidsbronnen, zou mogelijk dezelfde lokalisatieperformantie kunnen behaald worden met een verhoogde ruisonderdrukking.

Een ander belangrijk punt van verder onderzoek is het ontwerp van objectieve performantiematen die het effect van algoritmen op het binauraal horen beter voorspellen. Gedurende dit werk werden een aantal objectieve performantiematen beschreven en gebruikt. Vervolgens werden de algoritmen zowel objectief en perceptueel geëvalueerd. Performantiematen die gebaseerd zijn op bestaande en verder te verfijnen modellen van het menselijk lokalisatiemechanisme kunnen de nood aan tijdsintensieve perceptuele evaluaties sterk verminderen.

Chapter 1

Introduction

1.1 Motivation

Hearing aids offer hearing impaired subjects the ability to perceive and recognize sounds or speech signals. Hearing aids have been introduced a very long time ago, for an overview starting from the acoustical era around the 17th century see Lybarger (1988); Berger (1984), and have been evolving ever since. However, one of the main complaints of hearing aid users remains the lack of speech understanding in noisy environments (**section 1.2**). These complaints have led to a enormous amount of research done in the field of noise suppression algorithms for hearing aids (**section 1.3**).

When designing signal processing algorithms for hearing aids, different features of the human auditory system are taken into account. However, since most algorithms are developed monaurally, i.e. maximizing the performance for a single hearing aid, only a very limited amount of research has been done on the effects of these algorithms on binaural information, i.e. the information which can be derived from comparing the sound signals received at the left and the right eardrum. This information is essential for the localization (**section 1.4**) of sound sources (Hartmann, 1999; Makous and Middlebrooks, 1990) and for the auditory scene analysis (ASA) done by the human auditory system. ASA, the main principles of which have been described by Bregman (1993), McAdams (1993), and others (see Moore (1989) and Bregman (1999) for an overview), is the development of an internal representation of the acoustic environment around the listener. It is based on combining loudness cues, pitch information, binaural and monaural localization cues, gaps in speech or noise signals, acoustical reflections, visual inputs, knowledge of the sound sources, etc., to perceive and segregate audio streams. The spatial separation

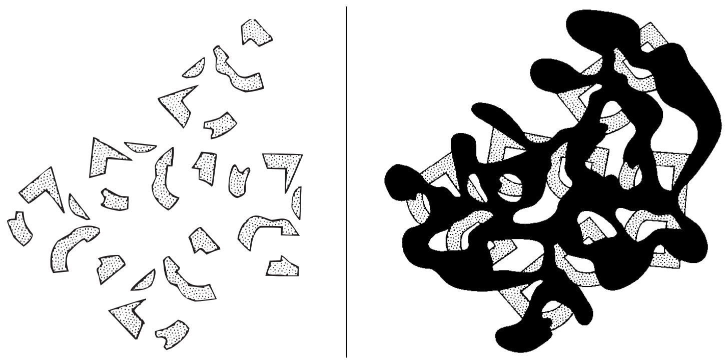

Figure 1.1 — Bregman's Bs as an illustration of the task of scene analysis. **Left**: even if the observer knows that the depicted objects are characters, it is hard to recognize them because they are partially missing. **Right**: When the lacking visibility of part of the characters is explained by some other shape, e.g. a spot of ink, it becomes easy to recognize them (Bregman, 1981).

of competing sound signals leads to differences in binaural information and often result in a surprisingly good speech understanding in very adverse listening conditions (Bronkhorst and Plomp, 1988, 1989; Peissig and Kollmeier, 1997; Drennan et al., 2003). This effect is generally known as spatial release from masking or 'the cocktail-party effect' (**section 1.5**). Providing a hearing aid user with two independently working hearing aids, having independent noise reduction schemes, could have a destructive effect on the binaural cues. Correspondingly, the hearing aid users localization performance and speech perception in a complex environment could be degraded.

A visual analogy of the complexity of scene analysis is given by Bregman (1981) and illustrated in Figure 1.1. The task in Figure 1.1 exists of identifying the characters which are partly visible in the picture. Although the overall signal to noise ratio (SNR) is higher in the left part of Figure 1.1, the characters are unreadable. When the lack in visibility of parts of the characters is explained by some other shape, as illustrated in the right part of Figure 1.1, it becomes easy to recognize them. This gives an idea on how noise reduction algorithms may affect scene analysis cues and hence may interfere with the natural segregation mechanisms.

In the last decade, the general awareness has grown that hearing aids in general and more specifically noise reduction algorithms should not be evaluated merely using monaural performance measures, such as SNR, but that their effect on binaural and other cues should be carefully studied (Desloge et al., 1997; Van den Bogaert et al., 2006; Keidser et al., 2006). Moreover, the recent analysis of large datasets of information provided by hearing impaired people and hearing aid users show that, besides speech perception in adverse listening conditions, the reduced spatial aspects of hearing and listening are often perceived as a main disability. (Noble et al., 1995). This is also demonstrated by the recently developed speech and spatial quality questionnaire (SSQ) (Gatehouse and Noble, 2004; Noble and Gatehouse, 2004; Noble, 2006) which, in contrast to classic questionnaires, integrates questions regarding spatial hearing into the perceptual evaluation of hearing aids. Room for improvement in binaural processing, such as directional hearing, is also found in the work of Kochkin (2005) in which 1511 hearing aid users were interviewed concerning their satisfaction of their bilateral hearing aids. Recent technological developments, which enable the use of a communication link between two hearing aids, have boosted the interest in the combination of hearing aids with binaural hearing even further, also from a commercial point of view.

Since more than a decade, the majority of hearing aid users is convinced of using a bilateral instead of a monaural hearing aid configuration (Kochkin and Kuk, 1997; Libby, 2007). This number has been rising ever since and reached its maximum in 2001 with 75% of the hearing aid users wearing a bilateral configuration. A similar evolution is found now with technological and commercial interests shifting towards binaural hearing aids instead of bilateral hearing aids (Deiss, 2002). The communication link between hearing aids in commercial products is, at the moment, still restricted to a very narrow bandwidth. This allows the transmission of a limited number of parameters between the left and the right hearing aid (e.g. Siemens Acuris). It is expected that transmitting one and even more microphone signals will soon become a realistic option in commercial hearing aid designs. Enabling the access to ipsi- and contralateral microphones offers new possibilities with respect to an improved noise reduction performance and the preservation of binaural cues by noise reduction algorithms.

The main focus of this project was to study the combination of noise reduction algorithms with the preservation of binaural information, also known as the binaural cues. First, commercially available, bilateral, noise reduction algorithms (**chapter 2**) were evaluated with respect to their influence on binaural cues. This was achieved by a theoretical and a perceptual evaluation, the latter being based on a localization experiment in the frontal horizontal hemisphere, which is a binaural task. Secondly, three binaural algorithms were developed and evaluated, i.e. the MWF, the MWF-ITF and the MWF-N (**chapter 4** to **chapter 6**). These algorithms aim at improving speech perception by performing noise reduction while preserving the binaural cues. Since the perceptual relevance of objective performance measures is not always trivial or even unknown, particularly during the study of binaural cue preservation, all of these algorithms were evaluated by using theoretical, objective and perceptual performance measures. The localization experiments to evaluate the MWF-based algorithms were performed under headphones using binaural room impulse-responses measured with a manikin. This methodology was first validated (**chapter3**).

1.2 Hearing impairment and hearing aids

A hearing impairment or hearing loss is a full or partial decrease in the ability to detect, discriminate and identify sounds. A hearing loss does not simply result in an attenuation of all sounds entering the ear, but also in distortions in the ear. Each of the different aspects of a hearing loss, i.e. a decreased audibility, dynamic range, frequency resolution and temporal resolution, can cause a reduction in speech intelligibility. Combined, they can cause a hearing impaired subject to understand speech much worse than a normal hearing person in the same situation, even when the hearing impaired is wearing a hearing aid (Humes, 1991; Baer and Moore, 1993; Moore, 2003). In the work of Plomp (1978), a model of the speech reception threshold (SRT) as a function of the noise level has been developed for hearing impaired persons. The SRT is defined as the SNR at which 50% of the speech can be correctly identified by the listener. The effect of the attenuation, the distortions and the combination of both on the SRT as a function of noise is depicted in Figure 1.2. Compared to the SRT of normal hearing listeners, the attenuation component of the hearing loss influences the SRT at low but not at high noise levels. Since the distortion component is independent of the noise level, hearing impaired listeners need a higher SNR than normal hearing persons. Whereas normal hearing subjects are capable of understanding speech in a noisy environment around a SNR of -5dB, people with a mild or severe hearing loss may require a SNR that is up to 15dB (on average 5dB) higher (Duquesnoy and Plomp, 1983; Plomp, 1978; Plomp and Mimpen, 1979).

Hearing aids are commonly used to overcome the deficits associated with hearing loss. Most hearing aids can be categorized into three different types. The first, illustrated in Figure 1.3, is the behind-the-ear (BTE) hearing aid. The microphone(s), the electronics and the receiver are mounted in a banana-shaped case which is placed on top of the ear. The second is the in-the-ear (ITE) hearing aid, which occupies the concha (deep center portion of the visible part of the ear) as well as about half of the length of the ear canal. The third type occupies a small portion of the external auditory canal of the ear and is referred to as an in-the-canal (ITC) hearing aid. In this research project we will focuss on BTE devices which offer the highest amplification levels since their large battery generates more power than those of other hearing aids, and which offer the largest flexibility in terms of signal processing algorithms due to their large and powerful DSP chip.

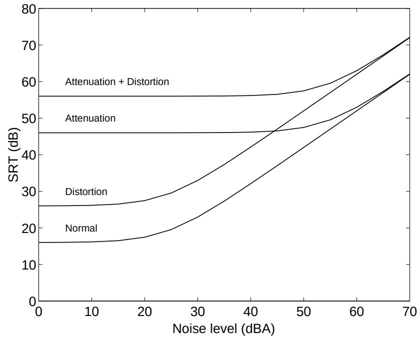

Figure 1.2 — Speech level required for a 50% sentence score, i.e. the SRT level, as a function of noise level. The lower curve represents normal hearing. The other curves represent the effect of attenuation, distortion and a combination of both.

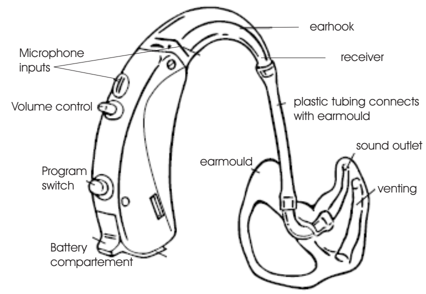

Figure 1.3 — A Widex Inteo BTE hearing aid with two microphone inputs. The hearing aid is connected to a custom made earmold by using a plastic tubing.

There are two basic functions of a hearing aid. The first function is to amplify the input signals, thereby compensating for the higher hearing thresholds of the hearing impaired. The second function consists of compressing the acoustical signals. This is necessary since an impaired ear has a reduced dynamic range due to the combination of higher hearing thresholds and similar or even lower uncomfortable loudness levels, which is also known as recruitment, compared to a normal functioning ear. Therefore, a compression scheme is implemented which limits the maximum output level (compression limiter) and/or which takes into account the reduced dynamic range of the impaired ear (dynamic range compression) (Lippman et al., 1981; Dillon, 2001*a*). Besides these basic functions a lot of additional features are present in modern hearing aids, such as noise reduction or feedback cancellation.

Combined with every BTE hearing aid is an earmold, shown in Figure 1.3. This is a flexible or custom made fixture, that fits the individuals ear and delivers the processed sound produced by the BTE to the eardrum. For the patient's comfort, a venting tube is present in the earmold with a diameter ranging from 1mm to almost an open fitting. This tube ventilates the ear canal and reduces the occlusion effect which is the often uncomfortable amplification of the users own voice and low frequencies due to the closing of the ear canal. Moreover, this tube also enables a direct sound component to reach the ear drum which may be used by the hearing aid user. This has to be taken into account when designing perceptual evaluations, especially when evaluating hearing impaired subjects with a mild hearing loss.

The first digital hearing aids were introduced in 1995, i.e. the Widex Senso. Since then, these devices are an overwhelming success (in 2005, 90% of the hearing aids sold in the United States were digital hearing aids (Kochkin, 2005)). As a result, almost all currently developed high end hearing devices contain a digital signal processor (DSP) offering more possibilities concerning data logging, the usage of a remote control, the development of much more complex algorithms which can be turned on or off manually or automatically by decision making routines running on the DSP, etc.. Despite the rapid developments in DSP technology, designing algorithms for hearing aids remains challenging since a compromise needs to be found between a good performance, a high robustness, a low power consumption and a relatively low complexity of the algorithm.

1.3 Noise reduction algorithms for hearing aids

Hearing aid users have great difficulty understanding speech in noisy environments (Duquesnoy and Plomp, 1983; Plomp and Duquesnoy, 1982; Plomp, 1986; Helfer and Wilber, 1990; Cox and Alexander, 1991). Therefore noise reduction algorithms have been developed which aim at reducing unwanted sounds and improving the SNR for the hearing aid user. After all, an improvement of 1dB in SNR around the SRT can generate an increase in speech understanding of 10-15% in every day communication (Plomp and Mimpen, 1979). Although the definition of noise reduction seems rather straightforward, in reality it is not. This, because the classification of a sound as being unwanted is often dependent on the individual perception. Signals such as speech and music are possibly a wanted or an unwanted sound. All the different noise reduction approaches can be classified into two categories, i.e. single and multichannel techniques.

1.3.1 Single channel noise reduction

Single channel techniques, i.e. single microphone techniques with a single connection to the outside world (this to exclude a directional microphone), are based on exploiting differences in physical characteristics, such as frequency content, temporal characteristics, etc., between speech and other sound sources. An overview can be found in the work of Bentler and Chiou (2006). The main single channel techniques used in hearing aids are: a high pass filter, spectral substraction techniques and a multiple band pass filter.

A high pass filter is the first noise reduction technique ever used in commercial hearing aids (Dillon and Lovegrove, 1993; Levitt, 2001). It is based on the hypothesis that noise, in contrast with speech, typically consists of a large amount of energy at the low frequencies. By reducing these frequencies, an improvement in overall SNR is obtained. Moreover, it also avoids that the noisy energy triggers the compression algorithm which, in older hearing aids, operates on the full frequency range.

Spectral substraction (Weiss, 1974; Boll, 1979) assumes that the short term noise spectrum can be obtained during pauses in the speech by using a voice activity detector (VAD), see section 1.3.4. Furthermore it is assumed that the noise is sufficiently stationary such that its estimate can be subtracted from the spectrum obtained during speech and noise periods.

A multiple band pass filter, proposed by Clarkson and Bahgat (1991), implements a filter bank to separate the input signal in different frequency channels. By examining the modulation frequency in each frequency band this technique determines whether this channel is more likely to be a noise or a speech signal. Each frequency band is then amplified accordingly. However, due to the fluctuating gains in the different frequency bins, typical distortions such as musical noise occur.

Although single channel systems do provide useful results in some other audio applications and although an increase in listening comfort is often reported by hearing aid users, they generally do not generate any benefit in terms of speech intelligibility (Levitt et al., 1993; Dillon and Lovegrove, 1993; Walden et al., 2000; Arehart et al., 2003; Moore, 2003; Bentler and Chiou, 2006). This indicates that an improvement in SNR does not automatically yield an increase in intelligibility. This is most likely due to the fact that improving speech intelligibility in noise with only a single input relies on the signal and the noise being sufficiently different in frequency or time to be separable by signal processing but not by a person with impaired hearing (Dillon, 2001*b*). Moreover, due to the spectral and temporal overlap present in many speech-in-noise conditions, it becomes extremely difficult for single channel techniques to sufficiently suppress the noise without introducing speech distortions and so called musical noise (Cappe, 1994; Spriet, 2004).

1.3.2 Multichannel noise reduction

Due to the miniaturization of microphones (Ouellette, 1999), two and even three (e.g. Siemens Triano 3) microphones can be integrated in commercial BTE and ITE hearing aids. In contrast with single channel systems, multichannel noise reduction has the ability to exploit not only spectral and temporal differences but also the spatial separation between sound sources to enhance the SNR. Hence, it is preferred over single channel systems.

A distinction can be made between systems with a directional characteristic which is invariant in time and those with the capacity to adapt to the environment in order to minimize the amount of noise at each time instant (Dillon, 2001*b*). All of these techniques can be classified based on beamforming, multichannel Wiener filtering or computational auditory scene analysis (CASA). At present, beamforming is the most commonly used commercial multi-microphone noise reduction approach in hearing aids.

Fixed Beamforming

Beamforming techniques linearly combine different microphone signals to enhance or suppress signals arriving from certain angles. In fixed beamforming, the filters that are applied to the microphone signals are fixed and hence data-independent. The filter coefficients are optimized to steer a focussed beam to the target direction, for hearing aids typically the forward field of view is used, while reducing the sounds arriving from other directions as much as possible. Fixed beamformers which have been considered for hearing aid applications are additive and subtractive arrays a.k.a. delay-and-add and delay-and-subtract beamformers, filter and sum beamformers and superdirective beamformers (Dillon, 2001*b*).

Subtractive arrays (Thompson, 2000), such as a conventional fixed directional microphone (Figure 2.1) (Soede, Bilsen and Berkhout, 1993; Ricketts and Henry, 2002) which will be discussed in chapter 2, subtract delayed microphone signals to produce a zero sensitivity for sounds arriving from the back hemisphere. Additive arrays, also called delay-and-add arrays, try, by adding delayed microphone signals, to produce maximal sensitivity for signals arriving from the front and less sensitivity for signals arriving from the back. However, additive arrays are not commonly used in typical hearing aid applications since they have limited performance and since they only become effective if the size of the microphone array is larger than, or equal to, a quarter wavelength of the frequency of interest (Van Compernolle and Van Gerven, 1995; Dillon, 2001*b*). Therefore it has only been successfully used in hearing devices based on a large microphone array such as the design based on a pair of spectacles in the work of Soede, J. and Bilsen (1993) and Soede, Bilsen and Berkhout (1993).

Instead of just delaying microphone signals before adding or subtracting them, it is also possible to first perform a more general filter operation. This is obviously the most general beamforming structure and is called a filter and sum beamformer (Johnson and Dudgeon, 1993). Using this structure, it is possible to design a fixed beamformer whose spatial directivity pattern optimally fits a predefined desired shape.

A distinct class of beamformers are superdirective beamformers (Cox et al., 1986; Kates and Weiss, 1996; Bitzer and Simmer, 2001). These are designed to maximize the directivity index (DI). This represents the directional sensitivity in the direction of the speech source for a known (diffuse) noise field. The DI (Beranek, 1954) is often-used to measure the performance of fixed directional noise reduction algorithms (Desloge et al., 1997; Ricketts, 2000). With θ the azimuth coordinate and ϕ the elevation coordinate, the DI equals:

$$DI(f) = \frac{4\pi \ |P(f,0,0)|^2}{\int_0^{2\pi} \int_0^{\pi} |P(f,\theta,\phi)|^2 \ |sin\theta| \ d\theta \ d\phi} \tag{1.1}$$

where the $|P(f,\theta,\phi)|^2$ is the mean squared sound pressure, at frequency f, of the output signal of the hearing aid when a single sound source is located at the coordinate (θ,ϕ). In other words the DI evaluates the sensitivity of the directional system to sounds arriving from the front compared to all other angles. It has been shown by Ricketts (2000) that a correlation exists between the DI and the improvement of speech intelligibility in noise.

Adaptive Beamforming

Since the location of jammer sources is unknown and may vary over time, adaptive beamforming strategies have been developed. These are able to adapt their filter coefficients and hence their spatial characteristic according to changes in the acoustic environment. Hence, adaptive beamformers generally exhibit a better noise suppression performance than fixed beamformers, especially if the number of interferers is relatively small, i.e. smaller than the number of microphones. A good overview of adaptive beamforming can be found in Van Veen and Buckley (1988), Van Compernolle and Van Gerven (1995) and Spriet (2004).

To avoid distortion of the speech signal, adaptive beamformers are typically constrained to preserve the signals arriving from the forward field of view. Hence they typically give rise to a constrained optimization problem. The least expensive implementation and therefore the most interesting for early hearing aid applications is the adaptive directional microphone (Figure 2.4) (Ricketts and Henry, 2002; Luo et al., 2002) which will be subject of further discussion in chapters 2, 5 and 6.

A more general and often a more performant implementation is the Griffiths-Jim beamformer (Griffiths and Jim, 1982) also known as the Generalised Sidelobe Canceller (GSC) (see Figure 1.4). The GSC translates the constrained optimization problem into an unconstrained problem through the combination of a fixed spatial preprocessor, i.e. a fixed beamformer and a blocking matrix, and an adaptive noise canceller (ANC). The fixed beamformer creates a so called speech reference. The blocking matrix, which is usually orthogonal to the fixed beamformer, creates so called noise references and avoids the distortion of the speech component by the ANC. The ANC which consists of a multichannel adaptive filter, removes the remaining noise in the speech reference. The optimal filters are typically computed during 'noise only' periods by minimizing the energy at the output of the algorithm. It is assumed that the noise source is sufficiently stationary such that during 'speech and noise' peri-

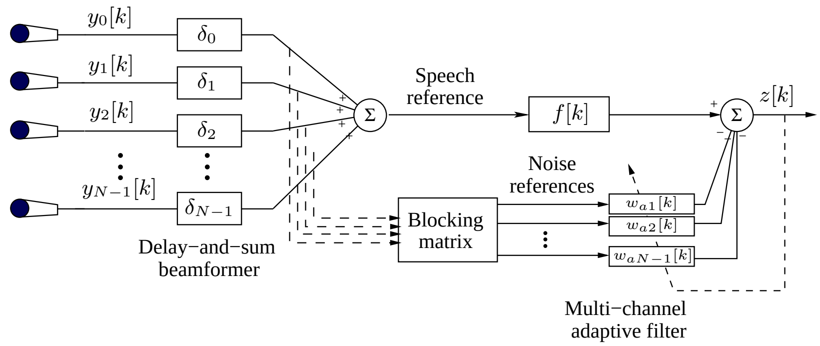

Figure 1.4 — Generalized sidelobe Canceller structure (GSC)

ods the optimal solution found during the last 'noise only' period is still valid. This approach, applied on a two microphone configuration, has proven its effectiveness for hearing aid and cochlear implant applications (VandenBerghe and Wouters, 1998; Wouters and VandenBerghe, 2001; Maj et al., 2006; Spriet et al., 2007).

Multichannel Wiener Filtering

A multichannel Wiener filter (MWF) is based on a statistical approach and provides a minimum mean square error estimate of a given reference signal or, when dealing with hearing aids, of the speech component received at a microphone input. In contrast with beamforming techniques, an MWF can take spatial and spectral differences between the speech and the noise component into account. Hence some speech distortion might be introduced by the algorithm.

Since the MWF, in contrast with a GSC or other array processing techniques, does not require any a priori information about the desired signal or the microphone characteristics (neither spatial nor spectral information), it is appealing for hearing aid applications. Physical evaluations of the MWF have shown promising results. In Doclo and Moonen (2002) and Spriet et al. (2001) it was even shown that the MWF can outperform the GSC strategy in adverse listening situations. However, due to the high complexity of the MWF, little effort has been done to test its applicability for hearing aids. Recently, this has changed due to low-cost subband and stochastic gradient implementations of the MWF making it feasible for commercial hearing aids (Spriet et al., 2004, 2005).

Since the MWF has some very interesting properties regarding the preservation of binaural cues, a binaural MWF complemented with two extensions, is developed and evaluated in this research project. An overview of the binaural MWF techniques and its binaural properties are presented in chapter 4.

Computational auditory scene analysis - CASA

Computational auditory scene analysis (CASA) is the study of auditory scene analysis by computational means. In essence, CASA systems are "machine listening" systems that aim to separate mixtures of sound sources in the same way that human listeners do. CASA differs from the field of blind signal separation in that it is, at least to some extent, based on the mechanisms of the human auditory system. Effort has been done to integrate some of these mechanisms into hearing aid applications, e.g. the work of Wittkop and Hohmann (2003). However, these systems are typically quite complex. At the moment no evidence is found that these systems can outperform the beamforming or MWF strategies taking into account the robustness and complexity issues present in hearing aid applications.

1.3.3 Binaural noise reduction

Changing from a bilateral to a binaural noise reduction algorithm, i.e. generating an output signal for both ears using all available microphone signals, may enhance the amount of noise reduction and may additionally increase the capability to preserve the binaural cues between the left and the right hearing aid. An important limitation of most noise reduction array systems studied thus far is that they are designed to produce a single, i.e. a monaural, output. Extending these algorithms to a binaural output is not always trivial. In section 1.1 it was stated that the amount of research on noise reduction algorithms for binaural applications is limited. However, some attempts have been made to combine noise reduction with the preservation of binaural cues.

The first class of techniques is based on CASA. Wittkop and Hohmann (2003) proposed a method in which the input signal is split into different frequency bands. By comparing the estimated binaural properties (e.g. coherence) of each frequency band with the expected properties of the signal component (typically it is assumed that the signal component arrives from the frontal area with ITD and ILD values close to 0μs and 0dB), these frequencies are either enhanced or attenuated. By applying identical gains to the left and the right hearing aid, binaural cues are preserved. However, the noise reduction performance of these methods is limited and typically, spectral enhancement artifacts such as 'musical noise' occur.

The second class of techniques is based on fixed or adaptive beamforming. In the studies of Desloge et al. (1997), Welker et al. (1997) and Zurek and Greenberg (2000), fixed and adaptive multi-microphone beamforming systems were studied. These were designed to optimize their directional response and to faithfully preserve the binaural cues. In Desloge et al. (1997), six different fixed beamforming systems were tested and compared with a reference system which consisted of two independent cardioid microphones. Two of these systems used all microphone inputs from both hearing aids to calculate the output. The first system was designed to limit the amount of ITD distortion at the output to $40\mu s$. The second system used a low/high pass filtering system and applied non-adaptive noise reduction on the higher frequencies (f>800Hz) of the signal. The low frequency component (f<800Hz) remained unprocessed. This approach is inspired by the observation that the ITD information, present at the lower frequencies, is a dominant localization cue compared to the ILD information, present at the higher frequencies (Wightman and Kistler, 1992). Both systems showed a significant SNR gain of 2.7 to 4.4dB in comparison with the reference system. Tests were performed in a diffuse noise source scenario with speech arriving from the front. In general, both systems provided the subjects with moderate localization capabilities when evaluating localization performance using a horizontal resolution of 30°. In Welker et al. (1997), the previously mentioned low/high pass scheme was employed in an adaptive noise reduction algorithm using two microphones, i.e. one at each ear. The high frequency part ($f > f_c$) of the signal was now processed in an adaptive way. It was clearly shown that f_c determined a trade-off between noise reduction and localization performance. When testing normal hearing subjects with $f_c = 500Hz$ a noise reduction performance of 3dB was obtained together with a localization accuracy of 70 percent. Tests of Zurek and Greenberg (2000), using hearing impaired subjects and $f_c = 1000Hz$, showed a small improvement in SNR of about 2dB.

The third class of techniques is based on blind source separation (BSS). Very recently, Aichner et al. (2007) proposed two methods to incorporate binaural cue preservation in BSS. The first method is based on using adaptive filters as a postprocessing stage after BSS. These filters remove the noise components, estimated by the BSS, from the reference microphone. By doing this at both sides of the head, the binaural cues of the speech component are preserved. Due to the fact that not all noise can be removed from the reference signal, it was claimed that the binaural cues of the remaining noise component are also preserved. The second method is based on constraining the BSS filters themselves, thereby avoiding distortion of the separated signals produced by the BSS. However, localization results were described very briefly using a quality rating on the output of the algorithm and so far no results have been published on the source separation performance of these methods.

A final class of systems, on which chapters 4 to 6 will focuss, is based on the multichannel Wiener filter (MWF). In general, the goal of a Wiener filter is to filter out noise corrupting a desired speech signal. By using the second-order statistical properties of the desired signal and the noise, the optimal Wiener filter can be calculated. It generates an output signal that optimally

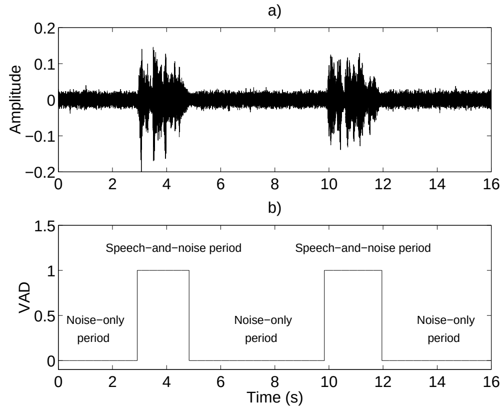

Figure 1.5 — a) Speech and noise signal. b) Ideal VAD output

approaches the desired signal in a mean-square-error (MSE) sense. The main advantages of using MWF-based strategies, as already mentioned in section 1.3.2, are that, in contrast with common beamforming strategies, the MWF does not require any a priori information of the location of the desired signal nor of the microphone characteristics. In Doclo and Moonen (2002) and Spriet et al. (2005) it was shown that the MWF can be used for monaural hearing aid applications. Chapters 4 to 6 presents three new binaural algorithms based on the MWF which aim at performing noise reduction while preserving the binaural cues.

1.3.4 Voice activity detector

Adaptive algorithms are often based on the assumption that the adaptation of the algorithm, which maximizes the amount of noise reduction at a certain time instant, can be done during periods in which only noise is present. If speech is present no adaptation is done to avoid distortions of the speech component. Hence, the noise is assumed to be sufficiently stationary such that the optimal solution during speech and noise periods is similar as during the 'noise only' periods. The classification of 'speech and noise' and 'noise only' periods is done by a voice activity detector (VAD) and is illustrated in Figure 1.5.

Different methods of VAD designs have been studied in the past. VADs are typically based on physical differences between a target speech signal and jammer signals. The main VAD designs are based on estimating features such as the zero-crossing rate, the periodicity, the energy or the inter-microphone correlation of the signals. Although very complex VADs exist, it is commonly known that very good results are only obtained at high SNRs and when using stationary noise conditions. Therefore, studying the influence of VAD errors on the performance and behavior of noise reduction algorithms is an important aspect of the evaluation of hearing aid algorithms (Maj et al., 2002). A thorough overview of the most common VAD designs can be found in Doclo et al. (2002).

This manuscript introduces three new noise reduction algorithms. Since the main focus lies on demonstrating and evaluating the potential of the different MWF algorithms in terms of combining noise reduction with the preservation of binaural cues, a perfect VAD algorithm has been used throughout the manuscript. The effect of VAD errors on these algorithms has only been briefly studied and is therefore not reported.

1.4 Localization of sound sources

The human auditory system is capable of localizing sound sources in a three dimensional environment. It does so by combining several audiovisual information streams and by integrating them over time and frequency. How these cues are integrated is still part of ongoing research. Although a lot of progress has been made, a fixed hierarchy of rules has not yet been found.

The main localization cues, responsible for localizing sound sources in the horizontal hemisphere, i.e. localizing azimuth, are concealed within the differences between the auditory signals received at the eardrums of the left and the right ear, i.e. the binaural cues a.k.a. the interaural cues. This information can be fully described by an interaural transfer function (ITF) which represents the difference in acoustical pathway of a sound source arriving at both ears.

Although humans are most sensitive for frequencies between 1 and 4kHz, their localization performance is best for frequencies below 1.5kHz and above 4kHz (Stevens and Newman, 1936). More than a century ago, Rayleigh (1907) proposed the duplex theory of binaural hearing stating that two different binaural mechanisms existed for localizing azimuth. The first mechanism is based on interaural time differences (ITD) and operates mainly at low frequencies, i.e. $f < 1.5kHz$. The other mechanism is based on interaural level differences (ILD) which are mainly present at high frequencies, i.e. $f > 3kHz$. This dichotomy arises simply due to the physical characteristics of binaural sound signals as will be described in the following subsections.

1.4.1 Interaural time information: ITD

Interaural time differences (ITDs) are defined as the difference in time of arrival of a sound signal arriving at both ears. For a human head, these time differences are within the range of -700μs to +700μs (Kuhn, 1977; Wightman and Kistler, 1992) with a time difference of 0μs when the sound source is positioned at 0°, i.e. in front of the listener. The sensitivity of the human auditory system to changes in ITD, or in other words its ITD resolution, is in the order of 10μs with the maximum resolution found in the area around $0°$.

While differences in the onset time could be useful for localizing sound sources, for ongoing sounds the human auditory system has to rely on differences in phase. These are called the interaural phase differences (IPDs) from which ITDs can be derived. By comparing the phase of the signal present at the left and the right ear, the human auditory system can determine the location of the sound source. However, if the wavelength of the sound signal becomes smaller than the diameter of the head, i.e. around $f = 1.5kHz$, then the IPD information becomes useless since the sound wave may have shifted more than a period from its counterpart at the other ear. Hence, when determining ITD information, the human auditory system ignores frequencies higher than $f = 1 - 1.5kHz$.

In contrast with the original duplex theory of Rayleigh (1907), research showed that stimuli containing only high frequencies could also generate useful ITD information for the auditory system. In these scenarios, ITD information can be derived from the low frequency envelope of the signal. However, these cues are typically much less dominant than the ITD and ILD cues described in the duplex theory (Henning, 1974; Bernstein and Trahiotis, 1985; Dreyer and Delgutte, 2006).

1.4.2 Interaural level information: ILD

Interaural level differences (ILDs) are defined as the difference in level between the signals present at the left and the right eardrum. When a sound source is positioned at the side of the head, the head will produce an acoustical shadow at the contralateral ear leading to a difference in amplitude between both ears. Depending on the sound source angle and the frequency content of the signal, these ILDs can reach, in absolute values, up to 20 à 25dB . Hence, by comparing the level of the signals at the left and the right ear the human auditory system can determine the location of the sound source. The sensitivity to changes in ILD or the ILD resolution is in the order of 0.5dB (Hartmann, 1999; Moore, 1997*a*).

Since the head does not produce a large acoustical shadow for frequencies smaller than f = 2 à 3kHz, the ILD processing by the auditory system is dominated by the ILD cues at higher frequencies. However, it is important to note that if ILDs are introduced at low frequencies, e.g. by a hearing aid algorithm, they will introduce a spatial percept. This is in contrast with ITD information, which, if introduced in the high frequency region, is ignored by the auditory system and does not introduce a spatial percept (Hartmann, 1999).

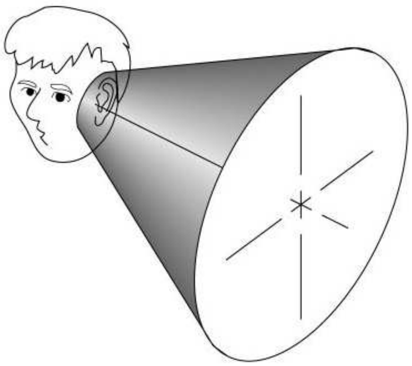

Figure 1.6 — Sound sources placed on the cone of confusion generate identical ITD and ILD cues.

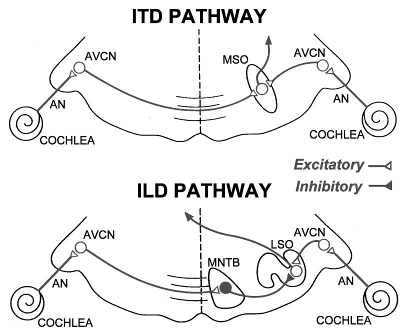

Figure 1.7 — Schematic (simplified) of the ITD and ILD pathway in a cat brain (Joris et al., 1990).

1.4.3 Spectral and other cues

The duplex theory might resolve the problem of localizing the azimuth of a sound source, however ITDs and ILDs do not contain any information concerning the elevation of the sound source. Moreover, no front back discrimination can be made when using ILD and ITD information since these cues are identical for sound sources placed at the front or the back hemisphere. The surface on which ILD and ITD information is identical (if one uses a sphere to model the head) is described by a cone, called the cone of confusion. The top of this cone is positioned at the center of the head, see Figure 1.6. Therefore, other auditory cues are used to resolve these issues.

One of these cues are the pinna cues or spectral cues (Hebrank and Wright, 1974; Musicant and Butler, 1984; Middlebrooks et al., 1989; Langendijk and Bronkhorst, 2002). Each human listener possesses two very strangely shaped, highly personalized pinnae. This shape reflects and diffracts sounds dependent on the angle of arrival of the signal. Hence, the spectrum of the sound arriving at the eardrum will contain a complex set of directional dependent features, i.e. the spectral cues, which allow to resolve front-back issues and determine the elevation of the sound source. These cues are called monaural cues since they can be used by each ear individually and no comparison between the sounds at both ears have to be made to deduce the relevant information. Since the cavities present in the pinnae are relatively small, pinnae effects are only present for frequencies in the range of 6kHz and higher (Hartmann, 1999).

Besides the spectral cues, the human auditory system also uses head movements (Fisher and Freedman, 1968; Mackensen, 2004) to determine the location of a sound source. The uncertainty created by the cone of confusion can be removed by turning the head and integrating ITD and ILD information over time.

An other cue that may be used when perceiving and localizing sound sources are visual cues. How and where a sound source is perceived can be heavily influenced by the visual information present at that time. The most notorious example being the ventriloquist effect (Bertelson, 1999). The McGurk effect on the other hand, demonstrates the influence of visual cues on the perception of speech (McGurck and Macdonald, 1976). Therefore visual information has to be taken into account when interpreting and comparing results of localization experiments (Perrett and Noble, 1995).

1.4.4 Anatomy and physiology

Anatomically, it has been observed that ITD and ILD cues follow a different path through the human auditory system. Figure 1.7 shows a highly simplified diagram of the neural circuits (in a cat brain) that thought to be involved in ITD and ILD processing. ITD cues are being processed in the medial superior olive (MSO) (Joris and Yin, 2007) which receives information from both the left and the right anteroventral cochleus nuclei (ACVN). ILD information is processed in the lateral superior olive (LSO) (Tollin, 2003) which receives information from the contralateral ACVN through the medial nucleus of the trapezoid body (MNTB). From the LSO and MSO, information is sent through to the inferrior colliculus (IC) which projects the information to the ipsi- and contralateral medial geniculate body, which projects on the auditory cortex. Research on how ITDs and ILDs are translated into neural activity and integrated is still in progress. Nonetheless a very short introduction is given in the following paragraphs.

The most important model regarding ITD cues was published in 1948 in the work of Jeffres (1948). In this work an answer was formulated on how the MSO structure could deduct ITD information from the inputs from the contra- and ipsilateral ear. The model consists of three important ingredients: phase-locking, delay lines and coincidence detectors. Phase locking is a well known phenomenon present in the cochlea which represents the fact that action potentials are fired on certain time instants within a stimulus period. Delay lines inside the MSO cell delay these pulses with a delay dependent on the length of the delay line. Coincidence detectors make sure that the EE (excitation-excitation) cell, present in the MSO becomes active if the action potential arriving from the left and right side arrive simultaneously. In other words, depending on the ITD or on the angle of arrival of the sound source a different EE cell will become active. Hence, ITD cues are translated into neural activity. One should be aware that these cells are specifically tuned for one particular frequency. The integration over frequency is done later. It is remarkable how well this system matches the mathematical cross correlation operation which is often used to estimate time delays between microphone signals. Evidence has been found in the brain of mammals for all of these three components (Joris et al., 1998). However, the Jeffres model remains a topic of discussion among researchers (Joris and Yin, 1995; Joris et al., 1998; Fitzgerald, 2002).

The processing of ILDs seems a bit more straightforward. The MNTB cells which deliver the information of the contralateral ear to the ipsilateral ear are inhibitory cells which reverse the polarity of the information of the contralateral ear. The information of the ipsilateral ear is received as an excitational signal. Therefore the magnitude of responses received by the LSO cells, mostly tuned on higher frequencies, are proportional to the ILD (Boudreau and Tsuchitani, 1968). For an extensive discussion on the LSO see Joris and Yin (1995).

Each of the different cues are typically processed in separate brainstem nuclei. However, this does not withhold the auditory system to obtain a single spatial percept. How the different cues are combined and weighted is still subject of ongoing research. It is generally assumed that the ITD information, present in the lower frequency region, is the most dominant factor for horizontal sound localization (Wightman and Kistler, 1992). However, this remains part of ongoing discussions.

1.4.5 Localization experiments

In this manuscript, research is presented on the effect of noise suppression algorithms on the binaural cues and, related to this, spatial awareness and speech recognition. Different experiments have been used to evaluate the spatial awareness induced by binaural cues. Three binaural experiments are often reported in literature: a lateralization experiment, a localization experiment and evaluating minimal audible difference, e.g. the minimal audible angle (MAA). For an extensive overview see Blauert (1997).

Headphone experiments are a popular research tool. They improve reproducibility of the experiment and allow to control ITD, ILD and other cues in an independent way. Therefore, even unnatural combinations of binaural cues can be presented to the subject. On the other hand, auditory events presented through headphones are often perceived as arriving from somewhere in the head with no externalization taking place due to an incomplete restoration of the full head related transfer functions (HRTFs) (e.g. due to the lack of spectral cues) or due to the absence of other realistic cues (e.g. visual cues). Therefore, the subject's task is usually to describe the lateral displacement of the auditory event. This is done on a line connecting the ear drums, i.e. the axis of the ears. The relationship between the attributes of the ear input signal and its lateral displacement is called lateralization. On the basis of such relationship, various hypothesis may be formed about the processes that generate this lateral displacement. There is some indication that some of these hypotheses may be generalized to spatial hearing in free field listening experiment but caution is advised. To which degree a lateralization corresponds to the perception of a sound source at a certain angle around the head is often unknown.

A localization experiment is commonly done by placing loudspeakers around the head. The subject has to report where a sound is heard. To investigate the effect of binaural cues the test should be balanced and level roving should be used. A balanced test is a test which has the same number of presentations (trials) arriving from the left and right side of the head. Level roving is a randomized attenuation or amplification of the stimuli presented to the subject. If the same stimulus is repeated over and over again without level roving the test subject will learn the level of the presented sound stimulus. Since the head casts an acoustical shadow over the contralateral ear which is dependent on the angle of arrival of the signal (see ILD), the perceived loudness of the signal could be used as a monaural localization cue. Since in reality the loudness of a sound source is often unknown, a test without level roving will not represent a real life situation. If one is interested in binaural cues head movements should be avoided since these offer extra localization cues (Van Wanrooij and Van Opstal, 2004).

One may also perform a localization experiment by using headphones. This can be done by convolving the stimuli with HRTFs of the subject measured at the angles of interest. Often a set of HRTFs measured on a artificial head, representing an average human head and torso, is used. However one should be aware that by using these impersonalized HRTFs, localization performance will typically drop (Møller et al., 1996, 1999; Minnaar et al., 2001). Important to note is that a localization task is a subject dependent process. It has even been proven that some people possess a set of HRTFs more appropriate for sound localization than others (Møller et al., 1999). Therefore the conclusions of this manuscript are mainly based on intra-subject evaluations.

Measuring the minimal audible angle (MAA) is commonly done by displacing a loudspeaker in location. The subject has to report when the displacement is heard. This displacement equals the MAA. This could be interpreted as measuring the resolution of the binaural hearing system. The same experiment can be performed under headphones, measuring the minimal audible differences in ITD or ILD. The results of these experiments point out that the human auditory system has a ITD resolution of up to $10\mu s$ and an ILD resolution of 0.5dB. (see section 1.4.1 and 1.4.2).

1.5 Spatial release from masking

Binaural cues are crucial for the correct localization of sound sources but also improve speech perception in noisy environments. Differences in binaural cues, due to the spatial separation of the competing sound sources, are known to improve the detection and recognition of sound signals. This is called spatial release from masking. The improvement in detecting sound signals is often referred to as the binaural masking level difference (BMLD). In normal hearing adults, BMLDs up to 19dB have been detected, depending on the type of stimuli (Webster, 1951; Bernstein and Trahiotis, 1992; Van Deun et al., 2007). The improved speech perception is often referred to as the binaural intelligibility level difference (BILD) (Bronkhorst and Plomp, 1988, 1989; Peissig and Kollmeier, 1997; Drennan et al., 2003). An overview of both of these extensively studied phenomena is given in Gelfand (1998).

In the work of Bronkhorst and Plomp (1988), BILDs were measured for normal hearing subjects. A free field condition was simulated by presenting recordings, made with a KEMAR manikin in an anechoic room, through earphones. Recordings were made with the speech signal arriving from the front of the subject and a noise signal arriving from seven angles in the azimuthal plane, ranging from 0° to 180° in steps of 30°. The main results of their measurements are given in Figure 1.8. During the experiment, a maximum gain in speech intelligibility, due to the spatial separation of the speech and the noise signal, of approximately 10dB was obtained. This can be observed by comparing the free field (FF) data in Figure 1.8 of the condition with noise at 0° with the condition with noise at 90°. The main portion of this gain can be attributed to a simple acoustic effect. The spatial separation of the target and masker generally increases the SNR at one of the two ears, which improves speech perception. This is a pure monaural effect. However, binaural spatial processing also provides an important additional improvement in performance. This effect is known as binaural unmasking or the binaural squelch effect. To separate both effects, Bronkhorst and Plomp (1988) isolated the ITD and ILD cues. The gain in speech intelligibility due to ILD is expected to depend mainly on monaural, 'best ear', effects whereas the unmasking due to ITD represents purely binaural interaction. The introduction of ILD information resulted in a maximum gain of 7.8dB. The gain due to ITD, representing binaural unmasking, reached a maximum value of 5.1dB. Other studies confirm these results and show that, on average, an improvement in SRT of 2 to 3dB can be expected purely due to the binaural processing of the human auditory system (Zurek, 1993). Since the intelligibility of sentence material improves by about 15% for each decibel of increase in SNR, these improvements can lead to a large benefit in speech intelligibility.

If bilateral noise reduction systems distort the binaural information, this may result in a sub-optimal speech intelligibility due to a decreased spatial release from masking.

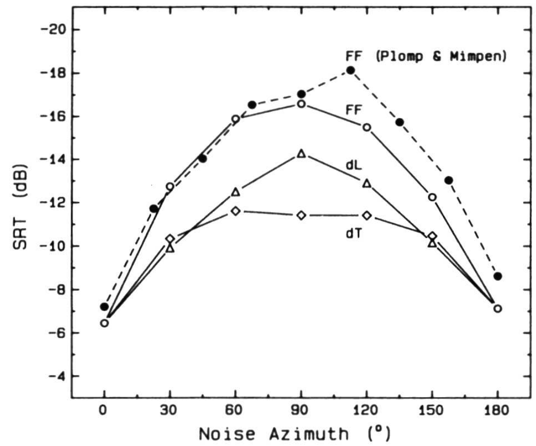

Figure 1.8 — Mean SRTs of normal hearing subjects for three different noise types: FF (free field), dL (ILD only), dT (ITD only) (Bronkhorst and Plomp, 1988). The closed data points represent results of Plomp and Mimpen (1981) obtained in free field. A maximum gain, due to the spatial separation of speech and noise source, of approximately 10dB for condition FF, 8dB for condition dL and 5dB for condition dT were observed.

1.6 Microphone signals and the acoustic environment

In hearing aid applications, the signals recorded at the microphones have been subjected to the room acoustics and the microphone characteristics present in that particular setup. Noise reduction algorithms are typically used to remove the noise component from the input signal. An adaptive echo canceller or speech de-reverberation algorithm may be used to remove the acoustic impulse response from the speech component of the signal. Speech understanding and the performance of noise reduction algorithms is highly dependent on the acoustic environment (e.g. see chapter 5). Therefore, it is important to define the acoustic environment in which the tests have been performed.

1.6.1 Acoustic environment

An acoustic impulse response between a loudspeaker and a microphone incorporates the reverberation of the acoustic environment. This can be defined by three main aspects: the direct sound, the early and the late reflections (see Figure 1.9). The direct sound is the sound which propagates through the shortest path from loudspeaker to microphone. The early reflections are the sounds which have been reflected by one or more surfaces. If these reflections are

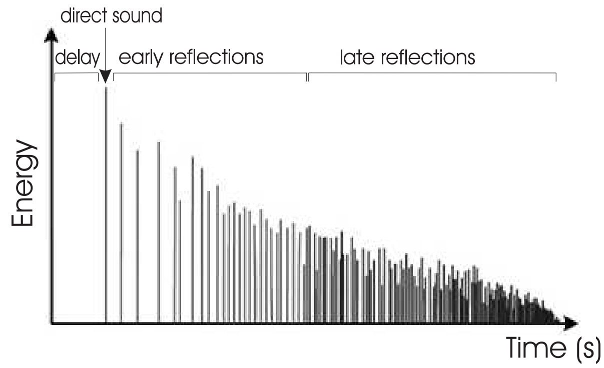

Figure 1.9 — Reverberation of a room: energy arriving at a microphone, distributed over time. Taken from Maj (2004).

delayed by more than 30ms they are perceived as echoes. The late reflections have been reflected many times and are therefore arriving from all directions to the microphone. Since sounds are absorbed at every reflection, they gradually fade away.

Two measures are often used to characterize the environment: the reverberation time, known as T_{60} and the critical distance. T_{60} is the amount of time needed for a sound, when switched off, to decrease 60dB in energy. The critical distance is the distance at which the level of reverberant sound equals the level of direct sound. Within the critical distance, the direct sound is the dominant sound component.

Since reflections and absorptions are a frequency dependent phenomenon, the room characteristic will also be frequency dependent. Therefore a reverberation time or critical distance should be calculated per frequency band. However, to facilitate the description of a room, an averaged T_{60} is often used. This averaging can be done in several ways, e.g. linearly or according to speech energy weighting criteria, depending on the signal or task of interest.

1.6.2 Microphone characteristics

A microphone is an acoustic to electric transducer or sensor that converts sound into an electrical signal. Acoustical signals picked up by a microphone are subject to its specific phase and amplitude characteristic. A microphone, like any type of sensor, will introduce a noise component due to thermal noise, shot noise and flicker noise. Internal noise can severely limit the performance of noise reduction systems. Moreover, the internal noise component may even be amplified due to the noise reduction algorithm.

When developing noise reduction algorithms it is often assumed that the characteristics of all microphones are identical in terms of gain, phase and spatial properties. However, this is not the case for real sensors. Different phase and gain characteristics may even have a large impact on noise reduction performance, especially for beamforming strategies. To overcome this issue, calibration procedures exists. This is a time consuming task since it has to be done for every sensor or every hearing aid. As an alternative hearing aid companies often buy more expensive, matched microphone pairs. These pairs are guaranteed to have a limited amount of mismatch. Nevertheless, two microphones will never have identical characteristics which will affect noise reduction performance and which, as will be shown in chapter 2, may have a large influence on ITD and ILD cues.

1.7 Outline of the thesis

1.7.1 Main research objectives

The main research objectives of this thesis are:

1. Evaluate and quantify the impact of currently used, independently operating, bilateral noise reduction systems on the binaural cues. Evaluate the perceptual consequences for the hearing aid user.
2. Design new noise reduction algorithms based on a binaural hearing aid configuration. In these designs, a link between the left and the right hearing aid is assumed. The goal of these algorithms is to improve speech intelligibility by performing noise reduction while preserving the binaural cues. The new algorithms should offer an improved combination of noise reduction performance with binaural cue preservation compared to the commercially used noise reduction systems.

These objectives are met by using theoretical (chapters 2 and 4), objective (chapters 2 and 5) and perceptual (chapters 2, 3, 5 and 6) evaluations and performance measures. Perceptual evaluations include speech perception in noise and localization in the frontal horizontal hemisphere experiments, performed by hearing aid users and normal hearing subjects. An adaptive directional microphone (ADM) is taken as a reference throughout the document since it is the most widespread noise reduction system implemented in hearing aids.

1.7.2 Chapter by chapter overview

In this section a chapter by chapter overview of the thesis is given. It represents an overview of the experiments and derivations done to achieve the research objectives summarized in the previous section.

In **chapter 2**, the two most commonly used noise reduction algorithms in hearing aids, a bilateral fixed directional microphone (FDM) and a bilateral adaptive directional microphone (ADM) are evaluated in terms of preserving binaural cues and spatial awareness. First, a general overview of the literature on localization with hearing aids is given. Secondly, the impact of a bilateral ADM and FDM is analyzed using theoretical and objective evaluations. Finally, the impact of commercially available hearing aids and a bilateral ADM on binaural cues are studied by using a localization experiment in the frontal horizontal hemisphere. These experiments are performed by hearing aid users on a newly developed test platform. A group of normal hearing subjects are evaluated as a reference condition.
The main findings of this chapter are published in Van den Bogaert et al. (2005), Van den Bogaert et al. (2006).

To evaluate new noise reduction algorithms it is often preferred to use headphone experiments which facilitate the development process. However, this is not straightforward for localization experiments. **Chapter 3** presents the validation of an experimental procedure which examines the localization performance of normal hearing subjects by using headphone experiments and recordings made with an artificial head. In the frame of a joint research project by the groups ExpORL, SISTA-SCD and Acoustics and thermal physics this research question was extended with the use of virtual acoustics to perform localization experiments. First a short overview is given on virtual acoustics. Afterwards the methodology to evaluate localization performance, which will be used in the proceeding chapters, is validated.
A publication presenting the main findings of this chapter is in preparation.

Chapter 4 changes the focus of the manuscript from bilateral to binaural noise reduction algorithms. It presents an overview of the research done on binaural noise reduction algorithms and it presents the development of three new noise reduction algorithms for binaural hearing aids. These algorithms are extensions of the monaural MWF algorithm first applied in hearing aids by Doclo and Moonen (2002) and Spriet et al. (2005). The following algorithms were developed: the binaural MWF, the binaural MWF with interaural transfer function extension (MWF-ITF) and the binaural MWF with partial noise estimation (MWF-N). These algorithms are described within a binaural framework which is first mathematically defined. A theoretical evaluation analyzes the noise reduction performance and the impact of the binaural MWF on binaural cues. To improve the preservation of the binaural cues of the noise component, two extensions are proposed: the MWF-N and the MWF-ITF. The binaural MWF and MWF-N are both further evaluated by using objective and perceptual performance measures in chapters 5 and 6. The MWF-ITF is an algorithm which is currently still under development. This chapter includes pilot experiments with the MWF-ITF illustrating the potential of this algorithm.
A full binaural algorithm requires the transmission of all contralateral micro-

phone signals to the ipsilateral hearing aid which comes at a large cost of bandwidth and power consumption. Therefore a number of reduced bandwidth algorithms are presented. These algorithms aim at maximizing the noise reduction performance while sending over a reduced number of signals. This is done by first combining the microphone inputs at the contralateral hearing aid before transmitting the signals to the ipsilateral hearing aid.
The main publications related to this chapter are Doclo et al. (2005), Doclo et al. (2006), Klasen et al. (2006), Klasen et al. (2007), Van den Bogaert et al. (2007), Van den Bogaert et al. (2008*a*), Doclo et al. (2008).

Chapter 5 focusses on speech intelligibility and speech in noise enhancement. It presents an evaluation of speech intelligibility and noise reduction performance when using the MWF and MWF-N algorithm. This is done by using objective and perceptual performance measures. A bilateral ADM algorithm and an unprocessed condition are used as reference conditions. Using contralateral microphone signals comes at the large cost of transmitting these signals to the ipsilateral hearing aid. Therefore, different microphone combinations are used to evaluate the impact of adding no, one or two contralateral microphone signals to the ipsilateral hearing aid. This is done for different spatial scenarios in different acoustic environments, ranging from a quasi anechoic environment to a realistic reverberant environment. The correlation between the objective and perceptual performance measures is also evaluated.
The main publication related to this chapter is Van den Bogaert et al. (2008*b*).

Chapter 6 evaluates the localization performance in the frontal horizontal hemisphere when using the MWF and MWF-N algorithm. This is done by performing a perceptual evaluation in a realistic reverberant environment. Both algorithms are compared with a bilateral ADM algorithm. An unprocessed condition is used as a reference condition. Two different experiments are performed. First, the speech and the noise component, which have been processed by the different algorithms, are presented separately to the subjects. By presenting both components separately, interactions between components are avoided (masking effects, localizing two sound sources is different from localizing one sound source, etc.). This provides the best insight in the behavior and the impact of the different algorithms. In the second condition, the speech and the noise component are presented simultaneously and the subject is asked to localize both components. This resembles more a steady-state real-life situation.
The main publication related to this chapter is Van den Bogaert et al. (2008*a*).

Finally, **chapter 7** presents the overall conclusions of the manuscript as well as some suggestions for further research.

Chapter 2

The effect of current bilateral hearing aid technology on binaural cues

In this chapter, current hearing aid technology is evaluated with respect to the preservation of binaural cues and spatial awareness. Objective and perceptual evaluations were performed to measure the influence of commonly used noise reduction schemes on the preservation of binaural cues and localization performance. In a localization experiment the ability of hearing impaired subjects, not wearing their hearing aids, to localize sound sources in the frontal horizontal hemisphere was evaluated. Since the main focus of this manuscript is on binaural cue processing, localization experiments were restricted to the frontal horizontal hemisphere. By doing so, front-back confusions, which are related to the processing of monaural spectral cues (see section 1.4.3), were avoided which could complicate the analysis of the data. Subsequently, these subjects were evaluated when using their hearing aids in order to compare and to quantify the influence of these systems on their localization performance. The hearing aids were evaluated by using a bilateral omnidirectional configuration (using no noise reduction) and by using a bilateral adaptive directional microphone (ADM) configuration. Tests with normal hearing subjects were carried out as a reference condition.

Section 2.1 defines the problem and places the performed research within a more general framework of studies done by other researchers.

Section 2.2 analyzes the effect of a fixed directional microphone (FDM) and an ADM on the binaural cues, more specifically ITDs and ILDs. This is done by using theoretical derivations and objective performance measures.

Section 2.3 presents a perceptual evaluation which quantified the impact of hearing aids in general and more specifically a bilateral ADM on a binaural task, namely a localization experiment in the frontal horizontal hemisphere. This experiment was carried out by hearing impaired subjects with and without their own hearing aids. Hearing aids were tested using an omnidirectional and a bilateral ADM configuration. A group of normal hearing subjects were evaluated as a reference.

Section 2.4 combines the results and conclusions of the objective and perceptual evaluations. Answers are formulated on the question whether hearing impaired subjects, with and without hearing aids, are still capable of using binaural cues and whether two of the most commonly used noise reduction algorithms in hearing aids affect the binaural cues.

This work has been published in Van den Bogaert et al. (2005) and Van den Bogaert et al. (2006).

2.1 Introduction

The available processing power in hearing aids increases as technology evolves. One of the main benefits is that more complex noise reduction algorithms can be implemented, improving speech understanding performance in acoustically challenging scenarios. In recent years, good results have been obtained using adaptive filtering techniques. These techniques adapt according to changes in noise scenario or acoustic condition, but are typically designed and evaluated monaurally. An important question is whether using these techniques bilaterally can have an impact on binaural processes such as horizontal localization.

The main mechanisms used for sound localization are fairly well known. Localization in the horizontal field involves binaural processing of very small differences in time (0-700μs), intensity (0-20dB) and spectrum between the two ears (Stevens and Newman, 1936; Gilkey and Anderson, 1997; Hartmann, 1999; Langendijk and Bronkhorst, 2002). Extensive psychoacoustical research has been done on localization: experiments to measure localization performance of normal hearing (Makous and Middlebrooks, 1990; Hofman and Van Opstal, 1998; Lorenzi et al., 1999*b*) and hearing impaired subjects (Hausler et al., 1983; Lorenzi et al., 1999*a*; Noble et al., 1994) with different stimuli and in different test conditions, experiments under headphones with isolated or conflicting binaural cues (Wightman and Kistler, 1992; Lorenzi et al., 1999*b*), comparing performance of a monaural hearing aid or cochlear implant configuration with a bilateral hearing aid or cochlear implant configuration (Van Hoesel and Tyler, 2003), and many others. Although a lot of work has been done on localization with normal hearing and hearing impaired subjects, little work has questioned the effect of a bilateral hearing aid system on the binaural potential of the hearing aid user. In the human auditory system, the acoustical input signals of both ears are linked to the binaural centers where the binaural cues are interpreted and processed. Adding independently working hearing aids, each using its own compression scheme, introducing its own time delay (in the order of 5 to 10ms) (Dillon et al., 2003) and having independent noise reduction schemes, could have a destructive effect on the binaural cues. Hence, the hearing aid user's localization performance and speech perception in a complex environment could also be degraded.

In the work of Hausler et al. (1983), the question was raised whether hearing aids could have an impact on sound localization performance. Noble and Byrne (1990) tested localization performance in the frontal horizontal and vertical planes with bilateral behind the ear (BTE), in the ear (ITE) and in the canal (ITC) hearing aids with omnidirectional microphone configurations. A normal hearing group was used as a reference. Intra subject analysis did not show significant differences between unaided and aided performance for all three groups. These analyses were done on an error measure in which both vertical and horizontal errors were included. No statistical analysis on only horizontal or on only vertical localization errors was presented in the study. However, Noble and Byrne state that for the control group, i.e. a group of 6 normal hearing subjects, horizontal localization performance dropped from nearly 100 percent correct unaided to 73 percent correct wearing BTE hearing aids. In the same study, it is mentioned that the hearing aid users tended to show poorer aided than unaided localization performance in the frontal horizontal plane, except for the ITE hearing aid users, when wearing their own hearing aids. The difference in horizontal localization performance was not quantified in the study.

Later, Noble et al. (1998) and Byrne et al. (1998) showed that better performance could be obtained by using open earmolds instead of closed earmolds for subjects with a moderate high-frequency (and a severe low-frequency loss) or a moderate low-frequency (and a severe high-frequency) hearing loss. By using open earmolds, the subject can use the direct soundfield in the region of the moderate hearing loss for localization. For subjects with a moderate high-frequency loss, improvement in the vertical plane was found. For subjects with a moderate low to mid-frequency hearing loss, improvement in the horizontal plane was found and performance was restored to unaided performance.

These studies suggest that bilateral BTE hearing aids do not preserve localization cues. In all three studies a broadband pink noise target stimulus was used and no jammer sources were present.

This chapter studies and quantifies the effect of current signal processing techniques on localization performance. By mathematically analyzing noise reduc-

tion strategies and by using broadband, low and high-frequency stimuli in a localization experiment carried out by hearing aid users, it tries to determine which cues are being affected by different signal processing strategies. The following research questions are addressed in this chapter:
(i) how well do bilateral hearing aid subjects, relative to normal hearing subjects, perform on a localization task using low-frequency, high-frequency and broadband signals and what is the influence of jammer sources on localization performance. This quantifies to which extent hearing impaired subjects are capable of using binaural cues.
(ii) do modern digital hearing aids preserve binaural cues and do they enable hearing aid users to benefit from the full potential of their binaural processing capabilities?
(iii) do noise reduction systems have an influence on horizontal localization performance?

These research questions are answered by combining objective and perceptual evaluations, investigating the possible binaural cue distortions generated by noise reduction algorithms and the influence of commercially available hearing aids on localization performance.

2.2 Theoretical analysis

The two most commonly used multi-microphone noise reduction systems in current commercial digital hearing aids are a FDM and an ADM (Luo et al., 2002; Maj et al., 2004). Both techniques are based on the assumption that the target signal arrives from the frontal field of view and that jammer signals arrive from the back hemisphere. The physical differences in time of arrival between the microphones are then used to improve the SNR by canceling out (steering a null in) the direction of the jammer signals. It has been proven that under this assumption the FDM and ADM combine a low complexity with a relatively good noise reduction performance. This section examines whether a bilateral FDM and ADM can affect the binaural cues.

2.2.1 Fixed directional microphone (FDM)

One of the basic building blocks of directional noise reduction systems in multi-microphone hearing aids is the FDM. This can be used as a stand alone noise reduction system or it can be integrated in more enhanced noise reduction schemes (e.g. an ADM, a two stage general sidelobe canceller (GSC), etc.) (Haykin, 1996; Teutsch and Elko, 2001; Maj et al., 2004).

Directional microphones are typically designed to be most sensitive to sounds arriving from the front and try to cancel (or to null) unwanted sounds arriving

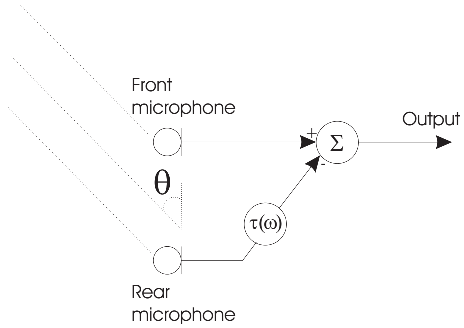

Figure 2.1 — A fixed software directional microphone (FDM). The sound is captured by two omnidirectional microphone ports. Both signals are then combined to create a directional pattern. Different patterns can be created by using a different time delay τ.

from a specific direction. The *hardware directional microphone* has two entry ports. Sounds enter both the front and the rear cavity which allows them to arrive on either side of the microphone diaphragm. If the delayed version (internal delay) of the sound, arriving through the rear port, reaches the diaphragm at the same time as the sound arriving through the front port, a cancellation of the sound occurs.

Based on this technique, *software directional microphones* have been developed, e.g. Figure 2.1. This method uses two omnidirectional microphones as the front and rear entry port. The output of the directional microphone is computed as the difference between the signal from the front microphone and the delayed signal of the rear microphone resulting in a response comparable to a hardware directional microphone. The parameter $e^{-j\omega\tau}$ determines the spatial characteristic of the directional microphone and varies as a function of internal time delay $\tau(\omega)$. The transfer function of a fixed directional microphone with time delay $\tau(\omega)$ and inter-microphone distance d equals

$$H_{FDM}(\omega,\theta) = (A_{front} - A_{back}e^{-j\omega t_0}) \tag{2.1}$$

$$\sim 1 - Ae^{-j\omega t_0} \tag{2.2}$$

with A_{front} and A_{back} representing the amplitude of the sound at the first and the second microphone. A represents the amplitude of the sound at the second microphone relative to the amplitude of the sound at the first microphone and t_0 representing the time delay of the second microphone signal due to the inter-microphone distance d and the fixed internal time delay τ, i.e. $t_0 = \tau + \frac{d}{c}\cos\Theta$. Θ represents the angle of the sound source relative to the hearing aid and c represents the speed of sound. The transfer function of the FDM can be written as:

$$H_{FDM}(\omega,\theta) \sim B_{FDM}(\omega,\theta)e^{j\Phi_{FDM}(\omega,\theta)} \tag{2.3}$$

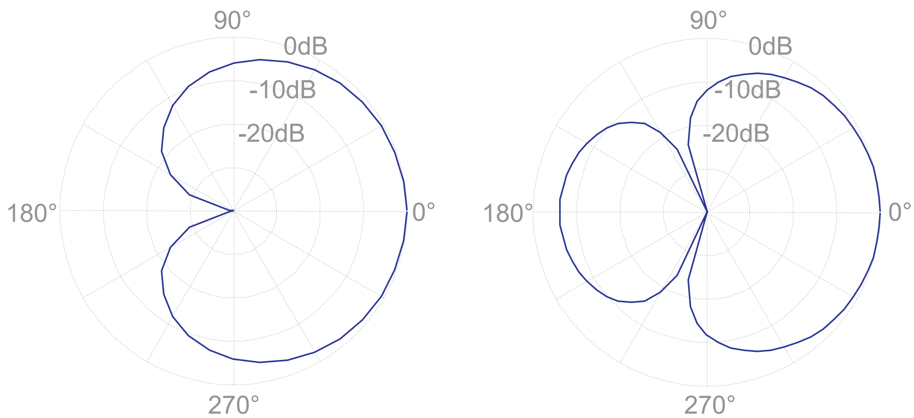

Figure 2.2 — Two different spatial characteristics of a FDM. **Left**: a cardioid polar pattern with a null at 180° generated by $\mathbf{B} = 1$. **Right**: a hypercardioid polar pattern with a null at 110° generated by $\mathbf{B} = 3$.

with

$$B_{FDM}(\omega,\theta) = \sqrt{1 + A^2 - 2A\cos(\omega t_0)} \quad (2.4)$$

$$\Phi_{FDM}(\omega,\theta) = \arctan(\frac{A\sin(\omega t_0)}{1 - A\cos(\omega t_0)}) \quad (2.5)$$

In an ideal system, i.e. if all microphones have identical amplitude characteristics, i.e. $A = 1$, the response of the FDM can be written as:

$$B_{FDM}(\omega,\theta) = \sqrt{2 - 2\cos(\omega t_0)} \quad (2.6)$$

$$\Phi_{FDM}(\omega,\theta) = \frac{\pi}{2} - \frac{\omega t_0}{2} \quad (2.7)$$

The spatial amplitude characteristic, also known as the polar pattern of the microphone, is dependent on $t_0 = \tau + \frac{d}{c}\cos\Theta$ or on the ratio between the external delay, $\frac{d}{c}$, and the internal delay τ. This ratio can be abbreviated by $\mathbf{B} = \frac{d}{\tau c}$. Two commonly used spatial characteristics of an ideal FDM in free field conditions are illustrated in Figure 2.2. A cardioid pattern, $\mathbf{B} = 1$, nulls out the sounds arriving from the back, i.e. 180°. A hypercardioid pattern, $\mathbf{B} = 3$, nulls out sounds arriving from 110°. It can be proven that for a two-microphone configuration $\mathbf{B} = 3$ maximizes the directivity index (DI), i.e. the ratio of the output power for a sound arriving from the front over the output power for a sound arriving from all other directions (see eq. 1.1) (Dillon, 2001*c*). However, the patterns shown in Figure 2.2 are an idealized representation of the performance of an FDM. In realistic conditions, a hypercardioid FDM typically offers a maximum noise reduction gain of around 4dB for single noise source scenarios, dependent on the angle of arrival of the signal, and around 2dB for a diffuse noise source scenario (Leeuw and Dreschler, 1991; Maj et al., 2004).

If the ideal FDM is used in a bilateral hearing aid configuration, i.e. an independent FDM at each side of the head and $A_{left} = A_{right} = A = 1$, then the interaural time information generated by the two hearing aid systems can be calculated by using the phase response, eq. 2.7, for both devices. The interaural time information at the output of the bilateral FDM, $\widehat{\Delta_{ITD}}$, can be calculated as:

$$\widehat{\Delta_{ITD}}=\left(\Delta_{ITD}-\frac{\pi}{2\omega}+\frac{\tau_{right}}{2}+\frac{d_{right}}{2c}\cos\Theta\right)-\left(-\frac{\pi}{2\omega}+\frac{\tau_{left}}{2}+\frac{d_{left}}{2c}\cos\Theta\right) \quad (2.8)$$

with Δ_{ITD} representing the natural ITD occuring between the front microphone inputs of the left and the right hearing aid. When using identical devices with ideally matched microphones, i.e. $d_{left} = d_{right}$, $\tau_{left} = \tau_{right}$ and $A_{left} = A_{right} = A = 1$, the previous equation is reduced to

$$\widehat{\Delta_{ITD}} = \Delta_{ITD} \quad (2.9)$$

which proves that an ideal bilateral FDM configuration does not distort interaural time information. Moreover, since $\tau_{left}=\tau_{right}$ the spatial patterns generated by both hearing aids will be identical and interaural level information will also be preserved.

In a non-ideal system, the parameter A, the ratio between the amplitude of the sound signal at the front and back microphone, will not be equal to 1 due to e.g. reflections or differences in microphone characteristics between the two microphones on each hearing aid. Microphone mismatch is a well known phenomenon and is present due to the manufacturing process, aging and effects of dirt and humidity with the latter factor producing the largest amount of mismatch. Since it is highly unlikely that, in a bilateral hearing aid configuration, these imperfections are equal for the left and right hearing aid, this will result in a different parameter A for both hearing aids (A_{left}, A_{right}) which leads to a different phase (eq. 2.5) and amplitude (eq. 2.4, Figure 2.3) response and therefore a distortion of interaural time and level cues.

The left part of Figure 2.3 shows the ITD distortion, $\Delta_{ITD} - \widehat{\Delta_{ITD}}$, generated by a bilateral FDM with on one side an ideal FDM, $A_{left} = 1$, and on the other side an FDM subject to imperfections. The parameter A_{right} is varied between 0.8 and 1.1 corresponding to a quite modest mismatch ranging from approximately -2dB to +1dB. A 500Hz sinusoid is used as a test signal. Ideally, if $A_{right} = A_{left}$, no ITD distortion is present since the phase responses of both systems are identical. However, when adding a small microphone mismatch, an absolute interaural time distortion is measured ranging from 0 to 500μs dependent on the angle of arrival of the signal and the value of A_{right}.

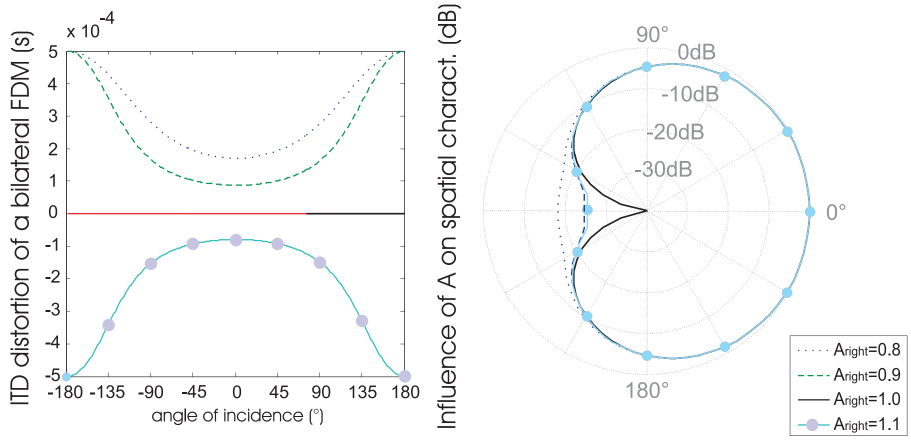

Figure 2.3 — **Left:** Distortion of ITD information by a bilateral FDM configuration. Measurements are made using a 500Hz sinusoid and a cardioid amplitude characteristic for both FDM's. The left hearing aid has a value $A_{left} = 1$, the parameter A of the right hearing aid is varied between 0.8 and 1.1. Absolute interaural time distortions generated by such a system are between 0 and 500μs depending on the angle of arrival of the sound signal and the value of A_{right}.
Right: The influence on the spatial characteristic of an FDM (in dB) with a cardioid amplitude characteristic. A is varied between 0.8 and 1.1. Measurements are made with a 3000Hz sinusoid. Varying A has a large influence on the response of the FDM only around the angle of maximal noise suppression. For both figures, a microphone distance of 2cm was used.

This illustrates that the amount of introduced time distortion is easily within the range used by the auditory system (0-700μs) and well above the minimal audible ITD (10μs). Secondly, this figure illustrates that the ITD distortion reaches a maximum around the angle of maximum noise suppression, i.e. 180° for a cardioid pattern. This is verified by taking the partial derivative of eq. 2.5 to A, representing the sensitivity of the phase response to changes in A. The partial derivative

$$\frac{\delta\Phi(\omega)}{\delta A} = \frac{sin(\omega t_0)}{1 - 2A\cos(\omega t_0) + A^2} \tag{2.10}$$

shows a very large sensitivity to changes in A if A approaches 1 and ωt_0 approaches 0 or in other words if the system approaches maximum noise suppression. This can be easily visualized by interpreting an FDM as the substraction of 2 vectors. The phase of the resulting output vector is very sensitive to changes in amplitude of one of the input vectors if both input vectors are nearly identical, or in other words in the region with maximal noise suppression.

The right part of Figure 2.3 shows the influence of changes in A on the spatial characteristic of an FDM. Large differences in output are only observed around the angle of maximum noise suppression. Therefore, for the FDM, severe ILD distortions are only expected around the angle of maximum noise suppression.

2.2.2 Adaptive directional microphone (ADM)

A commonly used, more advanced, noise reduction technique in hearing aids is an ADM, illustrated in Figure 2.4. It consists of two software FDM's and an adaptive part. One FDM has a forward oriented spatial characteristic, described by parameter τ_f, to create a so-called speech reference signal. The second FDM has a backward oriented spatial characteristic, described by parameter τ_b, to create the so-called noise reference. The adaptive part, which is in current hearing aids typically implemented as a one tap adaptive scalar, β, combines both spatial patterns to minimize the amount of noise at the output of the algorithm.

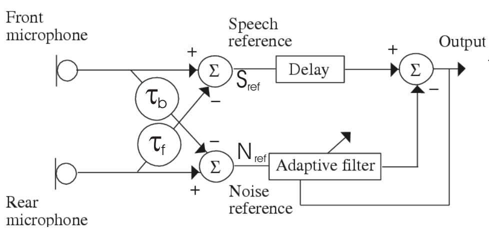

Figure 2.4 — An adaptive directional microphone.

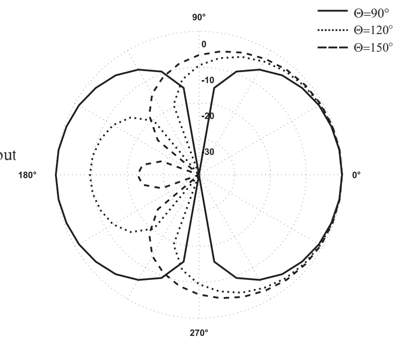

Figure 2.5 — Spatial characteristics of an ideal ADM for three different dominant noise source angles Θ.

The transfer function of the ADM can be written as

$$H_{ADM}(\omega,\theta) = 1 - Ae^{-j\omega(\frac{d}{c}\cos(\theta)+\tau_f)} - \beta Ae^{-j\omega(\frac{d}{c}\cos(\theta))} + \beta e^{-j\omega\tau_b} \tag{2.11}$$

The output of the system when a speech signal $S(w)$ and a noise signal $N(w)$ are present at the microphone inputs equals:

$$O_{ADM}(\omega,\theta) = S(\omega)H_{ADM}(\omega,\theta_s) + N(\omega)H_{ADM}(\omega,\theta_n) \tag{2.12}$$

with θ_s and θ_n representing the angle of incidence of respectively the speech and the noise source. If we assume that the internal delay of both FDM's are equal to each other, $\tau_f = \tau_b = \tau$, then the phase of the speech reference and the noise reference signals equal:

$$\Phi_{Sref} = \arctan\left(\frac{A\sin\omega(t_x+\tau)}{1-A\cos\omega(t_x+\tau)}\right) \tag{2.13}$$

$$\Phi_{Nref} = - \arctan\left(\frac{\sin\omega t_x - A\sin\omega\tau}{A\cos\omega t_x - \cos\omega\tau}\right) \tag{2.14}$$

with t_x the delay between the front and back microphone of the hearing aid ($t_x = \frac{d}{c}cos\Theta$) and τ the fixed delay in the directional microphones. The parameter τ is most commonly fixed to $\tau = \frac{d}{c}$, thereby creating a forward and a backward oriented pattern with respectively a null at 180° and 0°.

In an ideal one-tap system, $A = 1$ and $\tau_f = \tau_b = \tau$. The transfer function of the system can be written as

$$H_{ADM}(\omega,\theta) = B_{ADM}(\omega,\theta)e^{-j\Phi_{ADM}(\omega,\theta)} \tag{2.15}$$

with

$$B_{ADM}(\omega,\theta) = \left[sin\frac{\omega d(1+\cos(\Theta))}{2c} - \beta\sin\frac{\omega d(1-\cos(\Theta))}{2c}\right] \tag{2.16}$$

$$\Phi_{ADM}(\omega,\theta) = \frac{\pi}{2} - \frac{\omega\tau}{2} - \frac{\omega d}{2c} \tag{2.17}$$

The adaptive factor β is designed to create a single independent null angle Θ_0 in the direction of the dominant noise source. The relation between Θ_0 and β can be written as:

$$\Theta_0 = \arccos\left(\frac{2c}{\omega d}\arctan\left(\frac{\beta-1}{\beta+1}\tan\frac{\omega d}{2c}\right)\right) \tag{2.18}$$

Since the dominant noise source may be arriving from a different angle for both hearing aids this will lead to different spatial characteristics for the left and right hearing aid thereby distorting the ILDs (see eq. 2.16). Figure 2.5 shows different spatial characteristics of the ideal ADM after convergence to different noise source angles Θ and their corresponding factors β.

When analyzing the ITD behavior, eq. 2.17 shows that the adaptive part (β) of the system has no influence on the phase transfer function of the ADM, therefore it will have no influence on the ITD perception of a bilateral hearing aid user. This can also be derived from eq. 2.13 and eq. 2.14 which are identical under the assumption that $A = 1$. This implies that the noise reference signal is in phase with the speech reference signal which makes the 1-tap adaptive filter a scaling factor with no influence on the phase relationship between the in-and the output. Therefore, hearing aids with an ideal one tap ADM, $A = 1$ with $\tau_f = \tau_b$, do not distort interaural time information.

In a non-ideal one-tap system, A will be different for both hearing aids and not equal to 1. Two effects are observed. First, the interaural time and level

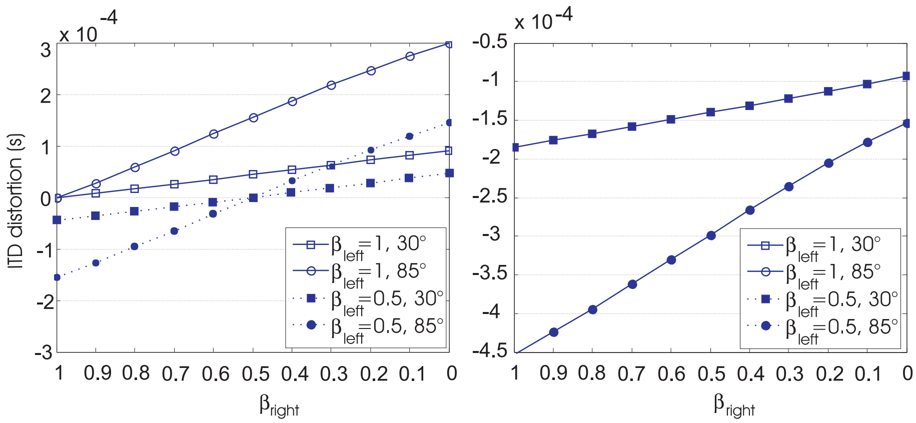

Figure 2.6 — ITD distortion, $\Delta_{ITD} - \widehat{\Delta_{ITD}}$, generated by a bilateral ADM as a function of β_{right}. Simulations are done with $\tau = \frac{d}{c} = 59\mu$s and an inter-microphone distance of 2cm for both hearing aids. A 500Hz sinusoid is used as a test signal. Data are given for two different angles of arrival, i.e. $\theta = 30^\circ$ and $\theta = 85^\circ$. β_{right} is varied between 0 and 1, meaning that a dominant noise source for the right hearing aid is varied between 180° and 90° respectively. A β of 0.5 resembles a dominant noise source around 110°.
Left: $A_{left} = A_{right} = 0.9$. Severe time distortions are observed if $\beta_{left} \neq \beta_{right}$ or in other words if both hearing aids are converged towards a different spatial pattern.
Right: $A_{left} \neq A_{right}$ and $\beta_{left} \neq \beta_{right}$. A_{left} is fixed to $A_{left} = 1$ and $A_{right} = 0.9$. A combination of distortions generated by A and β are observed. Severe ITD distortions, up to 450μs are generated.

distortion discussed in section 2.2.1, concerning the non-ideal FDM, will be present between the speech references of the left and right hearing aid. Second, since $A \neq 1$ the speech reference will no longer be in phase with the noise reference (eq. 2.13 and eq. 2.14). Hence, the phase at the output of each FDM will be dependent on the adaptive part (β) of that particular noise reduction system. This generates interaural time distortion since the angle of incidence of the dominant noise source may be different for both hearing aids leading to a different β for each hearing aid.

The results of a number of simulations, calculating the time distortions, $\Delta_{ITD} - \widehat{\Delta_{ITD}}$, generated by a bilateral ADM are shown in Figure 2.6 as a function of β_{right}. The parameters A and β were varied for the left and the right hearing aid. A 500Hz sinusoid was used as input signal. Results are shown for two different sound source angles ($\theta = 30^\circ$ and $\theta = 85^\circ$). In the left part of Figure 2.6, i.e. $A_{left} = A_{right} \neq 1$, it is shown that severe time distortions were generated if different adaptive factors β were present in both hearing aids. Time distortions up to 300μs were observed.

The right side of Figure 2.6 illustrates the combined effect of ITD distortions generated by A_{left} being different from A_{right} and by β_{left} being different from β_{right}. A_{left} was fixed to $A_{left} = 1$ while A_{right} was fixed to $A_{right} = 0.9$. This resembles a microphone mismatch of approximately 1dB. Since $A_{left} = 1$, the factor β_{left} resulted in a scaling factor which does not affect the phase response of the left hearing aid. Therefore, varying β_{left} had no influence on the ITD distortion (both curves of $\beta_{left} = 1$ and $\beta_{left} = 0.5$ are on top of each other). Time distortions up to 400μs were observed.

Figure 2.6 illustrates that only in very rare cases the ITD information between bilateral ADMs is preserved. As soon as e.g. microphone mismatch is present, ITD distortions occur which easily reach values of more than 100μs. These distortions are dependent on the values of A, β and the angle of arrival of the signal and are well within the range used by the human auditory system, i.e. 0 to 700μs.

Subband implementations of the ADM may offer a better noise reduction performance since they adapt to the angle of the dominant noise source in each frequency band. However, a different factor β for each frequency band will result in independent time delays. This will not only generate distortions of ITDs and ILDs but may, in addition, produce conflicting binaural cues over the different frequency bands.

2.2.3 Discussion

Two commonly implemented noise reduction strategies for bilateral hearing aids were evaluated theoretically. It was shown that these systems preserve binaural cues only when evaluated in very ideal conditions. Large binaural cue distortions were observed if realistic imperfections, such as a microphone mismatch of e.g. 1dB, were introduced in the bilateral hearing aid systems. The distortions typically became larger if the angle of arrival of the sound approached the angle of maximal noise suppression.

Since in reverberant, real life, conditions a monaural ADM or FDM suppresses sound signals in the range of 0 to 5dB depending on the angle of arrival of the signal (Maj et al., 2004), one may suggest that ILD distortions generated by such systems are limited but not insignificant. It was shown that under realistic conditions both the FDM and ADM can introduce large ITD distortions in the order of 100μs and more. This is well within the range used by the auditory system, i.e. between 0 and 700μs, and well above the minimal audible ITD value, i.e. 10μs.

One of the difficulties when interpreting these results is the plasticity of the human auditory system. Several studies have suggested that the human auditory system is capable of re-learning the spatial information when receiving altered binaural cues, for an overview see Wright and Zhang (2006). However, the adaptation to one set of altered binaural cues typically takes a week or more. Since, an ADM is capable of changing its respons in the order of several seconds, adjusting to these rapid changes seems impossible. Moreover, whether the outcome of these studies is based on the plasticity of the cells generating the ITD and ILD information (Javer and Schwarz, 1995) or on the adaptivity of the weighting factors used during the integration of the different, often redundant, cues (Van Wanrooij and Van Opstal, 2007) remains a topic of ongoing research. The current general tendency is the assumption that the internal representation of ILD cues and the monaural spectral cues can be re-learned while the internal representation of ITD cues remains fixed (Wright and Fitzgerald, 2001; Van Wanrooij and Van Opstal, 2005). This implies that only the fixed ILD distortions, generated by a bilateral FDM, might be re-learned by the human auditory system.

One should be aware that the imperfections discussed in this section are typically variable over time and frequency. Since the human auditory system integrates binaural information over these dimensions to produce a stable spatial representation of surrounding sound sources and due to the plasticity of the human auditory system it is uncertain what the perceptual effects of these distortions will be. Therefore, section 2.3 presents a localization experiment to quantify the perceptual influence of hearing aid systems on localization performance.

2.3 Perceptual evaluation

This section describes localization experiments carried out by normal hearing and hearing impaired subjects in a newly developed localization setup. This is done to examine the effects of bilateral hearing aids in general and more specifically a bilateral ADM on binaural cues and localization performance. Since the main focus of this manuscript is on binaural cue processing, localization experiments are restricted to the frontal horizontal hemisphere. Hearing impaired subjects were tested with and without hearing aids to quantify the effects of their hearing aids. Two different hearing aid settings were examined, one using a bilateral omnidirectional microphone setting (no noise reduction present) and one using a bilateral ADM.

2.3.1 Methods

Test setup

Tests were carried out in a reverberant room with dimensions 6m x 3m x 3.5m (length x width x height) and a reverberation time, T_{60}, of 0.54s as determined for a speech weighted noise spectrum. Test persons were placed inside an array of 13 newly designed single-cone speakers with a cone diameter of 10cm. The

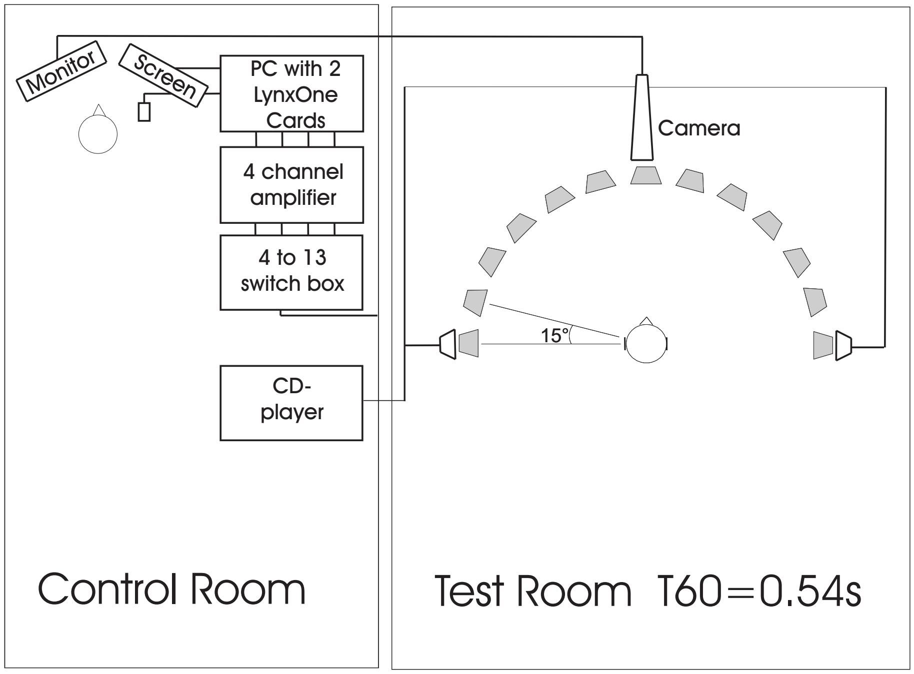

Figure 2.7 — An overview of the used test setup. The "4 to 13 switch box" is connected to the array of 13 (gray colored) loudspeakers and enables the test program to play target sounds from PC on the speakers (connections are not drawn for reasons of clarity). The CD players are connected to the 2 YAMAHA speakers located at +90° and -90° of the subject. These are used to create the noise scenario. A camera is used to monitor the subject.

speakers were located in the frontal horizontal plane at angles ranging from -90° to +90° relative to the subject, a spacing of 15° was used. The speakers were placed at a distance of 1 meter of the subject and were labelled 1 to 13. The target signal was played through one of the 13 speakers using a LYNX-ONE soundcard and a programmable electronic switch box. This switch box, together with the whole test procedure was controlled by the test software, referred to as "Advanced Localization Procedure"(ALP). The test operator loads the test into the program and enters the responses of the subject by clicking the appropriate buttons. The full test procedure is stored by the program, together with the calculated performance measures. Two other, YAMAHA CBX-S3 powered, speakers were present in the test room to evaluate localization performance in noise. These speakers were placed at a distance of 1 meter from the subject at an angle of -90° and +90° relative to the subject. The aim of the noisy scenario is to evaluate the localization performance when using the bilateral ADM in a noisy environment. An illustration of the test setup is given in Figure 2.7.

Subject	HA type	Right ear							Left ear						
		250	500	1k	2k	4k	6k	8k	250	500	1k	2k	4k	6k	8k
AP	Perseo 211	55	50	45	40	55	60		35	35	45	55	60	60	
BG	Canta 7	35	35	40	40	45	30		40	35	35	35	35	25	
BJ	Perseo 211	25	35	40	40	45	55		35	40	45	45	60	60	
CH	Perseo 111	45	35	35	50	80	100		40	30	35	70	80	105	
DH	Perseo 211	30	30	40	40	40	45		35	30	45	50	45	50	
MA	Canta 7	15	10	30	55	60	45		25	25	30	55	55	45	
ML	Diva	35	55	65	60	70	80		15	25	35	40	55	75	
SM	Canta 7	45	40	35	45	55	60		50	50	40	60	55	60	
VP	Perseo 111	25	30	35	35	50	60		20	20	30	35	45	55	
VM	Perseo 211	50	50	45	50	50	45		50	55	50	50	50	50	
Mean HI		36	37	41	46	55	58		35	35	39	50	54	59	
StDev HI		13	13	10	8	12	20		12	11	7	11	12	21	
Mean NH		-3	1	4	3	3		14	4	5	8	5	10		11
StDev NH		9	6	7	7	5		6	6	4	3	5	7		6

Table 2.1 — The audiometric data (in dBHL) of the 10 hearing impaired (HI) subjects and the mean audiometric data of the 10 normal hearing (NH) subjects.

The programmable electronic switch box, ALP and the speakers were designed and built as part of this thesis. Later, ALP was extended to support headphone experiments (chapter 3 and chapter 6) and a child-mode version was included. The latter allows researchers to evaluate the localization capabilities of young children, starting from the age of approximately 4 years old. This is used in a research project to evaluate localization capabilities of normal hearing children and children fitted with two cochlear implants, e.g. the work of Van Deun et al. (2007). Later, a second version of ALP was built enabling horizontal and vertical localization. In this version, the electronic switch box was replaced by two eight-channel sound cards and the speakers were replaced by commercially available loudspeakers. This version of ALP is used in chapter 3, section 4.5.2 and chapter 6. An exact replica of this design is now also used in the Academical Hospital of Maastricht.

Subjects

Prior to the study with the hearing impaired subjects, a similar study was performed with ten normal hearing subjects between 20 and 25 years old (average age: 22 years old). Mean audiometric data of the normal hearing subjects is shown in Table 2.1. These subjects had a maximum hearing threshold of 20dB HL on all octave frequencies starting from 125Hz up to 8000Hz with dB HL defined in ISO389-1 (1998) as the sound pressure level (SPL) relative to the normal hearing threshold. The relevant data of the normal hearing group will be described and used as a reference for the hearing impaired group.

Ten hearing impaired subjects, ranging from 44 to 79 years old, participated in this study. All of them were experienced bilateral hearing aid users. Six of them used Phonak Perseo hearing aids, three used GNResound Canta7 hearing aids and one of the test persons used Widex Diva hearing aids. The settings of their everyday hearing aids were copied into another pair of hearing aids of the identical brand and type and monaural spectral enhancement techniques were switched off in the GNResound and Perseo devices. Audiometric data of the hearing impaired subjects is given in Table 2.1. The mean absolute difference between the left and the right hearing loss was less than 10dB for all subjects, except for subject ML who had a larger asymmetrical hearing loss. The subjects used their own earmolds with a venting between one and three mm, except for subject ML who used an open venting on this patient's best ear for otological reasons. The amplification levels of all subjects did not show large asymmetrical settings (<7dB difference between the mean left and right amplification levels at an input level of 50dB SPL (G50) and <7dB difference for the mean left and right amplification levels at an input level of 80dB SPL (G80) at all frequencies), except for subject ML who had asymmetrical amplification levels to compensate for the asymmetrical hearing loss. However, this subject did not show a bias with (on average 5°) or without hearing aids (on average 2°), and showed similar results as all other subjects.

Stimuli

Earlier studies have shown that localization of high and low frequencies rely on different binaural processing strategies (see section 1.4). When localizing low-frequency sounds primarily ITD information is used. For high frequencies (f>1500Hz) the localization system is based on ILD information and ITD information of the low- frequency envelope of the signal. To obtain information about both binaural processing paths a 200-ms, 1/3-octave, low-frequency noise band centered at f=500Hz, and a 200-ms, 1/3-octave, high-frequency noise band were chosen as target stimuli. In the first study with normal hearing subjects, the high-frequency noise band was centered at 5000Hz. When testing hearing impaired persons this stimulus proved to be useless due to the inaudibility of the stimulus to some test persons. Therefore a 200-ms, 1/3-octave, high-frequency noise band centered at 3150Hz was used for the hearing impaired subjects. Nevertheless, the data of the 5000Hz test with the normal hearing subjects are given and compared with the 3150Hz test with the hearing impaired subjects. We believe that this comparison remains fair for three reasons. First, frequencies are well above 1500Hz, meaning that the same localization mechanisms are being used. Second, the center frequencies 3150Hz and 5000Hz are separated by less than one octave. Third, the minimum audible angle for frequencies in both stimuli are similar, with the 5000Hz noise band having a slightly smaller minimal audible angle (MAA) at the most left and right sides of the head (starting from 60°) Moore (1997*b*) which would give the normal hearing group a slight disadvantage. The third stimulus was a one-second broadband telephone ringing signal. This is the alerting signal of a telephone, it contains both low and high frequencies and includes a lot of transients which should make localization easier. The time structure and the

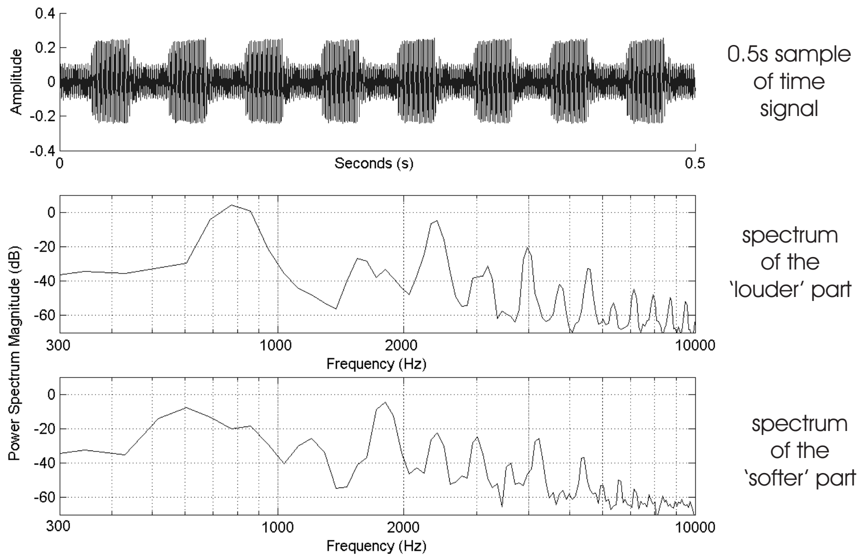

Figure 2.8 — **Up**: The time structure of a 0.5sec sample of the used telephone ringing signal which alternates between a 'softer' and a 'louder' fragment. **Middle**: the power spectral density of the 'louder' part of the telephone ringing signal. **Down**: the power spectral density of the 'softer' part of the telephone ringing signal.

spectrum of this signal is shown in Figure 2.8. An interesting fact is that human subjects are very familiar with and highly trained on localizing a telephone signal in their daily lives.

All target signals were cosine windowed with a rise and fall time of 50ms to avoid broadband clicks. The telephone stimulus was tested in silence as well as with a multitalker babble source located at the left and the right side of the subject. When tests were carried out when using hearing aids or by normal hearing subjects, stimuli were presented at 65dB SPL. For the noise scenario, the two noise sources were set at 62dB SPL giving an SNR of 0dB in the center of the speaker array. Sound level calibrations were done in absence of the subject.

To estimate the impact of signal processing in bilateral hearing aids, tests with and without hearing aids were compared. To rule out the effect of audibility, tests were carried out at equal sensation levels with and without hearing aids. The amplification level of the stimuli was corrected until the subject confirmed that an equal sensation level was obtained with and without hearing aids. Afterwards, the noise level was corrected to keep the SNR of 0dB. Because of the extra amplification in the unaided condition, the results do not reflect a "daily life" comparison between the aided and unaided condition. However, they should reflect the auditory ability of the subject to use binaural cues.

Test protocol

Subjects were seated inside the array of 13 speakers and the chair was elevated until the ears matched the height of the speaker array. For the first evaluations done by normal hearing subjects (the test with the low- and high-frequency noise band) four repetitions were used per speaker, resulting in 52 presented trials per test. Due to time restrictions, all the other tests (all tests with the hearing impaired and both tests with the telephone signal for the normal hearing) were done by using three repetitions on each speaker resulting in 39 presented trials per test. The stimuli were presented randomly and were roved with a roving level of 4dB (between -4dB and 0dB) to avoid the use of monaural loudness cues (see section 1.4.5). To avoid head movements which would facilitate the process of localizing sound sources, subjects were instructed to keep their head fixed and pointed towards 0° during stimulus playback. They were watched on a monitor. The task was to identify the speaker where the target sound was heard.

Hearing aid users were evaluated without their hearing aids, with their hearing aids using an omnidirectional microphone configuration and with their hearing aids using a bilateral ADM configuration. The broadband stimulus was evaluated under two different scenarios: in silence and with multitalker babble sources placed at the left and the right side of the subjects. All 4 test scenarios were performed twice for each test subject.

Two sessions were held on different days to evaluate the test setup on test-retest variations like e.g. learning effects. One test took about 5 minutes, and one session including all scenarios took less than one hour and a half. Subjects had a break after every 6 tests and could take a break whenever they felt tired. In total 320 individual test runs were completed (4x4x10x2).

Performance measures

Different error measures have been used in previous localization studies (Noble and Byrne, 1990; Lorenzi et al., 1999*b*; Van Hoesel et al., 2002). We focus on two commonly used error measures:

1. Root Mean Square Error (RMS)

$$RMS(^\circ) = \sqrt{\frac{\sum_{i=1}^{n}(stimulus - response)^2}{n}} \tag{2.19}$$

2. Mean Absolute Error (MAE)

$$MAE(^\circ) = \frac{\sum_{i=1}^{n}|(stimulus - response)|}{n} \tag{2.20}$$

with n the number of presented stimuli. When using MAE, all errors are weighted equally while in the RMS error large errors have a bigger impact than small errors. The smallest non-zero error a subject can make during one test run equals 2.40° RMS and 0.38° MAE (with 3 repetitions per speaker or $n = 39$ trials per test run). Statistical analysis was performed on both error measures, showing similar results. Throughout this chapter the data and the statistical analysis is reported in detail only using the RMS error measure. The mean and standard deviation of the MAE values will be given for the different subject categories for all tested conditions. This gives the reader the opportunity to compare with other work where MAE error measures have been used.

2.3.2 Results and analysis

First the data and statistical analysis of the normal hearing persons are presented, followed by the data and analysis of the hearing impaired subjects. The data shown for each subject is the average over test and retest condition. This is done because no statistical difference was found between test and retest condition for both the normal hearing and the hearing impaired subjects. All statistical analysis were performed using SPSS 10.0 with test and retest separated in a repeated measures analysis of variance (ANOVA). A standard significance level of 0.05 is used throughout this chapter. All pairwise comparisons of the different ANOVAs were Bonferroni corrected for multiple comparisons. This correction was necessary to reduce falsely significant results when doing multiple comparisons. All reported p-values are lower bound values.

The different test conditions are identified as follows: For the hearing-impaired group, the three hearing aid conditions are: no – without hearing aids; o – with hearing aids with an omnidirectional configuration; a – with hearing aids with a bilateral ADM configuration. The normal hearing group is identified as nh. The stimulus conditions are identified as follows: 500 Hz – 1/3-octave low-frequency band; 3150 Hz or 5000 Hz – 1/3-octave, high-frequency band; ts – wideband telephone signal in silent condition; tblr – telephone signal with babble jammer on left and right. The data and statistical analysis will motivate the discussion in section 2.3.3.

Normal hearing subjects

The data of the normal hearing subjects is given in Table 2.2 and represented in Figures 2.9, 2.10 and 2.11. Table 2.2 shows the individual RMS results and the mean RMS and MAE results of the normal hearing subjects. It shows that all normal hearing subjects performed better with the low-frequency narrow-band signal (average RMS error: 13.5°) than with the high-frequency narrow-band signal (average RMS error: 21.3°). Also, all subjects performed better with the

RMS-error($^\circ$)	ts	tblr	500Hz	5000Hz
TVS	10.5	12.5	14.3	22.9
SB	6.1	12.1	14.4	22.8
SVD	10.4	11.6	15.3	23.1
LD	5.3	13.6	15.8	21.5
LVDP	5.3	9.9	10.5	14.0
KH	3.3	11.6	11.0	16.4
HD	9.9	13.0	14.0	25.4
EBI	4.5	12.0	12.3	21.9
EBO	10.0	11.1	16.4	24.2
DVS	2.4	10.3	11.2	21.0
mean RMS	6.8	11.8	13.5	21.3
stdev	3.1	1.1	2.1	3.5
mean MAE	3.5	7.3	8.7	14.3
stdev	2.6	0.9	1.9	2.9

Table 2.2 — The individual RMS data with the mean MAE and RMS data of the normal hearing subjects for the 4 different test conditions. All data are in units of degrees ($^\circ$)

transient broadband signal (average RMS error: 6.8°) than with both narrow-band signals. Performance dropped for all tested subjects when noise was added to the scenario with the broadband telephone ringing signal (average RMS error: 11.8°).

These statements were confirmed by a two-way repeated-measures ANOVA. The ANOVA was carried out on the normal hearing data with the factors 'test signal' (500Hz, 5000Hz, telephone in silence and telephone with babble) and the factor 'test-retest'. A main effect for the factor 'test signal' was observed ($p < 0.001$). No main effect was observed for the factor 'test-retest' and no significant interaction between the two factors was found. All 4 test conditions were significantly different from each other (all $p \leq 0.003$), except for the 500Hz and the 'telephone with babble' condition (p=0.066).

The data can also be interpreted per angle of incidence. The white error bars in Figure 2.9 show the RMS error per speaker when accumulating all responses of the normal hearing subjects over the different test conditions. This illustrates that normal hearing subjects had a very good performance in the most frontal area of the horizontal plane (RMS error <10° and > 80% correct answers for every angle in the area from -30° and +30°). At the sides, sensitivity dropped and localization started to detoriate. The same tendency can be found in the left column of Figure 2.10 which shows an accumulation of all responses given by the normal hearing subjects under the different test conditions. However, both figures should be interpreted with caution because they represent

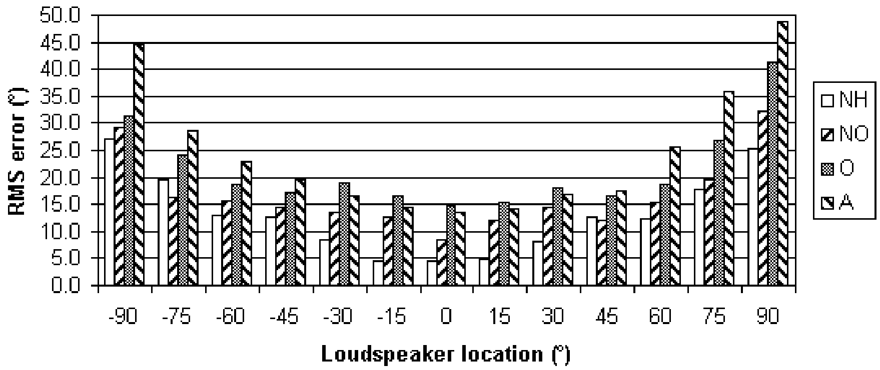

Figure 2.9 — The error bars show the RMS error per speaker when accumulating all responses of the different test conditions per stimulus location. This was done for the group of normal hearing subjects (NH), the group of hearing impaired subjects without hearing aids (NO) and with hearing aids using an omnidirectional configuration (O) and a bilateral ADM configuration (A).

an accumulation of the responses of the different test subjects on a localization task, which is a subject dependent process. The left column of Figure 2.11 illustrates the distribution of the mean response given by each subject for each speaker location and for each test condition. This figure shows that the similarity between the mean responses given by the different normal hearing subjects was relatively high for most test conditions. Only the high frequency stimulus showed larger dissimilarities, especially at the sides of the head.

Hearing impaired subjects

Hearing impaired subjects were evaluated in the same conditions as the normal hearing subjects except for the fact that the 5000Hz stimulus was replaced by a 3150Hz stimulus. All tests were performed with 6 stimuli in total per condition (3 for test and retest). Figures 2.9, 2.10 and 2.11 illustrate the data obtained from the experiments with the hearing impaired subjects. The individual results of the hearing impaired subjects are given in Table 2.3.

A two-way repeated-measures ANOVA was carried out on the hearing impaired data on the factors 'test signal' (500Hz, 3150Hz, telephone in silence and telephone with babble), the factor 'hearing aid' setting (no-hearing-aid, omnidirectional, and ADM) and the factor 'test-retest'. A main effect for the factor 'test signal' ($p=0.011$) and the factor 'hearing aid setting' ($p=0.001$) was found. No main effect was observed for the factor 'test-retest'. No significant interactions were found.

RMS (°)	ts			tblr			500Hz			3150Hz		
	no	o	a	no	o	a	no	o	a	no	o	a
AP	12.6	16.5	19.6	14.3	18.3	17.1	16.5	17.3	18.6	21.7	21.3	21.3
BG	18.2	15.6	24.4	15.7	18.4	26.8	14.8	20.2	16.4	28.3	23.4	37.4
BJ	15.3	19.4	16.3	18.0	24.6	26.4	16.8	14.4	16.4	29.3	30.9	35.7
CH	12.1	12.9	8.3	16.6	18.0	28.2	12.4	12.8	14.1	18.2	33.7	61.5
DH	8.1	15.0	10.6	9.7	28.9	31.1	11.6	13.8	13.1	18.2	23.4	22.8
MA	8.5	10.2	12.4	12.1	15.8	14.2	15.8	15.0	15.5	14.3	39.0	49.6
ML	14.6	14.8	15.0	15.4	16.9	19.8	27.1	38.4	35.0	16.9	22.6	18.3
SM	16.2	19.4	25.9	16.3	24.7	31.1	18.7	25.0	27.1	25.0	36.9	47.5
VP	13.3	17.5	17.5	14.4	21.1	25.4	14.5	15.6	18.8	31.9	26.8	29.2
VM	11.0	20.1	24.9	20.8	26.3	29.3	21.3	27.8	26.3	20.3	21.4	26.6
mean RMS	13.0	16.1	17.5	15.3	21.3	25.0	17.0	20.0	20.1	22.4	27.9	35.0
stdev	3.3	3.1	6.2	3.0	4.5	5.9	4.5	8.1	7.1	5.9	6.7	14.1
mean MAE	8.8	11.7	12.4	10.6	15.9	17.8	12.1	15.1	15.4	16.0	21.4	25.7
stdev	2.9	3.0	5.2	2.5	3.8	4.1	4.2	7.2	7.1	5.2	6.5	10.3

Table 2.3 — The individual RMS data with the mean RMS and MAE data of the hearing impaired subjects for the 4 different test signals and 3 different hearing aid settings.

Pairwise comparisons for the factor 'hearing aid setting' showed a significantly better performance without hearing aids (no) compared to both conditions with hearing aids. No significant difference was found between the omnidirectional (o) and ADM condition (a) although the p-value was very close to the 0.05 bound (p=0.053). The p-values of these pairwise comparisons are summarized in the column 'full' of Table 2.4.

To evaluate the impact of the hearing aid configuration on the two main horizontal localization mechanisms (ITD and ILD), a separate analysis on the data gathered for every signal was performed. For each signal a repeated measures ANOVA was calculated with a factor 'hearing aid setting' (no-hearing-aid, omnidirectional and ADM) and the factor 'test-retest'. The p-values of the pairwise comparisons are discussed in the next paragraphs and are summarized in the column 'full' of Table 2.4.

Low-frequency noise band: A main effect for the factor 'hearing aid setting' (p=0.037) was found. No interactions were present. Analysis showed that the condition without hearing aids outperformed the ADM condition significantly (on average by 3.1°). Although the mean difference between the omnidirectional configuration and the no-hearing-aid condition was 3.0°, no significant difference was found between these conditions. The mean performance of the omnidirectional and the ADM configuration was very similar for this test signal and no significant difference between these conditions was found.

High-frequency noise band: There was a main effect for the factor 'hearing aid setting' (p=0.042). No interaction was observed. Large differences in the mean results between the different conditions were observed, however no significant differences were found in the pairwise comparisons.

general (4 conditions)	45°/-45°	±60°/±90°	full				
no vs o	0.038*	0.002*	0.002*				
no vs a	0.014*	0.001*	0.001*				
o vs a	0.998	0.046*	0.053				
500Hz	45°/-45°	±60°/±90°	full	3150Hz	45°/-45°	±60°/±90°	full
no vs o	0.092	0.371	0.133	no vs o	1.000	0.105	0.284
no vs a	0.320	0.040*	0.032*	no vs a	0.441	0.071	0.098
o vs a	0.625	1.000	1.000	o vs a	1.000	0.333	0.120
ts	45°/-45°	±60°/±90°	full	tblr	45°/-45°	±60°/±90°	full
no vs o	0.014*	0.227	0.046*	no vs o	0.038*	0.004*	0.016*
no vs a	0.060	0.095	0.059	no vs a	0.012*	0.084	0.002*
o vs a	1.000	0.637	1.000	o vs a	0.458	0.009*	0.044*

Table 2.4 — p-values (with Bonferroni adjustment) of the pairwise comparison for hearing impaired subjects. Conditions without hearing aids (no), with hearing aids with an omnidirectional microphone (o) and with hearing aids with an adaptive directional microphone (a) are compared with each other. (*= significant for a significance level of 0.05)

Broadband telephone ringing signal in silence: There was a main effect for the factor 'hearing aid setting' (p=0.043). No interaction was observed. The no-hearing-aid condition was significantly better than the omnidirectional condition (on average 3.1°). No significant difference was found between the 'no-hearing-aid' and the ADM condition, although the p-value were close to the level of significance (p=0.059). No difference was found between the omnidirectional and the ADM condition (p=1.000).

Broadband telephone ringing signal with babble from left and right: A main effect was found for the factor 'hearing aid setting' (p=0.002). No interaction was observed. Statistical significant differences were found in all pairwise comparisons, with the no-hearing-aid condition performing better than both hearing aid conditions (on average 6.0° and 9.7°), and the omnidirectional configuration performing better than the ADM configuration (on average 3.7°).

One of the aspects that seems to differentiate the data obtained with the ADM configuration and the other configurations is the presence of left-right confusions for the extreme left and right angles (Figure 2.10). Several subjects experienced this problem for stimuli presented at ±90° when they were tested while using the bilateral ADM. This was observed for the high-frequency noise band (subjects SM, MA, VM and CH) and for the telephone signal in noise (subjects DH, VM, CH and BG).

An evaluation per angle of incidence is shown in Figure 2.9. This shows the RMS error per speaker location when accumulating the responses of the hearing impaired subjects on the different stimulus conditions and was done for the three different hearing aid conditions (omnidirectional, ADM and without hearing aids). This figure, together with Figure 2.10 illustrates that hearing

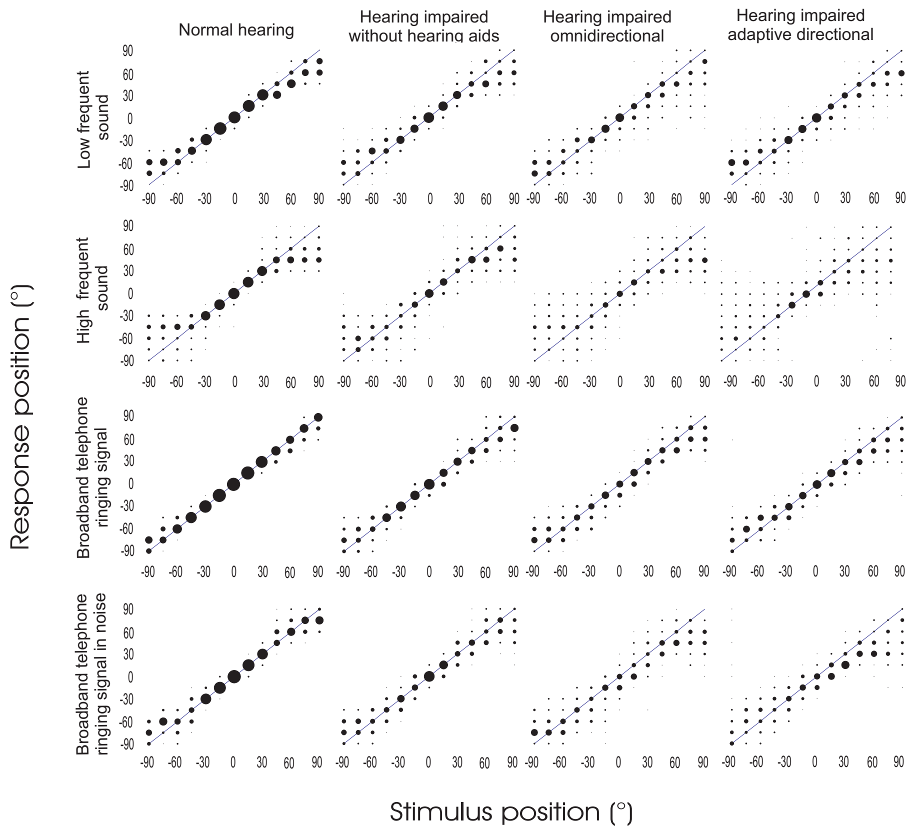

Figure 2.10 — All responses given by the normal hearing subjects and the hearing impaired subjects with and without hearing aids for the different stimuli and acoustical conditions. The surface of the circles is proportional to the amount of responses given by the subjects.

impaired subjects had a better localization performance in the frontal region compared to the region at the sides of the head, which is in agreement with the data of the normal hearing subjects. However, it is important to note that when using hearing aids, a decrease in performance was observed not only at the sides of the head, but also in front of the listener. These figures represent an accumulation of the responses given by the different hearing impaired subjects on a localization task. Since localization is a subject dependent process, these figures should be interpreted with care. Moreover, a large inter-subject variance is present in the hearing impaired data, especially when using hearing aids. This is illustrated in Figure 2.11 which gives for each test condition the mean response of each subject on each stimulus location.

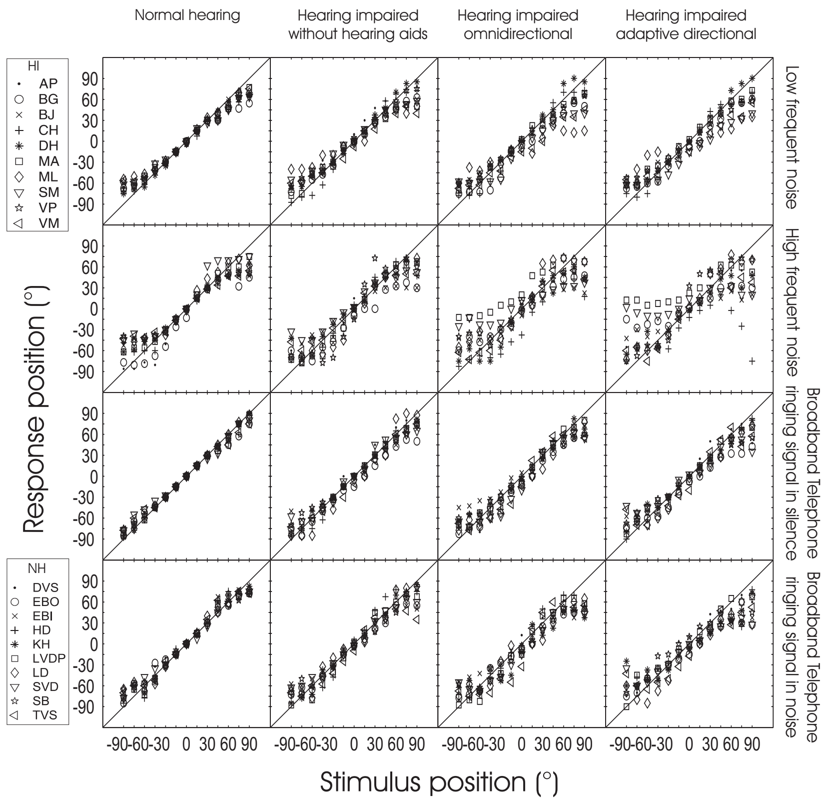

Figure 2.11 — The mean responses given by the different normal hearing subjects and the different hearing impaired subjects with and without hearing aids for the different stimuli and acoustical conditions.

Hearing impaired without hearing aids versus normal hearing

When evaluating the normal hearing and the hearing impaired subjects, one clearly observes differences in performance (Table 2.2 vs Table 2.3, Figure 2.9, 2.10 and 2.11). In this section, the data of the normal hearing group are compared with the best condition of the hearing impaired group, being the condition without hearing aids. The data of all four test signals (low-frequency, high-frequency, telephone ringing signal and telephone ringing signal in noise) were included in a repeated measures ANOVA. A between-subjects factor 'subjects' was introduced which separates the group of normal hearing and hearing impaired persons. There was a main effect present for the factor 'test signal' ($p<0.001$) and no effect was found for the factor 'test-retest'. No interactions were found in the ANOVA.

The results of the ANOVA show that the group of normal hearing subjects performed better than the group of hearing impaired subjects without hearing aids ($p=0.005$). This can also be seen when comparing the data of the normal hearing and the hearing impaired in Figure 2.10 and Figure 2.11. Another difference between the normal hearing and the hearing impaired subjects was the larger consistency between the different subjects in the normal hearing group compared to the hearing impaired group (Figure 2.11).

The pairwise comparisons of the factor 'test signal' showed the same results as described in the normal hearing section. A better performance was obtained when using the broadband stimulus compared to both narrow-band stimuli ($p<0.001$ for both the low and high-freq. stimulus) and compared to the broadband stimulus in noise ($p=0.001$). A better performance was also obtained when localizing the low-frequency stimulus compared to the high-frequency stimulus ($p<0.001$).

Microphone placement

In the previous paragraphs, clear differences were observed between the performance obtained by hearing impaired subjects with and without hearing aids. These results are further discussed in section 2.3.3. One of the factors which could explain these differences and which is not related to the signal processing by hearing aid algorithms is microphone placement. On a BTE hearing aid, the microphones are positioned relatively far from the eardrum, which may influence the binaural cues presented to the listener. Therefore extra measurements were made using the microphones of BTE hearing aid devices and in the ear microphones (ITE) of a manikin. This was done in an anechoic environment.

A 01dB CORTEX MK2 manikin, a dummy head with torso build according to the IEC 959 standard, was used with two G.R.A.S IEC 711 ear simulators. A G.R.A.S 40AG pressure microphone was located in each ear simulator. The BTE devices were two Canta7 dual microphone shells with direct microphone outputs. All recordings were made using an 8-channel G.R.A.S. 12AG preamplifier and 2 synchronized LYNXONE soundcards at a sampling rate of 48kHz. A broadband white noise signal was recorded simultaneously with all 6 microphones (2 CORTEX MK2 ITE and 4 BTE microphones). ITDs and ILDs were calculated between the two in the ear microphones (ITE) of the manikin, the two omnidirectional microphones located at the front of the BTE devices and between the two omnidirectional microphones located at the back of the BTE devices.

The measured ITDs were determined by calculating the delay generating the maximum value of the cross correlation function. The ITD estimations are given in the left part of Figure 2.12. In this figure, it is shown that the ITDs between the BTE microphone pairs did not fully agree with the data of the

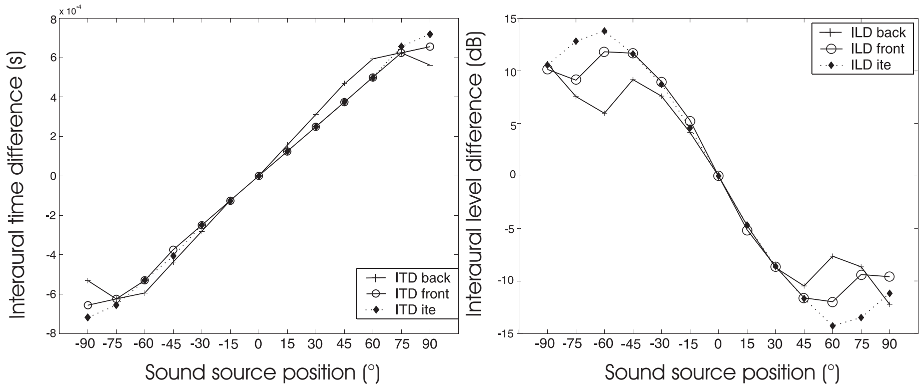

Figure 2.12 — ITD and ILD measurements with two Canta7 behind the ear (BTE) hearing aids prototypes on a CORTEX manikin in anechoic conditions. Interaural time and level differences were measured between the in-the-ear microphones of a CORTEX manikin (ITD ite, ILD ite), between the front omnidirectional microphones of the BTE devices (ITD front and ILD front) and between the back omnidirectional microphones of the BTE devices (ITD back and ILD back). Measurements were made using a broadband white noise stimulus.

ITE microphone pair, which represents the ITD information at the eardrums of a human listener. The front microphone pair showed a small distortion (in the order of 40μs) of ITD information at -90° and +90°. The back microphone pair generated larger ITD distortions (in the order of 100μs to 200μs). These distortions were only present in the area between -60° and -90° and between 60° and 90°.

The right part of Figure 2.12 shows the measured ILD cues generated by a broadband white-noise stimulus. The distortion generated by the back microphone pair was again more pronounced than the distortion by the front microphone pair and was mainly located at the most left and right angles. Since ILDs are highly dependent on frequency content, sub-band analyses were performed. These analyses showed that almost no ILD information and no ILD distortions were present for frequencies $f < 1kHz$. Higher frequencies, between 1 and 5kHz, which generate the majority of ILD information were subject to larger distortions (maximum measured difference of 7dB at an angle of 75°).

In general, it was observed that small binaural cue distortions may be introduced by microphone placement. However, no distortions were present in the area between -45° and +45°. To rule out the effect of microphone placement, the perceptual data was split into two parts: the data of the area where the impact of the microphone position is almost negligible (from -45° to 45°) and the remaining data for which the drop in localization performance might be (although unlikely) generated by the microphone position. The statistical analysis of section 2.3.2 was repeated for these two subareas, the p-values of the pairwise comparisons are summarized in Table 2.4.

Table 2.4 demonstrates that for both listening scenarios which generated a significant difference between the with and without hearing aid condition (o vs no), i.e. the telephone in silence and the telephone in noise, this difference is also found in the area from -45° to 45°. The general analysis on the four test conditions confirms that performance in the area between -45° and 45° was worse with than without hearing aids.

2.3.3 Discussion

Perceptual evaluations were performed to quantify the impact of a bilateral ADM in particular and bilateral hearing aids in general on localization performance in the horizontal hemisphere. The data and the analysis of the perceptual evaluations are now discussed with respect to the three research questions raised in the introduction of this chapter.

Normal hearing and hearing impaired performance

Section 2.3.2 described and quantified the localization performance of normal hearing and hearing impaired subjects in the frontal horizontal plane. Hearing impaired subjects were tested in three different conditions: without their hearing aids (no), with hearing aids with an omnidirectional microphone configuration (o) and with hearing aids with an ADM configuration (a). Overall, the average performance with the low-frequency stimulus (mainly ITD) was better than performance with the high-frequency (mainly ILD) stimulus. Test results improved when using the broadband stimulus compared to high or low-frequency stimuli. This can be explained by the possibility to use both ILD and ITD cues and by the time structure of the broadband signal. The length of the stimulus could also have affected localization performance. When using a one-second signal, slight head movements can occur during stimulus playback which would give the subject an extra advantage. When adding jammer sources with a SNR of 0dB, performance dropped significantly for both the hearing impaired subjects and the normal hearing subjects. This confirms the results of the study of Lorenzi et al. (1999*b*).

Statistical analysis (section 2.3.2) showed a better performance of the normal hearing subjects compared to the hearing impaired subjects which is also illustrated in Figure 2.9 and Figure 2.10. Although both groups were not age matched, 64 percent of the individual scores of the hearing impaired subjects without hearing aids were within two standard deviations, and 39 percent were within one standard deviation of the results of the normal hearing group (when comparing the data of the same test conditions). In addition to our findings, Lorenzi et al. (1999*a*) mentioned that a considerable percentage of their evaluated hearing impaired subjects (between 50 and 75%) could reach normal hearing performance on a binaural task. This concludes that some hearing impaired subjects were able to use binaural cues as well as normal hearing subjects. The other subjects had a significantly lower performance than the normal hearing group but were still able to use the binaural cues. This is important to motivate further research on binaural signal processing for hearing aids.

It was also shown that for both the normal hearing as the hearing impaired subjects, localization was more accurate in front of the listener than at the left and right side of the listener (Figure 2.9 and Figure 2.10) which agrees with the data of Makous and Middlebrooks (1990) and Carlile et al. (1997). However, in these studies the decrease in accuracy at the sides of the head were less pronounced which might be explained by the difference in the experimental setup.

Finally, it was observed that the hearing impaired subjects, who were using their own hearing aids, and were therefore highly trained on localizing sound sources with these devices in their daily lives, were not able to localize half of the targets correctly in the most frontal region (from -45° to +45°) whereas normal hearing subjects have a near 100 percent score in this region (section 2.3.2).

Do bilateral hearing aids preserve localization cues ?

To answer this question, tests were performed by hearing impaired subjects with and without hearing aids. By using equal sensation levels (for restoring audibility) a comparison could be made between performance when all binaural information was present and performance with hearing aids. Throughout the different tests, 4 out of 8 comparisons between unaided and aided conditions showed significant better performance unaided than aided for a significance level of $p=0.05$ (and 6 out of 8 for a significance level of 0.1) (see Table 2.4).

A general analysis confirmed that performance without hearing aids was significantly better than with hearing aids (section 2.3.2). This is in accordance with the data of Noble et al. (1998) and Byrne et al. (1998). An improvement in localization performance was obtained when using open earmolds, allowing the subjects to use binaural cues of the direct sound instead of the output of the hearing aid in frequency regions with a moderate hearing loss. Moreover, when comparing with normal hearing performance, only 36 percent of the individual test results with hearing aids fell within 2 standard deviations of the results of the normal hearing subjects and only 20 percent of the individual test results fell within 1 standard deviation of the results of the normal hearing subjects. These percentages are considerably smaller than the numbers shown in the previous section for the hearing impaired subjects without hearing aids.

The current findings confirm that hearing aid users localize better without their hearing aids than with their hearing aids.

One of the factors influencing localization performance when using hearing aids is the positioning of the microphones. To quantify this influence, recordings were made in an anechoic environment. These recordings suggest that no large binaural cue distortions were introduced by microphone placement, especially in the region between -45 and +45°. To exclude the influence of microphone placement, separate analyses were performed on the -45 to +45° data. These analyses suggest that the drop in localization performance cannot be explained only by microphone positioning. This confirms the results of the work of Noble and Byrne (1990) in which no significant differences in localization performance were found when comparing BTE, ITE and ITC hearing aids. Because microphone positioning seems insufficient to explain the difference in localization performance, the data strongly suggest that the signal processing of hearing aids introduces binaural cue distortion when used in a bilateral hearing aid configuration even if noise reduction algorithms are switched off. Further analysis of the different building blocks is needed to analyze this behavior.

To gain insight in the type of interaural distortion (ITD or ILD) experienced by hearing aid users, low-and high-frequency stimuli were used. When comparing the omnidirectional condition with the condition without hearing aids in Table 2.3, it shows that 4 out of 10 subjects had a large decrease in performance for the 500Hz stimulus when using hearing aids (subjects BG, ML, SM and VM) which could indicate distortion of time cues by the hearing aids. On the other hand half of the subjects showed a large decrease in performance for the 3150Hz stimulus (subjects CH, DH, MA, ML and SM) which could indicate distortion of level cues. Thus, some subjects seemed to experience problems with level cues, some with time cues and two subjects do not experience problems at all when localizing a high and low-frequency stimulus with an omnidirectional configuration (subjects AP and BJ). Surprisingly, subject AP, who had no decrease in performance with both low and high-frequency stimuli did show a decrease in performance with the transient broadband signal and the broadband signal with jammer sounds. The reasons for these inter-subject differences are, at the moment, unclear.

Do noise reduction systems have an influence on localization performance ?

Table 2.3 shows that in two out of four test conditions the ADM was outperformed by the omnidirectional microphone configuration. A general ANOVA did not confirm this difference, although the p-value was close to the significance level (p=0.053). When examining Table 2.4, it is observed that significant differences could be detected if the signal was presented from the left or the right side of the head. Table 2.4 shows that these differences were due to a degradation in localization performance when localizing a telephone ringing signal in the presence of noise sources. These noise sources were playing continuously throughout the experiment. Therefore the noise reduction system had plenty of time to adapt to a steady filtering operation. In the other tests only very short sounds were used which could have deprived the noise reduction system of sufficient time to adapt to a steady filtering operation. Since these tests were performed with commercial hearing aids, no knowledge is available on the behavior of the noise reduction systems during these short sound presentations.

The fact that only differences were observed at the side of the head can be explained by the mechanisms underlying the ADM. An ADM creates a directional pattern depending on the specific noise scenario. A null is put in the direction of the most dominant noise source. When testing with babble jammer sources at the left and right side of the subject, the ADM will try to cancel out these directions. Hence, it can be assumed that the ILD or ITD perception of the stimulus around $+90°$ and $-90°$ was degraded by this filtering operation. This was also shown in section 2.2 which demonstrated that the main ILD and ITD distortions are located around the angle with maximal noise suppression. It should be mentioned that none of the subjects made remarks on the inaudibility of the stimuli due to the presence of the noise reduction system.

When comparing the small-band data gathered with the omnidirectional and with the ADM configuration in Table 2.3, it was observed that only small differences were found for the 500Hz stimulus. For the 3150Hz much larger differences were found, especially for subjects BG, CH, MA, SM (differences $>10°$). This could indicate that, for these subjects, the extra distortions generated by the ADM were based on ILD distortions. However, as mentioned in one of the preceding paragraphs, it is uncertain how the ADM in each hearing aid did react on short stimuli presented in silent conditions which makes it difficult to interpret the data of these test conditions. Moreover, the extrapolation of results obtained with narrowband stimuli to broadband stimuli is not straightforward since modern hearing aids are multiband processing devices. Each of the subbands may therefore be processed differently. This would not only lead to the possibility of binaural cue distortion, but also to the possibility of generating interfering binaural cues over the different frequency channels. How the auditory system of the different subjects would react on these interfering cues is an unknown factor when interpreting the data.

2.4 Conclusions

Three main research questions have been addressed in this chapter.

The localization performance of normal hearing and hearing impaired subjects in the frontal horizontal plane was quantified (section 2.3). This was necessary to evaluate the localization performance obtainable by hearing impaired subjects relative to normal hearing subjects. It was shown that the group of hearing impaired subjects, without wearing their hearing aids, localized sounds slightly less accurately than the normal hearing subjects which might have been influenced by the age differences in the studied populations. However, the hearing impaired subjects were still able to use binaural cues which motivates further research on binaural hearing aid systems.

Current state-of-the-art bilateral hearing aids with their multi-microphone noise reduction systems switched off, have a negative impact on localization performance (section 2.3). The decrease in localization performance with hearing aids could not be fully explained by microphone placement which indicates that different signal processing blocks inside a hearing aid, e.g. compressions systems, distort binaural cues. More fundamental research should be done on these separate building blocks but is outside the scope of this research project.

It was shown that the two most commonly used noise reduction systems in commercial hearing aids, i.e. a FDM and ADM configuration, can distort binaural cues and that they have a negative impact on localization performance. An objective evaluation (section 2.2) showed that these systems preserve binaural cues only when evaluated in very ideal conditions. Large binaural cue distortions were observed if realistic imperfections, such as a microphone mismatch of e.g. 1dB, were introduced. These distortions typically became larger if the angle of arrival of the sound approached the angle of maximal noise suppression. Moreover, these distortions are typically variable in time and frequency. Since the human auditory system integrates information over these dimensions to produce a stable spatial representation of incoming sounds, a perceptual evaluation was done (section 2.3) to quantify the influence of an ADM on a binaural task, i.e. a localization experiment in the frontal horizontal hemisphere. This experiment showed that a bilateral ADM did have a significant additional negative impact on localization performance. The influence of the bilateral ADM was clearly dependent on the angle of arrival of the signal and on the noise scenario present during the experiment. Significant differences between the omnidirectional and ADM configuration were only observed when localizing sound sources around the area of maximum noise suppression.

We can conclude that using bilateral, i.e. independently operating, monaural directional multi-microphone noise reduction algorithms such as the ADM and FDM, and bilateral hearing aids in general, tend to distort binaural information. This leads to a degraded localization performance or spatial awareness of the hearing aid user. This might also lead to a worse speech understanding in noise due to sub-optimal spatial release from masking.

Chapter 3

Virtual acoustics for binaural hearing aid research

One of the major steps in algorithmic design for hearing aids involves the evaluation of these algorithms in realistic conditions. This is generally done by using objective or perceptual performance measures, preferably in different acoustic environments. The impact of signal processing algorithms is often strongly dependent on the acoustic environment. Therefore, these environments should be varied from quasi-anechoic, typically at the beginning of the validation, to more challenging and even very reverberant conditions. Setting up experiments in these environments, such as a localization experiment, is very time-consuming and requires the physical availability of the different rooms for the duration of the experiments. Performing evaluations by using advanced acoustic modelling of different virtual acoustic environments may offer a solution for this. This has been one of the main objectives of the VIRTAC-project ('Virtual Acoustics', FWO-Vlaanderen project G.0334.06) which is part of a cooperation between ExpORL, SISTA-SCD and the Acoustics and thermal physics group of the KULeuven.

One of the main goals of this dissertation was the development and evaluation of new binaural noise reduction algorithms (see chapters 4, 5 and 6). To demonstrate the potential of new algorithms, it is often preferred to perform evaluations off-line, thereby avoiding the time investment of writing real-time code and avoiding the specific difficulties that arise when testing real-time algorithms such as extra parameters that have to be tuned and investigated. As a consequence, evaluations are typically done using headphone experiments with pre-calculated filters or sound signals. One of the issues encountered during this project is whether localization under headphones can be done as accurately as localizing sound sources in a natural way. In the frame of the VIRTAC-project, this study was extended with the evaluation of localization accuracy when performing headphone experiments using stimuli generated by a commercial virtual acoustics software package, namely ODEON. As mentioned in section 2.3, this dissertation has focussed on the binaural cues. Therefore localization experiments were carried out only in the frontal horizontal hemisphere.

Section 3.1 gives a brief introduction to the literature on localization through headphones, summarizes the techniques implemented in ODEON and presents the specific research questions of this chapter.

Section 3.2 defines the different acoustic environments and the stimuli used to perform a frequency dependent validation of the localization experiments.

Section 3.3 presents the results and analysis of the perceptual evaluations.

Section 3.4 discusses the results presented in the previous chapter. It analyzes whether doing experiments using headphones and/or a virtual environment has a significant influence on localization performance.

Section 3.5 summarizes the conclusions of this chapter.

The data of this chapter are discussed in Rychtáriková et al. (2008) and Rychtáriková et al. (2007).

3.1 Introduction

Creating virtual acoustic environments can be done either by using headphone stimuli or by mimicking the acoustics of a virtual room in a second room. The latter is generally done by using a set of loudspeakers which generate a sound field according to the acoustical properties of the virtual room (Møller, 1989; Gierlich, 1992; Akeroyd et al., 2007). In this chapter, focus will be limited towards the use of headphones. Headphone experiments can be based on measuring the transfer functions between a sound source and the eardrums of an artificial head or a person, also known as binaural room impulse response (BRIR) measurements (Butler and Belendiuk, 1977; Møller et al., 1996; Minnaar et al., 2001), or on using advanced acoustic modelling of a room in simulation software (Hammershoi and Møller, 2002). The latter represents a more flexible approach for listening tests, since the room acoustics can be quickly changed without measuring impulse responses in this specific environment.

Generating audible information by using an acoustic room simulation is generally known as auralization. Virtual acoustics combines the properties of the sound source (directivity and power spectrum), the listener (HRTF) and the room (the room impulse response) to generate auralized sounds. The first auralization attempts date back to 1934 (Spandock, 1934). Computer based binaural auralization was introduced in the work of Posselt (1987) and was soon followed by others (Kutruff, 1993; Kleiner et al., 1993). In general, virtual modelling of binaural signals in a room is based on convolving the simulated BRIRs with anechoic sound samples. The BRIRs can be obtained in several ways. For an overview see Savioja et al. (1999). They can be computed using wave models: finite element methods (Kleiner et al., 1993) or boundary element methods (Katz, 2001), or by using particle methods: ray-tracing methods (RTM) (Krokstadt et al., 1968), cone tracing methods or image source methods (ISM) (Allen and Berkley, 1979), with currently the latter strategies being preferred over wave models due to complexity issues. The software evaluated in this chapter, i.e. ODEON, is based on a combination of ISM and RTM (Vorlander, 1989; Rindel, 2000).

ISM is based on the principle that a wave front originally arriving from a point source and which is reflected on an infinite plane, can be observed as if this reflection originates from an image source. The position of this image source is located at the mirrored position of the original sound source using the reflecting plane as the mirror plane. Also secondary image sources of the initial image sources can be introduced and reflections of the second-order, third-order etc. can be calculated. The more surfaces the acoustic model of the room contains, the more image sources have to be constructed. Obviously, for computational reasons, the exclusive usage of ISM is suitable only for smaller rooms with a not too complicated geometry and with a limited number of reflections (Allen and Berkley, 1979).

In RTM, a large number of rays are sent from a point source in a large number of directions, according to the user defined directivity of the sound source, e.g. a loudspeaker. The trajectory of reflections from the boundary surfaces are determined according to Snell's law. The intensity I of a ray decreases with its travel distance according to the classical geometrical attenuation of a point source ($I \sim \frac{1}{r^2}$) and is reduced at every reflection according to the absorption coefficient of the incident surface. Each ray is computed until its intensity is below a certain threshold or until it has been reflected a number of times. Scattering of sounds is introduced in these models by using a scattering coefficient, which is defined as the ratio between the sound energy reflected in non-specular reflections to the total reflected sound energy. To obtain a result related to a specific receiver position, an area is defined around the receiver which 'catches' the particles when travelling by (Rindel, 2000).

A hybrid calculation method based on the combination of ISM and RTM has already proven to be a useful tool in terms of assessing the general acoustic comfort by end users during architectural room design (Rindel, 2000, 2004).

However, the amount of research done on validating acoustic models in terms of evaluating localization performance in these simulated environments is non-existing or at least very limited. Therefore, part of this study aims at validating the hybrid calculation method implemented in the ODEON software.

The main research questions addressed in this chapter are:
(i) What is the impact on localization performance when subjects are localizing sound sources in a natural way compared to localizing stimuli presented under headphones, generated by measured impulse responses or by impulse responses constructed by a virtual acoustics software package, namely ODEON.
(ii) What is the difference in localization performance when using measured impulse responses or stimuli generated by ODEON.
(iii) What is the influence of reverberation and critical distance on localization performance. To study the effect of distance, two different distances between the loudspeakers and the subjects are used.

3.2 Methods

3.2.1 Test setup

Evaluations were performed in two extreme acoustic environments: an anechoic room and a very reverberant room. In both the anechoic and the reverberant

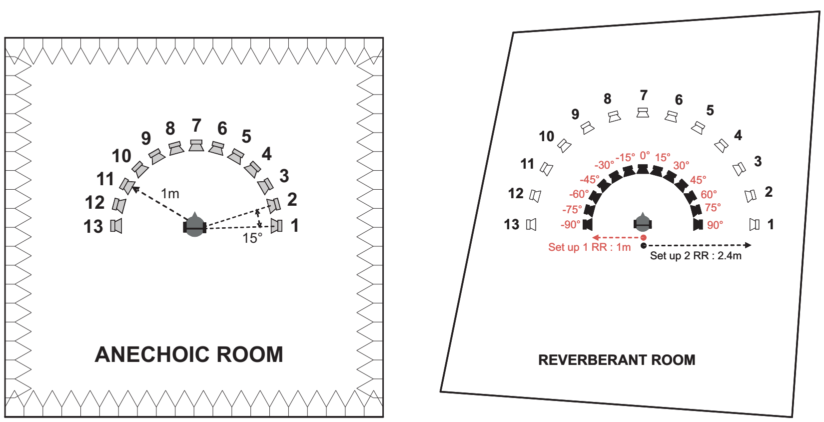

Figure 3.1 — Left: setup anechoic room. Right: setup in the reverberant room. An array of 13 speakers was positioned in the frontal horizontal hemisphere of the subject or manikin. The speakers were placed at 1m of the subject. In the reverberant room, a second array of loudspeakers was used with a radius of 2.4m. In all setups the impulse responses between the loudspeakers and a CORTEX MK2 manikin were measured. These impulse responses were used to generate stimuli which were presented under headphones. The impulse responses measured in the anechoic room were also used in ODEON to generate a second set of headphone stimuli.

room an array of 13 single cone FOSTEX 6103B speakers were placed in the frontal horizontal hemisphere at a distance of 1m around the subject. This was done in steps of 15°. To study the influence of distance on localization performance, a second loudspeaker array with a radius of 2,4m was also used in the reverberant room. Both of the rooms were acoustically shielded from outdoor noise (noise level $<$ 30dBA). The different test setups are illustrated in Figure 3.1. The reverberation time T_{60} of the reverberant room was experimentally determined by the T_{30} method described in ISO3382 (1997). This was done using an omni-directional point source, BK 4295, and an omni-directional microphone, BK 2642. The results are shown in Figure 3.2. During this measurement, equipment, such as the array of loudspeakers, and a test subject were kept in the room, in order to approach the acoustics present during the listening tests. The accompanying critical distance of the reverberant room for frequencies between 250Hz and 4000Hz varied between 0.3m and 0.5m respectively.

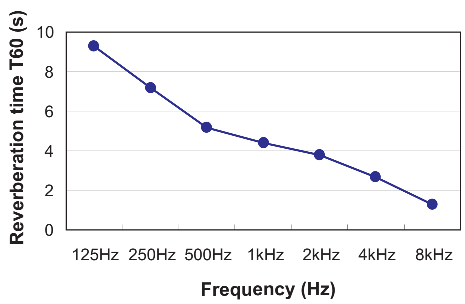

Figure 3.2 — Reverberation time, T_{60}, measured in the reverberant environment.

3.2.2 Subjects

Seven normal hearing subjects between 24 and 50 years old were evaluated. All subjects had maximum hearing thresholds of 15dB HL measured at all octave frequencies between 125Hz and 8kHz.

3.2.3 Stimuli

Three different stimuli were used, identical to the ones used in chapter 2: a 200-ms, 1/3-octave, low-frequency noise band centered around f=500Hz, a 200-ms, 1/3-octave, high-frequency noise band centered around $f = 3150Hz$ and a 1s broadband telephone ringing signal containing a lot of transient information (Figure 2.8). This was done to enable a frequency dependent analysis of the results.

BRIRs were measured between the array of loudspeakers, described in section 3.2.1, and both eardrums of an artificial head (CORTEX MK2) in all three test environments (anechoic room, 1m distance = AR1, reverberant room 1m distance = RR1, reverberant room 2.4m distance = RR2). The obtained impulse responses were convolved with the different stimuli to generate the first set of headphone stimuli which are abbreviated as HPM (headphone - measured). The anechoic HRTFs were also inserted in the ODEON software together with a model of the reverberant environment, enabling ODEON to generate the set of virtual stimuli (HPO, headphone - ODEON). Both sets of headphone stimuli were de-convolved with an earlier recorded transfer function between the headphones and the eardrum to avoid taking the transfer characteristic of the ear canal into account twice during headphone evaluations. To generate the stimuli a non-commercial, modified, version of ODEON 8.0 was used. The modifications made for these evaluations are now commercially available in ODEON 9.0. An overview of the different test conditions, used for each type of stimulus, is given in Table 3.1.

ODEON is based on combining ISM and RTM. The simulation of the room impulse response (RIR) is performed in two steps. The first part, which contains information about early reflections, is calculated by ISM. This is typically done due to the accuracy of ISM in finding reflection paths. However, the number of image sources grows exponentially as a function of the reflection order which leads to a computationally inefficient way to find the higher order reflections. The duration of the early part can be chosen by varying the so-called transition order which is the maximum number of image sources taken into account per initial ray. The second part of the RIR, the part containing the late reflections, is calculated by a modified ray-tracing algorithm that takes into account the scattering coefficient of the involved surfaces. At every reflection event, local diffuse secondary sources are generated that radiate sound with a directivity according to Lambert's cosine-law (Zeng et al., 2006). Finally, the BRIR is calculated by convolving the calculated RIR of each arriving ray with the anechoic HRTF for the corresponding direction. The anechoic HRTFs can be inserted by the user, enabling the use of individualized HRTFs, of HRTFs measured on an artificial head or of HRTFs generated by e.g. boundary element methods.

OE	Loudspeaker presentation, own ears condition
HPM	Headphones, manikin measured impulse responses
HPO	Headphones, ODEON generated impulse responses
AR1	Anechoic room, loudspeakers are at 1m distance
RR1	Reverberant room, loudspeakers are at 1m distance
RR2	Reverberant room, loudspeakers are at 2.4m distance

Table 3.1 — An overview of the different test conditions which were used for each type of stimulus.

3.2.4 Simulation of the Reverberant Room

A geometrical computer model of the reverberant room was built based on the measured dimensions of the room (volume V = 198 m^3). The acoustic model was calibrated by using the measurements of the reverberation time. The loudspeakers were simulated using the proper directivity and spectrum, measured in an anechoic room. The acoustical properties of the subjects were defined by the measured anechoic HRTFs on the artificial head. The simulations were performed using 6000 rays, a maximum reflection order of 2000 and a transition order of 2.

3.2.5 Protocol

A similar protocol was used as the one defined in chapter 2. Each individual subject was seated in the center of the array of 13 loudspeakers which were located in the frontal horizontal plane and were labelled by numbers 1 to 13 (Figure 3.1). Subjects were seated such that the height of the eardrums matched the height of the speaker array. A test-retest procedure was used with 3 repetitions of each stimulus in each test session. A roving level of 6dB was used to avoid monaural localization cues. Subjects were instructed to keep their head fixed and pointed towards 0° during stimulus playback. They were watched on a monitor. The task of the subject was to identify where the sound was perceived.

In the anechoic environment two perceptual evaluations were performed, i.e. one with the sound samples presented through the array of loudspeakers (OE, own ears) and one with the HPM stimuli presented over headphones, i.e. a SENNHEISER HD650 which is specifically designed for an accurate spatial representation. In the reverberant environment an extra evaluation was added which consists of localizing the stimuli generated by ODEON. All evaluations were done by using ALP (see section 2.3.1). The RMS error measure, defined in section 2.3.1, was used to evaluate the localization performance of the different subjects.

3.3 Results and analysis

Tables 3.2, 3.3 and 3.4 contain the individual and averaged localization results of the 7 normal hearing subjects in the different test conditions. The average values of the test and re-test condition are given since no significant difference between test and re-test were observed. The averaged results of all evaluations are summarized in Figure 3.3. All the data were analyzed using SPSS 15.0. For conciseness the term "factorial repeated measures ANOVA" is abbreviated by "ANOVA" and all reported pairwise comparisons were Bonferroni corrected for multiple comparisons. The reported p-values are lower bound values and

RMS-error (°) - Anechoic room, distance = 1m (AR1)

	500Hz		3150Hz		Telephone	
	OE	HPM	OE	HPM	OE	HPM
S1	12.9	12.9	19.3	17.9	8.0	6.8
S2	10.6	9.5	16.7	22.6	6.4	6.8
S3	12.1	10.8	17.0	21.8	6.6	9.9
S4	8.2	8.8	13.9	16.6	6.2	7.0
S5	8.4	14.8	16.0	20.7	7.2	9.8
S6	15.2	14.6	18.8	23.6	10.3	10.9
S7	9.0	8.3	18.9	24.7	6.8	3.3
average	10.9	11.4	17.2	21.1	7.3	7.8
stdev	2.6	2.7	1.9	2.9	1.4	2.6

Table 3.2 — The individual and averaged RMS data (°) of the normal hearing subjects for the different test conditions. Tests were done in an anechoic room with a distance between the loudspeakers and the subject of 1m (AR1).

RMS-error (°) - Reverberant room, distance = 1m (RR1)

	500Hz			3150Hz			Telephone		
	OE	HPM	HPO	OE	HPM	HPO	OE	HPM	HPO
S1	11.6	12.1	15.0	13.1	13.9	20.0	9.0	10.8	11.6
S2	10.4	10.0	9.8	17.1	22.9	27.6	8.1	5.6	6.9
S3	11.1	13.1	9.2	19.4	18.5	26.9	8.2	9.6	7.6
S4	7.6	8.1	10.0	14.3	16.0	17.5	5.5	7.0	9.3
S5	10.4	10.6	12.1	12.2	13.1	24.4	7.2	9.0	10.9
S6	15.1	18.4	18.2	11.1	18.4	25.3	11.0	10.3	17.4
S7	11.6	11.2	10.4	20.3	26.1	26.8	9.5	11.5	10.0
average	11.1	11.9	12.1	15.3	18.4	24.0	8.3	9.1	10.5
stdev	2.2	3.3	3.3	3.6	4.7	3.9	1.7	2.1	3.5

Table 3.3 — The individual and averaged RMS data (°) of the normal hearing subjects for the different test conditions. Tests were done in a highly reverberant room with a distance between the loudspeakers and the subject of 1m (RR1).

RMS-error (°) - Reverberant room, distance = 2,4m (RR2)

	500Hz			3150Hz			Telephone		
	OE	HPM	HPO	OE	HPM	HPO	OE	HPM	HPO
S1	14.5	13.6	11.8	19.1	28.5	20.1	9.0	8.5	11.7
S2	13.6	12.6	12.5	18.0	27.1	29.9	9.0	9.6	13.2
S3	13.2	13.6	13.5	17.6	25.9	22.3	9.2	10.6	14.0
S4	13.0	10.6	11.9	17.0	26.0	21.1	7.3	10.2	11.7
S5	11.9	11.6	10.6	17.8	22.4	19.1	8.6	10.8	13.8
S6	21.4	21.7	26.0	23.5	25.6	30.5	10.1	13.1	24.1
S7	18.9	17.1	16.7	17.7	25.9	25.4	16.1	17.6	17.6
average	15.2	14.4	14.7	18.7	25.9	24.0	9.9	11.5	15.1
stdev	3.5	3.8	5.3	2.2	1.9	4.6	2.9	3.0	4.4

Table 3.4 — The individual and averaged RMS data (°) of the normal hearing subjects for the different test conditions. Tests were done in a highly reverberant room with a distance between the loudspeakers and the subject of 2.4m (RR2)

a significance level of p=0.05 was used throughout this chapter. First a short overview of the presented data is given. Later the ANOVAs, used to examine the research questions, are presented.

Tables 3.2, 3.3 and 3.4 illustrate some of the observations already made in chapter 2. These tables show that the high-frequency stimulus, of which the localization performance is based on ILD information, was localized the least accurately. The broadband telephone stimulus was localized best and the performance when localizing the 500Hz noise band approached the performance obtained when localizing the telephone signal. Moreover, it is observed that the range of responses, especially those of the anechoic data, correspond very well with the normal hearing data presented in Table 2.2.

A second observation is that the localization accuracy of the normal hearing subjects was not drastically influenced neither by reverberation nor by how the stimuli were generated or presented to the subjects. This motivates the use of recorded or ODEON generated impulse responses during the first evaluation stages of a signal processing algorithm. Statistical analyses were performed to thoroughly evaluate the data. In the first analysis, the difference between natural localization and performing headphone experiments with recorded impulse-responses using a CORTEX MK2 manikin was examined. In the second analysis, the localization performance when using recorded impulse responses was compared to the condition in which impulse responses are generated by ODEON.

Natural localization vs. measurements CORTEX MK2 Manikin

The main question related to this dissertation is whether using impulse-responses measured in an acoustic environment with a CORTEX MK2 manikin has a large influence on localization performance. This was evaluated by examining the own ear (OE) data and the data of the measured impulse responses (HPM) gathered in the three different acoustic settings, abbreviated as AR1, RR1 and RR2. The natural localization data, also referred to as own ear data (OE), and the HPM data, were inserted in an ANOVA using the following main factors: stimulus type (3150Hz, 500Hz, telephone signal), acoustic environment (AR1, RR1 and RR2), stimulus presentation (OE and HPM) and the test-retest factor.

First, an interaction between the main factors stimulus type and stimulus presentation was observed (p=0.004). Therefore, separate ANOVAs were performed for each stimulus. In these three ANOVAs no interactions were observed. The results are summarized in Table 3.5. This table illustrates that, when evaluating the 3150Hz stimulus, a significant decrease in localization performance was detected when the stimuli were generated using the impulse responses measured with the artificial head. For the 3150Hz centered noise

	500Hz	3150Hz	Tel
AR1 vs RR1	1.000	0.354	nme
AR1 vs RR2	0.074	0.022*	nme
RR1 vs RR2	0.014*	0.047*	nme
OE vs HPM	0.744	<0.001*	0.067

Table 3.5 — p-values of the pairwise comparisons using the data OE and HPM for the three different acoustic settings. A significant difference between OE and HPM was only observed for the 3150Hz centered noise band. The term "nme" indicates that no main effect was found, hence no pairwise comparisons were performed.

	500Hz	3150Hz		Tel	
		RR1	RR2	RR1	RR2
RR1 vs RR2	0.015*	-	-	-	-
OE vs HPM	nme	0.126	0.001*	nme	0.044*
OE vs HPO	nme	0.003*	0.030*	nme	0.042*
HPM vs HPO	nme	0.021*	0.966	nme	0.094

Table 3.6 — p-values of the pairwise comparisons using the data OE, HPM and HPO in the reverberant environment. A significant difference between HPM and HPO was only observed when localizing the 3150Hz centered noise band arriving from 1m distance of the subject. The term "nme" indicates that no main effect was found, hence no pairwise comparisons were performed.

band, the OE condition significantly outperformed the HPM condition by 4.7° RMS, averaged of the three acoustic environments.

During this analysis no significant effect of reverberation was observed since the performance of AR1 was not significantly different from RR1 for all stimuli. However, the significant difference between RR1 and RR2, found for both small-band stimuli, do suggest an impact of loudspeaker distance on the localization performance. For the telephone signal, no main effect on acoustic environment was found and hence no pairwise comparisons were calculated.

Measurements CORTEX MK2 Manikin vs. ODEON software

A second research question motivating these experiments is whether a commercial software package, ODEON, is able to produce virtual sound signals which can be used to perform reliable localization experiments. This was analyzed by using the ODEON data gathered in the reverberant room. An ANOVA was performed using the main factors stimulus type (3150Hz, 500Hz, telephone signal), acoustic environment (RR1 and RR2), stimulus presentation (OE, HPM and HPO) and the test-retest factor.

Since interactions were found between stimulus type and stimulus presentation, separate ANOVAs were performed for each stimulus. In these ANOVAs interactions were found between stimulus presentation and acoustic environment for the 3150Hz centered noise band and the telephone signal, motivating separate ANOVAs for each acoustic environment for these stimuli. The results of all analyses are presented in Table 3.6.

Table 3.6 shows that a significant difference in localization performance between the HPM and the HPO condition was only found when localizing the high frequency stimulus in acoustic environment RR1. In all other experiments, both headphone stimuli yielded the same localization performance. The other pairwise comparisons are somewhat harder to interprete. It is observed that, in general, headphone experiments seem to reduce localization performance of the 3150Hz stimulus but not of the 500Hz stimulus which was also concluded in one of the previous paragraphs (OE vs. HPM and HPO). When localizing a telephone stimulus this trend was observed only in the condition RR2.

3.4 Discussion

Three research questions were formulated in section 3.1. The results and analysis of the previous section (Tables 3.2 to 3.6 which are summarized in Figure 3.3) are used to answer these questions.

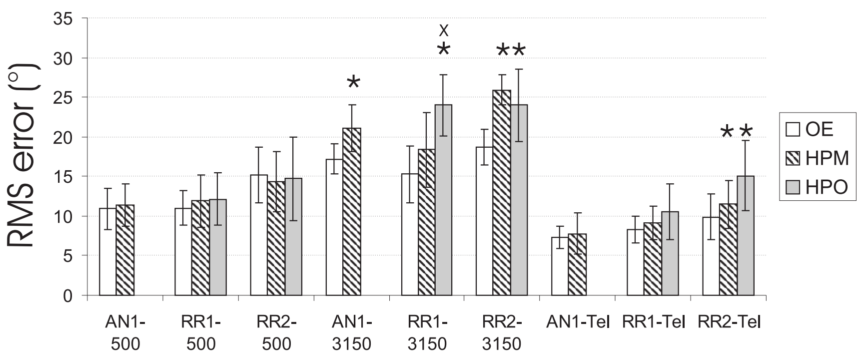

Figure 3.3 — Summary of the mean localization performance of the 7 normal hearing subjects together with their inter-subject standard deviation. The full data sets are given in Tables 3.2 to 3.4. A '*' depicts a significant difference with the OE condition. A 'x' depicts a significant difference in performance between both headphone conditions, HPM and HPO. Significant differences were mainly present when localizing high-frequent stimuli. Overall, the differences between conditions were fairly small, especially when localizing broadband or low frequency stimuli. The analyses done are intra-subject analyses (Tables 3.5 and 3.6), therefore they are not related to the shown inter-subject standard deviation.

What is the difference in localization performance when localizing sound sources naturally or when using headphones.

Two different headphone conditions were evaluated. In the first condition, condition HPM, impulse-responses between the loudspeakers and an artificial head were measured in the acoustic environment of the localization experiment. Afterwards, these impulse-responses were used to generate the headphone stimuli. In the second condition, condition HPO, commercial virtual acoustics software was used to generate the reverberant stimuli used in the headphone experiment. This included the combination of measuring HRTFs of an artificial head in an anechoic environment and modelling the room in the ODEON software package.

Table 3.5 summarizes the statistical analysis done on the data of the anechoic room (AR1) and both settings in the reverberant room (RR1 and RR2). This table indicates that the low frequency noise was localized equally well under headphones as in real-life. This is confirmed by the data in Table 3.6 which show that, for this stimulus, the performance of both headphone conditions was similar to the natural localization.
This suggests that the ITD cues measured between the microphones of the CORTEX MK2 and which are determined by the positioning of the ear simulators in the manikin, sufficiently approached the ITD cues normally used by the evaluated subjects.

The data of the high-frequency noise component on the other hand shows that, in general, a significant decrease in localization performance was present when localizing sounds under headphones compared to using own ears. This is observed in Table 3.5 and in 3 out of 4 pairwise comparisons in Table 3.6. When examining the data in Tables 3.2, 3.3 and 3.4 it is observed that the decrease in performance was subject dependent. This decrease was often in the order of 8 to 9° with a maximum decrease in performance of 14.2° (S6, Table 3.3).
Localization of the 3150Hz centered noise band is mainly based on using ILD information which is introduced by diffraction and reflections of sounds around and on the head and torso of a human listener. The observed significant decrease in localization performance may be explained by differences in acoustical properties between an artificial head and a human listener. These differences are due to differences in shape, in material (the artificial head is made out of a hard synthetic material) and due to a lack of clothing and hair (Treeby et al., 2007) which all may have a significant impact on ILD cues.

For a broadband telephone signal, it was observed that when taking all environments (Table 3.5) into account, no significant decrease in localization performance was found. When isolating the data of the reverberant environment (Table 3.6) a significant difference was observed, but only if the sound sources were placed at 2.4m distance from the subject. However, when analyzing the differences made in this condition (Table 3.4), it was observed that the intra-subject differences between conditions OE and HPM were only in the range of -0.5 to 3.0°. The intra-subject differences between OE and HPO were in the range of 1.5 to 5.2° RMS except for subject S6 which showed a decrease of 14.0° RMS. These differences may be interpreted as being acceptable, depending on the experiment one wishes to perform.

It can be concluded that significant differences can be present between naturally localizing sound sources or localizing sound sources under headphones. However, these differences are relatively small and are mainly present when localizing small-band high frequency stimuli. When using lower frequencies or broadband stimuli, no or smaller differences were observed. These differences seem to be originating from the use of an artificial head.
This conclusion is supported by two studies of Møller et al. In the first study (Møller et al., 1996) no significant difference in localization performance was found between natural localization and headphone experiments if individual HRTF recordings were used. In the second study (Møller et al., 1999), a significant decrease in localization performance was found when using recordings of 8 different artificial heads (not including the CORTEX MK2) in the same localization setup. By isolating the so-called out-of-cone errors, which are related to the horizontal localization performance studied in this chapter, 7 out of 8 artificial heads introduced a decreased localization performance. A second evaluation, in the same localization setup, by Minnaar et al. (2001), using more recent artificial heads, demonstrated that artificial head recordings were improving and were approaching real-life performance. However, artificial heads are designed to mimic an average human subject. Since the approximation of the shape of a human subject by an artificial head is highly subject dependent, a significant across-subject variance and small localization errors should be expected.

What is the difference in localization performance when generating stimuli with measured impulse-responses or with virtual acoustics software.

To resolve this question, evaluations were performed using HPM and HPO stimuli in a reverberant environment. Table 3.6 indicates that a significant difference between HPM and HPO was only observed when localizing high frequency sounds arriving from 1m from the subject. The introduced errors by the ODEON software for this specific condition were in the range of 0.7° to 11.3°. In contrast with condition RR1, ODEON did not introduce a large decrease in localization performance in condition RR2. Moreover, in this condition, performance even improved for 5 out of 7 subjects. This seems contradictory. However, when examining the HPM data of the 3150Hz noise band, which was the reference condition for this research question, it is observed that the RR1 condition significantly outperformed the RR2 condition. By using the ODEON software the performance of RR1 decreased to the level of RR2. In other words, ODEON significantly decreased localization performance of the narrow-band high frequency stimulus if a high performance was obtained at the baseline condition. When using low and broadband stimuli no significant differences were observed (Table 3.6).

It can be concluded that the ODEON software introduces an amount of binaural cue distortion in comparison with the HPM stimuli when generating narrow-band high-frequency signals. These distortions lead, to a certain extent, to a significant decrease in localization performance. If the originally obtained HPM localization performance is moderate, ODEON will not introduce an additional decrease in performance.

What is the influence of reverberation on the natural localization of sound sources or on the localization of sound sources under headphones.

During these experiments, two different extreme acoustic environments were studied: an anechoic room and a highly reverberant room. The data in Table 3.5 illustrates that, for all stimuli, reverberation time as such did not influence localization performance (AR1 vs RR1). However, interestingly a significant difference was observed when increasing the distance between the loudspeakers and the subject (RR1 vs RR2) in the reverberant room.

The fact that the reverberation time did not affect localization performance can be explained by the so called precedence effect which is also known as the law of the first wavefront. For an overview on the precedence effect see Litovsky et al. (1999). The precedence effect is based on the ability of the human auditory system to associate the direction of arrival of a sound source with the direction of arrival of the direct sound. The time interval in which this takes place is called the fusion zone. Outside this time interval (around 40ms depending on the frequency content), reflecting sounds are perceived as echo's which have their own direction of arrival. These evaluations demonstrate that the precedence effect can avoid significant differences in localization performance between two very extreme acoustic environments.

An explanation for the significant difference in performance between conditions RR1 and RR2 is less straightforward to give. Two possibilities come to mind. The first explanation might be that in this particular setup localizing sound sources positioned at 2.4m distance was a less accurate process than localizing sound sources at 1m distance. This can not be concluded from the gathered data since no evaluations were done in a 2.4m setup in the anechoic environment. Moreover, Brungart and Rabinowitz (1999) claim that ITD and ILD cues are virtually independent of distance if the sound sources are positioned beyond 1m distance from the listener. A second explanation is that due to the low direct-to-reverberant or direct-to-total sound ratio, the first wavefront in condition RR2 did not contain enough information for a correct localization of the sound source giving rise to a small but significant decrease in performance.

Additional measurements were made to quantify the effective direct-to-total ratio in the different test conditions. The direct-to-total ratio is defined as

$$D = \frac{\int_0^{t_{dir}} p^2(t)dt}{\int_0^{\infty} p^2(t)dt} \tag{3.1}$$

and was measured by isolating the direct sound from the total sound in the measured impulseresponses between the loudspeakers and the ears of the artificial head. The direct-to-total ratio, unlike the reverberation time which was given in Figure 3.2, takes into account the properties of the loudspeakers and the manikin used during the evaluations. Due to the directionality of the loudspeakers and the use of an artificial head to record the impulseresponses, the direct-to-total ratio is highly dependent on the positioning of the loudspeaker relative to the artificial head. Therefore, this measurement was done for all loudspeaker positions used during the different evaluations. The measurements made in RR1 and RR2 for two extreme angles of arrival, i.e. 0° and 90°, are illustrated in Figure 3.4. This figure shows a large drop in direct-to-total ratio when the distance between the loudspeaker and the artificial head was increased from 1m to 2.4m. This drop was present at all frequencies and at all the different angles of arrival of the sound source. This drop might have been large enough to significantly decrease localization performance. However, more fundamental research on localization in extreme reverberant conditions seems appropriate to validate this assumption.

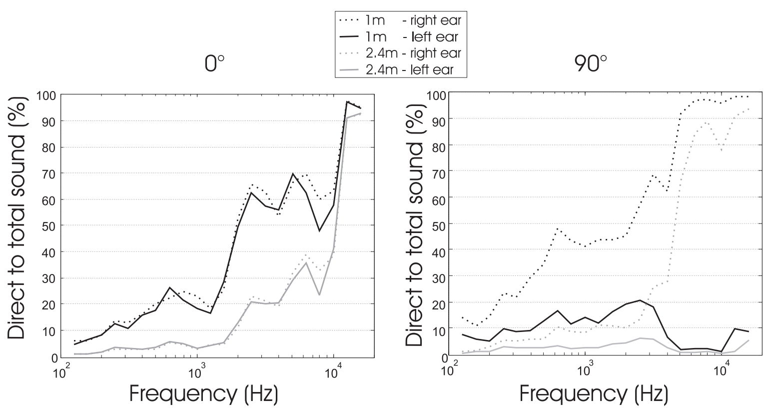

Figure 3.4 — The amount of direct sound relative to the total amount of sound. This is given for the loudspeaker placed in front of the subject (0° - left figure) and for the loudspeaker placed at the right side of the subject (90° - right figure). In both figures a large drop in the direct-to-total sound ratio is observed when the distance between loudspeaker and artificial head was increased from 1m to 2.4m. The direct-to-total ratio strongly depended on the positioning of the loudspeaker relative to the ears of the artificial head.

3.5 Conclusion

In this chapter, a comparison was made between localizing sounds in a natural way and localizing sound sources under headphones. Two different headphone conditions were evaluated. In the first condition impulse-responses were measured between loudspeakers and an artificial head. These impulse-responses were then used to generate the headphone stimuli (HPM). A second set of headphone stimuli were generated by using ODEON, a commercial software package enabling the acoustic modelling of virtual environments (HPO).

Three research questions were addressed in this chapter. First, it was observed that headphone experiments tend to preserve the localization performance when localizing low frequency signals or a broadband telephone signal. A significant decrease in localization performance is introduced when localizing narrow band high frequency stimuli. The use of an artificial head during these experiments may explain these differences. The acoustical properties of an artificial head are different from those of a human listener due to differences in material, shape (an artificial head only represents an average human), etc. thereby generating less individually suited ILD cues.

Second, it was observed that ODEON introduced a significant decrease in localization performance compared to measured impulse-responses in only one out of six test conditions. It was concluded that ODEON may introduce an amount of uncertainty when localizing high frequency sound sources. A drop in localization performance was only observed if the originally obtained localization performance was high enough.

Third, it was observed that due to the precedence effect reverberation time had no influence on localization accuracy. However, a decrease in performance for low and high frequency stimuli was observed if the distance between the loudspeakers and the subject was increased from 1m to 2.4m in the reverberant room. This might have been a consequence of the very low direct-to-reverberant ratio in the latter condition.

Finally, it can be concluded that the use of headphones may have a significant impact on localization experiments in the frontal horizontal hemisphere, especially when using narrow band high-frequency signals. However, in general, these differences tend to be small and are often, depending on the research question, acceptable. Since almost no significant impact was observed when using headphone experiments while localizing broadband stimuli, headphone experiments will be used in the remainder of this manuscript to evaluate newly developed algorithms in terms of localization performance (section 4.5.2 and 5).

Chapter 4

Preserving binaural cues with a multichannel Wiener filter approach: MWF, MWF-N and MWF-ITF

In this chapter, a theoretical framework for binaural noise reduction with multichannel Wiener filter (MWF) based algorithms in hearing aids is developed. Three different binaural algorithms are presented: the binaural MWF, the multichannel Wiener filter with partial noise estimation (MWF-N) and the multichannel Wiener filter with interaural transfer function extension (MWF-ITF). In Spriet et al. (2004), Spriet et al. (2005) and Doclo et al. (2007), it was illustrated that the MWF can be used as a noise reduction strategy for monaural hearing aids. By extending this algorithm to a binaural framework, the total number of microphones increases, which may improve noise reduction performance. Moreover, the communication between hearing aids may facilitate the preservation of binaural cues.

In this chapter, it is shown that the binaural MWF preserves the binaural cues of the speech component but not of the noise component. Therefore two different extensions of the MWF algorithm are presented which aim at preserving the cues of both the speech and the noise component, i.e. the MWF-N and the MWF-ITF. In chapters 5 and 6, the MWF and MWF-N are further validated through objective and perceptual evaluations. The MWF-ITF is currently still under further development. However, pilot experiments are reported in this manuscript illustrating its potential.

From a commercial point of view, the binaural link is preferred to be a wireless connection between both hearing aids. Transmitting all microphone signals is therefore highly demanding in terms of bandwidth requirements and power consumption. Some alternatives of the binaural MWF are described. These algorithms aim at approaching the performance of the binaural MWF while reducing the necessary bandwidth of the binaural link.

Section 4.1 briefly recapitulates the research goals which motivated the design of binaural MWF-based algorithms.

In **section 4.2**, the binaural framework in which the algorithms are developed is described together with the theoretical performance measures used to predict the noise reduction performance and the ITD and ILD errors generated by the different algorithms.

In **section 4.3**, the binaural MWF is described. It will be shown that the binaural MWF inherently preserves the binaural cues of the speech component but distorts the binaural cues of the noise component into those of the speech component.

In **section 4.4**, the binaural MWF-N is presented. The MWF-N aims at eliminating only a portion of the noise component. The remaining, unprocessed, part restores the binaural cues of the noise component of the signal at the output of the algorithm.

In **section 4.5**, the binaural MWF-ITF is presented. The MWF-ITF adds an extra term to the cost function of the MWF. This term constrains the Wiener solution, to some extent, to filters which preserve the ITF, hence also the binaural cues, of the noise component.

Section 4.6 presents an overview of binaural MWF strategies that aim at reducing the necessary bandwidth of the binaural link while preserving an optimal noise reduction.

The research presented in this chapter has been published in Doclo et al. (2005), Doclo et al. (2006), Klasen et al. (2006), Klasen et al. (2007), Van den Bogaert et al. (2007), Van den Bogaert et al. (2008*a*). A journal paper on the theoretical analysis of the different algorithms is in preparation. A journal paper on the reduced bandwidth algorithms is currently under revision (Doclo et al., 2008).

4.1 Introduction

Multi-microphone, typically adaptive, noise reduction algorithms currently used in hearing aids are designed to optimize the SNR in a monaural way, and not to preserve the binaural cues. As a consequence, as illustrated in chapter 2, hearing aid users often localize sounds better when switching off the adaptive directional noise reduction (Keidser et al., 2006; Van den Bogaert et al., 2006). This puts the hearing aid user at a disadvantage. In certain situations, such as traffic, incorrect localization of sounds may even endanger the user. In addition, binaural cues and spatial awareness are important for speech segregation in noisy environments due to spatial release from masking effects, a.k.a. the cocktail-party effect (Bronkhorst and Plomp, 1988, 1989).

In general, the goal of a Wiener filter is to filter out noise corrupting a desired speech signal. By using the second-order statistical properties of the desired signal and the noise, the optimal Wiener filter can be calculated. It generates an output signal which approaches the desired signal as well as possible in a mean-square-error (MSE) sense. The main advantages of using MWF-based strategies, as already mentioned in section 1.3.2, are that, in contrast with common beamforming strategies, the MWF does not require any a priori information of the desired signal (e.g. the location of the signal) nor of the microphone characteristics. In Doclo and Moonen (2002) and Spriet et al. (2005) it was shown that the MWF can be used for monaural hearing aid applications.

In this chapter it will be shown that the MWF can be extended to a binaural framework thereby using microphone signals of the contralateral and the ipsilateral hearing aid to generate an enhanced output signal for the ipsilateral hearing aid. Since the binaural MWF is designed to produce two outputs, $Z_L(\omega)$ and $Z_R(\omega)$, estimating the speech component at the front omnidirectional microphone of the left and the right hearing aid, respectively, it may be assumed that the binaural cues of the speech component are inherently preserved. This chapter theoretically proves this statement and further analyzes the noise reduction performance of the binaural MWF and its effects on binaural cues. Additionally, two extensions of the MWF are presented, the MWF-N and the MWF-ITF, which perform noise reduction and at the same time aim to preserve the binaural cues of both the speech and the noise component. The noise reduction and localization performance of the MWF and MWF-N are thoroughly evaluated and compared to a standard ADM in chapters 5 and 6. The evaluation and further analysis of the MWF-ITF is currently further being developed. Data of pilot experiments with the MWF-ITF are presented in this manuscript showing the potential of this algorithm.

4.2 Binaural framework

4.2.1 Microphone configuration and output signals

Consider the binaural hearing aid configuration in Figure 4.1, which consists of a left and a right hearing aid, both having a microphone array of respectively M_L and M_R microphones. The m-th microphone signal of the left ear, $Y_{L,m}(\omega)$,

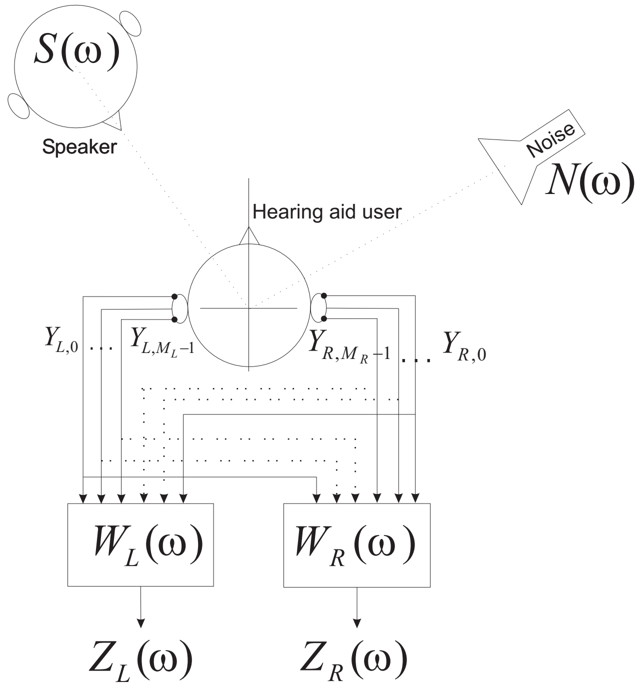

Figure 4.1 — Layout of a binaural noise reduction system which consists of two microphone arrays. In a binaural hearing aid design, microphone signals of the contralateral hearing aid may be used to enhance the SNR at the ipsilateral ear.

can be written in the frequency domain as

$$Y_{L,m}(\omega) = X_{L,m}(\omega) + V_{L,m}(\omega) \quad m = 0 \ldots M_L - 1, \tag{4.1}$$

where $X_{L,m}(\omega)$ and $V_{L,m}(\omega)$ represent the speech and the noise component at the m-th microphone input of the left hearing aid as a function of frequency ω. These components consist of the speech source signal, $S(\omega)$, and the noise source signal, $N(\omega)$, convolved with a room impulse response. $Y_{R,m}(\omega)$ is defined similarly for the right hearing aid. If a link between both hearing aids is present, all microphone signals from the ipsilateral and the contralateral hearing aid can be used to generate an output signal at the ipsilateral ear. The total number of microphones used is given by $M = M_L + M_R$. When assuming a typical binaural hearing aid setting, i.e. $M_L = M_R$, then for the left and the right hearing aid the M-dimensional input signal vector $\mathbf{Y}_L(\omega) = \mathbf{Y}_R(\omega) = \mathbf{Y}(\omega)$ can be written as

$$\mathbf{Y}(\omega) = \left[\begin{array}{cccccc} Y_{L,0}(\omega) & \ldots & Y_{L,M_L-1}(\omega) & Y_{R,0}(\omega) & \ldots & Y_{R,M_R-1}(\omega) \end{array}\right]^T. \tag{4.2}$$

This signal vector can be written as the sum of an M-dimensional vector representing the speech component in each microphone signal and an M-dimensional vector representing the noise component in each microphone signal,

$$\mathbf{Y}(\omega) = \mathbf{X}(\omega) + \mathbf{V}(\omega), \tag{4.3}$$

with $\mathbf{X}(\omega)$ and $\mathbf{V}(\omega)$ representing the M-dimensional input speech and noise components respectively. These vectors are defined in a similar way as $\mathbf{Y}(\omega)$. The output signal for the left and the right hearing aid, $Z_L(\omega)$ and $Z_R(\omega)$, is obtained by

$$\boxed{\begin{aligned} Z_L(\omega) &= \mathbf{W}_L^H(\omega)\mathbf{Y}(\omega) \\ Z_R(\omega) &= \mathbf{W}_R^H(\omega)\mathbf{Y}(\omega) \end{aligned}} \tag{4.4}$$

with $\mathbf{W}_L(\omega)$ and $\mathbf{W}_R(\omega)$ M-dimensional complex vectors representing the filters of respectively the left and the right hearing aid. The $2M$-dimensional stacked weighting vector $\mathbf{W}(\omega)$ is defined as

$$\mathbf{W}(\omega) = \begin{bmatrix} \mathbf{W}_L(\omega) \\ \mathbf{W}_R(\omega) \end{bmatrix} \tag{4.5}$$

By combining eq. 4.1 with 4.4, the output signal for the left ear can be written as

$$Z_L(\omega) = Z_{xL}(\omega) + Z_{vL}(\omega) = \mathbf{W}_L^H(\omega)\mathbf{X}(\omega) + \mathbf{W}_L^H(\omega)\mathbf{V}(\omega)\ , \tag{4.6}$$

where $Z_{xL}(\omega)$ and $Z_{vL}(\omega)$ represent the speech and the noise component at the output of the left hearing aid. Similarly, the output signal of the right hearing aid can be written as $Z_R(\omega) = Z_{xR}(\omega) + Z_{vR}(\omega)$. For conciseness, we will omit the frequency-domain variable ω in the remainder of this chapter.

4.2.2 Performance measures

Different performance measures are defined to theoretically evaluate the different algorithms. The ITF of the speech (noise) component is respectively defined as the ratio of the speech (noise) at the left and at the right hearing aid. The input and output ITF's of the speech and the noise component are defined as

$$\mathrm{ITF}_v^{in} = \frac{V_{L,r_L}}{V_{R,r_R}}, \qquad \mathrm{ITF}_v^{out} = \frac{Z_{vL}}{Z_{vR}} = \frac{\mathbf{W}_L^H\mathbf{V}}{\mathbf{W}_R^H\mathbf{V}} \tag{4.7}$$

$$\mathrm{ITF}_x^{in} = \frac{X_{L,r_L}}{X_{R,r_R}}, \qquad \mathrm{ITF}_x^{out} = \frac{Z_{xL}}{Z_{xR}} = \frac{\mathbf{W}_L^H\mathbf{X}}{\mathbf{W}_R^H\mathbf{X}} \tag{4.8}$$

with r_L and r_R defining the reference microphone, which is most commonly the front omnidirectional microphone, of the left and the right hearing aid respectively. Correspondingly, V_{L,r_L} is defined as $V_{L,r_L} = \mathbf{e}_{L,r_L}^H\mathbf{V}_L$ with

$$\mathbf{e}_{L,r_L} = [\ 0\ \ ...\ \ 0\ \ 1\ \ 0\ \ ...\ \ 0\ ]^T, \tag{4.9}$$

a vector containing a one at the r_L-th position and zeros elsewhere, or in other words $\mathbf{e}_{L,r_L}$ refers to the microphone which will be used as a reference microphone.

The impact on ITD cues is estimated by using the cross-correlation between the signals at the left and the right hearing aid, e.g. for the noise component

$$\begin{aligned} c_v^{in} &= \mathcal{E}\{V_{L,r_L} V_{R,r_R}^*\} = \mathbf{e}_{L,r_L}^H \mathbf{R}_{vv} \mathbf{e}_{R,r_R} \\ c_v^{out} &= \mathcal{E}\{Z_{vL} Z_{vR}^*\} = \mathbf{W}_L^H \mathbf{R}_{vv} \mathbf{W}_R \end{aligned} \tag{4.10}$$

with $\mathbf{R}_{vv} = \mathcal{E}\{\mathbf{V}\mathbf{V}^H\}$ the $M \times M$ correlation matrix of the noise component of the input signal. $\mathbf{R}_{vv}$ is defined as

$$\mathbf{R}_{vv} = \mathcal{E}\{\mathbf{V}\mathbf{V}^H\} = \begin{bmatrix} P_{v_0} & P_{v_0}P_{v_1} & \dots & P_{v_0}P_{v_{M-1}} \\ P_{v_1}P_{v_0} & P_{v_1} & \dots & P_{v_1}P_{v_{M-1}} \\ \dots & \dots & & \dots \\ P_{v_{M-1}}P_{v_0} & P_{v_{M-1}}P_{v_1} & \dots & P_{v_{M-1}} \end{bmatrix} \tag{4.11}$$

with P_{v_m} the power spectral density (PSD) of the noise component at the m-th microphone input of the hearing aid and P_{v_m,v_n} the cross power spectral density between the noise components at the m-th and the n-th microphone input of the hearing aid.

The ILD is defined as the power ratio between the signals at the left and the right hearing aid, e.g. for the noise component

$$\begin{aligned} L_v^{in} &= \frac{\mathcal{E}\{|V_{L,r_L}|^2\}}{\mathcal{E}\{|V_{R,r_R}|^2\}} = \frac{\mathbf{e}_{L,r_L}^H \mathbf{R}_{vv} \mathbf{e}_{L,r_L}}{\mathbf{e}_{R,r_R}^H \mathbf{R}_{vv} \mathbf{e}_{R,r_R}} \\ L_v^{out} &= \frac{\mathcal{E}\{|Z_{vL}|^2\}}{\mathcal{E}\{|Z_{vR}|^2\}} = \frac{\mathbf{W}_L^H \mathbf{R}_{vv} \mathbf{W}_L}{\mathbf{W}_R^H \mathbf{R}_{vv} \mathbf{W}_R}, \end{aligned} \tag{4.12}$$

The cross-correlation and ILD of the speech component are defined similarly.

The power transfer function of the noise and the speech component can be expressed as, e.g. for the left hearing aid

$$\begin{aligned} G_{vL} &= \frac{\mathcal{E}\{|Z_{vL}|^2\}}{\mathcal{E}\{|V_L|^2\}} = \frac{\mathbf{W}_L^H \mathbf{R}_{vv} \mathbf{W}_L}{\mathbf{e}_{L,r_L}^H \mathbf{R}_{vv} \mathbf{e}_{L,r_L}} \\ G_{xL} &= \frac{\mathcal{E}\{|Z_{xL}|^2\}}{\mathcal{E}\{|X_L|^2\}} = \frac{\mathbf{W}_L^H \mathbf{R}_{xx} \mathbf{W}_L}{\mathbf{e}_{L,r_L}^H \mathbf{R}_{xx} \mathbf{e}_{L,r_L}}. \end{aligned} \tag{4.13}$$

Hence, the SNR improvement at the left hearing aid can be expressed as

$$\Delta\mathrm{SNR}_L = \frac{G_{xL}}{G_{vL}}. \tag{4.14}$$

The power transfer functions (G_{vR} and G_{xR}) and the SNR improvement (ΔSNR_R) at the right hearing aid are defined similarly.

4.3 Binaural speech distortion weighted MWF: SDW-MWF

In general, the goal of a Wiener filter is to filter out noise that corrupts a desired signal. Using the second-order statistical properties of the desired signal and the noise, the optimal filter or Wiener filter can be calculated. It generates an output signal which optimally approaches the desired signal in an MSE sense. It is hence based on minimizing a cost function corresponding to the difference between the desired signal and the output of the algorithm. In contrast with a single channel approach, an MWF uses multiple input signals to compute a set of filters generating this output signal. This section presents the binaural MWF and describes its influence on binaural cues.

4.3.1 SDW-MWF solution

The aim of the binaural MWF is to produce, at each hearing aid, a minimum MSE estimate of the speech component arriving at a reference microphone. The MSE cost function J_{MSE} for the filter $\mathbf{W}_L$ estimating the speech component, X_{L,r_L}, in the reference microphone of the left hearing aid and the filter $\mathbf{W}_R$ estimating the speech component, X_{R,r_R}, in the reference microphone of the right hearing aid equals

$$\boxed{J_{MSE}(\mathbf{W}) = \mathcal{E}\left\{\left\|\begin{bmatrix} X_{L,r_L} - \mathbf{W}_L^H\mathbf{Y} \\ X_{R,r_R} - \mathbf{W}_R^H\mathbf{Y} \end{bmatrix}\right\|^2\right\}}. \tag{4.15}$$

with $\mathcal{E}$ the expected value operator. Minimizing $J_{MSE}(\mathbf{W})$ leads to the optimal filters $\mathbf{W}$ producing the best minimum MSE estimate of the speech component. Hence, since the speech component includes the room impulse response, no dereverberation is done.

The equation

$$J_{MSE}(\mathbf{W}_L) = \mathcal{E}\{|X_{L,r_L} - \mathbf{W}_L^H\mathbf{Y}|^2\} \tag{4.16}$$

can also be written as

$$J_{MSE}(\mathbf{W}_L) = \mathcal{E}\{|X_{L,r_L}|^2\}+\mathcal{E}\{\mathbf{Y}^H\mathbf{W}_L\mathbf{W}_L^H\mathbf{Y}\}-\mathcal{E}\{X_{L,r_L}\mathbf{Y}_L^H\mathbf{W}\}-\mathcal{E}\{\mathbf{W}_L^H\mathbf{Y}X_{L,r_L}^*\} \tag{4.17}$$

The optimal filter $\mathbf{W}_{MSE,L}$ is obtained by setting the derivative

$$\frac{\partial J_{MSE}(\mathbf{W}_L)}{\partial \mathbf{W}} = -2\mathcal{E}\{\mathbf{Y}X_{L,r_L}^*\} + 2\mathcal{E}\{\mathbf{Y}\mathbf{Y}^H\}\mathbf{W}_L \tag{4.18}$$

to zero. The optimal multi-dimensional Wiener filter is equal to

$$\mathbf{W}_{MSE,L} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{e}_{L,r_L} \tag{4.19}$$

with $\mathbf{R}_{yy} = \mathcal{E}\{\mathbf{Y}\mathbf{Y}^H\}$ the $M \times M$ correlation matrix of the noise component of the input signal and $\mathbf{R}_{yx} = \mathcal{E}\{\mathbf{Y}\mathbf{X}^H\}$ the $M \times M$ cross-correlation matrix of the input and the desired signal. Both are defined similar to R_{vv} (eq. 4.11).

At this point $\mathbf{R}_{yy}$ and $\mathbf{R}_{yx}$ are still unknown variables. Two assumptions are made to overcome this problem. First, it is assumed that the second-order statistics of the noise signal are sufficiently stationary. By using a robust VAD algorithm, noise-only observations can be made during speech pauses. As the noise is assumed sufficiently stationary, the noise correlation matrix, $\mathbf{R}_{vv}$, which can be estimated during noise only periods, can be used during subsequent speech and noise signals. Secondly, we assume that the speech and the noise signals are statistically independent, implying that

$$\mathbf{R}_{xv} = \mathcal{E}\{\mathbf{V}\mathbf{X}^H\} = \mathbf{0} \tag{4.20}$$

From this second assumption it can be verified that

$$\mathbf{R}_{yy} = \mathbf{R}_{xx} + \mathbf{R}_{vv} \quad \text{and} \quad \mathbf{R}_{yx} = \mathbf{R}_{xx} \tag{4.21}$$

such that the optimal filter can be written as

$$\mathbf{W}_{MSE,L} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{xx}\mathbf{e}_{L,r_L} \tag{4.22}$$

or

$$\boxed{\mathbf{W}_{MSE,L} = \mathbf{R}_{yy}^{-1}(\mathbf{R}_{yy} - \mathbf{R}_{vv})\mathbf{e}_{L,r_L}} \tag{4.23}$$

which can be solved since $\mathbf{R}_{yy}$ and $\mathbf{R}_{vv}$ can be estimated during 'speech and noise periods' and during 'noise only periods', respectively.

The optimal Wiener solution for the left and right hearing aid minimizing $J_{MSE}(\mathbf{W})$ (eq. 4.15) becomes

$$\boxed{\mathbf{W}_{MSE} = \mathbf{R}^{-1}\mathbf{r}} \tag{4.24}$$

with

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{yy} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{R}_{yy} \end{bmatrix} \tag{4.25}$$

and

$$\mathbf{r} = \begin{bmatrix} \mathbf{R}_{xx}\mathbf{e}_{L,r_L} \\ \mathbf{R}_{xx}\mathbf{e}_{R,r_R} \end{bmatrix} \tag{4.26}$$

Note that the optimal multi-dimensional Wiener filter takes into account both spatio-temporal and spectral information. This can be best demonstrated by assuming a homogeneous speech and noise sound field, i.e. the PSD of the speech and the noise components $P_{x_1} = \ldots = P_{x_{M-1}} = P_x$ and $P_{y_1} = \ldots = P_{y_{M-1}} = P_y$, then $P_{y_m} = P_y = P_x + P_v$ with $m = 0 \ldots M-1$, such that $\mathbf{R}_{yy} = P_y \mathbf{\Gamma}_y$ and $\mathbf{R}_{xx} = P_x \mathbf{\Gamma}_x$, with the coherence matrices $\mathbf{\Gamma}_y$ defined as

$$\mathbf{\Gamma}_y = \begin{bmatrix} 1 & \Gamma_{y_0 y_1} & \ldots & \Gamma_{y_0 y_{M-1}} \\ \Gamma_{y_1 y_0} & 1 & \ldots & \Gamma_{y_1 y_{M-1}} \\ \ldots & \ldots & & \ldots \\ \Gamma_{y_{M-1} y_0} & \Gamma_{y_{M-1} y_1} & \ldots & 1 \end{bmatrix} \tag{4.27}$$

and $\mathbf{\Gamma}_x$ defined similarly. Hence, the Wiener filter $\mathbf{W}_{WF}$ of 4.24 equals

$$\boxed{\mathbf{W}_{MSE,L} = \underbrace{\frac{P_x}{P_x + P_v}}_{spectral\ filtering} \underbrace{\mathbf{\Gamma}_y^{-1} \mathbf{\Gamma}_x e_{L,r_L}}_{spatial\ filtering} .} \tag{4.28}$$

SDW-MWF

By introducing an extra parameter μ, Doclo and Moonen (2002) and Spriet et al. (2004) introduced the monaural speech distortion weighted multi-channel Wiener filter (SDW-MWF). This minimizes the weighted sum of the residual noise energy and the speech distortion energy and hence provides a trade-off parameter between speech distortion and noise reduction. By using eq. 4.1 and introducing the parameter μ, the binaural SDW-MWF cost function can be written as

$$\boxed{J_{SDW-MWF}(\mathbf{W}) = \mathcal{E}\left\{ \left\| \begin{bmatrix} X_{L,r_L} - \mathbf{W}_L^H \mathbf{X} \\ X_{R,r_R} - \mathbf{W}_R^H \mathbf{X} \end{bmatrix} \right\|^2 + \mu \left\| \begin{bmatrix} \mathbf{W}_L^H \mathbf{V} \\ \mathbf{W}_R^H \mathbf{V} \end{bmatrix} \right\|^2 \right\}} \tag{4.29}$$

where the first term represents speech distortion energy and the second term represents the residual noise component. It can be seen that if all emphasis is put on noise reduction, i.e. if $\mu = \infty$, all noise will be removed. However, the solution in this particular case becomes $\mathbf{W} = \mathbf{0}_{2M}$ which implies that both the speech and the noise component are removed from the input signal. If $\mu = 0$ minimal speech distortion is present. However, no or only a minimal amount of noise reduction will be achieved. Note that if the trade-off parameter μ is set to $\mu = 1$, the SDW-MWF cost function (4.29) reduces to cost function (4.15).

The Wiener filter solution minimizing $J_{SDW-MWF}(\mathbf{W})$ equals

$$\mathbf{W}_{SDW-MWF} = \mathbf{R}^{-1}\mathbf{r} \tag{4.30}$$

with

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{xx} + \mu \mathbf{R}_{vv} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{R}_{xx} + \mu \mathbf{R}_{vv} \end{bmatrix} \tag{4.31}$$

and

$$\mathbf{r} = \begin{bmatrix} \mathbf{R}_{xx} \mathbf{e}_{L,r_L} \\ \mathbf{R}_{xx} \mathbf{e}_{R,r_R} \end{bmatrix} \tag{4.32}$$

with $\mathbf{e}_{L,r_L}$ and $\mathbf{e}_{R,r_R}$ vectors defining the reference microphone for the left and the right hearing aid (see 4.9).

Since the algorithms studied in this and in the following chapters are all based on the SDW-MWF, which is a generalized form of the MWF, the SDW-MWF algorithm will from now on be referred to as MWF for conciseness.

4.3.2 Theoretical analysis of the binaural MWF

As shown in the previous section, the monaural MWF can be extended in a straightforward way to the binaural MWF, thereby using microphone signals of the contralateral hearing aid to enhance the SNR at the ipsilateral hearing aid. This section will present a second advantage of the MWF. Since the MWF creates a MSE estimate of the speech component at the reference microphone of respectively the left and the right hearing aid, it inherently preserves the binaural cues of the speech component between these reference microphones, independent of the angle of arrival of the signal. However, it will be shown that the MWF changes the binaural cues of the noise component into those of the speech component.

Assuming that a single speech source is present, the speech signal vector $\mathbf{X} = \mathbf{A}S$, with the vector $\mathbf{A}$ containing the acoustic transfer functions between the speech source and the M microphones on the left and the right hearing aid (including head shadow effect, microphone characteristics and room acoustics) and S the speech signal. Hence, the speech correlation matrix

$$\mathbf{R}_{xx} = P_s \mathbf{A}\mathbf{A}^H \ , \tag{4.33}$$

is a rank-1 matrix with $P_s = \mathcal{E}\{|S|^2\}$ the power of the speech signal. Using the matrix inversion lemma,

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{CA}^{-1}\mathbf{B} + \mathbf{D}^{-1})^{-1}\mathbf{CA}^{-1}, \tag{4.34}$$

$(\mathbf{R}_{xx} + \mu\mathbf{R}_{vv})^{-1}$ can be written as

$$\begin{aligned} (\mathbf{R}_{xx} + \mu\mathbf{R}_{vv})^{-1} &= (P_s\mathbf{A}\mathbf{A}^H + \mu\mathbf{R}_{vv})^{-1} && (4.35) \\ &= \frac{1}{\mu}[\mathbf{R}_{vv}^{-1} - \frac{P_s\mathbf{R}_{vv}^{-1}\mathbf{A}\mathbf{A}^H\mathbf{R}_{vv}^{-1}}{\mu + \rho}] && (4.36) \end{aligned}$$

with the scalar ρ being defined as

$$\rho = P_s\mathbf{A}^H\mathbf{R}_{vv}^{-1}\mathbf{A} \tag{4.37}$$

such that

$$(\mathbf{R}_{xx} + \mu\mathbf{R}_{vv})^{-1}\mathbf{A} = \frac{\mathbf{R}_{vv}^{-1}\mathbf{A}}{\mu + \rho}. \tag{4.38}$$

By defining

$$\mathbf{P} = \mathbf{I}_M - \frac{P_s \mathbf{R}_{vv}^{-1} \mathbf{A} \mathbf{A}^H}{\mu + \rho}, \tag{4.39}$$

the expressions in 4.36 and 4.38 can be reformulated as

$$(\mathbf{R}_{xx} + \mu \mathbf{R}_{vv})^{-1} = \frac{\mathbf{P} \mathbf{R}_{vv}^{-1}}{\mu} \tag{4.40}$$

and

$$\mathbf{P} \mathbf{R}_{vv}^{-1} \mathbf{A} = \frac{\mu}{\mu + \rho} \mathbf{R}_{vv}^{-1} \mathbf{A}. \tag{4.41}$$

Using expression 4.31 and 4.40, the matrix $\mathbf{R}^{-1}$ can be written as

$$\mathbf{R}^{-1} = \frac{1}{\mu} \begin{bmatrix} \mathbf{P} \mathbf{R}_{vv}^{-1} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{P} \mathbf{R}_{vv}^{-1} \end{bmatrix} \tag{4.42}$$

such that, using

$$\mathbf{r} = \begin{bmatrix} \mathbf{R}_{xx} \mathbf{e}_{L,r_L} \\ \mathbf{R}_{xx} \mathbf{e}_{R,r_R} \end{bmatrix} = P_s \begin{bmatrix} \mathbf{A} A_{r_L}^* \\ \mathbf{A} A_{r_R}^* \end{bmatrix}, \tag{4.43}$$

together with eq. 4.41 and 4.42, the optimal filter in eq. 4.30 is equal to

$$\mathbf{W}_{MWF} = \mathbf{R}^{-1} P_s \begin{bmatrix} \mathbf{A} A_{r_L}^* \\ \mathbf{A} A_{r_R}^* \end{bmatrix} = \frac{P_s}{\mu + \rho} \begin{bmatrix} A_{r_L}^* \mathbf{R}_{vv}^{-1} \mathbf{A} \\ A_{r_R}^* \mathbf{R}_{vv}^{-1} \mathbf{A} \end{bmatrix}. \tag{4.44}$$

This leads to

$$\boxed{\begin{aligned} \mathbf{W}_{MWF,L} &= \frac{P_s \mathbf{R}_{vv}^{-1} \mathbf{A}}{\mu + \rho} A_{r_L}^*, \\ \mathbf{W}_{MWF,R} &= \frac{P_s \mathbf{R}_{vv}^{-1} \mathbf{A}}{\mu + \rho} A_{r_R}^* \end{aligned}} \tag{4.45}$$

with $A_{r_L}^*$ defined as the complex conjugate of the r_L-th element, which is the element of $\mathbf{A}$ referring to the reference microphone for the left hearing aid (see eq.4.9). $A_{r_R}^*$ is defined similarly.

Applying 4.13 for a single speech source scenario, the transfer function of the speech and the noise component are equal to

$$G_{vL} = \frac{P_s |A_{r_L}|^2 \rho}{P_{vL} (\mu + \rho)^2} \qquad G_{vR} = \frac{P_s |A_{r_R}|^2 \rho}{P_{vR} (\mu + \rho)^2} \tag{4.46}$$

$$G_{xL} = \frac{\rho^2}{(\mu + \rho)^2} \qquad G_{xR} = \frac{\rho^2}{(\mu + \rho)^2} \tag{4.47}$$

with

$$P_{vL} = \mathbf{e}_{L,r_L}^T \mathbf{R}_{vv} \mathbf{e}_{L,r_L}, \qquad P_{vR} = \mathbf{e}_{R,r_R}^T \mathbf{R}_{vv} \mathbf{e}_{R,r_R}, \tag{4.48}$$

such that, using 4.14, the SNR improvement (for frequency ω) in the left and right hearing aid is equal to

$$\boxed{\Delta\mathrm{SNR}_L = \frac{\rho}{P_s/P_{vL}|A_{r_L}|^2}, \quad \Delta\mathrm{SNR}_R = \frac{\rho}{P_s/P_{vR}|A_{r_R}|^2}.} \tag{4.49}$$

This implies that the SNR improvement in each frequency band is larger for the hearing aid with the lowest input SNR ($\frac{P_s}{P_v}$). It also implies that the output SNR in each frequency band is the same at both hearing aids and equal to ρ, i.e. the output SNR in each frequency band depends on the average input SNR and the spatial separation between the speech and the noise source.

When evaluating the influence of the MWF on the binaural cues, it can be seen that eq. 4.45 implies that

$$\boxed{\mathbf{W}_R^m = \alpha \mathbf{W}_L^m} \tag{4.50}$$

where $\alpha = A_{r_R}^*/A_{r_L}^*$ is the complex conjugate of the ITF of the speech component, i.e.

$$ITF_x^{in} = \frac{X_L}{X_R} = \frac{A_{r_L}}{A_{r_R}} \tag{4.51}$$

Hence, since the binaural MWF vectors for the left and the right hearing aid, $\mathbf{W}_{MWF,L}$ and $\mathbf{W}_{MWF,R}$ are parallel, the ITF of the output speech and noise components are the same and equal to ITF_x^{in},

$$\boxed{\begin{aligned} ITF_x^{out} &= \frac{\mathbf{W}_{MWF,L}^H \mathbf{X}}{\mathbf{W}_{MWF,R}^H \mathbf{X}} = ITF_x^{in}, \\ ITF_v^{out} &= \frac{\mathbf{W}_{MWF,L}^H \mathbf{V}}{\mathbf{W}_{MWF,R}^H \mathbf{V}} = ITF_x^{in} \end{aligned}} \tag{4.52}$$

implying that all components (including the noise components) are perceived as coming from the speech direction.

This can also be shown by calculating the cross-correlation and ILD information generated by the binaural MWF. By combining 4.45, 4.6 and $\mathbf{X} = \mathbf{A}S$, the speech and the noise components at the output of the left and the right hearing aid can be written as

$$\begin{aligned} Z_{xL} &= \frac{P_s}{\mu+\rho}\mathbf{A}^H \mathbf{R}_{vv}^{-1} \mathbf{A} X_{L,r_L} & Z_{xR} &= \frac{P_s}{\mu+\rho}\mathbf{A}^H \mathbf{R}_{vv}^{-1} \mathbf{A} X_{R,r_R} \\ Z_{vL} &= \frac{P_s}{\mu+\rho}\mathbf{A}^H \mathbf{R}_{vv}^{-1} \mathbf{V} A_{r_L} & Z_{vR} &= \frac{P_s}{\mu+\rho}\mathbf{A}^H \mathbf{R}_{vv}^{-1} \mathbf{V} A_{r_R} \end{aligned} \tag{4.53}$$

By using eq. 4.10 and 4.12, the input cross-correlation and the input ILD for the speech component can be written as

$$c_x^{in} = P_s\, A_{r_L} A_{r_R}^*, \qquad L_x^{in} = \frac{|A_{r_L}|^2}{|A_{r_R}|^2}\ . \tag{4.54}$$

The output cross-correlation and the ILD of the speech component at the output of the MWF are equal to

$$c_x^{out} = \frac{\left(\mathbf{A}^H\mathbf{R}_{vv}^{-1}\mathbf{A}\right)^2 P_s}{\left(\mathbf{A}^H\mathbf{R}_{vv}^{-1}\mathbf{A} + \frac{\mu}{P_s}\right)^2} A_{r_L} A_{r_R}^*, \qquad L_x^{out} = \frac{|A_{r_L}|^2}{|A_{r_R}|^2} \tag{4.55}$$

Since $L_x^{out} = L_x^{in}$ and $c_x^{out} \sim c_x^{in}$, the MWF perfectly preserves the ILD and the ITD of the speech component. However, since the output cross-correlation and the ILD of the noise component at the output of the MWF are equal to

$$c_v^{out} = \frac{\mathbf{A}^H\mathbf{R}_{vv}^{-1}\mathbf{A}}{\left(\mathbf{A}^H\mathbf{R}_{vv}^{-1}\mathbf{A} + \mu\frac{P_v}{P_s}\right)^2} A_{r_L} A_{r_R}^*, \qquad L_v^{out} = \frac{|A_{r_L}|^2}{|A_{r_R}|^2} \tag{4.56}$$

the ITD and the ILD of the output noise component are equal to the ITD and the ILD of the input speech component (and hence also the output speech component), which is obviously not desired.

4.3.3 Discussion

This section presented the theoretical framework of the binaural MWF (or SDW-MWF). It was theoretically shown that the MWF preserves the cues of the speech but not of the noise component. However, further evaluations are needed to validate the performance of the MWF in realistic conditions, i.e. including room acoustics, etc. Therefore, the MWF has been thoroughly evaluated regarding noise reduction performance and localization by using objective and perceptual performance measures. These evaluations are discussed in chapters 5 and 6.

Since the preservation of the binaural cues of both the speech and the noise component are important for a realistic spatial awareness and to benefit from spatial release from masking effects, two extensions of the MWF are proposed, i.e. the MWF-N and MWF-ITF. These algorithms are discussed in section 4.4 and section 4.5.

4.4 MWF with partial noise estimation: MWF-N

4.4.1 MWF-N solution

The rationale of the MWF-N is not to completely remove the noise component from the reference microphone signals but to remove only part of it. The unprocessed part may then be used for a correct sound source localization of the noise component. This corresponds to estimating the desired speech component summed with an unprocessed scaled version of the noise component. As a consequence, eq. 4.15 changes to

$$\boxed{J_{MWF-N}(\mathbf{W}) = \mathcal{E}\left\{\left\| \begin{bmatrix} X_{L,r_L} + \eta V_{L,r_L} - \mathbf{W}_L^H \mathbf{Y} \\ X_{R,r_R} + \eta V_{R,r_R} - \mathbf{W}_R^H \mathbf{Y} \end{bmatrix} \right\|^2 \right\}} \tag{4.57}$$

with η between 0 and 1. By using a small η, more emphasis is put on noise reduction and less emphasis is put on preserving the binaural cues of the noise component. If $\eta = 0$, the MWF-N reduces to the standard MWF and maximum noise reduction performance is obtained. If $\eta = 1$, no noise reduction is obtained and the binaural cues are perfectly preserved. Similar to the MWF, a trade-off parameter can be introduced by weighting the amount of speech distortion with the residual noise energy of the partial noise estimate. The cost function then becomes

$$\boxed{J_{MWF-N}(\mathbf{W}) = \mathcal{E}\left\{\left\| \begin{bmatrix} X_{L,r_L} - \mathbf{W}_L^H \mathbf{X} \\ X_{R,r_R} - \mathbf{W}_R^H \mathbf{X} \end{bmatrix} \right\|^2 + \mu \left\| \begin{bmatrix} \eta V_{L,r_L} - \mathbf{W}_L^H \mathbf{V} \\ \eta V_{R,r_R} - \mathbf{W}_R^H \mathbf{V} \end{bmatrix} \right\|^2 \right\}} \tag{4.58}$$

A simple relationship holds between the optimal filters obtained with the MWF and the MWF-N, i.e.

$$\begin{aligned} \mathbf{W}_{MWF-N,L}(\eta,\mu) &= \eta \mathbf{e}_{L,r_L} + (1-\eta)\mathbf{W}_{MWF,L}(\mu) \\ \mathbf{W}_{MWF-N,R}(\eta,\mu) &= \eta \mathbf{e}_{R,r_R} + (1-\eta)\mathbf{W}_{MWF,R}(\mu) \end{aligned} \tag{4.59}$$

or in terms of the filter output:

$$\begin{aligned} Z_{MWF-N,L}(\eta,\mu) &= \eta Y_{L,r_L} + (1-\eta) Z_{MWF,L}(\mu) \\ Z_{MWF-N,R}(\eta,\mu) &= \eta Y_{R,r_R} + (1-\eta) Z_{MWF,R}(\mu) \end{aligned} \tag{4.60}$$

In other words, the MWF-N solution is obtained by adding a portion of the unprocessed reference microphone signals (e.g. for the left hearing aid $\eta Y_{L,r_L}$) to the original MWF solution. This can be used to restore the binaural cues of the noise component in the processed signal. A similar principle is followed in the work of Noble et al. (1998) and Byrne et al. (1998) where localization performance was improved by using open instead of closed earmolds by bilateral hearing aid users with a moderate hearing loss in either the high or the low frequency region. In these studies the direct sound was used by the hearing aid user to improve sound localization. Obviously noise reduction performance will decrease when increasing η.

4.4.2 Theoretical analysis of the binaural MWF-N

In the case of a single speech source, the optimal filters, obtained from eq. 4.45 and eq. 4.59 are

$$\boxed{\begin{aligned} W_{MWF-N,L} &= (1-\eta)\frac{P_s\mathbf{R}_{vv}^{-1}\mathbf{A}}{\mu+\rho}A_{r_L}^* + \eta\mathbf{e}_{L,r_L} \\ W_{MWF-N,R} &= (1-\eta)\frac{P_s\mathbf{R}_{vv}^{-1}\mathbf{A}}{\mu+\rho}A_{r_R}^* + \eta\mathbf{e}_{R,r_R} \end{aligned}} \tag{4.61}$$

Using 4.13 and 4.59, the power transfer function of the noise component at the left hearing aid is found

$$G_{vL} = (1-\eta)^2\frac{P_s|A_{r_L}|^2\rho}{P_{vL}(\mu+\rho)^2} + 2\eta(1-\eta)\frac{P_s|A_{r_L}|^2}{P_{vL}(\mu+\rho)} + \eta^2, \tag{4.62}$$

$$= \frac{P_s|A_{r_L}|^2}{P_{vL}\rho}[(1-\eta)^2\frac{\rho^2}{(\mu+\rho)^2} + 2\eta(1-\eta)\frac{\rho}{\mu+\rho} + \eta^2\frac{P_{vL}\rho}{P_s|A_{r_L}|^2}] \tag{4.63}$$

$$= \frac{1}{\Delta SNR_L^0}\{(\frac{\eta\mu+\rho}{\mu+\rho})^2 + (\Delta SNR_L^0 - 1)\eta^2\} \tag{4.64}$$

with ΔSNR_L^0 the SNR improvement of the standard MWF. From eq. 4.13 and 4.59, the power transfer function of the speech component at the left hearing aid can be computed

$$G_{xL} = (1-\eta)^2\frac{\rho^2}{(\mu+\rho)^2} + 2\eta(1-\eta)\frac{\rho}{(\mu+\rho)} + \eta^2, \tag{4.65}$$

$$= (\frac{\eta\mu+\rho}{\mu+\rho})^2. \tag{4.66}$$

From this it follows that the SNR improvement at the left hearing aid is equal to

$$\Delta SNR_L = \Delta SNR_L^0\frac{(\frac{\eta\mu+\rho}{\mu+\rho})^2}{(\frac{\eta\mu+\rho}{\mu+\rho})^2 + (\Delta SNR_L^0 - 1)\eta^2} \tag{4.67}$$

Logically, if $\eta = 0$ the SNR improvement reduces to ΔSNR_L^0, whereas if $\eta = 1$ no SNR improvement is obtained, $\Delta SNR_L = 1$. A similar expression can be derived for the right hearing aid.

When evaluating the MWF-N with respect to the preservation of binaural cues, it is observed from eq. 4.61 that the binaural MWF-N filters are, in general, not parallel. Hence, the ITF of the speech and noise component at the output of the algorithm are typically different. Using 4.8, the ITF of the output speech component is

$$ITF_x^{out} = \frac{\mathbf{W}_L^H\mathbf{X}}{\mathbf{W}_R^H\mathbf{X}} = \frac{(1-\eta)\frac{\rho}{\mu+\rho}A_{r_L} + \eta A_{r_L}}{(1-\eta)\frac{\rho}{\mu+\rho}A_{r_R} + \eta A_{r_R}} = \frac{A_{r_L}}{A_{r_R}}, \tag{4.68}$$

Hence, the ITF of the speech component is preserved. Although it is not possible to easily formulate the ITF of the output noise component for a general noise scenario, it is possible to compute the cross-correlation and the ILD. Using 4.10, the cross-correlation of the output noise component is equal to

$$
\begin{aligned}
c_v^{out} &= \mathbf{W}_{MWF-N,L}^H \mathbf{R}_{vv} \mathbf{W}_{MWF-N,R} &(4.69)\\
&= [(1-\eta)^2 \frac{\rho}{(\rho+\mu)^2} + 2\eta(1-\eta)\frac{1}{\mu+\rho}] P_s A_{r_L} A_{r_R}^* + \eta^2 \mathbf{e}_{r_L}^H \mathbf{R}_{vv} \mathbf{e}_{r_R} &(4.70)\\
&= [(1-\eta)^2 \frac{\rho}{(\rho+\mu)^2} + 2\eta(1-\eta)\frac{1}{\mu+\rho}] c_x^{in} + \eta^2 c_v^{in}, &(4.71)
\end{aligned}
$$

which states that c_v^{out} is a weighted sum of the cross-correlation of the input speech, c_x^{in}, and the input noise component, c_v^{in}. When $\eta = 0$ the cross-correlation of the output noise component perfectly matches the cross-correlation of the output speech component, this corresponds to the behavior of the binaural MWF. If $\eta = 1$ the cross-correlation of the output noise component equals the cross-correlation of the input noise component, but no noise reduction is obtained.

Using 4.12 and 4.62, it can be shown that the ILD of the output noise component is equal to

$$
\begin{aligned}
L_v^{out} &= \frac{G_{vL} P_{vL}}{G_{vR} P_{vR}}, &(4.72)\\
&= \frac{[(1-\eta)^2 \frac{\rho}{(\rho+\mu)^2} + 2\eta(1-\eta)\frac{1}{\mu+\rho}] P_s |A_{r_L}|^2 + \eta^2 P_{vL}}{[(1-\eta)^2 \frac{\rho}{(\rho+\mu)^2} + 2\eta(1-\eta)\frac{1}{\mu+\rho}] P_s |A_{r_R}|^2 + \eta^2 P_{vR}} &(4.73)
\end{aligned}
$$

Again, if $\eta = 0$ this corresponds to the behavior of the MWF. If $\eta = 1$ the ILD of the output noise component corresponds to the ILD of the input noise component, but no noise reduction is obtained.

4.4.3 Discussion

This section presented a theoretical framework for the binaural MWF-N. It was shown that the MWF-N preserves the binaural cues of the speech component. The parameter η offers a trade-off between noise reduction and the preservation of the binaural cues of the noise component. If $\eta = 0$, the MWF-N reduces to the MWF and maximum noise reduction is achieved. If $\eta = 1$, the binaural cues of both the speech and the noise component are perfectly preserved, but no noise reduction is present.

Thorough evaluations have been performed to validate the performance of the MWF-N in realistic conditions. The results of these experiments are presented in chapters 5 and 6.

4.5 MWF with interaural transfer function extension: MWF-ITF

4.5.1 MWF-ITF solution

In order to control the binaural cues of both the speech and the noise component, the cost function in (4.29) is extended with one or two terms related to the ITF. The extra ITF term for preserving the binaural cues of the noise component is defined as the difference between the ITF at the output of the algorithm and the desired ITF, i.e the ITF of the noise component at the input of the algorithm. Hence,

$$J^v_{ITF}(\mathbf{W}) = \mathcal{E}\Big\{\Big|\frac{\mathbf{W}_L^H\mathbf{V}}{\mathbf{W}_R^H\mathbf{V}} - ITF^v_{des}\Big|^2\Big\}\,, \tag{4.74}$$

with ITF^v_{des} the desired ITF of the noise component. If this ITF is sufficiently stationary, e.g. for single noise source scenarios, it can be estimated in a least square sense using the cross-correlation matrices:

$$ITF^v_{des} = \frac{\mathcal{E}\{V_{L,r_L}V^*_{R,r_R}\}}{\mathcal{E}\{V_{R,r_R}V^*_{R,r_R}\}}\,. \tag{4.75}$$

By adding J^v_{ITF} to the cost function of the MWF using a weighting factor β, the filters $\mathbf{W}_L$ and $\mathbf{W}_R$ can be restricted to solutions which, to some extent (depending on the weight β), preserve the ITF of the noise component. A similar term can be derived for the speech ITF by replacing the noise component and the desired noise ITF with the speech component and a desired speech ITF, i.e. $J^x_{ITF}(\mathbf{W})$. Hence, a total cost function trading off noise reduction, speech distortion and binaural cue preservation of both the speech and the noise component can be defined as

$$\boxed{J_{MWF-ITF}(\mathbf{W}) = J_{SDW-MWF}(\mathbf{W}) + \alpha J^x_{ITF}(\mathbf{W}) + \beta J^v_{ITF}(\mathbf{W})} \tag{4.76}$$

where the weights α and β emphasize the binaural cue preservation of the speech and the noise component, respectively.

Eq. 4.74 can also be written as:

$$\boxed{J^v_{ITF}(\mathbf{W}) = \frac{\mathcal{E}\{|\mathbf{W}_L^H\mathbf{V} - ITF^v_{des}\mathbf{W}_R^H\mathbf{V}|^2\}}{\mathcal{E}\{|\mathbf{W}_R^H\mathbf{V}|^2\}} = \frac{\mathbf{W}^H\mathbf{R}_{vt}\mathbf{W}}{\mathbf{W}^H\mathbf{R}_{v1}\mathbf{W}}} \tag{4.77}$$

with

$$\mathbf{R}_{vt} = \begin{bmatrix} \mathbf{R}_v & -ITF^{v,*}_{des}\,\mathbf{R}_v \\ -ITF^v_{des}\,\mathbf{R}_v & |ITF^v_{des}|^2\,\mathbf{R}_v \end{bmatrix} \tag{4.78}$$

$$\mathbf{R}_{v1} = \begin{bmatrix} \mathbf{0}_M & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{R}_v \end{bmatrix}. \tag{4.79}$$

Since no closed-form expression is available for the filter minimizing $J_{MWF-ITF}(\mathbf{W})$, iterative optimization techniques should be used which are at present too computationally expensive for hearing aid applications. To reduce the complexity of the hearing aid algorithm a less computationally expensive quadratic cost function is derived from (4.77), i.e.

$$\boxed{J^v_{ITF}(\mathbf{W}) = \mathcal{E}\{|\mathbf{W}_L^H\mathbf{V} - ITF^v_{des}\mathbf{W}_R^H\mathbf{V}|^2\}} \tag{4.80}$$

This approximation is used throughout the remainder of this document.

The MWF-ITF seems to be a rather straightforward approach. However, problems arise when evaluating multiple noise source scenarios. If multiple, spectrally overlapping, noise sources are present, ITF^v_{des} will represent a combination of the ITFs of all noise sources present in that scenario. The introduction of the cost function J^v_{ITF} will preserve this ITF^v_{des} at the output of the algorithm. However, since not all noise sources are reduced equally in amplitude, ITF^v_{des} will not be the correct ITF-combination at the output of the algorithm. An extreme scenario is given as example to clarify this: if one noise source dominates a multiple noise source scenario, the ITF^v_{des} term, measured at the microphone inputs, will be dominated by this noise source. Therefore all noise sources present at the output of the algorithm will sound as arriving from the angle of this dominant noise source, even if the dominant noise source has been fully removed from the microphone signals by the noise reduction algorithm.

4.5.2 Objective and perceptual evaluations

Several pilot studies were performed analyzing the performance of the MWF-ITF using different parameter settings. More thorough theoretical, objective and perceptual evaluations are topic of current and future research. Nevertheless, a short summary of the results that have already been obtained are presented in this manuscript to illustrate the potential of this algorithm.

General settings

First, the impulse responses between a set of loudspeakers and all microphones present on two BTE hearing aids were measured using a CORTEX MK2 artificial head. Each BTE had an array of two omnidirectional microphones with a microphone spacing of 1cm. The loudspeakers were placed at a distance of 1m from the head. Recordings were made around the head in steps of 15° in the horizontal hemisphere. All measurements were done in a low reverberant room with a T_{60} linearly averaged over all 1/3th octave bands of $T_{60} = 0.21s$.

These impulse responses were used to artificially generate a spatial scenario, referred to as S_xN_y, with x the position of the speech source and y the position of the noise source. Dutch sentences (Versfeld et al., 2000) were used as speech material and the accompanying stationary speech-weighted noise was used as the competing noise signal. Based on the generated microphone signals, the correlation matrices of the speech and the noise component could be calculated off-line. The correlation matrices were estimated using a perfect VAD and were calculated over the full duration of the signal. They hence correspond to the correlation matrices obtained by a fully converged MWF-ITF. Since the speech and noise component were, on average, spectrally identical, the performance of the MWF-ITF solely depends on the spatial filtering capabilities of the algorithm. The speech correlation matrix was estimated from $\mathbf{R}_{xx} = \mathbf{R}_{yy} - \mathbf{R}_{vv}$ (eq. 4.21). Calculations were performed in the frequency domain using 256 points FFT's and with a trade-off parameter μ set to $\mu = 1$ throughout the evaluations (see eq. 4.29 and eq. 4.76).

During the objective and perceptual evaluations, the parameter α was most commonly fixed to 0 or 0.5, β was varied between the values 0, 0.1, 0.3, 1, 10 and 100. Only the most relevant subset of the gathered data is presented here.

Objective evaluation

In this section objective performance measures are defined to predict the improvement in SRT and to evaluate the preservation of binaural cues. Three different performance measures were used during the evaluation of the MWF-ITF: SNR improvement, ITD error and ILD error. These performance measures were applied on the output signals of the MWF-ITF.

The improvement in *speech intelligibility weighted SNR* ($\Delta\mathrm{SNR}_{SI}$), defined by Greenberg et al. (1993), was used to evaluate the noise reduction performance of the noise reduction algorithms. This is defined as the difference between the output $\mathrm{SNR}_{SI,out}$ and the input $\mathrm{SNR}_{SI,in}$, e.g. for the left hearing aid this is given by

$$\Delta\mathrm{SNR}_{SI,L} = \sum_i I(\omega_i)\mathrm{SNR}_{out,L}(\omega_i) - I(\omega_i)\mathrm{SNR}_{in,L}(\omega_i) \qquad (4.81)$$

with $\mathrm{SNR}(\omega_i)$ the SNR measured in the i-th third-octave band and $I(\omega_i)$ the importance of the i-th frequency band for speech intelligibility (Pavlovic, 1987), as defined by ANSI-SII (1997). $\Delta\mathrm{SNR}_{SI,L}$ is commonly used to quantify and predict the noise reduction performance or the gain in speech understanding when evaluating noise reduction algorithms, especially for monaural hearing aid configurations.

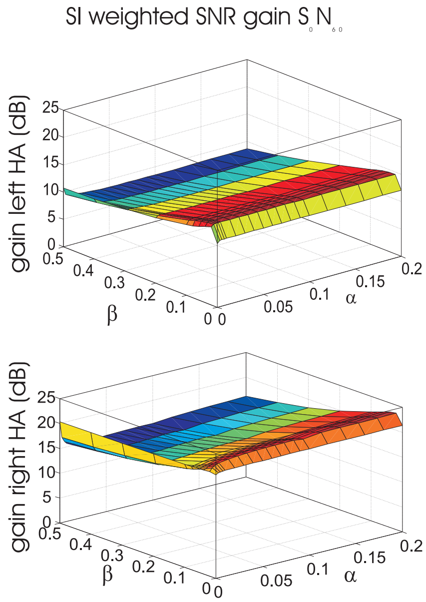

Figure 4.2 — SI weighted SNR improvement at the left and the right hearing aid when using an MWF-ITF with different parameters α and β in an environment with T_{60}=0.21s

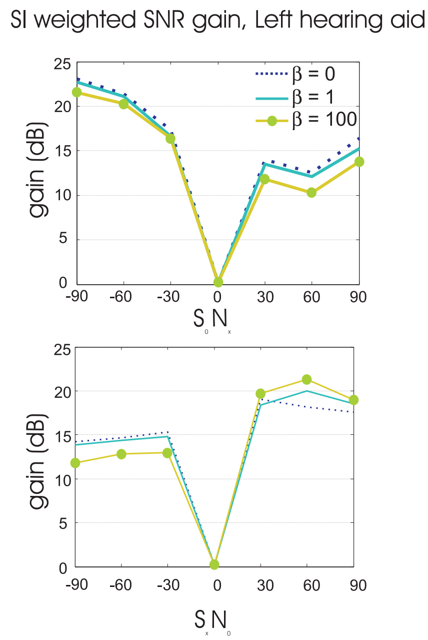

Figure 4.3 — SI weighted SNR improvement at the left hearing aid when using an MWF-ITF with $\alpha = 0$ and three different β-values in an environment with T_{60}=0.21s. **Upper plot**: the angle of the noise source was varied from -90° to +90° (S_0N_x). **Lower plot**: the angle of the speech source was varied (S_xN_0).

No standardized objective error measures were available from literature to evaluate the distortion of binaural cues by hearing aid algorithms. Therefore, two new objective error measures were defined: the ITD and ILD error. The *ITD error* of the speech or the noise component was calculated using the phase of the cross-correlation per frequency band. The ITD error of the noise component is computed as

$$\Delta ITD_v = \sum_i A_{ITD}(\omega_i)\frac{|\angle\mathcal{E}\{Z_{vL}(\omega_i)Z^*_{vR}(\omega_i)\} - \angle\mathcal{E}\{V_{L,r_L}(\omega_i)V^*_{R,r_R}(\omega_i)\}|}{\pi} \tag{4.82}$$

with $A_{ITD}(\omega_i)$ a weighting factor which only includes frequency bands below 1500Hz (see section 1.4.1). Note that this ITD error measure is always between 0 (no distortion) and 1 (maximum distortion or a phase shift of π). The ITD error of the speech component is defined similarly.

The *ILD error* generated by the noise reduction algorithm on the speech or on

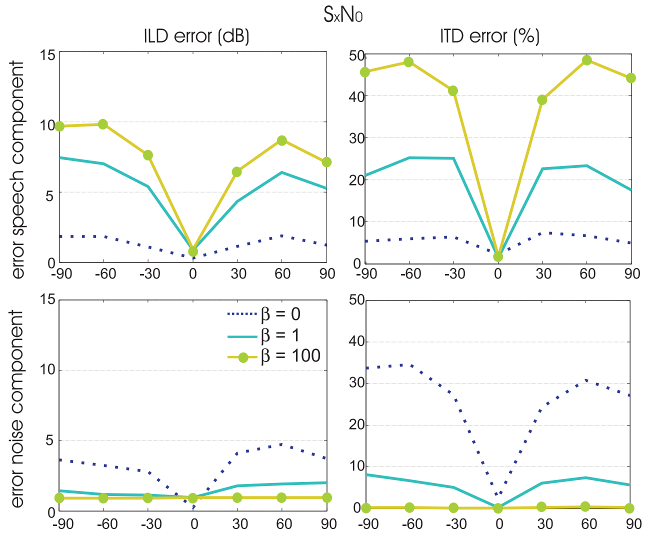

Figure 4.4 — Objective performance measures of ITD and ILD distortion when using the MWF-ITF with $\alpha = 0$ and three different β-values in an environment with T_{60}=0.21s. The noise source was fixed at 0°, the angle of the speech source was varied from -90° to +90° (S_xN_0). The figures at the top represents the ITD and ILD error of the speech component, at the bottom the ITD and ILD error of the noise component are shown.

the noise component was defined as the difference between the input (L^{in}) and the output ILD (L^{out}) of both components. Hence, the ILD error on the noise component was defined as

$$\Delta ILD_v = \sum_i A_{ILD}(\omega_i)|10\log_{10} L_v^{out}(\omega_i) - 10\log_{10} L_v^{in}(\omega_i)|\,. \tag{4.83}$$

with $A_{ILD}(\omega_i)$ a frequency dependent weighting function. In this section, it is assumed that the ILD errors in all frequency bands are equally important. Therefore the ILD error used is a linearly averaged ILD error, i.e. $A_{ILD}(\omega_i) = 1$. The ILD error of the speech component is defined similarly.

Figure 4.2 shows ΔSNR_{SI} for both ears for the spatial scenario S_0N_{60}. A slight decrease in noise reduction performance is observed when increasing α and/or β. Figure 4.3 shows the improvement in SNR_{SI} for the left hearing aid using a binaural MWF-ITF with three different values of β and $\alpha = 0$. In the upper figure, the speech source was fixed at 0° and the noise source was displaced in steps of 30° (S_0N_x). In the lower figure, the noise source was fixed at 0° and the speech source was moved in steps of 30° (S_xN_0). In general only very small losses in noise reduction were observed if the parameter β was increased. The noise reduction performance of the algorithm dropped if the spatial separation between the speech and the noise source decreased.

Figure 4.4 shows the ITD and ILD errors introduced by the MWF-ITF (eq. 4.82 and 4.83) for a spatial scenario with the noise source fixed at 0° and the location of the speech source varied in steps of 30° ($S_x N_0$). The top and bottom row represent the ITD and ILD error generated on the speech and the noise component respectively. If β was small, the MWF-ITF behaved as a standard MWF with small errors on the binaural cues of the speech component and large errors on the binaural cues of the noise component. If β was large, i.e. if a lot of emphasis was put on preserving the ITF of the noise component, the ITD and ILD errors of the noise component decreased. However, the ITD and ILD errors of the speech component increased drastically. Since the evaluation of scenarios $S_0 N_x$ led to similar conclusions as the $S_x N_0$ scenarios, these data is omitted.

A first question that arises is whether an appropriate value for β could be found which sufficiently preserves the binaural cues of both the speech and the noise component. Secondly, the perceptual relevance of the objective ITD and ILD error measures can be questioned. These measures calculate some sort of average ITD and ILD distortion over the different frequency bands. However, it is highly unlikely that a large distortion in only one band or several small distortions in several frequency bands will introduce the same perceptual effect. Moreover, since the physical ILDs are much larger at high frequencies than at low frequencies, it may be assumed that the introduction of a small ILD error in low frequency bands will generate a more severe perceptual distortion than the same error made in high frequency bands. To get more insight on these questions, perceptual evaluations were performed.

Perceptual evaluation

The binaural MWF-ITF was perceptually validated in terms of noise reduction performance and localization in the frontal horizontal hemisphere by five normal hearing subjects.

Noise reduction performance

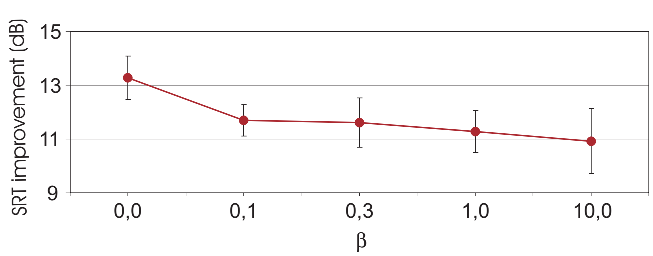

Figure 4.5 — SRT gain of the MWF-ITF algorithm using $\beta = 0, 0.1, 0.3, 1$ and 10 and $\alpha = 0$ in the condition $S_0 N_{60}$.

A spatial scenario S_0N_{60} was generated as described in section 4.5.2. The filter coefficients were calculated off-line for different SNRs and for different values of β. The parameter α was set to $\alpha = 0$. These filters were used in an adaptive SRT procedure which measures the SNR level at which the listener understands 50% of the speech correctly. Dutch VU sentences were used as speech material and the accompanying stationary speech weighted noise (Versfeld et al., 2000) was used as jammer sound. An unprocessed condition, using the front omnidirectional microphone signals of the left and the right hearing aid was used as a reference condition. The stimuli were presented under headphones.

The average SRT improvement of the MWF-ITF, relative to the unprocessed condition is shown in Figure 4.5. A maximum noise reduction of approximately 13dB was obtained with $\beta = 0$, i.e. when using the standard binaural MWF. A small decreasing trend in noise reduction performance was observed if more emphasis was put on the preservation of the binaural cues of the noise. This confirms the predicted gains by the objective evaluation in Figure 4.3. Figure 4.5 also shows that there was no indication of a gain in speech perception due to binaural unmasking which could take place if the binaural cues of both the speech and the noise source are preserved.

Localization performance

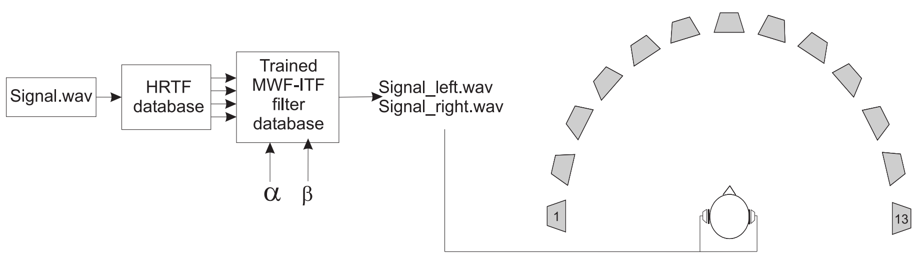

Figure 4.6 — Localization test setup

An overview of the localization experiment is given in Figure 4.6. During this experiment, the speech and the noise source were calibrated at an input SNR of 0 dBA. Correlation matrices and filters were generated for different spatial scenarios (S_xN_0) and different parameters α and β as discussed in section 4.5.2. A telephone signal, the same as used in section 2.3, was used to generate the artificial stimuli. These stimuli, arriving from a certain direction in the horizontal plane (direction x in case of the speech angle and direction 0° for the noise angle), were filtered by the pre-calculated MWF-ITF filter (trained on the condition $\mathrm{S}_x\mathrm{N}_0$). The subject was seated in the center of the loudspeaker array which was also used to record the impulse responses of the experiment (see Figure 4.6). The task of the subject was to localize the presented, i.e.

filtered, telephone signals. The possible responses were between -90° and +90° in steps of 15° and were marked by loudspeakers (see Figure 4.6), labelled 1 to 13. Stimuli were repeated three times and were presented in a random order. A level roving of 5dB was applied during the test procedure to avoid monaural loudness cues (see section 1.4.5). Parameter β was varied between the values 0, 0.1, 0.3, 1, 10 and 100. The parameter α was set to 0 or 0.5.

The results of the localization experiments are shown in Figure 4.7 and Figure 4.8. Figure 4.7 illustrates two extreme and one optimal setting of the parameters α and β. It shows the accumulation of responses for the five test subjects using α=0 and three different β's: two extreme values, i.e. $\beta = 0$ (standard binaural MWF) and $\beta = 10$, and one optimal β, i.e. $\beta = 0.3$. Figure 4.7 illustrates that the standard MWF technique ($\beta = 0$) moved the localization of the noise source towards the location of the speech source (left column). If $\beta = 10$ the noise component was correctly localized. However, the speech source was perceived as arriving from the location of the noise source (right column). This confirms the observations made during the objective evaluation in Figure 4.4. Again, the question arises if an optimal parameter setting can be found which sufficiently preserves the binaural cues of both the speech and the noise component. This is illustrated in Figure 4.8.

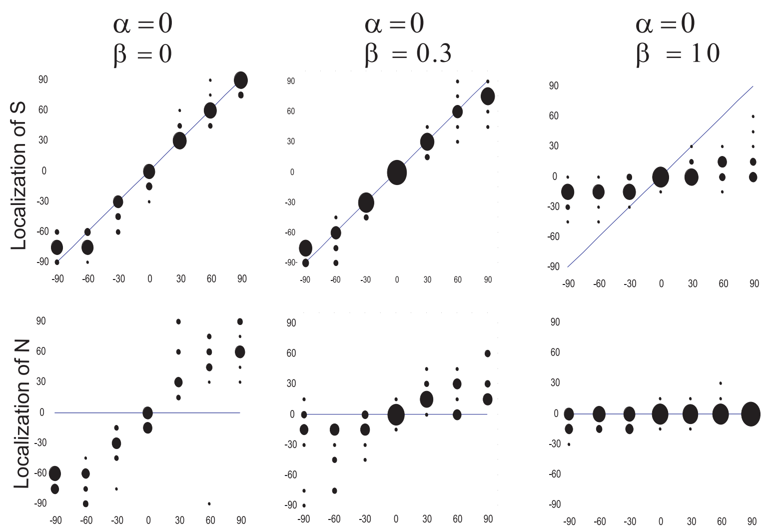

Figure 4.7 — Accumulation of the responses for 5 subjects at three β-values of the condition $S_x N_0$. The line represents the correct location of the presented stimulus. If β was (too) low, the speech component was correctly localized but the noise component was located at the place of the speech component (left column). If β was (too) high, the noise component was correctly localized but the speech component was located at the place of the noise component (right column).

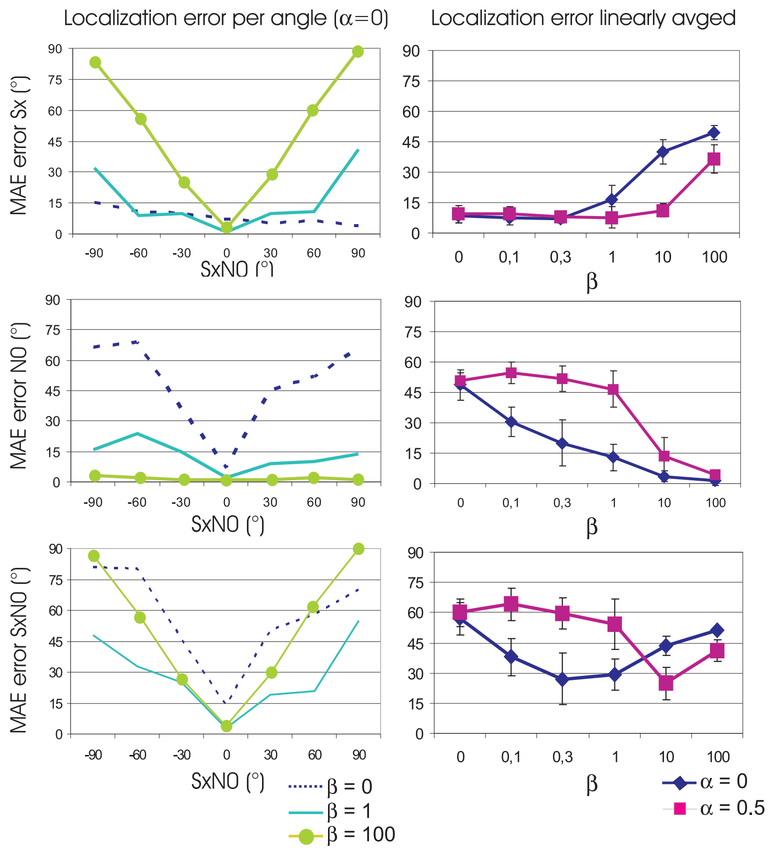

Figure 4.8 — Localization performance with the MWF-ITF using different parameters α and β. The MAE errors, averaged over 3 repetitions and 5 normal hearing subjects, of condition S_xN_0 are plotted in the left column. The top row shows the error made on the speech component which was varied per 30° around the subject. The middle row indicates the error made on the noise component which was positioned at 0°. The bottom row shows the accumulation of the errors made on the speech and the noise component. The right column shows the MAE localization error, linearly averaged over the different angles, i.e. one data point is the average of a curve given in the left column. Plots are shown for the different parameters α and β used in the perceptual evaluation. For reasons of clarity only a subset of the parameter settings are considered in the left column of the figure. The bottom right figure indicates that optimal localization performance was reached, when averaging over both the speech and the noise component and over the different angles, at the combination $\alpha = 0$, $\beta = 0.3$ or at $\alpha = 0.5$, $\beta = 10$.

Figure 4.8 summarizes the MAE localization errors (defined in eq. 2.20) made on the speech and the noise component, averaged over three repetitions per subject and five normal hearing subjects. The x-axis in the left column represents the location of the presented speech source. Low values of β introduced a small error on the speech component (first row), but a large error on the noise component (second row). A high value of β introduced a small error on the noise component but a large error on the speech component. If the assumption is made that each angle is equally important then the mean error of the speech and the noise component can be calculated over all angles. This is shown in the right column of Figure 4.8. In this way, a single number was obtained per parameter setting indicating the quality of this parameter combination for the spatial condition S_xN_0 with x varied between -90° and 90°. The bottom right figure shows that if $\alpha = 0$ an optimal setting for localizing sound sources was obtained for $\beta = 0.3$. For $\alpha = 0.5$ the optimal setting of β was around $\beta = 10$. The middle column of Figure 4.7 illustrates the responses made in one of the optimal parameter settings, i.e. $\alpha = 0$ and $\beta = 0.3$. When evaluating the S_0N_x conditions, the optimal parameter settings $\alpha = 0$, $\beta = 0.1$ and $\alpha = 0$, $\beta = 0.3$ were found with a similar performance for both parameter settings.

4.5.3 Discussion

In section 4.5 a second extension of the binaural MWF was presented, the MWF-ITF. By adding one or two extra term(s) to the cost function of the MWF, the Wiener solution can be restricted to filters which preserve the binaural cues of the noise component to some extent. By using extreme values of β, it was observed that if β is low, only the cues of the speech component are preserved. If β is very high only the cues of the noise component are preserved. By evaluating the parameter space, an optimal parameter setting could be found, generating a reduced localization error (Figure 4.8) compared to the MWF (β=0) without drastically reducing noise reduction performance (Figure 4.5).

Throughout the evaluations, different values for α and β were used. However, the reported evaluations of the MWF-ITF were often based on the assumption that $\alpha = 0$. This can be explained by interpreting eq. 4.76. As the binaural MWF inherently preserves the cues of the speech component but not of the noise component, the term J_{ITF}^{v} has been added to the cost function of the MWF to preserve the binaural cues of the noise component. However, it was observed that emphasizing this term too strongly results in a shift of the location of the speech component towards the location of the noise component. One may compensate for this by introducing an extra term J_{ITF}^{x} which tries to restore the balance. However, it is common sense to reduce the value of β instead of adding this extra term in the cost function. This can also be observed in Figure 4.8 which shows that an increase in α did not improve localization performance, but only resulted in an increase in β to restore the optimal performance of the MWF-ITF.

Despite the gain in localization performance compared to a binaural MWF (which equals an MWF-ITF with $\alpha = \beta = 0$), it is observed that the errors made by the subjects are still significant. Ongoing theoretical analysis shows that the MWF-ITF filters for the left and the right hearing aid are parallel vectors when using a simplified quadratic cost function (eq. 4.80) . This would result in an identical ITF for both the output speech and the output noise component. The size of β determines whether this transfer function is close to the one of the original speech component or to the one of the original noise component. Nonetheless a gain in localization performance can be obtained. This is due to the dependency of β on the output SNR. When using a modest β, the ITF of frequency bands with high output SNR stay relatively close to the speech ITF. Those with a low SNR tend to stay relatively close to the noise ITF. Hence, a gain in localization performance is obtained. This theoretical analysis and further improvements on the MWF-ITF are topics of ongoing research.

4.6 Reduced bandwidth algorithms

To conclude the chapter on the development of binaural MWF algorithms, an overview of MWF-based reduced bandwidth algorithms is appropriate. The binaural algorithms presented so far assume that all microphone signals are available at both sides through e.g. a wireless link between the hearing aids. However, this comes at a large cost of bandwidth and power consumption. Therefore, the question arises as to which extent it is feasible to reach $M_L + M_R$ microphone performance (for conciseness we will now assume that both hearing aids each have two microphones) by transmitting only one contralateral signal. Only a summary is presented here, more information can be found in Doclo et al. (2008).

The first option consists in transmitting only one, instead of all, contralateral microphone signals (MWF-front). This off course generates a sub-optimal solution. However, as will be shown in chapter 6, the localization behavior of the three-microphone MWF and MWF-N seems to match the predicted behavior of a fully binaural MWF. Moreover, due to the positioning of all microphones, it is rather straightforward to foresee that adding the first contralateral microphone signal to the ipsilateral hearing aid adds more new information, and hence more gain in noise reduction performance, to the algorithm than additionally adding the second contralateral microphone signal. This is also demonstrated in chapter 5 which compares a bilateral two microphone MWF and MWF-N with a binaural three and four microphone MWF and MWF-N. Other, more performant solutions combine the microphone signals at the contralateral hearing aid before transmitting them to the ipsilateral hearing aid.

The second and third option are based on transmitting a signal produced by a stand-alone noise reduction algorithm running on the contralateral hearing aid. Two options have been studied: transmitting a signal generated by a monaural superdirective beamformer and transmitting a signal generated by a monaural MWF. Both systems show an improvement in noise reduction performance of about 1.2 dB compared to the MWF-front with the MWF-front having an average noise reduction performance of about 14.5dB compared to an unprocessed condition for a single noise source scenario with speech at 0° and a noise source between 60° and 300° . This was measured in the same low reverberant environment as used in section 4.5.2 and chapter 5. However, these options introduce a large computational overhead since two independent noise reduction algorithms are running on each hearing aid, i.e. a three-channel MWF and the monaural algorithm generating the third signal for the contralateral hearing aid.

A final option is to perform an iterative distributed processing scheme (MWF-db). Basically the MWF-db computes the three-channel MWF solution at one hearing aid, e.g. the left hearing aid. Then, both microphones present on this hearing aid are combined by using their respective filters of the three-channel MWF solution. This signal is transmitted to the right hearing aid. The right hearing aid computes a new three-channel MWF solution and sends a combination of its microphone signals, using this new solution to the other side, and so on. To initialize the procedure it is possible to send the microphone signal of the front omnidirectional microphone to the contralateral hearing aid. In Doclo et al. (2008) it is proven that, if a rank-1 speech correlation matrix is used, the MWF-db converges to the full binaural MWF solution. In more general scenarios, this convergence is not guaranteed. However, simulations show that also in these conditions the MWF-db tends to converge to the optimal binaural solution. In general, the MWF-db seems to approach the optimal performance of the binaural MWF. One of the main limitations of the study performed is that it does not take into account the delay generated by the binaural link. Therefore further development is needed to prove the practical relevance of this technique.

4.7 Conclusions

This chapter presented the general mathematical framework in which binaural algorithms have been developed. It was shown that a monaural MWF, first presented by Doclo and Moonen (2002), can be easily extended to a binaural version. It was also proven that the binaural MWF inherently preserves the binaural cues of the speech component. However, the cues of the noise component are changed into those of the speech component. Since both cues are important for preserving the spatial awareness of the subject and to benefit from spatial release from masking, two extensions of the MWF were presented, the MWF-N and the MWF-ITF.

The rationale of the MWF-N is to partially (a portion η) preserve the noise component at the reference microphone. Hence, the unprocessed part may be used to localize the noise source. This obviously comes at the cost of noise reduction with the parameter η determining the trade-off between noise reduction performance and the preservation of binaural cues. When $\eta = 0$ the MWF-N reduces to the binaural MWF with maximal noise suppression. If $\eta = 1$ all binaural cues are preserved since no processing whatsoever is done. How the MWF and MWF-N perform in real life situations in terms of noise reduction and localization performance and how the parameter settings affect their performance will be discussed in chapters 5 and 6.

The rationale of the MWF-ITF is to add a term to the cost function of the binaural MWF to emphasize the preservation of the binaural cues of the noise component. The study of the MWF-ITF is topic of current and future research. However, pilot experiments showed that, when using the right parameter setting, localization performance can improve compared to the binaural MWF. Current research is showing that the ITF of both the speech and the noise component are identical at the output of the algorithm. The gain in localization performance can be explained by the interaction of the extra term in the cost function and the output SNR. Hence, the amount of spatial release from masking due to a spatial separation between both the speech and the noise component may be limited.

The binaural algorithms presented so far assumed that all microphone signals are available through e.g. a wireless link between the hearing aids. This comes at the large cost of power consumption. Therefore four strategies were studied which limit the amount of data transmitted to the contralateral hearing aid. In a first strategy, only one contralateral microphone signal, i.e. usually the front omnidirectional microphone signal, is transmitted to the ipsilateral hearing aid (MWF-front). The second and third strategy use a second monaural noise reduction algorithm running on the contralateral hearing aid, either a monaural superdirective beamformer or a monaural MWF. A fourth and final option is an iterative distributed processing scheme (MWF-db). Mathematical derivations and experimental results show that the MWF-db approaches the full MWF solution and reaches the highest noise reduction performance of all reduced bandwidth algorithms. The MWF-front is the easiest to implement but has the lowest noise reduction performance of all four evaluated reduced bandwidth algorithms.

Chapter 5

Noise reduction by the MWF and the MWF-N vs. an ADM

This chapter discusses the noise reduction performance and speech enhancement of the MWF and MWF-N algorithm. This is done by objective and perceptual evaluations. In the previous chapter an introduction to, and a theoretical evaluation of the MWF and MWF-N algorithm was presented. It was shown that these algorithms have promising theoretical properties with respect to combining noise reduction performance with the preservation of binaural cues. This chapter further evaluates the noise reduction performance of the MWF and MWF-N using objective and perceptual evaluations in different acoustical scenarios and using different parameter settings. Chapter 6 will further evaluate the preservation of binaural cues by the MWF and MWF-N algorithm by using a localization experiment in the frontal horizontal hemisphere.

Section 5.1 presents the research questions discussed in this chapter.

Section 5.2 specifies the stimuli, the acoustic environments and the different parameter settings of the different algorithms which were used during the objective and perceptual evaluations.

Section 5.3 presents the performance measured during the objective and the perceptual evaluations.

Section 5.4 is a discussion on the obtained noise reduction performance of the MWF and MWF-N algorithm in a large variety of spatial scenarios and two different acoustic environments. An answer is formulated to the specific research questions defined in the introduction.

Section 5.5 summarizes the conclusions of this chapter.

The results presented in this chapter are discussed in Van den Bogaert et al. (2008*b*).

5.1 Introduction

Chapter 4 states that a monaural MWF, as discussed in Doclo and Moonen (2002), can be extended in a straightforward way to a binaural framework. This increases the number of microphones used by the algorithm which typically enhances noise reduction performance. Moreover, it was mathematically proven that a binaural MWF inherently preserves the binaural cues of the targeted speech component. However, the cues of the noise component are distorted into those of the speech component. To preserve the binaural cues and hence the spatial awareness of the user, the MWF-N was introduced which aims at removing only part of the noise component. The unprocessed part can then be used to localize the noise component. Obviously this may come at the cost of noise reduction.

This chapter presents an evaluation of the noise reduction performance and speech enhancement of the MWF and MWF-N approach using different parameter settings in several spatial sound scenarios in different acoustic environments. Transmitting microphone signals between hearing aids comes at the cost of power consumption and bandwidth, especially since commercial manufacturers prefer a wireless connection between both devices. Therefore, a through evaluation was needed of the obtained gain in SRT when transmitting no, i.e. a bilateral configuration, one or all contralateral microphone signals. A commonly used ADM, already discussed in chapter 2, was used as a reference noise reduction system. The gain in SRT was calculated relative to an unprocessed condition. The algorithms discussed in this chapter were evaluated both monaurally and binaurally.

The main research questions addressed in this chapter are:
(i) What is the noise reduction performance of an MWF in comparison with a standard bilateral ADM in a monaural and a binaural hearing aid configuration.
(ii) What is the gain in noise reduction when evolving from a monaural hearing aid design to a binaural hearing aid design, i.e. adding a third and/or a fourth microphone, positioned at the contralateral hearing aid, to a MWF already using two microphones of the ipsilateral hearing aid.
(iii) What is the cost in noise reduction and speech perception when adding

partial noise estimation into the MWF-scheme, i.e. the MWF-N, which enables a correct sound localization of both the speech and the noise component (chapter 4 and chapter 6).
All three questions are discussed using objective performance measures (using semi-anechoic data and data of a realistic reverberant room) and perceptual performance measures (only for the realistic reverberant environment) in different single and multiple noise source scenarios. The correlation between both performance measures is also discussed.

By combining this chapter with chapter 6, which evaluates the localization performance of a binaural MWF and MWF-N in comparison with a bilateral ADM, an answer can be found to the question how the MWF and MWF-N combine noise reduction performance with the preservation of binaural cues in comparison with a state-of-the-art ADM approach in commercial digital hearing aids.

5.2 Methods

5.2.1 General

First, different sets of impulse responses were measured between a loudspeaker and the microphones on two BTE hearing aids worn by a CORTEX MK2 manikin. Loudspeakers were placed at 1m distance of the center of the head and impulse responses were measured in the horizontal plane in steps of 15 degrees. Measurements were done in a room with dimensions 5.50m x 4.50m x 3.10m (length x width x height) and acoustical curtains were used to change its acoustical properties. Two different acoustic environments were studied with a reverberation time, linearly averaged over all one-third octave bands between 100 and 8000Hz, of respectively $T_{60} = 0.21s$ and $T_{60} = 0.61s$ with the latter value corresponding to a realistic living room condition. The reverberation time was measured similar to the measurements done in section 3.2.1. The results, as a function of frequency for both acoustic environments is given in Figure 5.1.

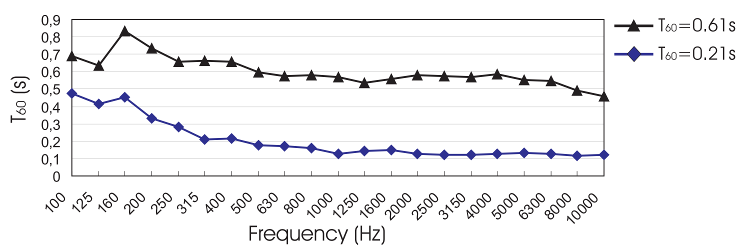

Figure 5.1 — Frequency dependent reverberation time of both acoustic environments.

Then, the impulse responses were convolved with the appropriate speech and noise material to generate the four input microphone signals for the different spatial scenarios used in the perceptual and the objective evaluations. The spatial scenario, with a target signal (S) arriving from angle x and one or multiple noise sources (N) arriving from angle(s) y, is defined as S_xN_y. The angles are defined clockwise with zero degrees being in front of the subject. The generated microphone signals were used as input for the different algorithms.

5.2.2 Noise reduction algorithms

Different microphone combinations were used to evaluate the influence of adding contralateral microphones to the MWF and MWF-N algorithm. A bilateral ADM and MWF system with each hearing aid independently using its own two microphone signals were first evaluated. In case of MWF based algorithms, this can also be interpreted as a binaural MWF configuration in which the communication link has broken down. The MWF systems were then extended by adding one or two microphone signals of the contralateral hearing aid (M_C). The six different MWF based systems are referred to as MWF_{2+M_C} and $\text{MWF}_{2+M_C}\text{-N}_\eta$ with $0 \leq M_C \leq 2$ and η the trade-off parameter between noise reduction and the preservation of binaural cues (eq. 4.58). The list of algorithms evaluated during this chapter is given in Table 5.1.

evaluated algorithms		spatial scenarios	
MWF_{2+0}	b+m	S_0N_x	x between 0° and 330°
MWF_{2+1}	b	$\text{S}_{90}\text{N}_{180}$	single noise source N at 180°
MWF_{2+2}	b	$\text{S}_{90}\text{N}_{270}$	single noise source N at 270° (=-90°)
$\text{MWF}_{2+0}\text{-N}_{0.2}$	b	$\text{S}_{45}\text{N}_{315}$	single noise source N at 315° (=-45°)
$\text{MWF}_{2+1}\text{-N}_{0.2}$	b	S_0N_{2a}	noise sources at -60° and +60°
$\text{MWF}_{2+2}\text{-N}_{0.2}$	b	S_0N_{2b}	noise sources at -120° and +120°
ADM	b+m	S_0N_{2c}	noise sources at 120° and 210°
unproc	b+m	S_0N_3	noise sources at 90°, 180° and 270°
		S_0N_{4a}	noise sources at 60°, 120°, 180° and 210°
		S_0N_{4b}	noise sources at 60°, 120°, 180° and 270°

Table 5.1 — The list of algorithms and spatial scenarios evaluated in this chapter. All algorithms were evaluated using objective performance measures (section 5.3.1). The second column represents whether bilateral/binaural (b) or monaural presentations (m) were used during the perceptual evaluation of the corresponding algorithm (section 5.3.2). The third and fourth column represents the list of spatial scenarios, S_xN_y, the evaluated during the objective evaluations. x represents the location of the speech source, y represents the location of the noise source(s). The conditions S_0N_{60}, $\text{S}_{90}\text{N}_{270}$ and S_0N_3 were also evaluated perceptually.

For both the MWF and the MWF-N algorithm, 'noise' and 'speech and noise' correlation matrices, $\mathbf{R}_{vv}$ and $\mathbf{R}_{yy}$, were calculated using a perfect VAD. The converged filters $\mathbf{W}$ were calculated from these correlation matrices. Calculating the correlation matrices and the accompanying filters was done off-line for each different parameter setting, i.e. a different acoustic environment, a different input SNR, a different spatial scenario, a different parameter setting of μ and η and so on. In this chapter only the most relevant set of the acquired data is presented with a filter length fixed to 96 taps per microphone channel and using only one set of speech and noise material, specified in sections 5.2.3 and 5.2.4.

A commercially used bilateral ADM, the general principles of which have been discussed in chapter 2, was used as a reference noise reduction system. Unlike the MWF-based algorithms, the ADM relies on the assumption that the target signal arrives from the frontal field of view and that jammer signals arrive from the back hemisphere. The ADM uses the physical differences in time of arrival between the microphones to improve the SNR by steering a null in the direction of the jammer signals. The bilateral ADM used in this chapter, is based on two omnidirectional microphones of the ipsilateral hearing aid. The adaptive parameter β of both ADM's was constrained between 0 and 0.5 and hence avoids noise reduction in the frontal hemisphere.

Besides the ADM, the MWF and the MWF-N algorithm with various parameter settings, an unprocessed condition (unproc), using the front omnidirectional microphone of each hearing aid, was evaluated and used as a reference. Evaluations were done after convergence of the filters for all algorithms. All simulations were done using a sampling frequency of $f_s = 20480Hz$.

5.2.3 Objective evaluation

The improvement in speech intelligibility weighted SNR (SNR_{SI}), defined by Greenberg et al. (1993), was used to evaluate the noise reduction performance of the algorithms. This measure was already defined in eq. 4.81.

Noise reduction performance was evaluated using an average speech spectrum of a Dutch male speaker (from VU test material, Versfeld et al. (2000)) as target sound (S) and multitalker babble (Auditec of St.Louis) as jammer sound (N). For multiple noise source scenarios, time-shifted versions of the same noise source were generated to obtain uncorrelated noise sources (Veit and Sander, 1987).

Since simulations are a time-efficient way to assess the performance of noise reduction algorithms, a large number of spatial scenarios were evaluated using one target signal and one, two, three and four noise sources (see Table 5.1 for a list of the studied single and multiple noise source scenarios). Simulations were done for both $T_{60} = 0.21s$ and $T_{60} = 0.61s$ to evaluate the influence of reverberation on the algorithms. Unless specified otherwise, the parameter μ was fixed to $\mu = 5$, which, as will be shown in Figure 5.3, proved to be the most appropriate value of μ. The input SNR during the simulations was fixed to 0dBA, which is well within the range in which noise reduction algorithms in hearing aids are typically used, i.e. from around -5dB to +5dB. Calibrations were always performed in the center of the loudspeaker setup, in absence of the subject or manikin.

5.2.4 Perceptual evaluation

SRTs were measured with ten normal hearing subjects using an adaptive test procedure (Plomp and Mimpen, 1979). The procedure adjusts the level of the speech signal in steps of 2dB to extract the 50% speech reception threshold. The level of the noise signal was calibrated with the SPL, averaged over the left and the right ear, equal to 65dBA. The male sentences of the VU test material (Versfeld et al., 2000) were used as speech material and a multitalker babble (Auditec of St.Louis) was used as noise source.

The seven algorithms were perceptually evaluated in a binaural way, presenting signals at both the left and the right ear. The MWF_{2+0} and the ADM were also tested for one ear only with a monaural presentation of the stimuli (for the full list of conditions see Table 5.1). In the monaural evaluation, signals were presented to the right ear of the subjects. In both the binaural and the monaural presentation an unprocessed condition was used as a reference, bringing the total of tested conditions to eleven. The gain in SRT achieved by each algorithm was calculated by subtracting the SRT score (in dB SNR) of the algorithm from the unprocessed SRT score, i.e.

$$\Delta SRT_{algo} = SRT_{unproc} - SRT_{algo}. \tag{5.1}$$

Tests were performed in a double walled sound booth under headphones (TDH-39) using an RME HAMERFALL MULTIFACE II soundcard and a TUCKER DAVIS HB7 headphone driver. The perceptual evaluations were carried out using the impulse responses of the acoustic environment with $T_{60} = 0.61s$, i.e. a realistic living room condition. Because of practical considerations, three spatial scenarios, selected from the list of scenarios tested in the objective evaluation, were perceptually evaluated, i.e. S_0N_{60}, $S_{90}N_{270}$ and a triple noise source condition $S_0N_{90/180/270}$.

5.3 Results and Analysis

5.3.1 Objective evaluation

First, the influence of input SNR and the parameter μ on the performance of the MWF is briefly discussed. Secondly, the noise reduction performance of the MWF is thoroughly evaluated and compared with the ADM. Finally, the MWF-N is discussed.

MWF

Figure 5.2 evaluates the influence of the input SNR on the MWF_{2+1} and the ADM. Two different input SNRs were evaluated which represent the range in which hearing aids are most commonly used. This figure shows that the noise reduction performance of both the MWF_{2+1} and the ADM were not influenced by the input SNR in the range of -5dB to +5dB. Similar results were obtained when evaluating the MWF_{2+0} and MWF_{2+2} algorithm. All further objective evaluations are therefore based on an input SNR of 0dBA.

As stated in section 5.2.3, the majority of evaluations were based on a trade-off parameter μ fixed to $\mu = 5$ (eq. 4.29 and eq. 4.58). Figure 5.3 shows the influence of μ on the SNR improvement achieved by the MWF_{2+1} at the left hearing aid in a subset of spatial scenarios. It can be seen that, when increasing μ to values higher than $\mu = 5$, no large gains in SNR improvement were observed. Due to the amount of allowed speech distortion, both the speech and the noise component decreased in intensity if $\mu > 5$. Listening tests confirmed that $\mu = 5$ was an appropriate value to further evaluate the MWF and the MWF-N.

Figure 5.4 shows the measured speech intelligibility weighted gain in SNR for a target speech source arriving from 0° and a single noise source arriving from x° (S_0N_x), for the left and the right hearing aid when using an ADM and three different MWF algorithms for a room with a low ($T_{60} = 0.21s$) and a realistic ($T_{60} = 0.61s$) reverberation time respectively. The performance in more challenging scenarios, with multiple noise sources or a non-zero speech source angle, are shown in Figure 5.5.

For both the single noise source data as well as for the more complex spatial scenarios it is observed that the acoustical parameters had a very large effect on the noise reduction performance of the algorithms. Due to the presence of reflections, the performance of all algorithms drastically decreased which is a well known effect from literature. In case of a low reverberant condition, gains of up to 23dB were obtained. In a more realistic environment, this performance dropped to 12dB for the same spatial scenario and the same hearing aid, i.e. the scenario S_0N_{120} at the right hearing aid.

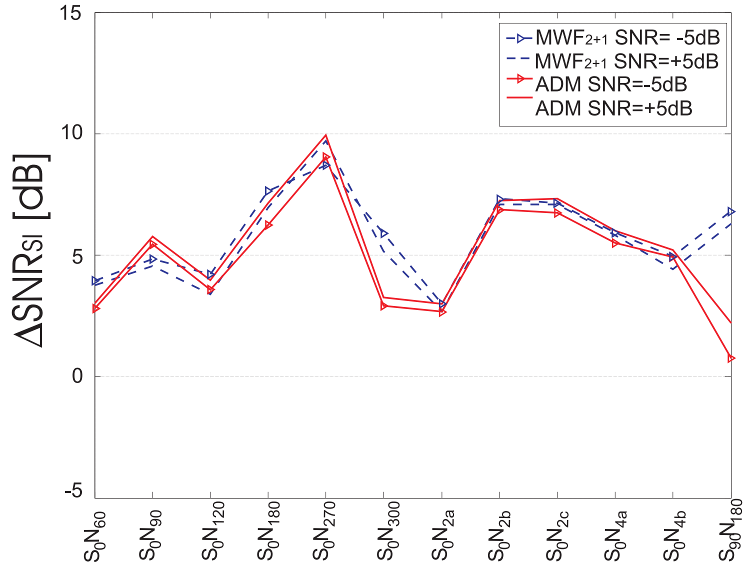

Figure 5.2 — The influence of SNR on the noise reduction performance obtained by an MWF$_{2+1}$ in $T_{60} = 0.61s$. The SNR improvement of the left hearing aid is shown with $\mu = 5$, no large deviations in performance were observed for both the MWF$_{2+1}$ and the ADM between -5 and +5dB which is the typical range in SNR in which noise reduction algorithms for hearing aids are functioning.

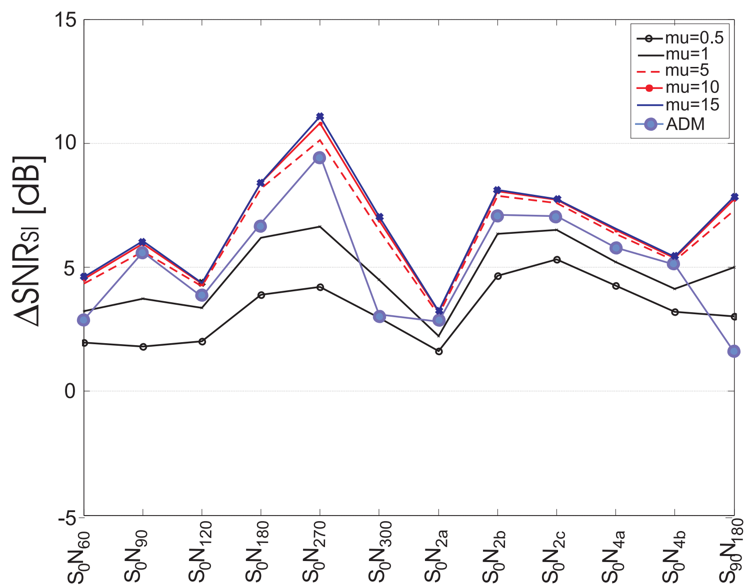

Figure 5.3 — The influence of μ on the noise reduction performance obtained by an MWF$_{2+1}$ in $T_{60} = 0.61s$. The SNR improvement of the left hearing aid is shown at an input SNR of 0 dBA. If $\mu > 5$, no large increases in SNR improvement were observed.

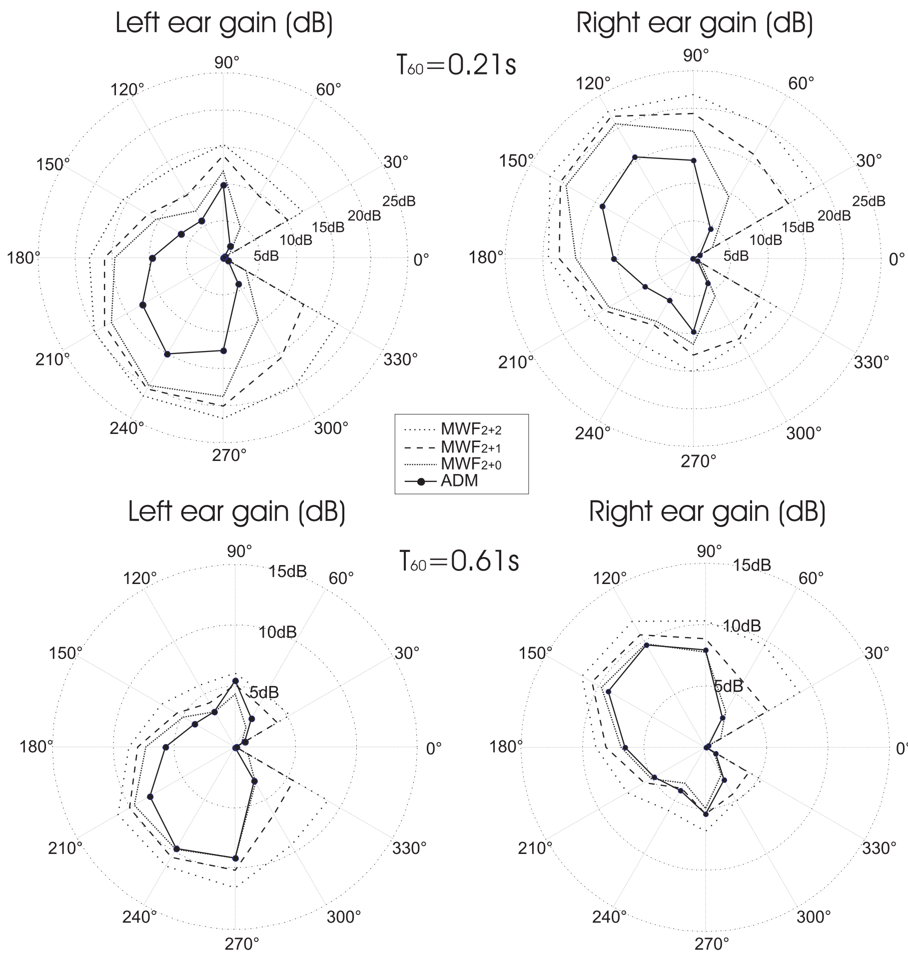

Figure 5.4 — Objective speech intelligibility weighted gain, ΔSNR_{SI}, of the MWF, using different microphone combinations, and the ADM for single noise source scenarios with speech arriving from 0° and noise arriving from x degrees (S_0N_x). The signals were calibrated to an input SNR of 0dBA. The output signals were analyzed every 30° in two different acoustic environments. If reverberation increased, the performance of all algorithms dropped significantly. Adding contralateral microphones to the MWF$_{2+0}$ significantly increased noise reduction performance, especially if speech and noise source were positioned within 60°. The MWF$_{2+0}$ outperformed the ADM only in low reverberant conditions. In realistic reverberant conditions they had a similar performance.

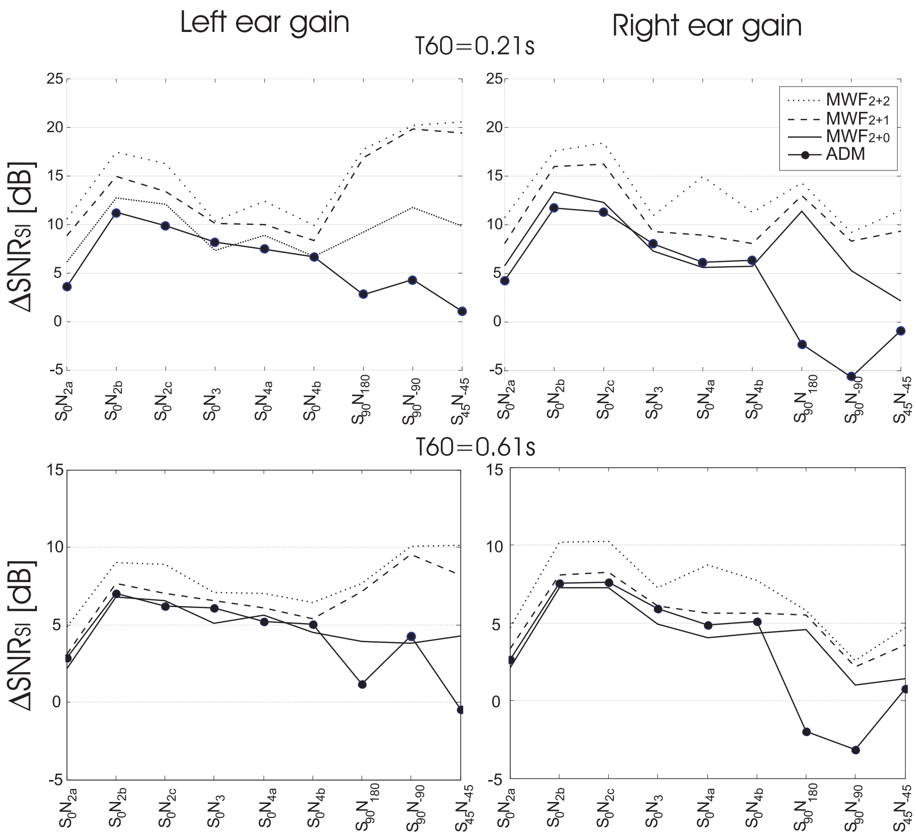

Figure 5.5 — Objective speech intelligibility weighted gain, ΔSNR_{SI}, of the MWF, using different microphone combinations, and the ADM for multiple noise sources. The abbreviations of the spatial scenarios are explained in Table 5.1. Two different acoustic environments were evaluated. The signals were calibrated at an input SNR of 0dBA. The ADM was outperformed by the MWF if speech was not arriving from the forward field of view. Adding contralateral microphones to the MWF_{2+0} scheme did, in general, significantly improve noise reduction performance.

It is also observed that, in single noise source scenarios (Figure 5.4), extending the MWF_{2+0} with contralateral microphone signals substantially increased noise reduction performance, especially if the speech and the noise source were positioned within 60° of each other. In these spatial scenarios, an additional gain of 7.5dB to 14dB in $T_{60} = 0.21s$ and of 3.1dB to 7.6dB in $T_{60} = 0.61s$ was obtained for the right hearing aid when going from the MWF_{2+0} to the MWF_{2+2}. In the other single noise source scenarios, the benefit was much more modest with an average difference between the MWF_{2+0} and respectively the MWF_{2+1} and the MWF_{2+2} of 1.4±0.7dB and 3.3±1.0dB for $T_{60} = 0.21s$ and of 0.8±0.3dB and 2.2±0.3dB for $T_{60} = 0.61s$. Interestingly the MWF$_{2+0}$ outperformed the ADM in low reverberant conditions. However, in a realistic environment both bilateral algorithms had a similar performance.

For the multiple noise source scenarios, as shown in Figure 5.5, the same trends are observed, with the MWF$_{2+2}$ outperforming the MWF$_{2+1}$, which in turn performed better than the MWF$_{2+0}$ and ADM. For both acoustic environments, both two-microphone algorithms, i.e. the ADM and the MWF$_{2+0}$, tended to have a similar performance. However, for the different noise source scenarios with the target signal not arriving from 0°, the ADM was easily outperformed by all MWF algorithms. In these scenarios, the ADM only showed very small improvements or even a decrease in SNR (up to -5dB and -2.5dB for $T_{60} = 0.21s$ and $T_{60} = 0.61s$ respectively).

For the multiple noise source scenarios, it is observed that the gain in noise reduction achieved by extending the MWF$_{2+0}$ with contralateral microphone signals was highly dependent on the spatial scenario. Very large gains in noise reduction performance were observed if the transmitted contralateral microphone signals had a better SNR compared to the SNR present at the ipsilateral hearing aid, e.g. a large noise reduction benefit was observed in the scenario $S_{90}N_{270}$ at the left hearing aid when extending the MWF$_{2+0}$ algorithm with one or two contralateral microphone signals. This large gain was not observed at the right hearing aid. It seems that sending over a signal with a relative high SNR enabled the ipsilateral hearing aid to drastically improve SNR.

MWF-N

As discussed in chapter 4, the MWF-N enables the user to correctly localize the speech and the noise component when used in a binaural hearing aid configuration (for localization experiments, see chapter 6). This is in contrast with other signal processing schemes for hearing aids, e.g. the ADM and partly (only for the noise component) the MWF. The parameter η controls the amount of noise which remains unprocessed by the algorithm (eq. 4.58). This restores the binaural cues of the noise component.

Figure 5.6 illustrates the influence of the parameter η on the estimated noise reduction performance of the MWF$_{2+2}$ and the MWF$_{2+0}$. This demonstrates that, when adding a partial noise estimate to the MWF algorithm (MWF-N$_\eta$), the loss in noise reduction is dependent on the parameter η, but also on the amount of noise reduction originally obtained by the MWF. Larger losses were observed if a high noise reduction performance was already obtained by the MWF-algorithm. As a consequence, the influence of the parameter η was more pronounced on the MWF$_{2+2}$ than on the MWF$_{2+0}$ algorithm. This can also be derived from eq. 4.67.

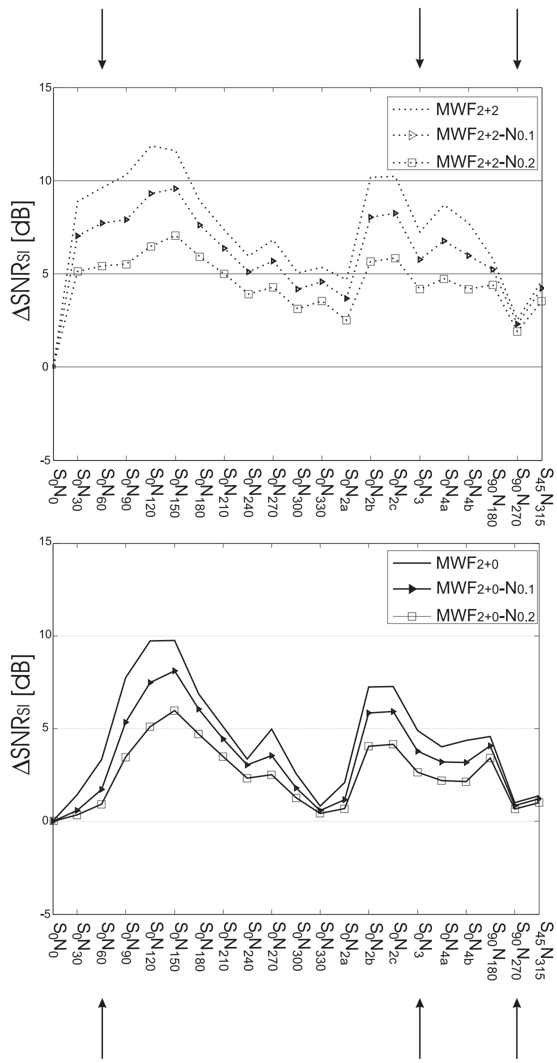

Figure 5.6 — The influence of eta on noise reduction performance: objective speech intelligibility weighted gain at the right hearing aid of the MWF and the MWF-N0.2 for $T_{60} = 0.61s$. A four-and two-microphone MWF based system were evaluated. The multiple noise source conditions have the same abbreviations as used in Figure 5.5. It can be seen that the drop in noise reduction when adding a partial (η) noise estimate was dependent on η and the originally obtained amount of noise reduction ($\eta = 0$) which was dependent on the spatial scenario. The arrows highlight which spatial scenarios were evaluated perceptually.

5.3.2 Perceptual evaluation

To validate the performance of the MWF and MWF-N, a number of perceptual evaluations were performed. Three spatial scenarios were selected to be further evaluated (see Table 5.2 or the arrows on Figure 5.6). The parameters μ and η were fixed to $\mu = 5$ and η=0.2. Table 5.2 shows the improvement in SRT relative to the unprocessed condition, i.e. the effective speech enhancement, averaged over 10 normal hearing subjects. The intelligibility weighted gains in noise reduction, calculated during the objective evaluations, were added to each condition for both the left and the right hearing aid. All statistical analyses were done using SPSS 15.0. For conciseness, the term 'factorial repeated measures ANOVA' is abbreviated to 'ANOVA' and pairwise comparisons discussed throughout this document were always Bonferroni corrected for multiple comparisons. The reported p-values of the pairwise comparisons are lower bound values.

Bilateral/binaural presentation

To compare the different algorithms, an ANOVA was carried out on the SRT data which was used to calculate the average gains (eq. 5.1) shown in Table 5.2. The ANOVA was carried out using the factors algorithm (7 algorithms and an unprocessed condition) and spatial scenario (3 spatial scenarios). An interaction was found between both factors (p=0.005). This was expected since the performance of the algorithms is clearly dependent on the location of the speech and the noise source(s). Therefore an ANOVA and pairwise comparisons were carried out for each spatial scenario. For all three spatial scenarios a main effect for the factor algorithm was found (p=0.002, p<0.001 and p<0.001 for respectively S_0N_{60}, $S_{90}N_{270}$ and $S_0N_{90/180/270}$).

First, comparisons were made with the unprocessed condition. A '*' was added to Table 5.2 if the algorithm generated a significant gain in SRT compared to the unprocessed condition. For the scenario S_0N_{60} a significant gain in noise reduction was achieved by all algorithms except for the ADM (p=0.155) and the $\text{MWF}_{2+0}\text{-N}_{0.2}$ (p=1.000). The highest significant gain was obtained by the MWF_{2+2} algorithm (4.3dB, p<0.001). The lowest significant gain was obtained when using the MWF_{2+0} (1.0dB, p=0.036). For the scenario $S_{90}N_{270}$ a significant speech enhancement was achieved only by the $\text{MWF}_{2+2}\text{-N}_{0.2}$ algorithm (2.0dB, p=0.047). When using the ADM in this scenario, a significant decrease in speech understanding was observed (-4.3dB, p<0.001). For the triple noise source scenario all MWF algorithms showed a significant gain in speech understanding ranging from 2.3dB for the $\text{MWF}_{2+0}\text{-N}_{0.2}$ (p=0.019) to 4.6dB for the MWF_{2+2} (p<0.001). The ADM showed no significant improvement compared to the unprocessed condition (p=0.435).

Bilat/bin SRT [dB]	S_0N_{60}			$S_{90}N_{270}$			$S_0N_{90/180/270}$		
	perceptual	Left	Right	perceptual	Left	Right	perceptual	Left	Right
ADM	2,1 ± 1,9	2,7	2,8	-4,3 ± 1,3*	4,3	-3,2	1,3 ± 1,4	6,0	5,9
MWF_{2+2}	4,3 ± 1,5*	4,9	9,6	0,7 ± 1,4	10,0	2,5	4,6 ± 0,8*	7,1	7,2
MWF_{2+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_{2+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_{2+2}-$N_{0.2}$	3,6 ± 1,4*	3,3	5,4	2,0 ± 1,4*	4,3	1,9	3,2 ± 0,8*	4,1	4,2
MWF_{2+1}-$N_{0.2}$	2,7 ± 1,3*	2,6	3,0	1,5 ± 1,6	3,9	1,6	3,4 ± 0,8*	3,7	3,3
MWF_{2+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
Monaural SRT[dB]									
ADM	5,4±2,0*		2,8	-5,4±1,2*		-3,2	3,4±2,3*		5,9
MWF_{2+0}	3,4±1,3*		3,3	-0,7±1,4		1,0	5,0±1,6*		4,9
SNR-unproc[dB]									
binaural	-6,2±1,8			-9,1±1,7			-7,2±1,6		
monaural	2,8±2,0			-8,0±1,7			-3,0±2,1		

Table 5.2 — The gain in SRT of the different algorithms, averaged over ten normal hearing subjects, in the three different spatial scenarios. The bottom rows show the SNRs at which the unprocessed reference SRTs were measured for the monaural and the binaural presentations. The inter-subject variance is also reported. A '*' depicts a significant gain in SRT relative to the unprocessed condition. For each spatial scenario and each algorithm, the objective noise reduction performance, calculated during the objective evaluation, which is commonly used to predict the SRT gain is added to the table for the left and the right hearing aid. Besides the condition $S_{90}N_{270}$, the MWF_{2+0} and ADM have approximately the same performance. Performance typically improves when adding contralateral microphone signals to the MWF. When adding a partial noise estimate to the MWF, its performance may drop significantly depending on the spatial scenario. Overall, a large correlation is observed between the perceptual evaluation and the predicted values at the hearing aid with the best input SNR.

Second, all MWF approaches were compared with the ADM. For the spatial scenario S_0N_{60}, only the MWF_{2+2} showed a significant gain compared to the ADM (2.2dB p=0.013), the MWF_{2+1} showed a non-significant gain of 1.6dB (p=0.061). The performance of the MWF_{2+0} showed no significant difference with the ADM (which is also a two microphone algorithm). For the scenario $S_{90}N_{270}$ all MWF and MWF-N algorithms showed a clear benefit (all p-values $p \leq 0.001$) compared to the ADM. This benefit was in the range of 3.1dB for the MWF_{2+0} to 6.3dB for the MWF_{2+2}-$N_{0.2}$. For the triple noise source scenario a significant benefit was found for the MWF_{2+2} (3.3dB, p<0.001), the MWF_{2+1} (2.7dB, p<0.001) and for the MWF_{2+1}-$N_{0.2}$ (2.1dB, p=0.002). Since the MWF_{2+1}-$N_{0.2}$ showed a significant gain compared to the ADM, it was expected that also the MWF_{2+2}-$N_{0.2}$, which has an extra microphone input, would show this benefit. However, no statistically significant difference was found between this algorithm and the ADM (1,9dB p=0.164).

Third, the influence of adding contralateral microphones to the original two microphone MWF-scheme was examined. For the spatial scenario S_0N_{60} both the MWF_{2+1} and MWF_{2+2} showed a significant increase in performance of respectively 2.8dB (p=0.022) and 3.3dB (p=0.001) compared to the MWF_{2+0}. The MWF_{2+2} and MWF_{2+1} were statistically not significantly different. For the MWF algorithms using $\eta = 0.2$, the same trends were observed, but these differences were not statistically significant (MWF_{2+1}-$\text{N}_{0.2}$ and MWF_{2+2}-$\text{N}_{0.2}$ showed an average improvement of respectively 1.7dB (p=0.341) and 2.5dB (p=0.125) compared to the MWF_{2+0}-$\text{N}_{0.2}$). For the spatial scenario $S_{90}N_{270}$ the same observations were made. With MWF_{2+1} and MWF_{2+2} performing statistically better than MWF_{2+0} (respectively 1.5dB, p=0.033 and 1.8dB, p=0.001) and no significant difference between MWF_{2+2} and MWF_{2+1}. Again both the MWF_{2+1}-$\text{N}_{0.2}$ and MWF_{2+2}-$\text{N}_{0.2}$ showed the same non significant trend compared to the MWF_{2+0}-$\text{N}_{0.2}$ (with respectively a gain of 1.5dB, p=0.454 and 1.9dB, p=0.215). For the triple noise source scenario only the MWF_{2+2} performed significantly better than the MWF_{2+0} (1.7dB, p=0.004). Again both the MWF_{2+1}-$\text{N}_{0.2}$ and MWF_{2+2}-$\text{N}_{0.2}$ showed a non significant improvement compared to the MWF_{2+0}-$\text{N}_{0.2}$.

The last comparisons examined the impact of introducing the partial noise estimate using $\eta = 0.2$ to the original MWF algorithm (MWF vs MWF-$\text{N}_{0.2}$). Over all three spatial scenarios only one significant difference was found when comparing the performance of MWF_{2+M_C} with MWF_{2+M_C}-$\text{N}_{0.2}$ with M_C ranging from 0 to 2. A significant increase in performance of 1.4dB was observed (p=0.016) when comparing the MWF_{2+2}-$\text{N}_{0.2}$ with the MWF_{2+2} in the triple noise source scenario. All other comparisons showed non-significant differences. However some trends were observed. In the triple noise source scenario and in scenario S_0N_{60}, the MWF_{2+M_C}-$\text{N}_{0.2}$ tended to decrease performance compared to the MWF_{2+M_C} condition, which was expected since the parameter $\eta = 0.2$ introduces an unprocessed noise component at the output of the noise reduction algorithm. Interestingly this trend was not observed in the scenario $S_{90}N_{270}$. In this scenario the MWF-N algorithms typically outperformed the MWF algorithms.

These trends were verified by using a different ANOVA. In this refined analysis, the factors algorithm (3 different MWF algorithms: MWF_{2+0}, MWF_{2+1} and MWF_{2+2}) and η ($\eta = 0$ and $\eta = 0.2$) were used per spatial condition. For all three ANOVAs no interactions were found between both factors. For the scenario S_0N_{60} no significant effect was observed. For the condition $S_{90}N_{270}$ a significant increase in performance of 1.3dB (p=0.002) was observed when comparing MWF_{2+M_C}-$\text{N}_{0.2}$ with MWF_{2+M_C}. For the triple noise source scenario a significant decrease in performance of 0.8dB (p=0.001) was observed when introducing $\eta = 0.2$. In all three of these ANOVAs a significant increase in performance was found when introducing one or two contralateral microphones, but no significant difference was observed between the 3 and 4 microphone algorithm, confirming the observations made in the previous paragraph.

Monaural presentation

The monaural SRT data, used to calculate the gains shown in the monaural section of Table 5.2, were analyzed using an ANOVA. Again the factors algorithm (2 algorithms and an unprocessed condition) and spatial scenario were used. Similar to the analysis in the bilateral/binaural configuration, an interaction was found between both factors ($p<0.001$) which led to a separate ANOVA and pairwise comparisons per spatial scenario.

In the scenario S_0N_{60}, both algorithms performed significantly better than the unprocessed condition with an average gain of 3.4dB for the MWF_{2+0}, $p<0.001$, and an average gain of 5.4dB for the ADM, $p<0.001$. Both algorithms were significantly different from each other with the performance of the ADM being 2.0dB better than the MWF_{2+0} ($p=0.007$).

For the scenario $S_{90}N_{270}$ the MWF_{2+0} was not significantly different from the unprocessed condition. The ADM showed a significant decrease in performance compared to both the MWF_{2+0} and the unprocessed condition (respectively 5.4dB and 4.7dB, both $p<0.001$).

In the triple noise source scenario both the MWF_{2+0} and the ADM showed a significant improvement compared to the unprocessed condition (respectively 5.0dB, $p<0.001$ and 3.4dB, $p=0.004$). The MWF_{2+0} showed a non-significant ($p=0.056$) gain compared to the ADM of 1.6dB.

Monaural vs. binaural presentation

The bottom rows of Table 5.2 show the SNRs at which the unprocessed reference SRTs were measured. It can be seen that if a bilateral/binaural configuration was used, subjects always benefited from the best ear advantage. This means that if both ears are available, one of the ears may have a much higher SNR than the other ear due to the headshadow effect and the positioning of the sound sources. The human auditory system will automatically focuss on the ear with the best SNR. In condition S_0N_{60}, the noise source was close to the right ear, i.e. the ear used in the monaural evaluation. Therefore, the SNR level at which 50% of the speech was understood was much worse in the monaural presentation compared to the binaural presentation. Overall, it is observed that accessing the signals from both ears always resulted in lower SRT levels. This has motivated the standard use of bilateral hearing aids in case of a bilateral hearing deficit (Libby, 2007).

Objective vs. perceptual evaluations

In Table 5.2, both the objective and perceptual evaluations are shown. Large correlations were present between the objective speech intelligibility weighted SNR improvements and the effective speech enhancement, i.e. the perceptual evaluation. It is observed that in the bilateral/binaural configuration, the perceptual gain in speech reception correlates best with the objective measurement done on the hearing aid with the best input SNR (e.g. for S_0N_{60} the left ear, for $S_{90}N_{270}$ the right ear, both ears for $S_0N_{90/180/270}$). This hearing aid was typically the device with the lowest noise reduction performance. In section 4.3 it was proven that the optimal solution of a binaural MWF tends to equalize the output SNR at the left and the right hearing aid. In other words the hearing aid with the worst input SNR will have the highest noise reduction performance. However, due to the best ear benefit, the objective gain in SNR at the hearing aid which had the best input SNR is the best prediction of the effective improvement in speech reception during perceptual evaluations. The more spectacular SNR improvements measured at the ear with the worst input SNR are typically unrealistic predictions.

Although, in general, large correlations between perceptual and objective performance measures were observed, some differences between both evaluations were still present (Table 5.2). When comparing the objective data of the hearing aid with the lowest gain in SNR and the perceptual evaluation in condition $\mathrm{S}_{90}\mathrm{N}_{270}$, it is observed that the objective performance measures of the ADM and all the MWF algorithms were systematically around 2dB higher than the effective perceptual speech enhancement. This might be due to the fact that speech distortion nor the preservation of binaural cues are taken into account in the objective performance measure. Since the binaural cues were preserved in the unprocessed condition but not in the MWF and ADM-processed conditions, the physical gain in SNR generated by the algorithms might not have resulted in an effective improvement in speech perception. The same overestimation was found for the MWF and ADM algorithm in the triple noise source condition.

The correlation between objective and perceptual evaluations illustrates that the intelligibility weighted SNR improvement is a useful tool for predicting noise reduction performance, even in a binaural configuration. However, it is purely based on monaural SNR information and the best ear advantage has to be taken into account when interpreting the data if sounds are presented to both ears of the listener. Moreover, the improvement in SNR does not take into account the preservation of binaural cues, the distortion of sounds, etc. which may explain differences found between the perceptual and the objective evaluations. Recently, Beutelmann and Brand (2006) published a method which aims at replacing the commonly used monaural speech intelligibility prediction methods by a binaural predictor. Besides taking into account the best ear advantage discussed earlier, the method also takes the binaural processing of the human auditory system into account. This is done based on the binaural equalization-cancellation model of Durlach (1963). Such a "binaural approach"seems more appropriate than monaural SNR improvement measures to realistically predict the effective gain in speech enhancement when evaluating noise reduction algorithms for bilateral or binaural hearing aids.

5.4 Discussion

This chapter presents a verification of the noise reduction performance of the MWF and MWF-N algorithm. A bilateral ADM was used as a reference noise reduction algorithm which is commonly implemented in current bilateral hearing aids to enhance speech perception in noise. Three research questions were raised in the introduction related to the combination of improving speech perception in noise while preserving sound source localization using multi-microphone noise reduction algorithms. The results and analysis from the previous sections will be used to answer these questions.

5.4.1 Noise reduction performance of the MWF

First the performance of the two-microphone MWF, i.e. MWF_{2+0}, will be compared with the two-microphone ADM. In section 5.4.2 the gain in noise reduction performance will be discussed when adding contralateral microphone signals to the MWF running on the ipsilateral hearing aid.

In section 5.3.1 the performance of the MWF was evaluated objectively in two different acoustic environments, i.e. $T_{60} = 0.21s$ and $T_{60} = 0.61s$. In the low reverberant condition, the MWF_{2+0} outperformed the ADM, especially in single noise source scenarios (Figure 5.4) and in conditions in which the target signal was not arriving from the forward field of view (the three right-most data-points of Figure 5.5). Performance of all the adaptive algorithms dropped significantly in a more realistic acoustic environment. This phenomenon is often demonstrated in literature. The simulations suggested that the MWF_{2+0} outperformed the ADM only if the speech source was not arriving from the forward field of view. In all other spatial scenarios, both two-microphone algorithms had approximately the same performance. The perceptual evaluations, carried out with $T_{60} = 0.61s$, support this conclusion. In condition $\text{S}_{90}\text{N}_{270}$, the MWF_{2+0} outperformed the ADM. If the speech signal arrived from $0°$, no significant differences were apparent between the ADM and the MWF_{2+0} when using a bilateral configuration (section 5.3.2). Still, unlike the ADM, the MWF preserves the binaural cues of the speech component independent of the angle of arrival of the signal (see chapters 4 and 6).

Why the ADM caught up with the performance of the MWF in more reverberant conditions can be explained by the MWF, unlike an ADM, not performing any de-reverberation. The MWF is designed to estimate the speech component, X, present at a reference microphone with $X = \mathbf{A}S$ (see section 4.3.2). Hence, no de-reverberation is performed. The ADM, on the other hand, is designed to preserve signals arriving from the frontal hemisphere. In other words, reflections arriving from the back hemisphere are reduced in amplitude. However, this also implies that the ADM will reduce speech perception if the target signal arrives from the side or the back of the head. Therefore the ADM is significantly outperformed by the MWF in these spatial scenarios. This was validated by the perceptual evaluation in which all MWF based algorithms outperformed the ADM in the condition $S_{90}N_{270}$.

5.4.2 Adding contralateral microphone signals

Adding contralateral microphone signals to the ipsilateral hearing aid clearly comes at the cost of transmitting and processing those signals. To evaluate this trade-off, different microphone combinations were evaluated.

The objective evaluations showed that, in single noise source scenarios, with speech arriving from 0°, adding contralateral microphones introduced a large gain in noise reduction performance if the speech and noise source were relatively close to each other (Figure 5.4), i.e. within 60°. In other words the directional pattern generated by the MWF became more narrow when more microphones were used. This effect is well-known in sensor array processing. Typically a large impact is obtained if additional sensors, in our case the contralateral front microphone, are placed sufficiently far away from the original sensors thereby enhancing the size of the array. Extreme examples of this phenomenon are, in the specific case of hearing aids, often referred to as tunnel-hearing (Stadler and Rabinowitz, 1993). Soede, J. and Bilsen (1993) proved that very narrow beams in the horizontal hemisphere could be created when using several (4 to 17) microphones positioned on eyeglasses. In single noise source scenarios with speech arriving from 0° and well separated from the noise source, objective evaluations did not show large noise reduction improvements when adding more microphones (Figure 5.4). This is due to the fact that in single noise source scenarios, the MWF only has to create a single null pointed towards the location of the noise source. As a consequence, adding more degrees of freedom, i.e. more microphones, to a two-microphone system did not tremendously improve noise reduction performance.

Significant gains in noise reduction performance were also obtained for asymmetric single noise source scenarios during the objective evaluations. In these conditions, i.e. $S_{90}N_{270}$, $S_{45}N_{315}$ and $S_{90}N_{180}$ a significant improvement in performance was observed at the ear with the worst input SNR, i.e. the left ear. This is due to the asymmetrical setting of the speech source. Since the microphone inputs of the left hearing aid have a low input SNR, due to the headshadow effect, it will produce a non-optimal estimate of the speech component. However, if a contralateral microphone, which has a higher SNR, is added to the system, a better estimate of the speech component can be generated and noise reduction performance will increase. One may interpret this as introducing the best ear benefit, used by our own auditory system, into the noise reduction algorithm.

One should be aware that the effect of this increased performance at the hearing aid with the worst SNR, i.e. the left ear, will probably be limited in daily life. As mentioned before, the human auditory system mainly uses the ear with the best SNR to listen to speech. The hearing aid with the large gain in SNR, obtained at the ear with the worst input SNR, will produce a similar output SNR as the hearing aid on the other side of the head. Therefore, perceptual evaluations will not confirm the large predicted gain in SNR in a bilateral/binaural hearing aid configuration.

For the multiple noise source scenarios (Figure 5.5), objective evaluations demonstrated that adding more microphones or more degrees of freedom does result in a significant gain in noise reduction. For the very asymmetrical condition, i.e. $S_0N_{60/120/180/210}$ (S_0N_{4a}), it was again observed that a larger benefit was obtained at the ear with the worst input SNR, i.e. the right ear.

In the perceptual evaluations, it was observed that in scenarios S_0N_{60} and $S_{90}N_{270}$ the MWF$_{2+1}$ and MWF$_{2+2}$ outperformed the MWF$_{2+0}$. These observations confirm the results of the objective evaluation, discussed earlier. In the triple noise source scenario only the MWF$_{2+2}$ significantly outperformed the MWF$_{2+0}$ which can be explained by taking into account the degrees of freedom needed to reduce three noise sources. The grouped analysis of the perceptual data of the MWF and MWF-N showed that, in general, a three microphone system, consisting of two ipsilateral and one contralateral microphone outperformed the two-microphone system. Adding a fourth microphone did, in general, not add a significant improvement over the three microphone system. Intuitively this can be explained by the fact that adding a third microphone placed at the other side of the head will introduce a lot of new information to the noise reduction system. The fourth microphone will definitely increase the degrees of freedom of the system, but its impact will be much smaller since it is located very close to the third microphone.

5.4.3 Noise reduction performance of the MWF-N

The binaural MWF-N was presented in chapter 4 to preserve the binaural cues of both the speech and the noise component. This is important for hearing aid users in terms of spatial awareness and release from masking. However, this comes at the cost of some noise reduction. Figure 5.6 of section 5.3.1 illustrated the influence of the parameter η on the estimated noise reduction performance of the MWF$_{2+0}$ and MWF$_{2+2}$ in an environment with a realistic reverberation. It shows that the loss in noise reduction due to the increase in η is dependent on its original noise reduction performance. This can also be seen in the relationship found between the SNR improvement of the MWF and the MWF-N in eq. 4.67.

It was also observed that when adding a partial noise estimate with $\eta = 0.2$, the predicted performance can drop for some spatial scenarios below the performance of the ADM. This can be interpreted as the cost to preserve the binaural cues of the speech and the noise component. However, in the perceptual evaluation, the ADM showed no significant gain in noise reduction performance when compared to the MWF-N(2+2)-$N_{0.2}$ or even with the MWF-N(2+0)-$N_{0.2}$. Moreover, the ADM showed a significant decrease in performance compared to the MWF-$N_{0.2}$ if the speech source did not arrive from 0° for reasons already discussed in the previous section.

The lack of a significant difference in SRT improvement between the ADM and the MWF_{2+0}-$N_{0.2}$ during the perceptual evaluation can be explained by spatial release from masking. The MWF-$N_{0.2}$ preserved the binaural cues of both the speech and noise component which resulted in an improved speech perception in noise. This is not taken into account in the SI-weighted SNR performance measure and speech reception threshold were better than expected. The same effect explains why the MWF_{2+M_c}-$N_{0.2}$ outperformed the MWF_{2+M_c} during the perceptual evaluations in the condition with the largest spatial separation between the speech and the noise source, i.e. $S_{90}N_{270}$. In this condition, the MWF_{2+M_C}-$N_{0.2}$ produced a worse ΔSNR_{SI} compared to the MWF_{2+0}, but since it preserves the users' ability to localize both the speech and the noise component correctly, a significantly better SRT could be obtained. This also illustrates that although ΔSNR_{SI} is a useful tool for predicting noise reduction performance, other factors such as binaural cues should be taken into account when evaluating speech enhancement by noise reduction algorithms in hearing aids.

5.5 Conclusion

In chapter 4, it was shown that MWF-based noise reduction approaches are interesting in terms of preserving binaural cues and hence spatial awareness for hearing aid users. Unlike other noise reduction approaches, it is capable of using multi-microphone information, it can easily integrate contralateral microphone signals and it inherently preserves the binaural cues of the speech component, independent of the angle of arrival of the signal. By preserving part of the noise component, i.e. the MWF-N, the ability to localize both the speech and the noise component can be preserved. This chapter has presented an evaluation of the noise reduction performance of the MWF and MWF-N in comparison with an unprocessed condition and an ADM, which is a commonly used noise reduction system in commercial digital hearing aids. Three different research questions were addressed.

First, it was shown that a two-microphone MWF has approximately the same performance as an ADM. However, it does so while preserving the binaural cues of the speech component. Since the MWF operates independently of the angle of arrival of the signal it easily outperformed the ADM if the speech signal was not arriving from the forward field of view. Moreover, in these scenarios the ADM may even reduce the speech perception of the hearing aid user compared to the unprocessed condition. This was observed during the perceptual evaluation of both the monaural (-5.4dB) and the bilateral (-4.3dB) ADM configuration in the spatial scenario $\mathrm{S}_{90}\mathrm{N}_{270}$. Large differences were observed when comparing the monaural with the bilateral data. It was observed that a bilateral presentation leads to an improved speech perception in noisy environments. This confirms, although tests were performed with normal hearing subjects, the common practice of using bilateral/binaural hearing aids for a bilateral hearing impaired subject.

Second, different microphone combinations were evaluated. A significant gain in performance was found if one contralateral microphone signal was added to the ipsilateral hearing aid. This shows that transmitting microphone signals can result in a significant gain in noise reduction, especially in multiple noise source scenarios or if the speech and the noise source(s) are placed asymmetrically around the head. Adding a second contralateral microphone signal to the ipsilateral hearing aid, however, did not show a significant SRT improvement in the perceptual evaluations.

Finally, it was demonstrated that adding a partial noise estimate, large enough to sufficiently preserve the binaural cues to restore the spatial awareness to the MWF, i.e. the MWF-$\mathrm{N}_{0.2}$ (this will be discussed in the next chapter, chapter 6), reduces the amount of noise reduction in a limited way. Moreover, perceptual evaluations showed that in some conditions, i.e. $S_{90}N_{270}$, the MWF-$\mathrm{N}_{0.2}$ even outperformed the MWF which might be due to improved release from masking effects. The parameter η controls the amount of noise reduction. Therefore it may be used as a control mechanism to maximize or to limit the amount of noise reduction if necessary. This can be done adaptively using sound classification algorithms, which are currently already implemented in some hearing aids.

The study also showed that carefully selected objective performance measures can be very useful to predict the performance of noise reduction algorithms. However, one has to take into account psychophysical properties of the auditory system for a correct interpretation of these objective measures, e.g. the best ear benefit, spatial release from masking effects, etc.. Integrating binaural information in binaural objective performance measures, such as the one suggested by Beutelmann and Brand (2006), may lead to a better prediction of the effective speech enhancement of noise reduction algorithms.

In conclusion, it seems that the binaural MWF based algorithms offer a valid alternative for standard beamforming. Unlike beamforming approaches, the MWF does not rely on the direction of arrival of the speech signal nor on assumptions on the microphone characteristics of the hearing aids. In this chapter, it was shown that the bilateral and the binaural MWF are capable of offering a good noise reduction performance in an environment with realistic acoustical parameters. The impact of the MWF and MWF-N on the binaural cues in realistic conditions is studied in the next chapter by using a localization experiment in the frontal horizontal hemisphere.

Chapter 6

Localization with the MWF and MWF-N vs. an ADM

In this chapter the preservation of binaural cues by the MWF and MWF-N algorithm using a perceptual evaluation, i.e. a localization experiment in the frontal horizontal hemisphere is studied. In chapter 4, a general binaural framework was presented together with an introduction to, and a theoretical evaluation of, the MWF and MWF-N algorithm. It was shown that these algorithms show promising theoretical properties with respect to combining noise reduction performance with the preservation of binaural cues. Chapter 5 further evaluated the noise reduction performance of these algorithms. This chapter presents a localization experiment which intends to estimate how well the algorithms preserve binaural cues. This is important for the preservation of spatial awareness of the subject and to improve speech perception in adverse listening situations.

Section 6.1 presents the specific research questions discussed in this chapter.

Section 6.2 specifies the broadband stimuli, the methodology (verified in chapter 3) and the parameter settings of the different algorithms used during the localization experiment. Three different spatial scenarios were evaluated together with two speech in noise conditions: speech and noise playing separately (S,N) and simultaneously (S+N).

Section 6.3 presents the data obtained during the localization experiment. A statistical analysis is carried out which aims at evaluating the differences between the MWF, the MWF-N, the ADM and the unprocessed condition.

Section 6.4 is an interpretation of, and discussion on, the obtained localization performance when using the MWF and MWF-N algorithm in a realistic environment. An answer is formulated to the specific research questions defined in the introduction.

Section 6.5 summarizes the conclusions made in this chapter.

The results presented in this chapter are discussed in Van den Bogaert et al. (2008*a*).

6.1 Introduction

In chapter 4 and Doclo et al. (2006), it was mathematically proven that a binaural version of the MWF generates filters which perfectly preserve the binaural cues of the speech component but changes the binaural cues of the noise component into those of the speech component. To optimally benefit from spatial release from masking effects and to optimize spatial awareness of the hearing aid user, it would be beneficial to also preserve the binaural cues of the noise component.

Hence, an extension of the MWF was proposed, i.e. the MWF-N, first described by Klasen et al. (2007), which aims at eliminating only part of the noise component. The remaining, unprocessed, part of the noise signal then restores the binaural cues of the noise component of the signal at the output of the algorithm. This is similar to the work of Noble et al. (1998) and Byrne et al. (1998), in which improvements in localization were obtained when using open instead of closed earmolds. The open earmolds enable the use of the direct, unprocessed, sound at frequencies with low hearing loss to improve localization performance.

The main purpose of this chapter is to study the effect of noise reduction algorithms on the ability to localize sound sources in binaural hearing aids. It evaluates two binaural noise reduction techniques, namely the MWF and the MWF-N, as well as a widely used bilateral noise reduction system, namely a bilateral ADM. An unprocessed condition is used as a reference. The evaluation was performed in a room with a realistic reverberation time ($T_{60} = 0.61s$) at two different SNRs. The focus of this chapter is on the localization performance in the horizontal plane for which the interaural time (ITD) and interaural level differences (ILD) are the main cues (section 1.4).

The main research questions addressed in this chapter are:
(i) What is the influence of a commonly used noise reduction algorithm, namely an ADM in a bilateral hearing aid configuration on the ability to localize sound sources in a realistic environment?
(ii) What is the influence of the binaural MWF in a binaural hearing aid con-

figuration on the ability to localize sound sources.
(iii) Does the MWF-N improve localization performance in comparison with the MWF.
By combining the results of chapters 5 and 6 a perceptual validation can be made on how the different algorithms combine noise reduction performance and the preservation of binaural cues. This validation is also discussed for each algorithm in the discussion section.

6.2 Methods

6.2.1 Test setup

Experiments were carried out in a reverberant room with dimensions 5.20m x 4.50m x 3.10m (length x width x height) and a reverberation time T_{60}, averaged over 1/3th octave frequencies from 100 to 8000Hz, of 0.61s. This is the same room as the one used to record the impulse responses in chapter 5. A frequency dependent T_{60} was shown in Figure 5.1. Stimuli were generated off-line (see section 6.2.3) and presented through headphones (SENNHEISER HD650) using an RME Hamerfall DSPII soundcard. Subjects were placed inside an array of 13 FOSTEX 6301B single-cone speakers. The speakers were located in the frontal horizontal plane at angles ranging from -90° to +90° relative to the subject with a spacing of 15°. The speakers were placed at a distance of 1 meter from the subject and were labelled 1 to 13. Since the stimuli were presented through headphones, speakers were used only for visualization purposes. The task was to identify the speaker where the target sound was heard.

6.2.2 Noise reduction algorithms

Three different noise reduction algorithms were evaluated. The first two algorithms were the MWF and the MWF-N with $\eta = 0.2$ (MWF-N$_{0.2}$). Both of these algorithms were implemented using two omnidirectional microphones on the ipsilateral hearing aid and the front microphone of the contralateral hearing aid to generate an output for the ipsilateral ear, i.e. MWF$_{2+1}$ and MWF$_{2+1}$-N$_{0.2}$. Simulations suggested to fix the trade-off parameter in eq. 4.29 and eq. 4.58 to $\mu = 5$ (see chapter 5). The third algorithm was a bilateral ADM, similarly defined as the one used in chapter 5. The outputs of the algorithms were calculated and stored off-line. ADM and MWF filters were trained on the specific spatial scenario and were fixed after convergence. For the MWF, a perfect VAD was used to calculate the filters.

6.2.3 Stimuli

The algorithms were evaluated using a steady speech-weighted noise signal from the VU test material (Versfeld et al., 2000) as speech component (S) arriving from angle x. This signal is a noise signal weighted by the average speech spectrum of a Dutch male speaker. A multitalker babble (Auditec) was used as the jammer sound (N) arriving from angle y, defining the spatial scenario S_xN_y. The spectrum of the speech and noise source are depicted in Figure 6.1. Three different spatial scenarios were evaluated, S_0N_{60}, $S_{90}N_{270}$ and $S_{45}N_{315}$, with two of these conditions being identical to the ones perceptually evaluated in chapter 5. The length of each stimulus was 1s and stimuli were cosine windowed with a rise and fall time of 50ms.

To generate the input signals for all algorithms, stimuli were convolved with the appropriate impulse responses measured between the loudspeakers of the loudspeaker array and the microphones on two behind-the-ear (BTE) hearing aids worn by a CORTEX MK2 manikin. The manikin was placed at the position of the test subjects. The BTE devices used in this study are identical to the ones described in chapter 5. Pilot testing suggested that the MWF filters behaved differently at different input SNRs. Therefore stimuli were generated at two different input SNRs, 0dBA and -12dBA, with the input SNR calculated in the absence of the head.

When testing performance in the unprocessed condition (unproc), the front omnidirectional microphone signals from the left and right hearing aids were presented to the subject.

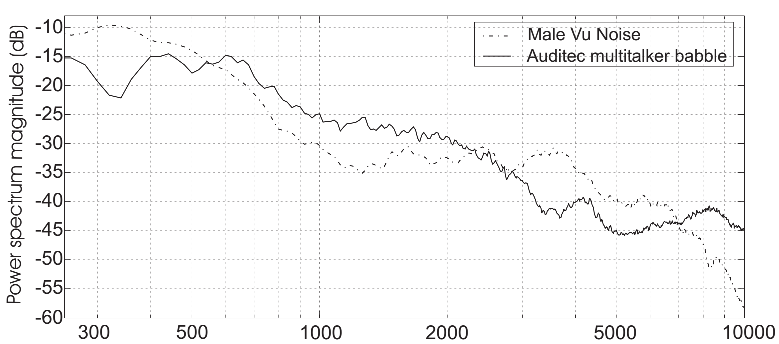

Figure 6.1 — Average power spectrum of the speech and noise signals used to train the filters of the algorithms. The speech source was a steady-state noise with a speech weighted spectrum of a male speaker (VU-male speaker), the noise source was a multitalker babble noise (Auditec). The graph is made at an overall SNR of 0dBA.

6.2.4 Protocol

In the first test condition (S,N), the speech and the noise component were filtered by the fixed filters and presented separately to the subjects. By presenting the two components separately, interactions between components were avoided (masking effects, localizing two sounds is different from localizing one sound source, ...). In the second condition (S+N) the speech and noise component were presented simultaneously and the subject was asked to localize both components. This resembled a converged steady-state real-life situation.

Subjects were instructed to keep their head fixed and pointed towards the 0° direction during stimulus playback and were supervised by the test leader. The task was to identify the speaker where the target sound was perceived. Although only the locations -90°,-45°,0°,45°,60° and 90° were used to generate the stimuli, subjects were free to use all given speaker positions ($-90°$ to $+90°$ in steps of 15°) to identify where the sound was perceived. Tests were restricted to the frontal hemisphere to avoid front-back confusions which would complicate the analysis of the results and which are more related to spectral cues than to binaural cues. None of the subjects experienced major problems with this restriction. Subjects were clearly instructed that the test could be unbalanced. The 5 subjects were all normal hearing subjects working in the Department of ExpORL and were used to performing listening tests.

Pilot testing showed that the presented stimuli might sound diffuse or even arriving from two different angles instead of one clear direction. Therefore subjects were asked to give comments on how the sound was perceived using the following classification: the sound arrives from a point source with one clear direction in space (point), the sound arrives from a wider area (wide), the sound arrives from everywhere (diffuse) or more than one sound source is perceived (dual). If the subjects perceived multiple components at different locations, subjects were asked to report both locations and to report to which direction they would look when hearing this stimulus. This direction was then used as the response to the presented stimulus. Only for the condition S+N the subjects were explicitly asked to report two angles of arrival within one trial, one for the speech and one for the noise component.

The two different sound conditions (S,N, S+N) were presented in different test sessions with the angle of arrival, input SNR and type of algorithm randomized throughout the test. Each stimulus was presented three times and an overall roving level of 6dB was used (ranging from 0dB to -6dB). The presented stimuli were equalized in dBA-level by adjusting the sound level, averaged over the left and right channel, to the same level for all generated stimuli. The stimuli were then presented at a comfortable level chosen by the subject. Because the task was quite hard, the subject had the possibility to repeat the same stimulus over and over again until a clear answer could be given to the test leader, who entered all responses and comments. The test leader had no information on the location of the stimulus nor on the type of stimulus that was played and no feedback was given to the subjects. Typically one session took somewhat more than one hour and several hours were in between different tests. If fatigue or low concentration were observed, breaks were taken during the test.

6.2.5 Performance measures

Different error measures have been used in previous localization studies (Noble and Byrne, 1990; Lorenzi et al., 1999*b*; Van Hoesel et al., 2002). Two commonly used error measures are the root mean square error (RMS) and the mean absolute error (MAE). In this study, a performance measure was calculated per presented source angle. This is different from section 2.3 in which an error measure was used based on the responses given for a full loudspeaker array. Analysis was done in this way since it was observed that the performance of the studied algorithms was largely dependent on the sound source angle. Since this implies that only a very limited number of repetitions were used in each error measure, i.e. only 3 repetitions, and since very large errors were observed, i.e. up to 180°, the MAE error measure was used throughout this chapter. This measure provides a more intuitive insight, since it is a linear error measure, of how a noise reduction algorithm is behaving when stimuli are presented from a certain angle. The MAE was already defined by eq. 2.20 in section 2.3.1.

The smallest non-zero error a subject could make during one test run equals 5° MAE (1 error of 15° made during the three repetitions of the stimulus, n=3). This may seem rather large, however, the statistical analysis in section 6.3 will show that this resolution was already sufficient to demonstrate large effects of, and large differences between, the algorithms in the different spatial scenarios, which was the goal of this study.

6.3 Results and analysis

First the data and analysis of the condition S,N are presented, followed by the data and analysis of the condition S+N. All statistical analysis was done using SPSS 15.0. For conciseness, the term 'factorial repeated measures ANOVA' is abbreviated by 'ANOVA'. Bonferroni corrections were applied to all reported pairwise comparisons.

6.3.1 Condition S,N

Localization data for the condition with the speech and the noise component presented separately to the listener are given in Table 6.1. Table 6.1 indicates where the stimulus was perceived by the different subjects, averaged over the

S_0N_{60}	S_0							N_{60}						
	unproc	ADM		MWF		MWF-$N_{0.2}$		unproc	ADM		MWF		MWF-$N_{0.2}$	
		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12
T	0	0	-5	-15	0	-15	0	90	90	75	0	50	0	85
J	0	0	0	0	0	0	0	90	90	90	0	90	80	90
H	0	0	0	-5	0	0	0	65	60	70	0	45	35	65
L	-5	-5	-5	-15	-10	-15	-10	90	80	85	-5	45	75	85
O	0	20	25	5	35	-10	0	85	80	90	10	80	65	80
Loc (avg)	-1	3	3	-6	5	-8	-2	84	80	82	1	62	51	81
MAE (avg)	1	5	7	8	9	8	2	24	20	22	59	28	25	21
min-max MAE	0-5	0-20	0-25	0-15	0-35	0-15	0-10	5-30	0-30	10-30	50-65	20-35	5-60	5-30
effect		nme	nme	nme	nme	nme	nme		p=0.687	nme	p=0.027	nme	p=1.000	nme
$S_{90}N_{-90}$	S_{90}							N_{-90}						
	unproc	ADM		MWF		MWF-$N_{0.2}$		unproc	ADM		MWF		MWF-$N_{0.2}$	
		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12
T	90	0	0	85	80	85	90	-80	-15	0	80	-55	-90	-85
J	90	0	0	90	80	90	90	-90	0	0	45	-90	-90	-90
H	70	20	15	65	55	70	75	-85	-75	-60	-25	-75	-80	-90
L	80	0	15	80	75	85	75	-70	-35	-35	80	-60	-75	-80
O	75	50	50	70	55	70	75	-75	-20	-10	80	55	30	-65
average	81	14	16	78	69	80	81	-80	-29	-21	52	-45	-61	-82
MAE (avg)	9	76	74	12	21	10	9	10	61	69	142	45	29	8
min-max MAE	0-20	40-90	40-90	0-25	10-35	0-20	0-15	0-20	15-90	30-90	65-170	0-145	0-120	0-25
effect		p=0.038	p=0.035	p=0.423	p=0.056	p=1.000	p=1.000		p=0.687	nme	p=0.027	nme	p=1.000	nme
$S_{45}N_{-45}$	S_{45}							N_{-45}						
	unproc	ADM		MWF		MWF-$N_{0.2}$		unproc	ADM		MWF		MWF-$N_{0.2}$	
		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12		SNR0	SNR-12	SNR0	SNR-12	SNR0	SNR-12
T	50	60	50	80	65	65	80	-50	-60	-85	75	-90	-70	-75
J	60	90	90	90	45	90	50	-45	-90	-90	75	-70	-75	-90
H	45	40	30	45	45	45	45	-75	-70	-70	-50	-75	-70	-75
L	75	70	50	75	75	85	60	-60	-45	-50	-25	-50	-60	-60
O	75	70	75	70	75	75	70	-80	-60	-75	80	75	25	-65
average	61	66	59	72	61	72	61	-62	-65	-74	31	-42	-50	-73
MAE (avg)	16	23	20	27	16	27	16	17	20	29	78	45	37	28
min-max MAE	0-30	5-45	5-45	0-45	0-30	0-45	0-35	0-35	0-45	5-45	5-125	5-120	15-90	15-45
effect		nme	nme	nme	nme	nme	nme		p=0.687	nme	p=0.027	nme	p=1.000	nme

Table 6.1 — Response location (°), averaged over three repetitions, together with the average, minimum and maximum MAE across subjects for the three different spatial scenarios (S_0N_{60}, $S_{90}N_{-90}$, $S_{45}N_{-45}$) and the different processing schemes (unprocessed, ADM, MWF, MWF-$N_{0.2}$) at two different SNRs (0 dB, -12 dB). The speech and the noise sources were presented separately through headphones (S,N). The rows labelled 'effect' show whether a significance difference from the unprocessed condition was found. P-values of pairwise comparisons are shown. If no main effects were found the term 'nme' is used. The individual MAE data are presented in Table 6.4 of Appendix 6.A.

three stimulus repetitions, together with the minimum, maximum and averaged MAE values across subjects. The individual MAE values, on which the statistical analysis is based, are shown in Table 6.4 of Appendix 6.A.

To compare the different algorithms, an ANOVA was carried out on the MAE data (Table 6.4). The factors algorithm (ADM, MWF, MWF-$N_{0.2}$), target (speech or noise component), SNR (0dB and -12dB) and angle (S_0N_{60}, $S_{90}N_{270}$, $S_{45}N_{315}$) were used. As expected, many interactions were found between these factors, e.g. the differences between algorithms were dependent on evaluating the data of the speech or the noise component (algorithm*target p=0.004). This was expected since it was proven in chapter 4 that the MWF tends to behave differently for the speech and the noise component. To disentangle these interactions, separate ANOVAs were carried out for the speech and noise components.

1. *Speech component*: An interaction was found between the factors angle and algorithm (p=0.019). This can be explained by the fact that especially the ADM behaved differently if the angle of arrival of the stimulus equaled 90° compared to the other angles. Hence, separate ANOVAs were carried out for each spatial scenario.

   For S_0N_{60} and $S_{45}N_{315}$ no main effects were found (p=0.470 and p=1.000 respectively for the factor algorithm). For the scenario $S_{90}N_{270}$ a main effect for the factor algorithm was found (algorithm p=0.009, SNR p=0.070, SNR*algorithm p=0.137). Pairwise comparisons showed a significant lower performance for the ADM than for the MWF (difference averaged over both SNRs equaled 58° MAE, p=0.039) and the MWF-$N_{0.2}$ (difference averaged over both SNRs equaled 65° MAE, p=0.019). Table 6.1 shows that for scenario $S_{90}N_{270}$ none of the subjects was capable of localizing the speech component correctly (Table 6.4 of Appendix 6.A) when using the ADM, and sounds were most commonly localized around 0° (4 out of 5 subjects). The MWF-$N_{0.2}$ just failed to give significantly better performance than the MWF (difference of 7° MAE, p=0.057).

   When comparing the algorithms with the unprocessed condition, no main effects were found for scenarios S_0N_{60} and $S_{45}N_{315}$ (p>0.252). For the scenario $S_{90}N_{270}$ a main effect was found. Pairwise comparisons showed that only the ADM performed significantly more poorly compared to the unprocessed condition (a difference of 67° MAE, p=0.038 for SNR=0dB and a difference of 65° MAE, p=0.035 for SNR=-12dB). No significant differences in performance were observed between the unprocessed condition and both the MWF and the MWF-N algorithm in all spatial scenarios when analyzing the data of the speech component.

   Interestingly, subjects also reported a clear difference in sound quality between the different algorithms. Table 6.2 shows the percentage of responses claiming a clear directional sound image during the subjective classification of the stimuli. For the speech component, the combination

	level	unproc	ADM	MWF	MWF-$N_{0.2}$		level	unproc	ADM	MWF	MWF-$N_{0.2}$
S_0	0dB	67	87	93	100	N_{60}	0dB	89	80	80+7du	27+53du
	-12dB		53	87	93		-12dB		93	47+40du	67+27du
S_{90}	0dB	75	13+53di	100	100	N_{270}	0dB	89	7+27di+40br	27+73du	7+93du
	-12dB		27+60di	60	100		-12dB		53+20di+20br	7+87du	13+87du
S_{45}	0dB	58	87	100	100	N_{315}	0dB	89	93	20+67du	7+87du
	-12dB		87	87	73		-12dB		100	27+67du	13+80du

Table 6.2 — Percentage of stimuli which were perceptually classified as a sound arriving from one clear direction in space, averaged over 5 subjects, for the three different spatial scenarios and the different processing schemes. The speech and the noise source were presented separately through headphones (S,N). In the conditions in which most sounds were not categorized as arriving from one clear direction, the percentage of diffuse sounds (di), dual sounds (du) or very broad source (br) is added. The table which contains the full classification by the subjects is given in Table 6.6 of Appendix 6.A.

of ADM and $S_{90}N_{270}$ led to a severely degraded performance compared to all other combinations. Interestingly, these stimuli were often perceived as being diffuse (53% for 0dB and 60% for -12dB). Subjects reported that when perceiving a diffuse sound, 0° was often picked as the direction from where the sound was heard since it is the neutral position in the middle of the sound array. Therefore, these 0° responses should be interpreted carefully. For the MWF and MWF-N algorithm, the speech component was almost always categorized as arriving from one clear direction in space and no large differences with the classification of the unprocessed sounds were observed.

2. *Noise component*: Due to an interaction with SNR (p=0.050), separate ANOVAs were carried out for each SNR. Since the speech and noise component were presented separately and presentation level for both components was calibrated to a comfortable level, the obtained results for the unprocessed stimuli are independent of SNR. Therefore the unprocessed data was incorporated in the ANOVA of each SNR.

   For SNR=0dB a main effect of the factor algorithm was observed (p=0.012). Pairwise comparisons showed significantly lower performance for the MWF than for all other strategies (vs unprocessed p=0.027, vs ADM p=0.017, vs MWF-$N_{0.2}$ p=0.049). This can also be observed in Table 6.1 (and in Table 6.4) which shows that the noise component at the output of the MWF was generally localized at the location of the speech component. No significant differences were found between the unprocessed condition, the ADM and the MWF-$N_{0.2}$ ($p \geq 0.687$).

   For SNR=-12dB, no interactions or main effects were found (angle*algorithm p=0.115, angle p=0.443, algorithm p=0.156), meaning that all algorithms, including the MWF, performed equally well at this SNR.

   Interestingly, no interaction was found at both SNRs between the factors algorithm and angle, although the results in Table 6.1 suggest that the ADM distorted the localization of the noise component in the scenario $S_{90}N_{270}$ (which was also observed in the data of the speech component). Table 6.1 shows that only one out of five subjects, subject H, localized the noise component with the ADM equally well as in the unprocessed condition. The lack of a significant interaction could be due to the limited amount of data collected per condition or due to the large variances observed during this study.

   The subjective classification, shown in Table 6.2, showed a clear drop in performance for almost all spatial scenarios for the MWF and MWF-$N_{0.2}$ compared to the unprocessed condition. This was quite surprising for the MWF-$N_{0.2}$ and for the MWF at SNR=-12dB since their MAE values were relatively modest in these conditions and not statistically different from those for the unprocessed condition. Interestingly, the output of these algorithms were often classified as being a "dual sound". Averaged over the three spatial scenarios, there were 49% and 65% of such cases for the MWF and 78% and 65% of such cases for the MWF-$N_{0.2}$ at 0dB and -12dB respectively. When dual sounds were reported, the sound was perceived as having two components, each arriving from a different angle. Subjects reported that one part arrived approximately from the position of the original noise component, whereas the other part arrived from around the position of the speech component. When using the MWF in the SNR=0dB condition, the sound arriving from the original noise position was typically described as being silent, lower in frequency and less distorted than the other part. For the SNR=-12dB condition, the part arriving from the original noise position was reported as being louder than the distorted part arriving from the speech position.

6.3.2 Condition S+N

S_0N_{60}	S_0 unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12	N_{60} unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12
T	-10	0	-40	-10	-20	-5	-10	90	90	90	80	75	90	70
J	0	0	-25	0	0	0	0	90	90	90	80	75	90	90
H	0	0	0	0	0	0	-5	75	75	75	75	75	80	70
L	-15	-10	-15	-15	-15	-15	-20	90	90	80	90	90	85	90
O	0	50	-5	0	-5	0	-20	70	85	85	85	75	85	70
average (°)	-5	8	-17	-5	-8	-4	-11	83	86	84	82	78	86	78
MAE avg (°)	5	12	17	5	8	4	11	23	26	24	22	18	26	18
min-max	0-15	0-50	0-40	0-15	0-20	0-15	0-20	10-30	15-30	15-30	15-30	15-30	20-30	10-30

$S_{90}N_{270}$	S_{90} unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12	N_{270} unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12
T	75	0	-60	90	90	85	70	-60	-60	0	-75	-65	-85	-70
J	90	15	75	90	85	90	90	-90	-70	-80	-85	-85	-90	-90
H	65	20	20	65	55	75	45	-85	-60	-65	-65	-75	-75	-90
L	90	-5	85	90	80	90	75	-75	-55	-25	-65	-60	-70	-80
O	70	50	75	85	60	75	65	-90	-70	-35	-65	-40	-60	-60
average (°)	78	16	39	84	74	83	69	-80	-63	-41	-71	-65	-76	-78
MAE avg (°)	8	74	51	6	16	7	21	10	27	49	19	25	14	12
min-max	0-25	40-95	5-150	0-25	0-35	0-15	0-45	0-30	20-35	10-90	5-25	5-50	0-30	0-30

$S_{45}N_{315}$	S_{45} unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12	N_{315} unproc	ADM SNR0	ADM SNR-12	MWF SNR0	MWF SNR-12	MWF-$N_{0.2}$ SNR0	MWF-$N_{0.2}$ SNR-12
T	75	75	55	75	80	60	75	-80	-60	-60	-90	-60	-90	-75
J	85	70	90	90	90	90	90	-80	-90	-80	-80	-65	-90	-90
H	45	45	45	45	45	40	50	-60	-60	-70	-75	-50	-70	-80
L	80	75	90	75	65	70	80	-65	-45	-45	-65	-55	-65	-60
O	85	90	80	75	70	75	80	-65	-60	-75	5	-10	-50	-40
average (°)	74	71	72	72	70	67	75	-70	-63	-66	-61	-48	-73	-69
MAE avg (°)	29	26	27	27	25	24	30	25	20	21	36	19	30	26
min-max	0-40	0-45	0-45	0-45	0-45	5-45	5-45	15-35	0-45	0-35	20-50	5-45	15-45	5-45

Table 6.3 — Response location (°), averaged over three repetitions, together with average, minimum and maximum MAE data over the different subjects for the three different spatial scenarios (S_0N_{60}, $S_{90}N_{270}$, $S_{45}N_{315}$) and the different processing schemes (unprocessed, ADM, MWF, MWF-$N_{0.2}$) at two different SNRs (0dB, -12dB). The speech and the noise source were presented simultaneously through headphones (S+N). The individual MAE data, used to perform the statistical analysis are presented in Table 6.5 of appendix 6.A.

Whereas in the first experiment the goal was to gain understanding of how the filtering operations perceptually affect the localization cues, the second experiment was more related to real-life performance. In this experiment, speech and noise components were presented simultaneously which resembled more a real-life listening situation. Subjects were asked to localize both the speech and noise component. Table 6.3 shows the individual data indicating where the stimuli were perceived, averaged over three repetitions, together with the minimal, maximal and averaged MAE values of the tested subjects. The full set of MAE data measured during this condition is reported in Table 6.4 in Appendix 6.A.

In most conditions no differences were found between the data for condition S+N and condition S,N, leading to the same differences between algorithms as discussed for condition S,N. This was assessed for the unprocessed data, the ADM data, the MWF-$N_{0.2}$ data and for the speech component data of the MWF by an ANOVA on all MAE data (S,N and S+N). For the noise component data of the MWF a significant effect of the factor stimulus presentation (S,N vs S+N) was found for the 0dB data (p=0.006) but not for the -12dB data (p=0.233). An ANOVA comparing the 0dB data of condition S+N demonstrated, in contrast with the S,N data, no significant difference between the MWF and all other conditions (factor algorithm, p=0.322). The data in Table 6.3 show that, for both SNRs, the performance of the MWF approached that for the unprocessed condition for the noise component for all three spatial scenarios. The 0dB data of the MWF contrast with the results obtained when speech and noise components were presented separately (Table 6.1).

6.4 Discussion

Three research questions were raised in the introduction related to preserving sound source localization using multi-microphone noise reduction algorithms. The results and analysis from the previous sections will be used to answer these questions. Moreover, by combining the data of this and the previous chapter a perceptual validation can be made on the combination of noise reduction performance and the preservation of binaural cues by the different algorithms. This is also discussed when evaluating the different algorithms in this section. However, first a short discussion on the unprocessed localization performance is appropriate since this condition was used as a reference condition throughout this chapter.

6.4.1 Discussion of reference condition

For the condition S,N, the average localization responses in the unprocessed condition were relatively accurate (Table 6.1) with average MAE values between 1° and 24°, depending on the spatial scenario. Although localization was not perfect, these values are within the expected localization performance. In chapter 2 it was shown that when evaluating the localization performance of normal hearing subjects with a broadband stimulus, a MAE, averaged over all angles, of around 8° can be expected. Large errors, up to 30°, typically occurred at the sides of the head. A slightly poorer performance was expected here since localization experiments were done using headphones (chapter 3) and since the unprocessed stimuli were generated by using front omnidirectional microphone signals of both hearing aids. Therefore, the signals will somewhat sound unnatural with slightly different ITDs and ILDs than normally occurring at the eardrums and with no relevant information of height and no externalization (pinnae effect). However, this condition was taken as a reference since an evaluation was made of the influence of noise reduction algorithms on the localization of sound sources for hearing aid users, which also have an omnidirectional microphone as reference condition. Since the allowed responses were limited to the frontal hemisphere, for reasons already explained, responses at the sides of the head might be slightly biased towards the front. However, this is true for all conditions and does not explain differences observed between algorithms.

For the unprocessed condition, a similar localization performance was found for conditions S,N and S+N. Since the data presented here was limited to only three repetitions for each spatial scenario with a limited number of subjects, one should be careful with generalizing this observation. Other researchers have demonstrated that localizing a sound source can be affected by the absence or presence of other sound signals (Lorenzi et al., 1999*b*).

6.4.2 Evaluation of the bilateral ADM

As a reference noise reduction algorithm for evaluating two MWF-based noise reduction strategies for hearing aids, a bilateral ADM was used. This technique is commonly implemented in current bilateral hearing aids to enhance speech perception in noise. In section 6.3.1 and 6.3.2 it was observed that the localization performance using the ADM was comparable to that for the unprocessed condition for spatial scenarios S_0N_{60} and $S_{45}N_{315}$. However, a large degradation was found for scenario $S_{90}N_{270}$ (Tables 6.1 and 6.3), which was statistically verified for the speech component (section 6.3.1). This can also be seen in Figure 6.3 and Figure 6.4. These figures illustrate the accumulation of responses given by all subjects over the different test conditions. It is observed that when sounds were played from ±90°, the ADM often distorted the perceived location towards 0°. Perceptual evaluation showed that in spatial scenario $\mathrm{S}_{90}\mathrm{N}_{-90}$, the signals generated by the ADM were often described as being diffuse with no directional information present in the signal. Neither the perceptual data, nor the statistical analysis showed a significant impact of SNR on localization performance.

The negative impact of adaptive and fixed directional microphones was also observed in chapter 2 (Van den Bogaert et al., 2006) and in the work of Keidser et al. (2006) respectively. In chapter 2, hearing-impaired users showed a significant decrease in localization performance when using their ADM systems compared to using omnidirectional microphones. This was observed when localizing a broadband stimulus in a noisy environment with the noise sources positioned at $\pm 90°$. A separate analysis showed that this was due to localization errors made when stimuli were presented from the sides, between $\pm 60°$ and $\pm 90°$, of the head. When testing the ADM in silence with a broadband stimulus, no significant decrease in localization performance was observed. This may be explained by the fact that when continuous noise sources were used, the ADM will certainly have converged towards a certain directional pattern. When presenting isolated 200ms or 1s signals to hearing aid users using their commercial hearing aids in silence, it is unknown how the ADM is reacting.

Keidser et al. (2006) tested the influence of multichannel wide dynamic range compression, single channel noise reduction and directional microphones on localization performance. They observed that directional microphone settings had the largest influence on localization performance. The aspect of different directional microphone characteristics for the left and right hearing aids was assessed, using an omnidirectional pattern in both hearing aids as a reference condition. Combining a cardioid pattern at one ear with a figure-8 pattern at the other ear produced the largest decrease in localization performance. It was suggested that this could be an extreme hearing aid setting when using independent ADMs at both sides of the head.

The reason why a significant decrease in localization performance was found only in spatial scenario $S_{90}N_{270}$ can be explained by the fact that the ADM, used in hearing aids, is typically constrained to avoid noise reduction and distortion of signals arriving from the front. For the scenarios $S_{45}N_{315}$ and S_0N_{60}, both the speech and the noise source were within or close to this area. Therefore the binaural cues of both the speech and the noise component remained unchanged. However, due to this constraint, noise reduction performance will also be limited in these scenarios. The low noise reduction performance of a bilateral ADM in these scenarios was illustrated in Table 5.2 of chapter 5 which showed a non-significant noise reduction performance of the bilateral ADM in condition S_0N_{60}. Moreover, in chapter 2.2 (Van den Bogaert et al., 2005) it was shown that the distortion of ITD cues by an ADM or FDM is in relation with the amount of noise reduction. This confirms the good localization performance, but low noise reduction performance of the bilateral ADM in scenarios S_0N_{60} and $S_{45}N_{315}$. Outside the forward field of view, sounds are suppressed. Therefore, both the speech and the noise source were suppressed in scenario $S_{90}N_{270}$. In this spatial scenario, the suppression of both components led to a low or even a negative noise reduction performance (Table 5.2 of chapter 5) and in a decreased localization performance for both the speech and the noise component (Table 6.1 and Table 6.3). These conclusions are supported by chapter 2 and the work of Keidser et al. (2006) in which ITD and ILD measurements on directional microphones showed large ITD and ILD distortion at angles around 90° and much lower distortion between $+50^\circ$ and -50°.

Another observation was that localization performance for the ADM was independent of SNR. This can be explained by the fact that noise reduction is based on exploiting physical differences in time of arrival between the microphones on the hearing aid, and these are independent of SNR. Since the coherence between microphone signals is used to attenuate the strongest source with the angle of arrival in the back hemisphere, the most coherent part of the noise signal is removed. This would explain the classification of the output as sounding "diffuse".

6.4.3 Evaluation of the binaural MWF

In chapter 4 it was mathematically proven that a binaural version of the MWF perfectly preserves the binaural cues of the speech component but changes the cues of the noise component into those of the speech component. This was also observed in ITD-error simulations, used to predict localization performance, in the work of Klasen et al. (2007). As a consequence, large localization errors of the noise component were expected in the subjective evaluation. These errors were indeed observed and statistically confirmed for the SNR=0dB condition when the filtered speech and noise component were presented separately to the subjects (S,N). However, they were not observed when SNR=-12dB (section 6.3.1) nor when speech and noise sources were presented simultaneously to the subject (S+N) (section 6.3.2). This can also be seen in Figure 6.3 and Figure 6.4. It is observed that in condition S,N (Figure 6.3), the noise component was often localized at the location of the speech component, especially when SNR=0dB. In condition S+N, this happened much less frequently. Since large inter-subject variances were present in the data, these figures should be interpreted with care and together with the ANOVAs and other Tables and Figures given in this chapter.

The differences found between the SNR=0dB and the SNR=-12dB condition, can be explained by using the subjective classification in Table 6.2. Despite the good localization performance for the SNR=-12dB condition, Table 6.2 suggests a decrease in sound quality for both SNRs. Subjects reported that the noise component of the MWF algorithm sounded as if it was produced by sound sources at two different positions, one at the original noise position which sounded relatively clear and one at the speech position which sounded more distorted. In the SNR=-12dB condition, subjects preferred the sound arriving from the original noise location often resulting in a correct localization of the noise component. In the SNR=0dB condition, subjects preferred the sound arriving from the original speech location. However, individual subjects did not always follow this general trend, e.g., for spatial scenarios $S_{90}N_{-90}$ and $S_{45}N_{-45}$, subject O preferred the sound arriving from the original speech location when using an MWF at a SNR=-12 dB which demonstrates the variability between subjects due to the dual sound phenomenon.

The reason for the dual sounds can be found in the filter generation of the MWF. Since the speech correlation matrix is estimated as $\mathbf{R}_{xx} = \mathbf{R}_{yy} - \mathbf{R}_{vv}$ (eq. 4.21), where $\mathbf{R}_{yy}$ and $\mathbf{R}_{vv}$ were computed during different time periods, this estimate will be poor at a low SNR. Hence, in the frequency region with high SNR (in our case between 3000 and 5500Hz, see Figure 6.1) a good estimate was obtained such that the binaural cues of the noise component were changed into those of the speech component, as illustrated by Figure 6.2 and proven in chapter 4. On the other hand, in the frequency region with low SNR (in our case between 500 and 2500Hz, see Figure 6.1), a poor estimate was obtained such that the output contained binaural cues corresponding to the original position of the noise source. Because of this different behavior for different frequency regions, a dual sound was created. For the low overall SNR, i.e. SNR=-12dB, a large proportion of the noise component contained the binaural cues of the original noise angle, which resulted in a correct localization of the noise component (Table 6.1).

Figure 6.2 shows the cross-correlation function and the ILD between the left and right ear signal of the unprocessed speech and noise component and of the noise component processed by the MWF and MWF-N algorithms. These are given for the spatial scenario $S_{90}N_{270}$ at SNR=-12dB and 0dB. The ILDs were calculated using third order butterworth filters with cut-off frequencies proposed by the Bark scale (Hartmann, 1997). The cross-correlation functions, used to interpret ITD information, were calculated for the low-pass filtered left and right ear signals and were normalized to a maximum value of 1 for identical signals. A cut-off frequency of 1000Hz was used since the most relevant ITD information for the human auditory system is present at these frequencies, e.g. Hartmann (1999). The ITD is approximated by the delay for which the cross-correlation function reaches its maximum.

For SNR=0dB, the ITD of the MWF noise component was shifted towards the ITD of the original speech component. Also, the amplitude of the cross-correlation, the amount of coherence between the left and right signal, and the width of the curve totally agree with the curve of the original speech component. It is observed that also the ILDs of the MWF processed noise component are shifted towards those of the speech component for SNR=0dB, except for a small area around 1000Hz which could be due to the low input SNR in this region (Figure 6.1). These findings agree with the mathematical findings of chapter 4.

For SNR=-12dB, the cross-correlation function of the processed noise component was shifted towards that for the speech component. However, a second peak was present around $-500\mu s$. Also the curve was somewhat more flat, meaning that the ITD information is less coherent than for the SNR=0dB con-

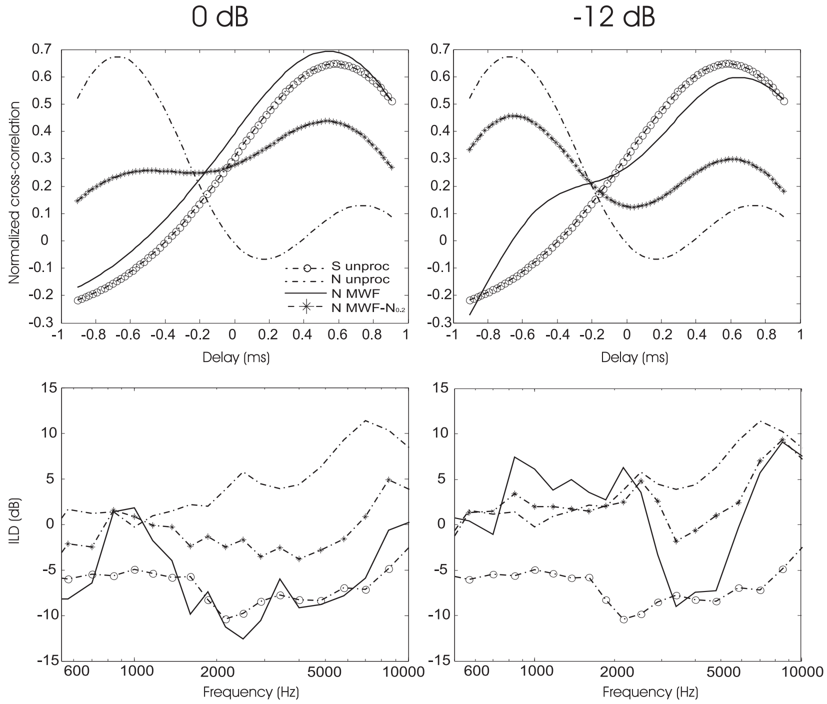

Figure 6.2 — ILD and cross-correlation function between the left and right ear signal for the unprocessed speech and unprocessed noise component together with the MWF and MWF-N$_{0.2}$ processed noise components. This is shown for spatial scenario $S_{90}N_{270}$. ILDs were calculated using a critical band analysis (bark-bands). For SNR=0dB the ITD, or the delay at which one observes a maximum cross-correlation, of the MWF noise component was moved towards the ITD of the original speech component. The ILD was shifted towards the ILD of the unprocessed speech component for almost all frequencies. This explains the localization of the processed noise component at the original location of the speech component. For SNR=-12dB, the correlation curves showed a second peak around -500μs and only frequencies between 3000 and 6000Hz were shifted towards the ILDs of the unprocessed component. The ILD of the other frequencies remained around the ILD values of the original noise component. This could explain the better localization with the MWF algorithm at SNR=-12dB compared to 0dB. The MWF-N$_{0.2}$ moved the binaural cues back to the original values of the unprocessed noise component. As a result a better localization performance of the noise component was obtained.

dition. The ILD plot shows that only ILDs for frequencies between 3000 and 5500Hz (the region with a high input SNR), were shifted towards the ILDs for the unprocessed speech component. These observations, especially at SNR=-12dB, illustrate the dual sound phenomenon and explain the good localization performance when using the MWF at SNR=-12dB compared to the SNR=0dB condition.

The dual sound phenomenon also explains the good localization performance when the speech and noise sources were played simultaneously (S+N). In this condition, the speech component masked parts of the frequency spectrum of the noise component at the output of the algorithm. The noise component was masked mostly in frequency regions with a good noise reduction performance. This is exactly the region where the binaural cues of the noise component were shifted towards the binaural cues of the speech component. When the sounds were played simultaneously, the part of the noise component with the incorrect cues was masked by the speech component. As a result the noise source was localized at the correct position in condition S+N.

This chapter has confirmed that the MWF, unlike the ADM, preserves the binaural cues of the speech component independent of the angle of arrival of the signal. The binaural cues of the noise component, which is reduced in intensity, are distorted when using the MWF or the ADM algorithm. In terms of noise reduction performance (chapter 5) it was observed that the MWF equaled or outperformed, depending on the acoustics, the number of microphones used, the angle of arrival of the signals etc., the ADM. Moreover, the MWF is robust against microphone-mismatch and allows a very straightforward extension to a binaural hearing aid framework. Therefore, when combining the data of chapter 5 and chapter 6, it is observed that the MWF has a greater potential in terms of combining noise reduction performance with the preservation of binaural cues than the bilateral ADM configuration.

The significant effect of SNR and presentation format (S,N or S+N) illustrates that testing algorithms on localization performance in lab conditions is not straightforward and results should be interpreted carefully when generalizing to real world situations. Both presentation formats (S,N, S+N) could be relevant for real-life situations. Speech and noise presented simultaneously could be relevant for situations with converged filters and a speech and noise source playing continuously. Speech and noise component presented separately could be relevant in the gaps of the speech or noise component. How these algorithms would eventually be rated perceptually by hearing aid users in real life situations is subject for further research.

6.4.4 Evaluation of the binaural MWF-N

In chapter 4, it was proven that the binaural MWF tend to distort the binaural cues of the noise component towards the binaural cues of the speech component. In the same chapter, it was proven that the binaural cues of the noise component will be restored when increasing the parameter η of the MWF-N. In Klasen et al. (2007) it was already shown that, in anechoic conditions, the ITD error of the noise component generated by the MWF algorithm could be decreased by extending this algorithm to the MWF-N. The perceptual relevance of the MWF-N was proven in section 6.3.

Large and significant improvements were observed when using the MWF-N compared to the MWF for all spatial scenarios when the speech and noise components were presented separately (S,N), especially at an input SNR=0dB. In the other conditions less or no room for improvement was available due to the reasons explained in the evaluation of the MWF algorithm (masking effects, errors in estimating the speech correlation matrix at low SNR), hence no statistical evidence of improvement was found for these conditions. This can also be observed in Figure 6.3 and Figure 6.4. However, even in the S+N data non-significant trends were sometimes observed (section 6.3.2) and subsets of the data did show significant improvement of the MWF-N vs the MWF algorithm, e.g. a significant improvement in the localization of the noise component at both SNRs was found in the condition S+N for scenario $S_{90}N_{270}$.

Although the MWF-$\mathrm{N}_{0.2}$ algorithm improved localization performance of the MWF algorithm to that of the unprocessed condition, a difference in perceptual evaluation between both conditions remained. When presenting speech and noise components separately, output signals of the MWF-$\mathrm{N}_{0.2}$ algorithm were still described as arriving from two different directions (Table 6.2). Adding more of the unprocessed signal (e.g. MWF-$\mathrm{N}_{0.3}$) would probably improve the sound quality but would further decrease noise reduction performance (Figure 5.6).

Figure 6.2 illustrates the binaural information present in the MWF and MWF-N processed noise component. When comparing the MWF curves with those of the MWF-N, it is observed that the distorted ILD and ITD cues at the output of the MWF were corrected towards those of the unprocessed noise component when using the MWF-N. This was true for both SNR=-12dB and SNR=0dB. Still, both the cross-correlation and ILD graphs illustrate that not all cues were corrected. The cross-correlation curves of the MWF-N algorithm still show a local maximum around the peak generated by the original speech component and the ILD cues of some frequency regions remain close to the ILD cues of the original speech component. This was observed more for the condition SNR=0dB than for SNR=-12dB since the MWF introduced larger distortions at high SNRs, meaning that a larger correction factor η is needed in this condition. This is consistent with the dual sound phenomenon which was observed, despite the good localization performance, when using the MWF-N in the S,N condition.

In chapter 5, the perceptual speech enhancement when using the MWF-N was examined and compared with the performance when using the MWF. The ANOVAs which compared the algorithms (section 5.3.2) showed that only in scenario $\mathrm{S}_0\mathrm{N}_{90/180/270}$, the MWF significantly outperformed the MWF-N and this only for the 4-microphone implementation. A refined analyses showed that in scenario $\mathrm{S}_0\mathrm{N}_{90/180/270}$ the MWF indeed outperformed the MWF-N, but that in scenario $\mathrm{S}_{90}\mathrm{N}_{270}$ the MWF-N significantly outperformed the MWF. In scenario S_0N_{60} no significant differences were found. During these perceptual evaluations, the observed differences between the MWF and MWF-N were typically smaller than expected from the objective SNR performance measures. It was suggested that this could be due to the preservation of the location of both sound sources by the MWF-N which enabled an improved speech understanding in noise due to an improved release from masking. This chapter supports this assumption by showing that the MWF-N preserves the binaural cues better than the MWF and by showing that the MWF-N always preserves the ability to localize both the speech and the noise source. This is not always true when using the MWF.

6.5 Conclusion

In this chapter, three research questions were addressed which are related to the influence of noise suppression techniques for hearing aids on the localization of sound sources. By combining the data of this and the previous chapter conclusions can be drawn regarding the combination of noise reduction performance and the preservation of binaural cues by the different algorithms.

First, the localization performance of normal hearing subjects was quantified when using a bilateral noise suppression system, namely a bilateral ADM which is commonly used in current high-end hearing aids. The ADM led to a significant drop in localization performance when sounds were presented from outside the forward field of view. If the target sound was presented from the front, localization performance was similar than for the unprocessed condition. However, under these conditions, noise reduction performance will be low or the ADM will even have a negative impact on speech perception. It can be concluded that the ADM cannot combine noise reduction with the preservation of binaural cues of both the speech and the noise component.

As a second and third research question, two binaural noise reduction algorithms were evaluated in terms of localization performance. The binaural MWF showed a good localization performance for the speech component. For the noise component large errors were expected since it was mathematically proven in chapter 4 that the binaural cues of the noise component are shifted towards those of the speech component when using the MWF algorithm. Interestingly, subjective evaluations did not always show this behavior. Localization of the noise component was often better than expected due to errors in the estimation of the speech correlation matrix and due to masking effects. Since the MWF preserves the binaural cues of the speech component independent of the angle of arrival of the signal and since the MWF equals or outperforms the performance of the ADM in terms of noise reduction performance (chapter 5), the MWF can be preferred over the ADM as a noise reduction system for hearing aids.

The evaluation of the binaural MWF-N showed that by adding part of the unprocessed signal ($\eta = 0.2$) to the output of the MWF algorithm unaided localization performance could be reached for both the speech and the noise component in all scenarios. This comes at the cost of noise reduction. However, chapter 5 demonstrated that the degradation in speech perception is often smaller than expected. If the speech and the noise source are well separated, speech understanding in noise even improved when using the MWF-N compared to the MWF algorithm. This may be due to the preservation of all binaural cues by the MWF-N which enabled a better speech understanding in noise due to an improved spatial release from masking.

The full data set consisted of many conditions: two different ways of presenting stimuli, three spatial scenarios, four different algorithms and two SNRs. Therefore, the subset of data per condition became relatively small. Due to this limitation small differences between algorithms could remain undetected. However, even these limited subsets of data were already sufficient to prove some very large effects of, and differences between, noise reduction algorithms. Moreover, it shows that interpreting results of localization experiments with noise reduction systems is not straightforward since they are dependent on spatial scenario, SNR and masking effects. Further research, involving larger data sets per condition, may reveal smaller differences present between algorithms. In addition, further research may clarify how MWF based algorithms are perceived by hearing aid users in real-life conversations, which contains a mixture of speech and noise presented separately (in speech and noise gaps) and presented simultaneously, in terms of localization and speech perception in noise.

6.A Additional Tables

S_0N_{60}	S_0 unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{60} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	0	0	5	15	0	15	0	30	30	15	60	30	60	25
J	0	0	0	0	0	0	0	30	30	30	60	30	20	30
H	0	0	0	5	0	0	0	5	0	10	60	25	25	5
L	5	5	5	15	10	15	10	30	20	25	65	35	15	25
O	0	20	25	5	35	10	0	25	20	30	50	20	5	20
average (°)	1	5	7	8	9	8	2	24	20	22	59	28	25	21
$S_{90}N_{270}$	S_{90} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{270} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	0	90	90	5	10	5	0	10	75	90	170	35	0	5
J	0	90	90	0	10	0	0	0	90	90	135	0	0	0
H	20	70	75	25	35	20	15	5	15	30	65	15	10	0
L	10	90	75	10	15	5	15	20	55	55	170	30	15	10
O	15	40	40	20	35	20	15	15	70	80	170	145	120	25
average (°)	9	76	74	12	21	10	9	10	61	69	142	45	29	8
$S_{45}N_{315}$	S_{45} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{315} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	5	15	5	35	20	20	35	5	15	40	120	45	25	30
J	15	45	45	45	0	45	5	0	45	45	120	25	30	45
H	0	5	15	0	0	0	0	30	25	25	5	30	25	30
L	30	25	5	30	30	40	15	15	0	5	20	5	15	15
O	30	25	30	25	30	30	25	35	15	30	125	120	90	20
average (°)	16	23	20	27	16	27	16	17	20	29	78	45	37	28

Table 6.4 — Individual and average Mean absolute error ((°) MAE), averaged over three repetitions, for the three different spatial scenarios and the different processing schemes at both SNRs. The speech and the noise source were presented separately through headphones (S,N). A summary of this Table is given in Table 6.1

MAE														
S_0N_{60}	S_0 unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{60} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	10	0	40	10	20	5	10	30	30	30	20	15	30	10
J	0	0	25	0	0	0	0	30	30	30	20	15	30	30
H	0	0	0	0	0	0	5	15	15	15	15	15	20	10
L	15	10	15	15	15	15	20	30	30	20	30	30	25	30
O	0	50	5	0	5	0	20	10	25	25	25	15	25	10
average (°)	5	12	17	5	8	4	11	23	26	24	22	18	26	18
$S_{90}N_{270}$	S_{90} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{270} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	15	90	150	0	0	5	20	30	30	90	15	25	5	20
J	0	75	15	0	5	0	0	0	20	10	5	5	0	0
H	25	70	70	25	35	15	45	5	30	25	25	15	15	0
L	0	95	5	0	10	0	15	15	35	65	25	30	20	10
O	0	40	15	5	30	15	25	0	20	55	25	50	30	30
average (°)	8	74	51	6	16	7	21	10	27	49	19	25	14	12
$S_{45}N_{315}$	S_{45} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12	N_{315} unproc	ADM SNR0	SNR-12	MWF SNR0	SNR-12	MWF-N$_{0.2}$ SNR0	SNR-12
T	30	30	10	30	35	15	30	35	15	15	45	15	45	30
J	40	25	45	45	45	45	45	35	45	35	35	20	45	45
H	0	0	0	0	0	5	5	15	15	25	30	5	25	35
L	35	30	45	30	20	25	35	20	0	0	20	10	20	15
O	40	45	35	30	25	30	35	20	25	30	50	45	15	5
average (°)	29	26	27	27	25	24	30	25	20	21	36	19	30	26

Table 6.5 — Individual and average Mean absolute error ((°) MAE), averaged over three repetitions, for the three different spatial scenarios and the different processing schemes at both SNRs. The speech and the noise source were presented simultaneously through headphones (S+N). A summary of this Table is given in Table 6.3

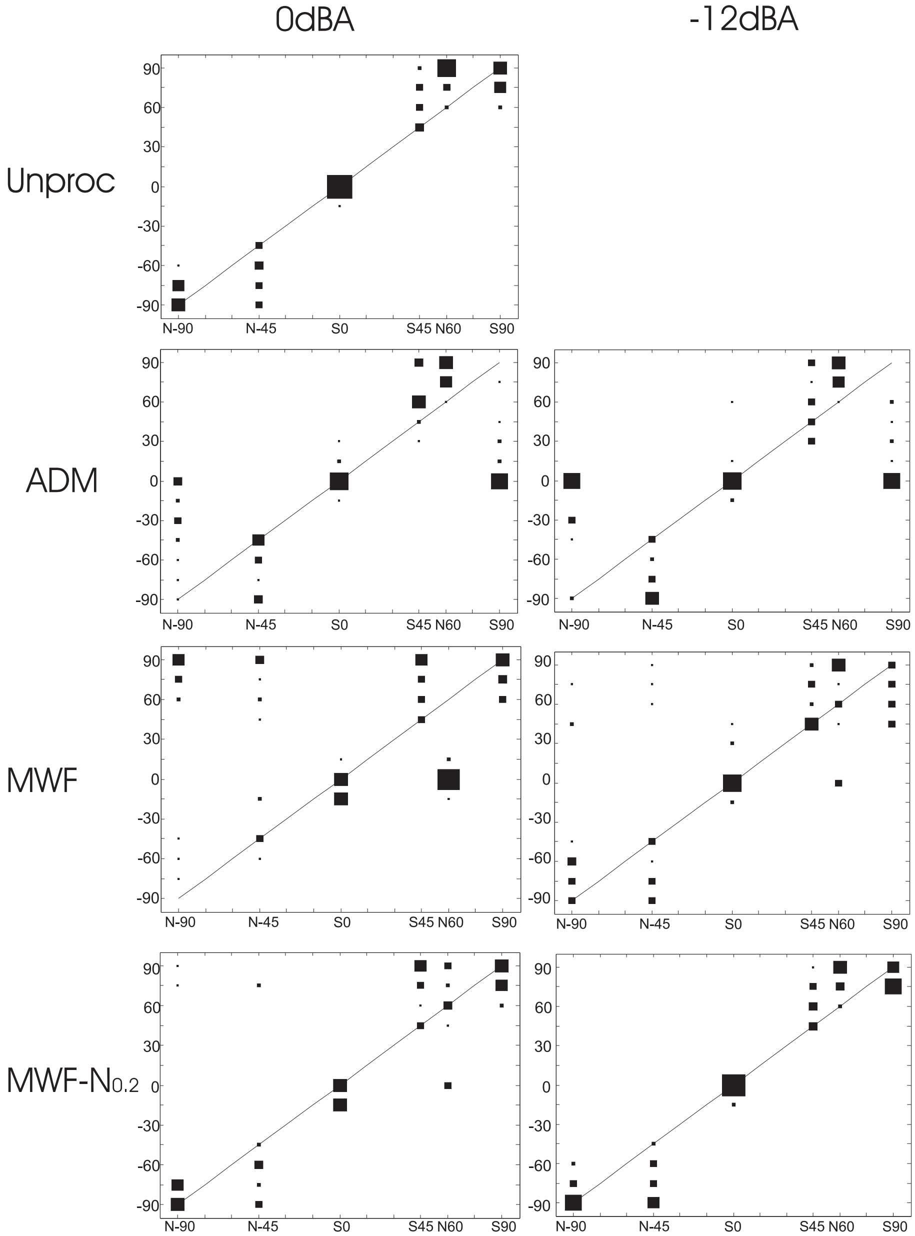

Figure 6.3 — Accumulated localization responses for all subjects in the condition S,N. The x-axis depicts the given stimulus position with the prefix S or N indicating whether this stimulus was a speech or a noise component. The y-axis represents the reported location by the listeners. The size of the squares are proportional to the number of responses at this point. When localization would be perfect, all responses would lie on the diagonal. If the angle of arrival was $\pm 90°$, the ADM distorted the perceived location towards $0°$. The MWF distorted the localization of the noise component towards the location of the speech component, especially for SNR=0dB. The MWF-N restored the localization performance towards the unprocessed performance.

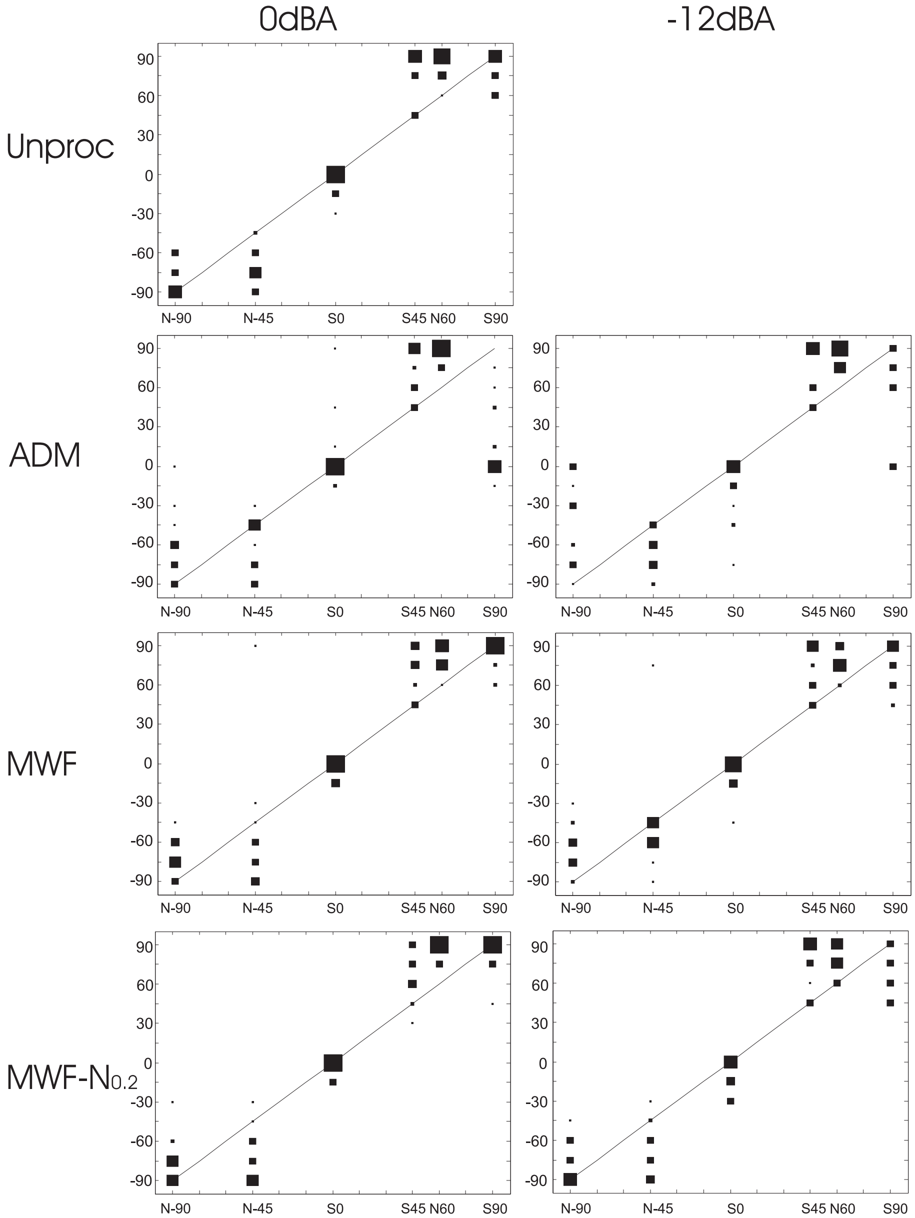

Figure 6.4 — Accumulated localization responses for all subjects in the condition S+N. The x-axis depicts the given stimulus position with the prefix S or N indicating whether this stimulus was a speech or a noise component. The y-axis represents the reported location by the listeners. The size of the squares are proportional to the number of responses at this point. When localization would be perfect, all responses would lie on the diagonal. If the angle of arrival was $\pm 90°$, the ADM distorted the perceived location towards $0°$. This was less pronounced than in condition S,N. The noise component was almost never localized at the location of the speech component when using the MWF. Localization with the MWF-N was similar to localization in the unprocessed condition.

			S_0N_{60}		$S_{90}N_{270}$		$S_{45}N_{315}$	
			SNR-12	SNR0	SNR-12	SNR0	SNR-12	SNR0
unproc	Speech	broad		8		0		25
		dual sound		17		25		0
		diffuse		8		0		17
		clear		67		75		58
	Noise	broad		11		11		11
		dual sound		0		0		0
		diffuse		0		0		0
		clear		89		89		89
ADM	Speech	broad	13	0	7	20	7	0
		dual sound	13	7	7	13	7	13
		diffuse	20	7	60	53	0	0
		clear	53	87	27	13	87	87
	Noise	broad	0	7	20	40	0	7
		dual sound	7	13	7	27	0	0
		diffuse	0	0	20	27	0	0
		clear	93	80	53	7	100	93
MWF	Speech	broad	13	7	0	0	0	7
		dual sound	0	0	33	0	0	7
		diffuse	0	0	7	0	0	0
		clear	87	93	60	100	100	87
	Noise	broad	7	7	7	0	7	13
		dual sound	40	7	87	73	67	67
		diffuse	7	7	0	0	0	0
		clear	47	80	7	27	27	20
MWF-$N_{0.2}$	Speech	broad	7	0	0	0	27	0
		dual sound	0	0	0	0	0	0
		diffuse	0	0	0	0	0	0
		clear	93	100	100	100	73	100
	Noise	broad	7	0	0	0	7	0
		dual sound	27	53	87	93	80	87
		diffuse	0	20	0	0	0	7
		clear	67	27	13	7	13	7

Table 6.6 — Percentage of stimuli, averaged over 5 subjects, perceptually classified as being a) a sound arriving from a point source with one clear direction in space (clear) b) a diffuse sound with no directional information (diffuse) c) sounds arriving from two directions in space (dual sound) or d) a very broad sound source (broad). This is done for the three different spatial scenarios and the different processing schemes The speech and the noise source were presented separately through headphones (S,N). A summary of this table is given in Table 6.2

Chapter 7

Conclusions and further research

Hearing impaired subjects have great difficulty understanding speech in noisy environments. Due to their hearing impairment they often need a SNR which is 5 to 10dB higher to perceive the same amount of speech as normal hearing subjects (section 1.2). Therefore, noise reduction algorithms have been applied in hearing instruments (section 1.3). These algorithms are typically evaluated and optimized in a monaural way, i.e. for one ear only. However, the human auditory system is a binaural system (section 1.4) which combines and compares the signals received at both ear drums to create a single and improved sound percept.

This study has evaluated the impact of common noise reduction algorithms, i.e. a bilateral fixed and adaptive directional microphone, on binaural information (chapter 2) (Research objective 1, section 1.7.1). Due to recent technological developments a communication link with reasonable bandwidth between hearing aids may soon be implemented. This enables new noise reduction strategies, based on binaural processing, i.e. using microphone signals of both hearing aids to generate a signal for both ears. This study presented and evaluated three new binaural algorithms based on a multichannel Wiener filter (MWF) approach which show promising potential in terms of combining noise reduction performance with the preservation of binaural cues (chapters 4 to 6). During the evaluations of these algorithms a bilateral adaptive directional microphone (ADM) was used as a reference noise reduction system since this is the most commonly used adaptive multi-microphone noise reduction system in commercial digital hearing aids (Research objective 2, section 1.7.1).

Section 7.1 summarizes the main conclusions of this study.

Section 7.2 provides some suggestions for further research.

7.1 Conclusions

7.1.1 Current bilateral hearing aids

In **chapter 2**, an evaluation was made on the impact of commercial noise reduction systems on the binaural cues. A setup to perform localization experiments in the frontal horizontal hemisphere was developed to perceptually quantify these effects.

In section 2.2, we have shown that commercial bilateral noise reduction algorithms, i.e. a bilateral ADM and a bilateral FDM, can distort the binaural cues of a sound signal. During this analysis it was proven that realistic imperfections, such as non-identical microphone characteristics which are known to affect noise reduction performance, can greatly affect the binaural cues produced by a bilateral FDM and ADM. The impact of both algorithms is typically large if the angle of arrival of the sound approaches the angle of maximal noise suppression.

A perceptual evaluation (section 2.3) confirmed these conclusions. By testing hearing impaired subjects we have shown that a better localization performance was achieved when hearing aid users were not wearing their hearing aids. Switching on their adaptive noise reduction system, i.e. a bilateral ADM, significantly degraded localization performance even further. It was observed that this additional degradation was mostly due to errors made when localizing sound sources at the sides of the head and not from the forward field of view. This agrees with the findings of section 2.2 in which it was shown that binaural cue distortion is maximized if the angle of arrival approached the angle of maximal noise suppression. The subjective evaluation also demonstrated that hearing impaired subjects are still capable of using binaural cues which further motivated this research.

7.1.2 Binaural MWF based algorithms

Recently the interest on binaural cues and binaural processing in general has risen since the awareness of the importance of binaural processing has grown (Keidser et al., 2006; Van den Bogaert et al., 2006; Noble, 2006) and since technological developments are enabling the use of a wireless communication channel between both hearing aids (e.g. the Siemens Acuris). It is expected that soon a full link will be available enabling the use of binaural noise reduction algorithms. These systems use the ipsi- and contralateral microphone signals to compute the output signals for both hearing aids. This may enhance noise reduction performance and may improve the preservation of binaural cues (see also section 1.1).

Chapter 4 introduced three new binaural algorithms. First a binaural MWF approach was described (section 4.3). This is an extension of the monaural algorithm developed by Doclo and Moonen (2002); Spriet et al. (2004). The MWF has the important advantage that, unlike beamforming strategies, no a priori knowledge is required on the location of the target signal nor of the microphone characteristics of the hearing aids. A theoretical analysis of the binaural MWF (section 4.3.2) proved that the MWF perfectly preserves the interaural transfer function (ITF) and hence the binaural cues of the speech component, independent of the angle of arrival of the signal. However, it was also proven that the ITF of the remaining noise component at the output of the algorithm are distorted into the ITF of the speech component. To preserve the spatial awareness of the user and to optimally benefit from spatial release from masking, the binaural cues of both the speech and the noise component should be preserved. Therefore two extensions of the MWF were presented, i.e. the MWF-N and the MWF-ITF.

The MWF-N (section 4.4) is based on removing only part of the noise component. Hence, the unprocessed part may be used to restore the binaural cues of the noise component. A theoretical analysis proved that the MWF-N preserves the ITF of the speech component. The preservation of the ITD and ILD cues of the noise component are dependent on the parameter η. If $\eta = 0$, the MWF-N is reduced to the MWF and maximum noise reduction performance is obtained without preserving the cues of the noise component. If $\eta = 1$, all binaural cues are preserved but no noise reduction is performed. Objective and subjective evaluations further evaluated the MWF-N and the MWF algorithm and compared these algorithms with a standard bilateral ADM approach (see section 7.1.3).

The MWF-ITF (section 4.5) is based on adding an extra term into the original cost function of the binaural MWF. This term, with a certain weight β, constrains to some extent (dependent on β) the optimal Wiener solution to filters which preserve the ITF of the noise component. In section 4.5.1 a quadratic cost function of the MWF-ITF was derived, valid for single noise source scenarios. Objective and perceptual pilot experiments with a single noise source scenario (section 4.5.2), proved that introducing this extra term preserved the binaural cues of the noise component. However, if β was chosen too large, the binaural cues of the speech component were distorted into those of the noise component. The subjective evaluation demonstrated that a parameter setting of β can be found which enables a gain in localization performance compared to the binaural MWF without losing noise reduction performance.

All the presented binaural MWF based algorithms assumed the availability of all microphone signals. However, since commercial manufacturers prefer a wireless communication channel between hearing aids, transmitting all contralateral microphone signals comes at the large cost of power consumption. In section 4.6, a brief overview was given of strategies which try to approach the noise reduction performance of a full binaural configuration by transmitting only one signal which is a linear combination of the microphone signals of the contralateral hearing aid. Several combinations have been studied and it was described that an optimal iterative distributed processing scheme exists which converges to the optimal full binaural solution. However, this evaluation did not take the delay of the communication channel into account and further research has to be done to make it commercially applicable.

7.1.3 Evaluation of the MWF, the MWF-N and the ADM

As mentioned before, extensive objective and perceptual evaluations were performed to quantify the noise reduction performance and the preservation of binaural cues of the MWF and MWF-N algorithm. These evaluations were presented in chapters 5 and 6 respectively. The bilateral ADM, which is the most commonly used adaptive multi-microphone noise reduction system in commercial hearing aids was used as a reference noise reduction strategy.

In **chapter 6**, objective performance measures were used in several spatial scenarios, using different microphone combinations and different parameter settings in two different acoustic environments described by a reverberation time of $T_{60} = 0.21s$ and $T_{60} = 0.61s$ respectively. The latter environment, representing a realistic living room, was also used in the perceptual evaluations of chapters 5 and 6.

From the data of chapter 6, it was concluded that a two-microphone bilateral MWF and ADM have approximately the same performance in realistic scenario's. Also both algorithms have a stable performance at an input SNR ranging from -5 to +5dBA. The MWF, unlike the ADM, does so while preserving the binaural cues of the speech component. Moreover, in low reverberant conditions or if the speech source was not positioned in the frontal field of view, the ADM was outperformed by the MWF. In the latter condition the ADM even degraded speech perception. It was also observed that, by using a bilateral or binaural hearing aid configuration, a lower SNR is needed to understand the same amount of speech compared to a monaural configuration. This confirms the common practice of using bilateral hearing aids for bilateral hearing impaired subjects.

The gain in noise reduction performance due to the transmission of one or two contralateral microphone signals was quantified by examining different microphone combinations. Objective evaluations in a large set of sound scenarios showed that performance improved if more microphone signals are available. The perceptual evaluations demonstrated that transmitting only one contralateral microphone signal to the ipsilateral hearing aid indeed resulted in an improved noise reduction performance. When sending over a second contralateral microphone signal, the perceptual evaluations did not show an additional significant improvement in noise reduction performance. The observed gains in noise reduction due to transmitting contralateral microphone signals shows the potential of binaural noise reduction algorithms.

Chapter 5 also examined the loss in noise reduction performance when adding a partial noise estimate to the MWF, i.e. the MWF-N. The drop in performance observed during the perceptual evaluations was often lower than expected from the objective evaluations. Moreover, in some scenarios, e.g. $S_{90}N_{270}$, it was even observed that speech perception improved when using the MWF-N compared to the MWF algorithm. This can be explained by an improved spatial release from masking due to the preservation of the binaural cues of both the speech and the noise component by the MWF-N algorithm.

Finally, the study on noise reduction performance showed that carefully selected objective performance measures allow the prediction of the noise reduction performance quite accurately, even for a binaural hearing aid configuration. However, psycho-acoustical properties of the human auditory system (e.g. the best ear benefit, spatial release from masking effects, etc.) have to be taken into account when interpreting the results. These properties can introduce differences between the objectively predicted and the perceptually observed performance. Other objective performance measures were recently developed by other researchers which aim at integrating these psycho-acoustical properties. This may improve the correlation between objective and perceptual performance measures.

Chapter 5 presented a localization experiment in the frontal horizontal hemisphere to evaluate the MWF and MWF-N algorithm in terms of preserving binaural cues. It was shown that the ADM, used as a reference noise reduction algorithm, preserves the binaural cues if the signal is positioned within the forward field of view. If signals are presented from outside this area, e.g. the sides of the head, large localization errors occur. This confirms the conclusions of chapter 2 and demonstrates that a bilateral ADM is not able to perform noise reduction ànd to preserve the binaural cues of both the speech and the noise component. If the noise is reduced in intensity, the binaural cues are distorted.

Due to the theoretical analysis it was predicted that the MWF would preserve the location of the speech component but that it would dislocate the perception of the noise component towards the location of the speech component. The data of the localization experiment confirmed the correct localization of the speech component. The localization of the noise component, however, was often better than expected. Afterwards it was proven that this could be explained by errors in the estimation of the speech correlation matrix and due to masking effects. Since the MWF preserves the binaural cues of the speech component independent of the angle of arrival of the signal (unlike the ADM) and since the MWF equals or outperforms the performance of the ADM in terms of noise reduction performance (chapter 6), the MWF is preferred over the ADM as a noise reduction system for bilateral/binaural hearing aids

When adding a partial noise estimate to the MWF, it was proven that the localization performance of the MWF improved. Moreover, no significant difference in localization performance was observed compared to the unprocessed condition when using $\eta = 0.2$. This obviously comes at the cost of noise reduction. However, chapter 5 concluded that the decrease in speech perception is often lower than expected. Since the MWF-N sufficiently preserves the binaural cues of both the speech and the noise component, it enables the human auditory system to improve the use of spatial release from masking. Therefore the drop in noise reduction or SNR does not necessarily imply a drop in speech perception. Moreover, chapter 5 demonstrated that in some conditions speech perception even significantly improved when using the MWF-N algorithm compared to the MWF algorithm.

7.1.4 Overall conclusions

This manuscript studied the interaction of noise reduction algorithms with the preservation of binaural cues. First, it was observed that hearing impaired subjects are still capable of using binaural cues which are important for a correct sound localization and to improve speech understanding in noisy environments due to spatial release from masking. This manuscript demonstrated that two of the most frequently used multi-microphone noise reduction systems in high-end commercial digital hearing aids are able to distort the binaural cues of a sound signal. Afterwards, binaural MWF-based algorithms were studied. It was shown that a binaural hearing aid design, i.e. transmitting contralateral microphone signals to the ipsilateral hearing aid, does significantly increase noise reduction performance and speech perception in noise. Moreover, it was demonstrated that the binaural MWF, the binaural MWF-N and the binaural MWF-ITF provide a better combination of noise reduction performance and preservation of binaural cues compared to the bilateral ADM algorithm.

7.2 Suggestions for further research

MWF-ITF

Due to the fact that the derivation of the MWF-ITF cost function (eq. 4.77) is only valid for single noise source scenarios, no extensive validations of this algorithm have yet been performed. Current research, further analyzing the MWF-ITF, shows that the quadratic cost function implies that the output speech and noise component have an identical output ITF (not reported in the manuscript). The parameter β determines in each frequency band whether this ITF approximates the ITF of the original speech component or of the original noise component. The gain in localization performance is obtained by the interaction between output SNR and the parameter β. When using a correct value of β, the ITF of frequency bands with a high output SNR tend to stay at the ITF of the speech component. The ITF of frequency bands with a low output SNR tend to stay at the ITF of the noise component. This explains the behavior of the MWF-ITF found in the pilot experiments. However, further research has to be carried out to evaluate this algorithm both theoretically as well as perceptually in multiple noise source scenarios and in realistic acoustic environments.

Perceptually motivated masking

The MWF-N algorithm is based on leaving part of the original noise component unprocessed which 'masks' the incorrect binaural cues of the noise component generated by the algorithm. Currently, this masking is performed linearly over all frequency bins. However, one might argue that only the frequency bins in which a noise signal is sufficiently present should be masked by an unprocessed component. By masking only the appropriate frequency bands or the frequency bands with the lowest SNR, the same localization performance could be obtained while reducing the loss in noise reduction performance of the MWF-N compared to the MWF. Moreover, the frequency bins dominating speech perception are different from the frequency bands dominating localization performance which further motivates frequency dependent masking algorithms.

Reduced bandwidth algorithms

Transmitting all microphone signals to the contralateral hearing aid comes at the large cost of power consumption. In section 4.6 it was mentioned that an iterative distributed processing scheme, which only transmits a single signal to the contralateral hearing aid, allows convergence towards the full binaural solution (Doclo et al., 2008). However, this evaluation did not take the delays introduced by the wireless link into account yet. Experiments including these delays and including situations in which the binaural link breaks down are necessary to prove the algorithm's applicability for hearing aids.

Adaptive MWF algorithms

This manuscript has proven the potential of binaural MWF based algorithms. However, the evaluations discussed in this manuscript have been based on using a perfect VAD and calculating cross-correlation matrices and their appropriate converged filters off-line. The real time behavior of the binaural MWF-based algorithms will have to be studied subsequently: convergence speed, the impact of VAD errors, the influence of head movements on binaural cues, the influence of own voice on the preservation of binaural cues, the impact of moving sound sources on the algorithm, and so on. This is necessary before they can be implemented in commercial hearing aids. To reduce the number of VAD-errors, the monaural VAD may be extended towards a binaural VAD, e.g. by using a combination of energy and correlation based strategies. It should be mentioned that a real-time implementation of a monaural MWF-based algorithm is currently studied at our lab within the framework of the EU-project HearCom (2004-2009). Perceptual evaluations are showing promising results and no substantial problems have occurred when using a real-time VAD or when allowing head movements.

Fundamental psycho-acoustic research

One of the main problems encountered during the evaluations of the different algorithms is the lack of decent objective performance measures in terms of localization performance or preservation of binaural cues. Several measures have been used during this dissertation such as the cross-correlation function (Figure 6.2, Klasen et al. (2007)) or the ITD and ILD error defined in eq. 4.82 and eq. 4.83. All of these measures have been very useful throughout this dissertation to roughly compare algorithms or parameter settings, or to explain observations made during perceptual evaluations. However, a strict correlation between objective and perceptual performance measures such as the one found during the speech perception in noise evaluations was not found. At the moment, not enough information is available on the integration of binaural cues over time and frequency by the human auditory system to create a perfect error measure. However, studying advanced localization models might lead to more appropriate objective, perceptually motivated, error measures.

Other research

This work has been restricted to the impact of noise reduction strategies on the binaural cues. However, the perceptual evaluation in chapter 2 demonstrated that even when noise reduction algorithms of hearing aids are switched off, localization performance can be reduced relative to unaided performance. Other building blocks of a hearing aid such as compression algorithms are known to affect binaural cues. To maximize comfort and performance, a thorough evaluation of the impact of other processing strategies on binaural cues is also necessary.

During this dissertation perceptual evaluations have been restricted to the frontal horizontal hemisphere because of the focuss on ITD and ILD cues. Future research could evaluate the influence of noise reduction algorithms on front-back confusions or elevation which are probably more related to spectral cues and hence microphone placement than binaural processing.

Bibliography

Aichner, R., Buchner, H., Zourub, M., and Kellerman, W. (2007), Multichannel source separation preserving spatial information, in *Proceedings of the 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu, HI, USA, pp. I5–I8.

Akeroyd, M. A., Chambers, J., Bullock, D., Palmer, A. R., Summerfield, A. Q., Nelson, P. A., and Gatehouse, S. (2007), "The binaural performance of a cross-talk cancellation system with matched or mismatched setup and playback acoustics.," *Journal of the Acoustical Society of America*, 121(2), 1056–1069.

Allen, J. B., and Berkley, D. A. (1979), "Image method for efficiently simulating small-room acoustics.," *Journal of the Acoustical Society of America*, 65, 943–950.

ANSI-SII (1997), "American National Standard Methods for Calculation of the Speech Intelligibility Index ANSI S3.5-1997," *Acoustical Society of America*, .

Arehart, K. H., Hansen, J. H. L., and Gallant, S. (2003), "Evaluation of an auditory masked threshold noise suppression algorithm in normal hearing and hearing impaired listeners.," *Speech communication*, 40(4), 575–592.

Baer, T., and Moore, B. C. J. (1993), "Effects of spectral smearing on the intelligibility of sentences in noise," *Journal of the Acoustical Society of America*, 94(3), 1229–1241.

Bentler, R., and Chiou, L. K. (2006), "Digital noise reduction: an overview," *Trends in amplification*, 10(2), 67–82.

Beranek, L. L. (1954), *Acoustics* New York: McGraw Hill Electrical and Electronic Engineering Series.

Berger, K. (1984), *The hearing aid: Its operation and development*, third edn MI: National Hearing Aid Society.

Bernstein, L. R., and Trahiotis, C. (1985), "Lateralization of low-frequency, complex waveforms: the use of envelope-based temporal disparities.," *Journal of the Acoustical Society of America*, 77, 18681880.

Bernstein, L. R., and Trahiotis, C. (1992), "Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies," *Journal of the Acoustical Society of America*, 91(1), 306–316.

Bertelson, P. (1999), "Cognitive contributions to the perception of spatial and temporal events.," Elsevier Science, chapter Ventriloquism: A case of cross-modal perceptual grouping, pp. 347–362.

Beutelmann, R., and Brand, T. (2006), "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *Journal of the Acoustical Society of America*, 120(1), 331–341.

Bitzer, J., and Simmer, K. U. (2001), "Microphone Arrays: Signal Processing Techniques and Applications," Springer-Verlag, chapter 2-Superdirective microphone arrays, pp. 19–38.

Blauert, J. (1997), "Spatial hearing, the psychophysics of human sound localization,", revised edn, : The MIT Press, chapter 2.4 Evaluating non-identical ear input signals, pp. 137–176.

Boll, S. F. (1979), "Suppression of acoustic noise in speech using a spectral subtraction," *IEEE Transaction on Acoustics, Speech and Signal Processing*, ASSP-27(2), 113–120.

Boudreau, J. C., and Tsuchitani, C. (1968), "Binaural interaction in the cat superior olive segment.," *Journal of Neurophysiology*, 31, 442–454.

Bregman, A. S. (1981), "Perceptual Organization,", Hillsdale, N.J.: Lawrence Erlbaum Associates, chapter 4-Asking the what for question in auditory perception, pp. 99–118.

Bregman, A. S. (1993), "Thinking in sound: The cognitive psychology of human audition.,", : Oxford UP, chapter 2-Auditory scene analysis: Hearing in complex environments, pp. 10–36.

Bregman, A. S. (1999), *Auditory scene analysis*, first paperback edition edn, : The MIT press.

Bronkhorst, A. W., and Plomp, R. (1988), "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *Journal of the Acoustical Society of America*, 83(4), 1508–1516.

Bronkhorst, A. W., and Plomp, R. (1989), "Binaural speech intelligibility in noise for hearing impaired listeners," *Journal of the Acoustical Society of America*, 86(4), 1374–1383.

Brungart, D. S., and Rabinowitz, W. M. (1999), "Auditory localization of nearby sources. Head-related transfer functions.," *Journal of the Acoustical Society of America*, 106(3), 1465–1479.

Butler, R. A., and Belendiuk, K. (1977), "Spectral cues utilized in the localization of sound in the median sagittal plane," *Journal of the Acoustical Society of America*, 61, 1264–1269.

Byrne, D., Sinclair, S., and Noble, W. (1998), "Open earmold fittings for improving aided auditory localization for sensorineural hearing losses with good high-frequency hearing," *Ear and Hearing*, 19(1), 62–71.

Cappe, O. (1994), "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Transactions Speech, Audio Processing*, 2(2), 345–349.

Carlile, S., Leong, P., and Hyams, S. (1997), "The nature and distribution of errors in sound localization by human listeners," *Hearing Research*, 114, 179–196.

Clarkson, P., and Bahgat, S. (1991), "Envelope expansion methods for speech enhancement," *Journal of the Acoustical Society of America*, 89(3), 1378–1382.

Cox, H., Zeskind, R. M., and Kooij, T. (1986), "Practical supergain," *IEEE Transactions on Speech and Audio Processing*, ASSP-34(3), 393–398.

Cox, R. M., and Alexander, G. C. (1991), "Hearing Aid Benefit in Everyday Environments," *Ear and Hearing*, 12(2), 127–139.

Deiss, A. (2002), A Low Power 200 MHz Receiver for Wireless Hearing Aid Systems, PhD thesis, ETH Zurich.

Desloge, J. G., Rabinowitz, W. M., and Zurek, P. M. (1997), "Microphone-array hearing aids with binaural output-part I: Fixed processing systems," *IEEE Transactions Speech and Audio Processing*, 5(6), 529–542.

Dillon, H. (2001*a*), "Hearing Aids," Boomerang Press, Sydney, chapter 6-Compression systems in hearing aids, pp. 159–186.

Dillon, H. (2001*b*), "Hearing Aids," Boomerang Press, Sydney, chapter 7-Advanced signal processing schemes in hearing aids, pp. 187–208.

Dillon, H. (2001*c*), "Hearing Aids,", : Boomerang Press, Sydney, chapter 2.2 Hearing aid components - Microphones, pp. 21–28.

Dillon, H., Keidser, G., and H., S. (2003), "Sound quality comparisons of advanced hearing aids," *The hearing journal*, 56(4), 1–6.

Dillon, H., and Lovegrove, R. (1993), "Acoustical factors affecting hearing aid performance," Allyn and Bacon, chapter Single microphone noise reduction systems for hearing aids: A review and an evaluation.

Doclo, S., Dong, R., Klasen, T. J., Wouters, J., Haykin, S., and Moonen, M. (2005), Extension of the multi-channel Wiener filter with localisation cues for noise reduction in binaural hearing aids, in *Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC)*, pp. 221–224.

Doclo, S., Eneman, K., and Moonen, M. (2002), Voice activity detection,, Internal report 01-62, K.U.Leuven.

Doclo, S., Klasen, T. J., Van den Bogaert, T., Wouters, J., and Moonen, M. (2006), Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions, in *Proceedings International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Paris, France, pp. 1–4.

Doclo, S., and Moonen, M. (2002), "GSVD-Based Optimal Filtering for Single and Multi-Microphone Speech Enhancement," *IEEE Transactions on Signal Processing*, 50(9), 2230–2244.

Doclo, S., Spriet, A., Moonen, M., and Wouters, J. (2007), "Frequency-Domain Criterion for the Speech Distortion Weighted Multichannel Wiener Filter for Robust Noise Reduction," *Speech Communication*, 49(7-8), 636–656.

Doclo, S., Van den Bogaert, T., Moonen, M., and Wouters, J. (2008), "Reduced-bandwith and distributed MWF-based noise reduction algorithms for binaural hearing aids.," *IEEE Transactions on audio, speech and language processing*, . Accepted.

Drennan, W. R., Gatehouse, S., and Lever, C. (2003), "Perceptual segregation of competing speech sounds: The role of spatial location," *Journal of the Acoustical Society of America*, 114(4), 2178–2189.

Dreyer, A., and Delgutte, B. (2006), "Phase Locking of Auditory-Nerve Fibers to the Envelopes of High-Frequency Sounds: Implications for Sound Localization," *Journal of Neurophysiology*, 96, 2327–2341.

Duquesnoy, A. J., and Plomp, R. (1983), "The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise," *Journal of the Acoustical Society of America*, 73(6), 2166–2173.

Durlach, N. I. (1963), "Equalization and cancellation theory of binaural masking-level differences," *Journal of the Acoustical Society of America*, 35(8), 1206–1218.

Fisher, H., and Freedman, S. (1968), "The role of the pinna in auditory localization.," *Journal of Auditory Research*, 8, 15–26.

Fitzgerald, R. (2002), “Inhibition in the brain plays a key role in sound localization.,” *Physics today*, pp. 13–14.

Gatehouse, S., and Noble, W. (2004), “The Speech, Spatial and Qualities of Hearing Scale (SSQ),” *International Journal of Audiology*, 43, 85–99.

Gelfand, S. A. (1998), “Hearing: an introduction to psychological and physiological acoustics,”, 3rd edn, New York: Marcel Dekker, chapter 13-Binaural hearing, pp. 369–400.

Gierlich, H. W. (1992), “The application of binaural technology.,” *Applied acoustics*, 36, 219–243.

Gilkey, R. H., and Anderson, T. R. (1997), *Binaural and spatial hearing in real and virtual environments*, : Lawrence erlbaum associates.

Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (1993), “Intelligibility-weighted measures of speech to interference ratio and speech system performance,” *Journal of the Acoustical Society of America*, 94(5), 3009–3010.

Griffiths, L. J., and Jim, C. W. (1982), “An alternative approach to linearly constrained adaptive beamforming,” *IEEE Transaction on Antennas Propagation*, AP-30, 27–34.

Hammershoi, D., and Møller, H. (2002), “Methods for binaural recording and reproduction.,” *Acta Acustica United With Acustica*, 88(3), 303–311.

Hartmann, W. M. (1997), “Signals, sound and sensation,”, : American institute of Physics, chapter 10 Auditory filters, pp. 251–252.

Hartmann, W. M. (1999), “How we localize sound,” *Physics today*, 52(11), 24–29.

Hausler, R., Colburn, S., and Marr, E. (1983), “Sound localization in subjects with impaired hearing,” *Acta Oto-Laryngologica*, 400, 1–62.

Haykin, S. (1996), *Adaptive filter theory*, third edn, : Prentice Hall, Inc.

HearCom (2004-2009), “The HEARCOM project,”, www.hearcom.org.

Hebrank, J., and Wright, D. (1974), “Spectral cues used in the localization of sound sources on the median plane.,” *Journal of the Acoustical Society of America*, 56(6), 1829–1834.

Helfer, K. S., and Wilber, L. A. (1990), “Hearing loss, aging and speech preception in reverberation and noise,” *Journal of speech and hearing research*, 33(1), 149–155.

Henning, G. B. (1974), “Detectability of interaural delay in high-frequency complex waveforms.,” *Journal of the Acoustical Society of America*, 55, 84–90.

Hofman, P. M., and Van Opstal, J. (1998), "Spectro-temporal factors in two-dimensional human sound localization," *Journal of the Acoustical Society of America*, 103(5), 2634–2648.

Humes, L. E. (1991), "Understanding the speech understanding problems of the hearing impaired," *Journal of the American Academy of Audiology*, 2, 59–69.

ISO3382 (1997), "Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters," *International Organization for Standardization*, .

ISO389-1 (1998), "Acoustics – Reference zero for the calibration of audiometric equipment – Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones," *International Organization for Standardization*, .

Javer, A. R., and Schwarz, D. W. (1995), "Plasticity in human directional hearing," *Journal of Otolaryngology*, 24, 111–117.

Jeffres, L. A. (1948), "A place theory of sound localization," *Journal of Comparitive and Physiological Psychology*, 41, 35–39.

Johnson, D. H., and Dudgeon, D. E. (1993), "Array signal processing: concepts and techniques,", Englewood Cliffs, NJ: Prentice Hall, chapter 4-Beamforming, pp. 111–198.

Joris, P., Smith, P. H., and Yin, T. C. T. (1998), "Coincidence Detection Minireview in the Auditory System: 50 Years after Jeffress.," *Neuron*, 21, 1235–1238.

Joris, P. X., Yin, T. C. T., and Smith, P. H. (1990), "Mechanisms of azimuthal sound localization in the central nervous system of the cat," *NAG journal*, 104, 23–35.

Joris, P., and Yin, T. C. T. (1995), "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences.," *Journal of Neurophysiology*, 73(3), 1043–62.

Joris, P., and Yin, T. C. T. (2007), "A matter of time: internal delays in binaural processing," *Trends in neurosciences*, 30(2), 70–78.

Kates, J., and Weiss, M. (1996), "A comparison of hearing-aid array processing techniques," *Journal of the Acoustical Society of America*, 99(5), 3138–3148.

Katz, B. F. G. (2001), "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation.," *Journal of the Acoustical Society of America*, 110(5), 2440–2448.

Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., and Convery, E. (2006), "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.

Klasen, T. J., Doclo, S., Van den Bogaert, T., Moonen, M., and Wouters, J. (2006), Binaural multi-channel Wiener filtering for hearing aids: preserving interaural time and level differences, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. V, Toulouse, France, pp. 145–148.

Klasen, T. J., Van den Bogaert, T., Moonen, M., and Wouters, J. (2007), "Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues," *IEEE Transactions on Signal Processing*, 55(4), 1579–1585.

Kleiner, M., Dalenback, B. I., and Svensson, P. (1993), "Auralization- An Overview.," *Journal of the Audio Engineering Society*, pp. 861–875.

Kochkin, S. (2005), "Customer satisfaction with hearing instruments in the digital age," *The Hearing Journal*, 58(9), 30–39.

Kochkin, S., and Kuk, F. (1997), "The binaural advantage: evidence from subjective benefit and customer satisfaction data," *The Hearing Review*, 4(4), 29–34.

Krokstadt, A., Strom, S., and Sorsdal, S. (1968), "Calculating the acoustical room response by the use of a ray tracing technique.," *Journal of Sound and Vibration*, 8, 118–125.

Kuhn, G. F. (1977), "Model for the interaural time differences in the azimuthal plane," *Journal of the Acoustical Society of America*, 62(1), 157–167.

Kutruff, K. H. (1993), "Auralisation of Impulse Responses Modeled on the Basis of Ray-Tracing Results.," *Journal of the Audio Engineering Society*, 41, 876–880.

Langendijk, E. H. A., and Bronkhorst, A. W. (2002), "Contribution of spectral cues to human sound localization," *Journal of the Acoustical Society of America*, 112(4), 1583–1596.

Leeuw, A. R., and Dreschler, W. A. (1991), "Advantages of directional hearing aid microphones related to room acoustics.," *Audiology*, 30(6), 330–344.

Levitt, H. (2001), "Noise reduction in hearing aids: A review," *Journal of Rehabilitation Research and Development*, 38(1), 111–121.

Levitt, H., Bakke, M., Kates, J., Neuman, A., Schwander, T., and Weiss, M. (1993), "Signal processing for hearing impairment," *Scandinavian Audiology Supplementum*, 38, 7–19.

Libby, E. R. (2007), "The search for the binaural advantage revisited," *The Hearing Review*, pp. 1–4.

Lippman, R. P., Braida, L. D., and Durlach, N. I. (1981), "Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss," *Journal of the Acoustical Society of America*, 69(2), 524–534.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999), "The precedence effect.," *Journal of the Acoustical Society of America*, 106(4), 1633–1654.

Lorenzi, C. S., Gatehouse, S., and Lever, C. (1999*a*), "Sound localization in noise in hearing impaired listeners," *Journal of the Acoustical Society of America*, 105(6), 3454–3463.

Lorenzi, C. S., Gatehouse, S., and Lever, C. (1999*b*), "Sound localization in noise in normal hearing listeners," *Journal of the Acoustical Society of America*, 105(3), 1810–1820.

Luo, F.-L., Yang, J., Pavlovic, C., and Nehorai, A. (2002), "Adaptive Null-Forming Scheme in Digital Hearing Aids," *IEEE Transactions on Signal Processing*, 50(7), 1583–1590.

Lybarger, S. F. (1988), "Handbook of hearing amplification,", Vol. I, : Boston: college hill press, pp. 1–29.

Mackensen, P. (2004), Auditive Localization: Head movements, an additional cue in Localization, PhD thesis, Technical University Berlin.

Maj, J. B. (2004), Adaptive noise reduction algorithms for speech intelligibility improvement in dual microphone hearing aids, PhD thesis, K.U.Leuven.

Maj, J. B., Moonen, M., and Wouters, J. (2002), "A robust voice activity detector for SVD-based noise reduction in hearing aids," *Signal Processing Symposium (SPS)*, March 21-22, Leuven, Belgium, 189–192.

Maj, J. B., Royackers, L., Moonen, M., and Wouters, J. (2006), "Comparison of adaptive noise reduction algorithms in dual microphone hearing aids," *Speech Communication*, 48, 957–970.

Maj, J. B., Wouters, J., and Moonen, M. (2004), "Noise reduction results of an adaptive filtering technique for dual microphone behind the ear hearing aids," *Ear and Hearing*, 25(3), 215–229.

Makous, J. C., and Middlebrooks, J. C. (1990), "Two dimensional sound localization by human listeners," *Journal of the Acoustical Society of America*, 87(5), 2188–2200.

McAdams, S. (1993), “Thinking in sound: The cognitive psychology of human audition.,”, : Oxford UP, chapter 6 Recognition of sound sources and events, pp. 147–197.

McGurck, H., and Macdonald, J. W. (1976), “Hearing lips and seeing voices,” *Nature*, 264(746-748).

Middlebrooks, J. C., Makous, J. C., and Green, D. M. (1989), “Directional sensitivity of sound-pressure levels in the human ear canal,” *Journal of the Acoustical Society of America*, 86, 89–108.

Minnaar, P., Olesen, S., Christensen, F., and Møller, H. (2001), “Localization with binaural recordings from artificial and human heads.,” *Journal of the Audio Engineering Society*, 49(5), 323–336.

Møller, H. (1989), “Reproduction of artificial-head recordings through loudspeakers.,” *Journal of the Audio Engineering Society*, 37, 30–33.

Møller, H., Hammershoi, D., Jensen, C. B., and Sørensen, M. F. (1999), “Evaluation of artificial heads in listening tests,” *Journal of the Audio Engineering Society*, 47(3), 83–100.

Møller, H., Sørensen, M. F., Jensen, C. B., and Hammershoi, D. (1996), “Binaural technique: do we need individual recordings,” *Journal of the Audio Engineering Society*, 44(6), 451–469.

Moore, B. C. J. (1989), “An Introduction to the Psychology of Hearing,”, third edition edn, : Academic Press limited, chapter 7 Auditory pattern and object recognition, pp. 229–253.

Moore, B. C. J. (1997*a*), “An introduction to the psychology of hearing,”, : Academic Press, chapter 8 Spatial hearing and related phenomena, pp. 297–345.

Moore, B. C. J. (1997*b*), “An introduction to the psychology of hearing,”, : Academic Press, chapter 7 Space perception, pp. 223–267.

Moore, B. C. J. (2003), “Speech processing for the hearing-impaired: successes, failures, and implications for speech mechanisms,” *Speech Commununication*, 41(1), 81–91.

Musicant, A. D., and Butler, R. A. (1984), “The influence of pinnae-based spectral cues on sound localization.,” *Journal of the Acoustical Society of America*, 75, 1195–1200.

Noble, W. (2006), “Effects of bilateral versus unilateral hearing aid fitting on abilities mesured by the Speech, Spatial and Qualities of Hearing scale (SSQ),” *International Journal of Audiology*, 45, 172–181.

Noble, W., and Byrne, D. (1990), “A comparison of different binaural hearing aid systems for sound localization in the horizontal and vertical planes,” *British Journal of Audiology*, 24, 335–346.

Noble, W., Byrne, D., and Lepage, B. (1994), “Effects on sound localization of configuration and type of hearing impairment,” *Journal of the Acoustical Society of America*, 95(2), 992–1005.

Noble, W., and Gatehouse, S. (2004), “The Speech, Spatial and Qualities of Hearing Scale (SSQ),” *International Journal of Audiology*, 43, 100–114.

Noble, W., Sinclair, S., and Byrne, D. (1998), “Improvements in aided sound localization with open earmolds: observations in people with high-frequency hearing loss,” *Journal of the American Academy of Audiology*, 9(1), 25–34.

Noble, W., Ter-Horst, K., and Byrne, D. (1995), “Disabilities and handicaps associated with impaired auditory localization,” *Journal of the American Academy of Audiology*, 6(2), 129–140.

Ouellette, J. (1999), “The incredible shrinking microphone,” *The industrial physicist*, 5(4), 7–9.

Pavlovic, C. V. (1987), “Derivation of premary parameters and procedures for use in speech intelligibility predictions,” *Journal of the Acoustical Society of America*, 82(2), 413–422.

Peissig, J., and Kollmeier, B. (1997), “Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners,” *Journal of the Acoustical Society of America*, 101(3), 1660–1669.

Perrett, S., and Noble, W. (1995), “Available response choices affect localisation of sound,” *Perception and psychophysics*, 57(2), 150–158.

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 and Hearing Research*, 29(2), 146–154.

Plomp, R., and Duquesnoy, A. J. (1982), “A model for the speech-reception threshold in noise without and with a hearing aid,” *Scandinavian Audiology Supplementum*, 15, 95–111.

Plomp, R., and Mimpen, A. M. (1979), “Improving the reliability of testing the speech reception threshold for sentences,” *Audiology*, 18(1), 43–52.

Plomp, R., and Mimpen, A. M. (1981), "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech reception threshold for sentences," *Acustica*, 48, 325–328.

Posselt, C. (1987), Binaurale Raumsimulation fr Kopfhrerwiedergabe., in *Fortschritte der Akustik - DAGA'87*, pp. 725–728.

Rayleigh, L. (1907), "On our perception of sound direction," *Philosophical Magazine*, 13, 214–232.

Ricketts, T. (2000), "Directivity quantification in hearing aids: fitting and measurement effects," *Ear and Hearing*, 21(1), 45–58.

Ricketts, T., and Henry, P. (2002), "Evaluation of an adaptive, directional-microphone hearing aid," *International Journal of Audiology*, 41(2), 100–112.

Rindel, J. H. (2000), "The use of computer modeling in room acoustics," *Journal of Vibroengineering*, 3(4), 219–224.

Rindel, J. H. (2004), Evaluation of room acoustic qualities and defects by use of auralization., in *Proceedings of the 148th Meeting of the Acoustical Society of America*.

Rychtáriková, M., Van den Bogaert, T., Vermeir, G., and Wouters, J. (2007), Virtual acoustics for localisation of the speaker in a real and virtual reverberant room, in *Proceedings of the International Congress and Exposition on Noise Control Engineering - Internoise*, Istanbul.

Rychtáriková, M., Van den Bogaert, T., Vermeir, G., and Wouters, J. (2008), "Validation of binaural sound source localization in virtual acoustical environments," *Journal of the Acoustical Society of America*, . Submitted.

Savioja, L., Huopaniemi, J., Lokki, T., and Vaananen, R. (1999), "Creating interactive virtual acoustic environments," *Journal of the Audio Engineering Society*, 47(9), 675–705.

Soede, W., Bilsen, F. A., and Berkhout, A. J. (1993), "Assessment of a directional microphone array for hearing-impaired listeners," *Journal of the Acoustical Society of America*, 94(2), 799–808.

Soede, W., J., B. A., and Bilsen, F. A. (1993), "Development of a directional hearing instrument based on array technology," *Journal of the Acoustical Society of America*, 94(2), 785–798.

Spandock, F. (1934), "Raumakustische. Modellversuche," *Annalen derPhysik*, V(20 Heft 4), 345–360.

Spriet, A. (2004), Adaptive filtering techniques for noise reduction and acoustic feedback cancellation in hearing aids, PhD thesis, K.U.Leuven.

Spriet, A., Moonen, M., and Wouters, J. (2001), "Robustness analysis of GSVD based optimal filtering and generalized sidelobe canceller for hearing aid applications," *IEEE Workshop on Applications on Signal Processing to Audio and Acoustics (WASPAA 2001)*, October 21-24, New Paltz, USA, 31–34.

Spriet, A., Moonen, M., and Wouters, J. (2004), "Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction," *Signal Processing*, 84, 2367–2387.

Spriet, A., Moonen, M., and Wouters, J. (2005), "Stochastic gradient-based implementation of spatially preprocessed speech distortion weighted multichannel Wiener filtering for noise reduction in hearing aids," *IEEE Transactions on Signal Processing*, 53(3), 911–925.

Spriet, A., Van Deun, L., Eftaxiadis, K., Laneau, J., Moonen, M., Van Dijk, B., van Wieringen, A., and Wouters, J. (2007), "Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the Nucleus Freedom cochlear implant system," *Ear and Hearing*, 28, 62–72.

Stadler, R. W., and Rabinowitz, W. M. (1993), "On the potential of fixed arrays for hearing aids," *Journal of the Acoustical Society of America*, 94(3), 1332–1342.

Stevens, S. S., and Newman, E. B. (1936), "The localization of actual sources of sound," *American Journal of Psychology*, 48, 297–306.

Teutsch, H., and Elko, G. W. (2001), First- and Second-order Adaptive Differential Microphone Arrays,, 7th International Workshop on Acoustic Echo and Noise Control, Darmstadt, Germany, pp. 35–38.

Thompson, S. C. (2000), "Directional patterns obtained from two or three microphones.," *Technical report, Knowles electronics*, .

Tollin, D. J. (2003), "The lateral superior olive: a functional role in sound source localization.," *The neuroscientist*, 9(2), 127–143.

Treeby, B. E., Pan, J., and Paurobally, R. M. (2007), "The effect of hair on auditory localization cues," *Journal of the Acoustical Society of America*, 122(6), 35863597.

Van Compernolle, D., and Van Gerven, S. (1995), "COST 229: Applications of difital signal processing to Telecommunications,", chapter Beamforming with microphone arrays, pp. 107–131.

Van den Bogaert, T., Doclo, S., Moonen, M., and Wouters, J. (2007), Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. IV, Honolulu, HI, USA, pp. 565–568.

Van den Bogaert, T., Doclo, S., Wouters, J., and Moonen, M. (2008*a*), "The effect of multi-microphone noise reduction systems on sound source localization in binaural hearing aids," *Journal of the Acoustical Society of America*, 124(1). In Press.

Van den Bogaert, T., Doclo, S., Wouters, J., and Moonen, M. (2008*b*), "Speech in noise enhancement using multi-microphone binaural hearing aids: evaluation of a binaural multichannel Wiener filter," *Journal of the Acoustical Society of America*, . Submitted.

Van den Bogaert, T., Klasen, T. J., Moonen, M., and Wouters, J. (2005), Distortion of interaural time cues by directional noise reduction systems in modern digital hearing aids, in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, pp. 57–60.

Van den Bogaert, T., Klasen, T. J., Van Deun, L., Wouters, J., and Moonen, M. (2006), "Localization with bilateral hearing aids: without is better than with," *Journal of the Acoustical Society of America*, 119(1), 515–526.

Van Deun, L., van Wieringen, A., Van den Bogaert, T., Scherf, F., Offeciers, F. E., Van de Heyning, P. H., Desloovere, C., Dhooge, I. J., Deggouj, N., and Wouters, J. (2007), "Sound localization, sound lateralization and binaural masking level differences in young children with normal hearing," *Ear and hearing*, . Under revision.

Van Hoesel, R. J. M., and Tyler, R. S. (2003), "Speech perception, localization and lateralization with bilateral cochlear implants," *Journal of the Acoustical Society of America*, 113(3), 1617–1630.

Van Hoesel, R., Ramsden, R., and O'Driscoll, M. (2002), "Sound direction identification, interaural time delay discrimination, and speech intelligibility advantages in noise for a bilateral cochlear implant user," *Ear and Hearing*, 23(2), 137–149.

Van Veen, B. D., and Buckley, K. M. (1988), "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, 5(2), 4–24.

Van Wanrooij, M. M., and Van Opstal, J. (2007), "Sound localization under perturbed binaural hearing," *Journal of Neurophysiology*, 97, 715–726.

Van Wanrooij, M. M., and Van Opstal, J. A. (2004), "Contribution of Head Shadow and Pinna Cues to Chronic Monaural Sound Localization.," *The Journal of Neuroscience*, 24(17), 4163–4171.

Van Wanrooij, M. M., and Van Opstal, J. A. (2005), "Relearning sound localization with a new ear," *The Journal of Neuroscience*, 25(22), 5413–5424.

VandenBerghe, J., and Wouters, J. (1998), "An adaptive noise canceller for hearing aids using two nearby microphones," *Journal of the Acoustical Society of America*, 103(6), 3621–3626.

Veit, I., and Sander, H. (1987), "Production of spatially limited diffuse sound field in an anechoic room," *Journal of the Audio Engineering Society*, 35, 138–142.

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000), "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *Journal of the Acoustical Society of America*, 107(3), 1671–1684.

Vorlander, M. (1989), "Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing and image-source algorithm.," *Journal of the Acoustical Society of America*, 86, 172–178.

Walden, B. E., Surr, R. K., Cord, M. T., Edwards, B., and Olson, L. (2000), "Comparison of benefits provided by different hearing aid technologies," *Journal of the American Academy of Audiology*, 11(10), 540–560.

Webster, F. A. (1951), "The influence of interaural phase on masked thresholds. I. The role of interaural time-deviation," *Journal of the Acoustical Society of America*, 23(4), 452–462.

Weiss, M. (1974), "Processing speech signals to attenuate interference," *IEEE Symposium on Speech Recognition*, .

Welker, D. P., Greenberg, J. E., Desloge, J. G., and M., Z. P. (1997), "Microphone-array hearing aids with binaural output-part II: A two-microphone adaptive system," *IEEE Transactions Speech and Audio Processing*, 5(6), 543–551.

Wightman, F. L., and Kistler, D. J. (1992), "The dominant role of low-frequency interaural time differences in sound localization," *Journal of the Acoustical Society of America*, 91(3), 1648–1661.

Wittkop, T., and Hohmann, V. (2003), "Strategy selective noise reduction for binaural digital hearing aids," *Speech Communication*, 39, 111–138.

Wouters, J., and VandenBerghe, J. (2001), "Speech recognition in noise for cochlear implantees with a two- microphone monaural adaptive noise reduction system," *Ear and Hearing*, 22(5), 420–430.

Wright, B. A., and Fitzgerald, M. B. (2001), "Different patterns of human discrimination learning for two interaural cues to sound-source localization," *Proceedings of the National Academy of Sciences*, pp. 12307–12312.

Wright, B., and Zhang, Y. (2006), “A review of learning with normal and altered sound-localization cues in human adults,” *International Journal of Audiology*, 45(Supplement 1), 92–98.

Zeng, X. C., L., C. C., and Rindel, J. H. (2006), “Practical methods to define scattering coefficients in a room acoustics computer model.,” *Apllied acoustics*, 67(8), 771–786.

Zurek, P. M. (1993), “Acoustical Factors Affecting Hearing Aid Performance,”, 2nd edn, Baltimore: University Park Press, chapter 15-Binaural advantages and directional effects in speech intelligibility, pp. 255–276.

Zurek, P. M., and Greenberg, J. E. (2000), Two-microphone adaptive array hearing aids with monaural and binaural outputs, in *Proceedings of the ninth IEEE DSP Workshop*, Hunt, TX, USA.

List of publications

Publications in International Journals

1. Wouters J., Geurts L., Laneau J., Luts H., Maj J.B., Moonen M., Royackers L., Spriet A. Van den Bogaert T., van Wieringen A.: ”Digital hearing aids and future directions for hearing aids,” *Acta Oto-Rhino-Laryngologica Belgica,* Vol.56, p357-361, 2002.

2. Van den Bogaert T., Klasen T.J., Moonen M., Van Deun L., Wouters J.: ”Horizontal localization with binaural hearing aids: without is better than with,” *Journal of the Acoustical Society of America,* vol.119, no.1, p515-526, Jan. 2006.

3. Klasen T.J., Van den Bogaert T., Moonen M, Wouters, J.: ”Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues,” *IEEE Trans. on signal processing,* vol.55, no.4, p1579-1585, April 2007.

4. Van Deun L., van Wieringen A., Van den Bogaert T., Scherf F., Offeciers F.E., Van de Heyning P.H., Desloovere C., Dhooge I.J., Deggouj N., Wouters J., ”Sound localization, sound lateralization and binaural masking level differences in young children with normal hearing,” *Ear and hearing*, Revision 2 submitted Dec 2007.

5. Van den Bogaert T., Doclo S., Moonen M., Wouters J.: ”Speech in noise enhancement using multi-microphone binaural hearing aids: Evaluation of a binaural multichannel Wiener filter,” *Journal of the Acoustical Society of America*, Submitted, March 2008

6. Doclo S., Van den Bogaert T., Moonen M., Wouters J.: ”Reduced bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids,” *IEEE Trans. on signal processing*, Accepted, April 2008.

7. Rychtáriková M., Van den Bogaert T., Vermeir, G. and Wouters J.: "Validation of binaural sound source localization in virtual acoustical environments" *Journal of the Acoustical Society of America,* Submitted, April 2008

8. Doclo S., Cornelis B., Van den Bogaert T., Moonen M. and Wouters J.: "Theoretical analysis of binaural multi-microphone noise reduction techniques" *IEEE Trans. on signal processing*, Submitted, June 2008

9. Van den Bogaert T., Doclo S., Moonen M., Wouters J.: "The effect of multi-microphone noise reduction systems on sound source localization in binaural hearing aids," *Journal of the Acoustical Society of America*, vol.124, no.1, 14 pages, In press, scheduled for July 2008.

International Conference Papers

1. Klasen, T., Van den Bogaert, T., Moonen, M. and Wouters, J., "Preserving interaural time delay cues during noise reduction in hearing aids", *Proc. of the Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA)*, Piscataway, NJ, USA, pp. 1-2, March 2005.

2. Klasen, T., Moonen, M., Van den Bogaert, T. and Wouters, J., "Preservation of interaural time delay for binaural hearing aids through multichannel wiener filtering based noise reduction", *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, USA, vol. 3, pp. 29-32, March 2005.

3. Klasen, T., Van den Bogaert, T., Moonen, M. and Wouters, J., "Binaural noise reduction for hearing aids: preserving interaural time delay cues", *Proc. of the IEEE Benelux/DSP Valley Signal Processing Symposium (SPS-DARTS)*, Antwerp, Belgium, pp. 23-26, April 2005.

4. Klasen, T., Van den Bogaert, T., Moonen, M. and Wouters, J., "Preventing ITD distortion in binaural hearing aids during noise reduction using multi-channel wiener filtering", *Proc. of the IEEE International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Eindhoven, The Netherlands, pp. 149-152, Sept. 2005.

5. Van den Bogaert, T., Klasen, T., Moonen, M. and Wouters, J., "Distortion of interaural time cues by directional noise reduction systems in modern digital hearing aids", *Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, pp. 57-60, Oct. 2005.

6. Wouters, J., Klasen, T., Moonen, M., Spriet, A., Van den Bogaert, T., Van Deun, L. and van Wieringen, A., "Effects of "directional"signal processing in hearing instruments on speech understanding in noise and localization", *Proc. of the Danavox Symposium*, Kolding, Denmark, pp. 93-111, 2005.

7. Klasen, T., Doclo, S., Van den Bogaert, T., Moonen, M. and Wouters, J., "Binaural multi-channel Wiener filtering for hearing aids: Preserving interaural time and level differences", *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, pp. V145-V148 May 2006.

8. Doclo, S., Klasen, T., Van den Bogaert, T., Wouters, J. and Moonen, M., "Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions", *Proc. of the IEEE International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Paris, France, Sept 2006.

9. Doclo, S., Van den Bogaert, T., Wouters, J. and Moonen, M., "Comparison of reduced-bandwidth MWF-based noise reduction algorithms for binaural hearing aids", *Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, , Oct. 2007.

10. Van den Bogaert, T., Doclo, S., Moonen, M. and Wouters, J., "Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter", *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, USA, vol. 4, pp. 565-568, April 2007.

11. Rychtáriková, M., Van den Bogaert, T., Vermeir, G. and Wouters, J., "Virtual acoustics for localisation of the speaker in a real and virtual reverberant room", *Proc. of the International Congress and Exposition on Noise Control Engineering (Internoise)*, Istanbul, Turkey, pp. 1-8, 2007.

12. Cornelis, B., Van den Bogaert, T., Moonen, M. and Wouters, J., "Analysis of localization cue preservation by Multichannel Wiener Filtering based binaural noise reduction in hearing aids"*European Conference on Signal Processing 2008*, Accepted.

Abstracts without conference proceedings

1. Van den Bogaert, T., Moonen, M., Wouters, J., "Localization of normal hearing and hearing aid users", *Belgian Society of Audiology (BSA 2003)*, Luik, Belgium, Dec. 2003

2. Wouters, J., Doclo, S., Klasen, T.J., Maj, J.-B., Moonen M., Royackers L., Spriet A. and Van den Bogaert T., "Noise reduction approaches for improved speech perception", *International Hearing Aid Research Conference (IHCON 2004)*, Lake Tahoe CA, USA, Aug. 2004

3. Van den Bogaert, T., Moonen, M., Wouters, J., "Localisation with binaural hearing aids: without is better than with", *International Hearing Aid Research Conference (IHCON 2004)*, Lake Tahoe, California (IHCON), USA, Aug. 2004

4. Van den Bogaert, T., Klasen, T., Van Deun, L., Moonen, M., Wouters, J., "Horizontal localization with bilateral hearing aids: without hearing aids is better than with?", *Conference of the International Collegium for Rehabilitative Audiology (ICRA)*, Gainesville, FL, USA, March 2005

5. Van Deun, L., van Wieringen A., Van den Bogaert, T., Scherf, F., Van de Heyning, P., Zarowski, A., Dhooge, I., Deggouj, N., De Raeve, L., Desloovere, C., Wouters, J., "Sound localization in bilaterally implanted children below 7 years of age", *Conference of Implantable Auditory Prostheses CIAP-2005*, Assilomar, California, USA, August 2005

6. Wouters J., Van den Bogaert T., Klasen T.J., Moonen M.,"Directional hearing with bilateral hearing aids: better signal processing is needed", *International Binaural Symposium*, Manchester, UK, Oct. 2005

7. Van den Bogaert, T., Moonen, M., Wouters, J., "Binaural cue distortion in hearing aids," *Vergadering werkgemeenschap auditief systeem (WAS-dag)*, Netherlands, Nov. 2005

8. Doclo, S., Klasen, T., Van den Bogaert, T., Moonen, M., Wouters, J., "Physical and perceptual evaluation of the Interaural Wiener filter algorithm", *International Hearing Aid Research Conference (IHCON)*, Lake Tahoe, CA, USA, Aug. 2006

9. Van den Bogaert, T., Moonen, M., Wouters, J., "Preservation of interaural cues using a binaural multichannel Wiener filter (MWF-ITF): Localization performance," *International Hearing Aid Research Conference (IHCON 2006)*, Lake Tahoe, California, USA, Aug. 2006

10. Van Deun, L., van Wieringen, A., Van den Bogaert, T., Scherf, F., Van de Heyning, P., Deggouj, N., De Raeve, L., Dhooge, I., Zarowski, A., Desloovere, C., Wouters, J., "Sound localization and lateralization in NH and bilateral CI Children", *9th International Cochlear Implant Conference*, Vienna, Austria, July 2006

11. Wouters J., Doclo S., Eneman K., Leijon A., Moonen M., Spriet A., Van den Bogaert T. "Advanced signal processing for hearing instruments and cochlear implants", *28th International Congress of Audiology (ICA 2006)*, Innsbruck, Austria, Sept. 2006

12. Wouters J., Doclo S., Moonen M., Van den Bogaert T.,"The use of both ears: noise reduction and localization with bilateral hearing aids", *8th European Federation of Audiology Societies Congress (EFAS 2007)*, Heidelberg, Germany, June 2007

13. Wouters J., Van den Bogaert T., Doclo S., Moonen M.,"Speech-in-noise enhancement and sound localization with improved binaural hearing instruments", *ASA-EAA joint conference on Acoustics 2008*, Paris, France, June 2008

14. Rychtáriková, M., Vermeir G., Van den Bogaert T. , Corveleyn S., Eneman K., Lauriks W., Moonen M., Wouters J.,"The use of virtual Acoustics in the evaluation and development of binaural hearing aid algorithms", *ASA-EAA joint conference on Acoustics 2008*, Paris, France, June 2008

15. Van den Bogaert, T., Doclo, S., Wouters, J., Moonen, M., "Improvements in speech perception and sound localization in hearing aids using binaural multichannel Wiener filtering", *International Hearing Aid Research Conference (IHCON)*, Lake Tahoe, CA, USA, Aug. 2008, Submitted.

Curriculum Vitae

Tim Van den Bogaert was born in Kapellen, Belgium, in 1978. In 2000 he received his Industrial Electronics Engineering degree from the Karel de Grote Hogeschool, Antwerp, Belgium. In 2002 he received his Elektrotechnical Engineering degree from the Katholieke Universiteit Leuven, Belgium. From 2002 he has been employed as a research assistant in a joint research project of the Laboratory of Experimental Oto-Rhino-Laryngology (Dept. Neurosciences) and the Signals, Identification, System Theory and Automation group (Dept. Electrical Engineering), both of the KULeuven, under supervision of Prof. dr. J. Wouters and Prof. dr. ir. M. Moonen. His research interests are in the area of applied signal processing techniques for speech, video and audio applications. One of his main research activities aimed at studying the impact of signal processing strategies on the human auditory system.

From 2003 till 2007 he was funded by the Innovation through Science and Technology in Flanders (IWT-Vlaanderen). From 2002 till 2003 and in 2008 he received funding from the KULeuven-ExpORL.