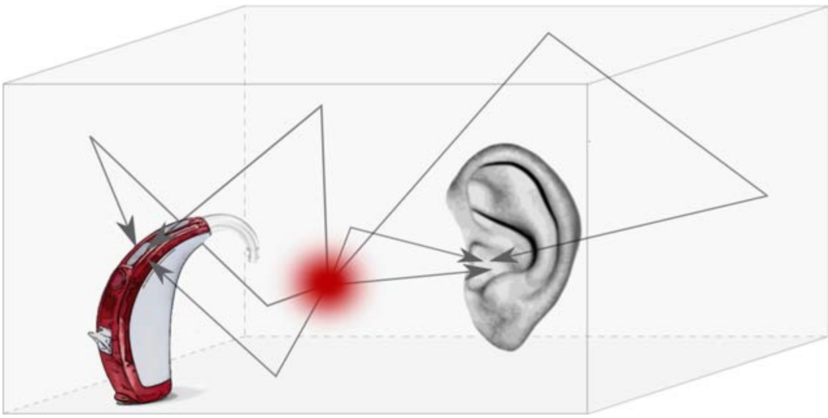
Technical University of Denmark

CONTRIBUTIONS TO
HEARING RESEARCH

Volume 11

Iris Arweiler

Processing of spatial sounds in the impaired auditory system



Centre for Applied Hearing Research
Department of Electrical Engineering

Processing of spatial sounds in the impaired auditory system

PhD thesis by

Iris Arweiler

Technical University of Denmark
2011

Cover illustration by Gilles Pigasse.
Printed in Denmark by Rosendahls-Schultz Grafisk A/S.
The defense was held on July 6, 2011.

This PhD-dissertation is the result of a research project at the Centre for Applied Hearing Research (CAHR), Department of Electrical Engineering, Technical University of Denmark (Kgs. Lyngby, Denmark). The work was completed between January 2008 and April 2011.

The project was financed by a DTU stipend, the Danish graduate school SNAK (Sense organs, neural networks, behavior, and communication) and Phonak AG. An external research stay at the National Acoustic Laboratories (NAL) in Sydney was supported by grants from Balslev Fonden and Oticon Fonden.

Supervisors

Main supervisor

Associate Prof. Jörg M. Buchholz
Centre for Applied Hearing Research
Department of Electrical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark

Co-supervisor

Prof. Torsten Dau
Centre for Applied Hearing Research
Department of Electrical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark

Preface

In 2005, I had the opportunity to come to Denmark and to join the Centre for Applied Hearing Research (CAHR). At that time, the CAHR was still a relatively small group. First, I was involved in the European Project HearCom. After the project had finished, I continued research and started my PhD studies. Over the years more people joined the CAHR, others left, and suddenly it is time to look back and to remember all the people that have contributed in one way or the other to this PhD thesis.

The biggest contribution came undoubtedly from my supervisors, Jörg and Torsten. Thanks for your support, patience, motivation and trust throughout the project. I have learned a lot from you and I truly appreciated your way of guiding me through the project. The way you both create a relaxed working environment combined with high-level research is remarkable.

Thanks to all my friends and colleagues for the good times in and outside the office. Special thanks to Jens Bo, Olaf, Tobias, Eric, Helen, Sébastien, Sarah and Sylvain. Thank you also, Sylvain, for your help with the SpaceLab.

Many thanks to Ellen for providing the Dantale II program and to Claus for translating the abstract into Danish.

I would also like to thank Harvey and Gitte for having me over at the National Acoustic Laboratories in Sydney. It was a great experience.

Thanks to Schelle and Stef from Phonak for your support and to Widex for producing the ear molds.

Special thanks go to Torben who is always approachable for questions as well as for a private 'snak'.

There would not be any results without the hearing-impaired and normal-hearing listeners who participated in my experiments. Thank you very much for repeating nonsense sentences over and over again.

Merci Gilles for always being there and for taking my mind off work.

Iris Arweiler, April 29, 2011

Contents

Abstract

Understanding speech in complex acoustic environments presents a challenge for most hearing-impaired listeners. In conditions where normal-hearing listeners effortlessly utilize spatial cues to improve speech intelligibility, hearing-impaired listeners often struggle. In this thesis, the influence of two such cues on speech intelligibility was studied. First, the benefit from early reflections (ER's) in a room was determined using a virtual auditory environment. ER's were found to be useful for speech intelligibility, but to a smaller extent than the direct sound (DS). The benefit was quantified with an intelligibility-weighted "efficiency factor" which revealed that the spectral characteristics of the ER's caused the reduced benefit. Hearing-impaired listeners were able to utilize the ER energy as effectively as normal-hearing listeners, most likely because binaural processing was not required for the integration of the ER's with the DS. Different masker types were found to have an impact on the binaural processing of the overall speech signal but not on the processing of ER's. Second, the influence of interaural level differences (ILD's) on speech intelligibility was investigated with a hearing aid research platform. ILD's are considered important for localizing sounds and for the perceptual separation of competing sound sources. Bilateral hearing aids with independent compression algorithms typically decrease ILD's, such that the perception of spatial sounds becomes distorted. Hearing aids that are binaurally linked can utilize the signals at both ears and preserve the ILD's through co-ordinated compression. Hearing-impaired listeners received a small, but not significant advantage from linked compared to independent compression. It was concluded that, for speech intelligibility, the exact ILD information is not crucial. The results from an additional experiment demonstrated that the ER benefit was maintained with independent as well as with linked hearing aid compression. Overall, this work contributes to the understanding of ER processing in listeners with normal and impaired hearing and may have implications for speech perception models and the development of compensation strategies in future generations of hearing instruments.

Resumé

For de fleste personer med høretab er det en udfordring at forstå tale i komplekse akustiske omgivelser. Under omstændigheder, hvor normalthørende lyttere ubesværet bruger rumlige cues til at forbedre taleforståelsen, er det ofte vanskeligere for mennesker med nedsat hørelse. Formålet med dette projekt var at undersøge hvordan to af de rumlige cues indvirker på taleforståeligheden. For det første blev fordelen ved de tidlige refleksioner i et rum bestemt ved hjælp af et virtuelt auditivt lyttemiljø. De tidlige refleksioner viste sig at være nyttige for taleforståeligheden, dog i mindre grad end den direkte lyd. Fordelen blev kvantificeret med en sprog-vægtet "effektivitetsfaktor" som viste, at de spektrale karakteristika af de tidlige refleksioner forårsagede den reducerede fordel. Personerne med høretab var i stand til at udnytte energien fra de tidlige refleksioner lige så effektivt som de normalthørende, sandsynligvis fordi binaural processering ikke var nødvendig for at integrere de tidlige refleksioner med den direkte lyd. Forskellige typer af støj viste sig at have en indvirkning på den binaurale processering af det samlede talesignal, men ikke på processeringen af de tidlige refleksioner. For det andet blev inflydelsen af interaurale niveauforskelle på taleforståeligheden undersøgt ved hjælp af en høreapparat-forskningsplatform. Interaurale niveauforskelle anses for at være vigtige for lokalisering af lyde og for at adskille forskellige lydkilder perceptuelt. Bilaterale høreapparater med uafhængige komprimeringsalgoritmer reducerer typisk de interaurale niveauforskelle således at opfattelsen af rumlige lyde bliver forvrænget. Høreapparater, der er forbundet binauralt kan udnytte signalerne ved begge ører og dermed bevare de interaurale niveauforskelle gennem en koordineret komprimering. Personerne med høretab viste en lille, men ikke signifikant fordel af den koordinerede komprimering i forhold til den uafhængige komprimering. Konklusionen var, at den nøjagtige interaurale niveauforskel ikke er afgørende for taleforståeligheden. Resultaterne fra et yderligere forsøg viste, at fordelen ved de tidlige refleksioner blev opretholdt med uafhængig såvel som med koordineret komprimering. Samlet set bidrager dette arbejde til at forstå processeringen af de tidlige refleksioner i personer med normal og nedsat hørelse, hvilket kan have betydning for udviklingen af modeller som kan simulere opfattelsen af tale. Desuden kan arbejdet have betydning for udvikling af fremtidige kompensationsstrategier i nye generationer af høreapparater.

Related publications

Journal articles

- Arweiler, I., Buchholz, J. M. (2011). The influence of spectral characteristics of early reflections on speech intelligibility. *J. Acoust. Soc. Am.*, 130, 996–1005.

- Arweiler, I., Buchholz, J. M., Dau, T. (2011). Speech intelligibility with binaurally linked hearing aids. *Int. J. Audiol.*, under review

- Arweiler, I., Buchholz, J. M., Dau, T. (2011). The influence of masker type on early reflection processing and speech intelligibility (L). *J. Acoust. Soc. Am.*, under review

Conference papers

- Arweiler, I., Buchholz, J. M., Dau, T. (2010). Monaural and binaural benefit from early reflections for speech intelligibility. *Fortschritte der Akustik DAGA'10, 36th German Convention on Acoustics*, Berlin, Germany, March 2010.

- Arweiler, I., Buchholz, J. M., Dau, T. (2010). Speech intelligibility enhancement by early reflections. *International Symposium on Auditory and Audiological Research (ISAAR), Binaural Processing and Spatial Hearing*, Helsingør, Denmark, August 2009.

Published abstracts

- Arweiler, I., Buchholz, J. M., Dau, T. (2010). The influence of spectral and spatial characteristics of early reflections on speech intelligibility. *International Hearing Aid Research Conference (IHCON)*, Lake Tahoe, CA, August 2010.

- Arweiler, I., Buchholz, J. M., Dau, T. (2010). Speech intelligibility enhancement by early reflections for normal-hearing and hearing-impaired listeners *International Society of Audiology, XXX International Conference of Audiology*, São Paulo, Brazil, March 2010.

List of abbreviations

ANOVA	Analysis of variance
AT	Attack time
BTE	Behind-the-ear
CI	Confidence interval
CIC	Completely-in-the-canal
CR	Compression ratio
CT	Compression threshold
DS	Direct sound
ER	Early reflection
FMB	Fluctuating masker benefit
F0	Fundamental frequency
HARP	Hearing aid research platform
HATS	Head and torso simulator
HL	Hearing level
HRTF	Head related transfer function
IG	Insertion gain
ILD	Interaural level difference
ITD	Interaural time difference
ITE	In-the-ear
LoRA	Loudspeaker Room Auralization
MT	Multi-talker
RIR	Room impulse response

rms	Root-mean-square
REAG	Real ear aided gain
RECD	Real-ear-to-coupler difference
REIG	Real ear insertion gain
REUG	Real ear unaided gain
revTT	Reversed two-talker
RT	Release time
SII	Speech Intelligibility Index
SNR	Signal-to-noise ratio
SPL	Sound pressure level
SRT	Speech reception threshold
SSN	Speech-shaped noise
WDRC	Wide dynamic range compression

1

General introduction

From an evolutionary point of view, it has always been important to use our hearing for a precise analysis of the acoustic environment. It was vital for our hunting ancestors to identify the location of the prey and to judge from which direction a predator was attacking. Things have changed in modern society; however, our sense of spatial hearing is still indispensable, for example, when we decide not to cross the road because we can hear a car approaching before we can even see it. Another example is our ability to listen to a person, while other people around us are talking at the same time and to notice that the telephone is ringing in the other room. Acoustic environments are becoming increasingly complex and the auditory system needs to be able to extract useful information and, at the same time, suppress information that is considered unimportant.

Several mechanisms in the auditory system contribute to the ability to localize sounds and to separate between them. Signals presented from right in front of or right behind the listener can be distinguished by the different diffraction patterns introduced by the pinna for these two directions. For sound sources located in the horizontal plane around the listener, each ear receives a slightly different signal, one being delayed and attenuated with respect to the other. The differences between these signals are used to determine the precise location of the source and are referred to as the interaural time difference (ITD) and the interaural level difference (ILD). However, sometimes these spatial cues can be ambiguous. In a room, for example, the ITD's and ILD's of a signal are constantly changing, because the various directions of the reflections lead to different spatial cues. In such a condition, the auditory system analyzes only the direct sound (DS) of the signal, i.e. the sound that arrives first at the listener's ears.

This ability is called the precedence effect (Blauert, 1997; Litovsky *et al.*, 1999). The precedence effect suppresses the direction of the reflections, such that the sound source can still be localized reliably. Reflections in a room are still noticeable and contribute considerably to the coloration of a room. Reflections that reach the listener shortly after the DS (approx. within 50 ms) are furthermore useful for speech intelligibility. In contrast to late reflections (or reverberation), these early reflections (ER's) are fused with the DS (Haas, 1951) and do not cause overlap masking.

Spatial processing of sounds is inevitably linked to binaural listening. ILD's and ITD's cannot be utilized when listening with only one ear. The suppression of a masker noise that is spatially separated from the target signal is facilitated when both ears are available. Even the coloration produced by reflections is less noticeable with two ears.

For the communication in noisy environments it is essential that the auditory system can accurately receive, translate and transmit the spatial cues that are available in the target speech signal and the interfering masker signal. While the auditory system of normal-hearing listeners typically accurately processes these cues and thus provides a clear analysis of the acoustic environment, the impaired auditory system often has problems to utilize them. A high-frequency hearing loss, for example, leads to a decreased advantage from the head shadow effect. The ILD cues introduced by the head shadow cannot be used effectively by the hearing-impaired listener because they are most effective in frequency regions above 1.5 kHz, exactly where their hearing sensitivity is lowest. Binaural processing and thus the ability to separate the speech and the masker is also often reduced in hearing-impaired listeners (Moore, 2007).

Hearing aids are therefore being developed to restore the natural perception of sounds as closely as possible. Signal processing algorithms in hearing aids have mainly focused on the audibility of the sound, such that soft speech is amplified above threshold while loud signals do not become uncomfortably loud. Due to the limited dynamic range in which hearing-impaired listeners are able to perceive sounds, amplitude compression is necessary to achieve this goal. While audibility is important, it does not ensure adequate speech intelligibility in noise. Other hearing aid algorithms, like noise suppression or directional microphones, were introduced to address this problem. However, many hearing-impaired listeners still struggle with understanding speech in complex acoustic environments. Recent technological advances have resulted in the development of algorithms that transfer signals *between* the two hearing aids. With these binaural hearing aids, it is expected to restore spatial hearing and binaural cues that are not available to the hearing-impaired listener with conventional bilateral hearing aids.

Before a binaural hearing aid algorithm can provide a measurable benefit for speech intelligibility, it has to be investigated *if* such a benefit can be expected and how large this potential benefit might be. This thesis has focused on the benefit from ER's and preserved ILD's. All experiments were performed in a virtual auditory environment (VAE) that was specifically developed for the presentation of realistic spatial sound environments (Favrot and Buchholz, 2010).

Chapter 2 investigates the benefit from ER's for normal-hearing and hearing-impaired listeners. The experiments are based on a study by Bradley *et al.* (2003) who found that DS energy and ER energy are equally important for speech intelligibility. Their study is extended here by using realistic ER's with preserved spectral, spatial and temporal characteristics. The ER benefit is then quantified by the introduction of an 'efficiency factor' and it is shown that the efficiency of the ER's, compared to the DS, is reduced due to less energy in frequency regions important for speech intelligibility. Furthermore, the assessment of monaural *and* binaural speech intelligibility provides an indication of the underlying mechanisms involved in ER processing. Both, the influence of the precedence effect and binaural processing abilities are considered.

Chapter 3 is an extension of the experiments described in Chapter 2. Here, it is investigated if a directional and/or temporally fluctuating masker, as opposed to a diffuse masker, alters the processing of ER's and thus the benefit for speech intelligibility. Typically, understanding speech in the presence of an interfering speaker, represents the most challenging listening situation for a hearing-impaired listener. While a normal-hearing listener takes advantage of the temporal gaps in a speech-like masker, this ability is reduced in most hearing-impaired listeners. The benefit from masker fluctuations in monaural and binaural listening conditions is analyzed in normal-hearing and hearing-impaired listeners.

In **Chapter 4**, a hearing aid research platform (HARP) is used to examine the influence of hearing aid signal processing algorithms on speech intelligibility. In particular, binaurally linked hearing aids that preserve ILD's are compared with conventional compression algorithms that alter the spatial cues. Frequency dependent amplification is provided to the hearing-impaired listeners, such that ILD cues become accessible. Speech intelligibility is then measured for two different maskers.

In **Chapter 5**, the HARP is employed to investigate how the ER benefit is affected by hearing aid signal processing. If both ears are required for appropriate ER processing, then a binaurally linked hearing aid that preserves spatial cues might provide an increased benefit from ER's for speech intelligibility compared to conventional bilateral hearing aids.

Finally, **Chapter 6** summarizes the main findings of this thesis, discusses implications for auditory processes of spatial sounds and suggests future research perspectives in the field of spatial hearing.

2

The influence of spectral characteristics of early reflections on speech intelligibility *

Abstract

The auditory system takes advantage of early reflections (ER's) in a room by integrating them with the direct sound (DS) and thereby increasing the effective speech level. In the present paper the benefit from realistic ER's on speech intelligibility in diffuse speech-shaped noise was investigated for normal-hearing and hearing-impaired listeners. Monaural and binaural speech intelligibility tests were performed in a virtual auditory environment where the spectral characteristics of ER's from a simulated room could be preserved. The useful ER energy was derived from the speech intelligibility results and the efficiency of the ER's was determined as the ratio of the useful ER energy to the total ER energy. Even though ER energy contributed to speech intelligibility, DS energy was always more efficient, leading to better speech intelligibility for both groups of listeners. The efficiency loss for the ER's was mainly ascribed to their altered spectrum compared to the DS and to the filtering by the torso, head and pinna. No binaural processing other than a binaural summation effect could be observed.

* This chapter is based on Arweiler and Buchholz (2011).

2.1 Introduction

Early reflections (ER's) of a sound in a given environment are characterized by arriving at the listener's ears shortly (approx. within 50 ms) after the direct sound (DS). They are integrated with the DS in the auditory system, i.e. within a certain time window their energy is added to the energy of the DS. With regards to speech intelligibility the DS and the ER's form the useful part of the speech signal whereas late reflections are considered detrimental for speech intelligibility. Thus, the effective level of a speech signal depends on the energy of the DS and the energy of the ER's at the listener's ears. ER energy increases the effective speech level and has been demonstrated to improve speech intelligibility (Lochner and Burger, 1964; Nábělek and Robinette, 1978; Soulodre *et al.*, 1989; Parizet and Polack, 1992; Bradley *et al.*, 2003). For example, Lochner and Burger (1964) used a single ER to measure (binaural) word intelligibility with a loudspeaker setup. The reflection was a delayed and attenuated copy of the DS. The intelligibility score depended on the reflection delay and the reflection level. When the reflection had the same energy as the DS and arrived within 30 ms after the DS, the energy of the two single sounds was perfectly added, i.e. the ER increased the effective speech level by 3 dB which increased speech intelligibility correspondingly. When the level of the ER was 5 dB below the DS level, there was still a perfect integration of the ER with the DS, but now up to a delay time of 40 ms. The findings of Lochner and Burger (1964) were important to understand the integration of a single reflection with the DS without any background noise. This is, however, not a realistic listening scenario.

The effect of a single reflection in background noise was studied by Parizet and Polack (1992). Here, no full integration of the reflection with the DS was found, i.e. speech intelligibility was always lower than what would have been expected from a perfect integration of the reflection with the DS. Soulodre *et al.* (1989) further extended the experiment of Lochner and Burger (1964) to investigate the combined effect of several ER's and background noise on speech intelligibility. They used 13 ER's arriving within 40 ms after the DS. The signal-to-noise ratio of the DS and the noise was 0 dB. The level of each ER was 5 dB below the DS level. According to Lochner and Burger's (1958) results, the increase in word intelligibility should have corresponded to a 7 dB increase of the DS level, but only an effect corresponding to a 3 dB increase was found by Soulodre *et al.* (1989). They concluded that, in the presence of a background noise, multiple ER's are not fully integrated with the DS and therefore speech intelligibility could not be predicted from a simple addition of the ER energy and the DS energy.

In contrast to Parizet and Polack (1992) and Soulodre *et al.* (1989) the more recent study by Bradley *et al.* (2003) found that increased ER energy had the same effect on speech intelligibility scores as increased DS energy. Bradley et al. presented seven ER's within 50 ms after the DS using an array of eight loudspeakers. One loudspeaker located at 0° azimuth produced the DS and the remaining loudspeakers produced the ER's. Word intelligibility was measured in ambient noise that had a constant level of 47.6 dB(A). Two conditions were considered: In the first condition, the speech signal consisted of the DS only and different signal-to-noise ratios (SNR's) were achieved by varying the level of the DS. In the second condition, the speech signal consisted of the DS plus ER's. Different SNR's were achieved by keeping the DS level constant and varying the level of the ER's. With this setup the benefit of increased DS energy could directly be compared with the benefit obtained by the same increase in ER energy. Bradley *et al.* (2003) did not find a difference in speech intelligibility between the two conditions, suggesting that ER energy was as efficient for speech intelligibility as DS energy. Furthermore, they repeated the experiments with hearing-impaired listeners and the benefit from ER's was found to be similar to that obtained in the normal-hearing listeners. Nábělek and Robinette (1978) also used hearing-impaired listeners to investigate word identification with a single ER as a function of different delay times of the ER. The results were similar to those obtained in normal-hearing listeners.

The discrepancy between the studies mentioned above is not easy to explain, because they used different speech intelligibility tests and different numbers and levels of ER's. In the present study one of the main goals was to investigate and quantify the ER benefit in a listening scenario as close as possible to a real world listening scenario. Bradley *et al.* (2003) used multiple reflections in background noise, but speech intelligibility was tested based on consonant recognition for a fairly narrow range of intelligibility scores (between 80-100% correct intelligibility). A sentence test was chosen here as a more realistic speech signal and a wider range of intelligibility scores was considered in order to expect more general conclusions. Furthermore, none of the previous studies has considered the spectral characteristics of ER's due to wall absorptions in a real room, which are neglected when only using delayed and attenuated copies of the DS. The spectrum of the ER's from the walls, floor and ceiling in a room differs from that of the DS, whereby the spectral filtering depends on the absorptive characteristics of these room boundaries. This filtering should have an influence on speech intelligibility, if frequency regions important for speech intelligibility are affected. Therefore ER patterns were derived from a simulation of a real room and reproduced in a loudspeaker-based virtual auditory environment. In this way the spectral characteristics of the ER's from the simulated room could be preserved. These realistic ER's and the DS were then manipulated in the same way as in the two conditions of Bradley *et al.* (2003).

Another aspect that has not been addressed so far, but which might be particularly relevant for hearing-impaired listeners, is whether the integration of the ER's with the DS is a monaural process or whether binaural processes are involved. Since it has been shown that hearing impairment can reduce binaural processing abilities (Moore, 2007) their benefit from ER's in speech intelligibility tasks might be reduced if ER's were processed binaurally. Thus, a third condition was added in which the ER's were not distributed spatially but presented from the same direction as the DS. Speech intelligibility measurements were then carried out monaurally and binaurally with normal-hearing and hearing-impaired listeners. Bradley *et al.* (2003) used hearing-impaired listeners with a mild hearing loss and most of them were unaware of the impairment before the experiment. Nábělek and Robinette (1978) used a broader range of hearing losses (the average hearing loss at 0.5, 1 and 2 kHz was between 32 and 82 dB HL) but only a single ER. The present study uses moderately and steeply sloping hearing losses and multiple ER's.

The three main goals of the study were the investigation (i) of the usefulness of ER's with a spectrum close to natural ER's compared to the DS, (ii) of the impact of the spatial distribution of the ER's, and (iii) whether the integration of the ER's

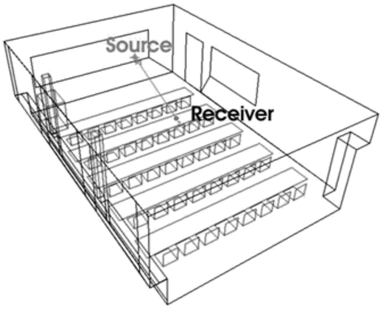


Figure 2.1: Model of the considered classroom with talker (source) and listener (receiver) positions as simulated with Odeon.

with the DS is a monaural or a binaural process and the consequences for hearing impairment.

2.2 Methods

2.2.1 Sound field simulations

The speech intelligibility measurements took place in an acoustically dampened room with 29 loudspeakers arranged symmetrically around the listener (Favrot and Buchholz, 2010). Sixteen of the loudspeakers were placed in the horizontal plane at the height of the listener's head, 7 loudspeakers were located at the ceiling and 6 loudspeakers on the floor. All loudspeakers were equalized to a flat frequency response between 0.2 – 10 kHz measured with an omni-directional microphone in the center of the array. This playback room had a reverberation time of $T_{30} < 100$ ms (for frequencies above 200 Hz) and it's ER's were attenuated by at least 10 dB at the listener's position. The room impulse response (RIR) used to create a realistic sound field was taken from a classroom modeled with the room acoustic software Odeon (version 9.1; Naylor, 1993) as shown in Fig. 2.1. The classroom had a volume of 170 m^3 and the source-receiver distance was 3.5 m. The source was at a height of 1.70 m with an omni-directional directivity and the receiver was at a height of 1 m. The room

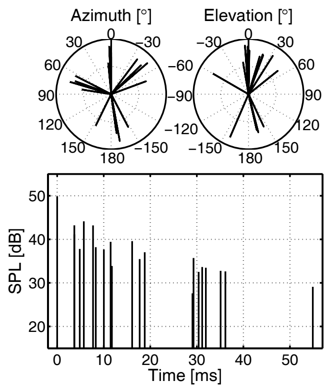


Figure 2.2: Reflectogram for the simulated classroom.

absorption properties were defined in Odeon by choosing the appropriate materials for the walls (smooth unpainted concrete or clinker concrete), ceiling (smooth brick work), floor (rough concrete) and the furniture (e.g. plywood paneling for the chairs). The simulated room was not occupied by an audience. The ER pattern (reflectogram) in this classroom is shown in Fig. 2.2. It illustrates delay times relative to the DS and the spatial distribution of the 20 ER's used in this experiment. From the simulated RIR, all reflections up to the second order were included, the last reflection arriving about 55 ms after the DS. Reflections arriving later than 55 ms after the DS were discarded. The direction of each component in the reflectogram was adjusted to match the position of the closest loudspeaker in the 29-loudspeaker array. The RIR was then processed with the Loudspeaker Room Auralization System (LoRA) toolbox (Favrot and Buchholz, 2010), resulting in a 29-channel RIR, which was then convolved with the speech signal and played back via the loudspeaker array. The listener was seated in the middle of the array at a distance of 1.8 m from the loudspeakers in the horizontal plane.

Speech intelligibility was measured binaurally, monaurally left and monaurally right (cf. Sec. 2.2.5) for three conditions. In the first condition, only the DS of the speech was presented from 0° azimuth (DS_{only}). The speech signal level was varied by changing the level of the DS. In the second condition, the DS of the speech was presented from 0° azimuth together with the spatially distributed ER's from Fig. 2.2 ($DSER_{spatial}$). In the third condition, both the DS of the speech and the ER's were presented from 0° azimuth ($DSER_{frontal}$). In the latter two conditions, the DS level was kept constant and the level of the ER's was varied. In all conditions the speech level was measured with an omni-directional microphone at the location of the center of the listener's head with the listener absent.

2.2.2 Speech material

The Danish sentence test Dantale II (Wagener *et al.*, 2003) was used to measure speech intelligibility. It is based on the Hagerman sentence test (Hagerman, 1982) where each sentence consists of five words with a fixed syntactical structure (name-verb-number-adjective-object). There are 10 alternatives for each sentence element. The listeners responded via a Matlab user interface by choosing the words they had heard from the 50 words presented on a hand-held touch screen.

2.2.3 Background noise

A diffuse stationary speech-shaped noise (SSN) was used as interferer. It was created from the Dantale II sentence material by repeatedly superimposing sentence sequences (Wagener *et al.*, 2003). The SSN was cut into 29 uncorrelated noise signals and each of them was played from one individual loudspeaker simultaneously with all the other loudspeakers. A gated-noise procedure was used, i.e. the noise started 1 s before each sentence with a 0.6 s onset ramp and ended 0.5 s after each sentence with a 0.3 s offset ramp. The interferer was presented at a fixed level of 60 dB SPL for the normal-hearing listeners and at 70 dB SPL for the hearing-impaired listeners. A higher level was chosen for the hearing-impaired listeners to ensure that as much of the spectrum of the noise as possible was audible without using an uncomfortably loud presentation level. The levels were measured with an omni-directional microphone at the location of the center of the listener's head with the listener absent.

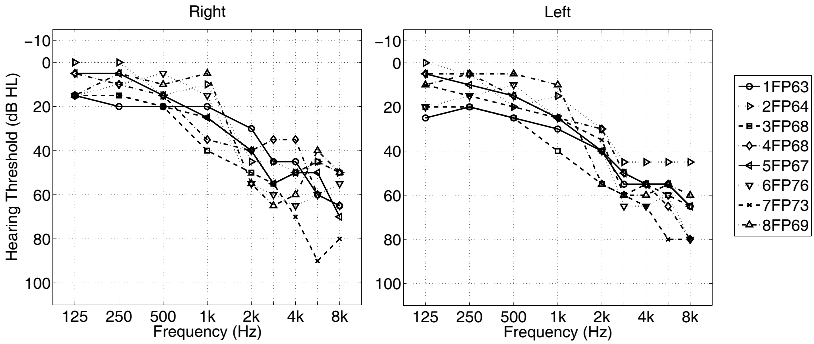


Figure 2.3: Hearing thresholds for 8 hearing-impaired listeners. The last two numbers indicate the age of the test subject at the time of the listening experiment.

2.2.4 Listeners

Nine normal-hearing and eight hearing-impaired listeners participated in the experiment. The thresholds of the normal-hearing listeners were 20 dB HL (ISO, 2004) or better for both ears at octave frequencies between 0.25 and 6 kHz. The normal-hearing listeners' age was between 23 and 46 years with a median age of 24 years. The thresholds for the hearing-impaired listeners are shown in Fig. 2.3. All hearing-impaired listeners had normal or close-to-normal thresholds in the low frequency range up to 1 kHz and a moderately or steeply sloping high-frequency hearing loss. All hearing losses were of sensorineural origin and symmetric (interaural threshold differences, averaged across frequency from 0.25 to 8 kHz, were $<$ 9 dB for all listeners). The hearing-impaired listeners were 63 to 76 years old with a median age of 68 years. The experiments were split in two sessions, each lasting about 1.5-2 hours per person. Before each session, the listeners performed a training with 30 sentences. Listeners were paid on an hourly basis for their participation.

2.2.5 Experimental procedure

In the first part of the experiment, the speech reception threshold (SRT) of each listener was determined with 20 sentences using the adjusted RIR with spatial ER's. The SRT reflects the SNR at 50% speech intelligibility. The overall level of the speech signal, i.e. the level of the DS plus ER's, was varied adaptively with a maximum likelihood procedure (Brand and Kollmeier, 2002) depending on the number of words understood correctly. Afterwards, in the main experiment, all speech intelligibility scores were measured relative to each individual listener's SRT, i.e., the sensitivity of each listener was normalized. This was done in order to facilitate the comparison between the results for the normal-hearing and those for the hearing-impaired listeners who might need very different SNR's to achieve the same speech intelligibility. At the SRT, the contribution of spatial ER's to the overall speech level was 6 dB. The reference point for all conditions was thus set to the overall speech level at the SRT minus 6 dB (i.e. a "dry" condition with no ER's). From this reference point, the SNR was increased stepwise by either adding DS energy or ER energy. For each SNR, the speech intelligibility was measured with 10 sentences per person. For the monaural speech intelligibility measurements, an ER2 earphone (Etymotic Research) was inserted in one ear at a time which provided a minimum of 30 dB sound attenuation between 0.125 and 8 kHz. In addition, white noise was presented through the earphone at a level of 75 dB SPL. In this way, the non-test ear was completely masked and no headphones disturbed the sound field at the test ear. Binaural speech intelligibility was measured in a first session followed by a second, monaural session where the listeners either started with the left or the right ear. The insert earphone was not removed until the monaural measurement was finished. The conditions and the order of measurements within each condition were randomized. Before the actual measurement, the listeners were instructed to look at the front loudspeaker and to hold their head upright during sentence presentation.

2.3 Results

2.3.1 Speech intelligibility data

The speech intelligibility scores averaged across the nine normal-hearing listeners are shown on the left side of Fig. 2.4 and those averaged across the eight hearing-impaired listeners are shown on the right side (hearing-impaired listeners were only tested with the left ear, see below). Error bars indicate ± 1 standard deviation. The upper panel shows the binaural speech intelligibility scores, the middle panel indicates the monaural scores obtained with the left ear only and the lower panel shows the scores obtained with the right ear only. The results for the DS only condition are indicated by circles. The squares represent the condition with the DS and spatial ER's and the diamonds show the condition with the DS and frontal ER's. In each condition speech intelligibility increased with increasing SNR and ranged from about 20% to almost 100% intelligibility. A logistic function $p(\Delta SNR)$, adapted from the logistic function used by Wagener *et al.* (2003), was fit to the data given by:

$$p(\Delta SNR) = \frac{1-\alpha}{1+e^{(4\cdot s_{55}\cdot(SRT_{55}-\Delta SNR-SNR_0))}} + \alpha \tag{2.1}$$

where SRT_{55} is the SNR in dB at 55% correct speech intelligibility, s_{55} is the slope at SRT_{55} and α is the chance level of 10% ($\alpha = 0.1$). The SRT_{55} was used instead of SRT_{50} because due to the chance level the dynamic range of the psychometric function was reduced to 90%. ΔSNR represents the SNR increase in dB when adding DS energy or the effective SNR increase when adding ER energy (cf. Sec. 2.3.2). SNR_0 is the SNR in dB for the DS_{only} condition at the reference point (cf. Sec. 2.2.5). The function was fit using the efficiency factor concept, which is explained in detail in Sec. 2.3.2.

The average SRT, measured in the first part of the experiment, was -12.9 dB SNR for the normal-hearing listeners and -7.71 dB SNR for the hearing-impaired listeners with a standard deviation of 0.76 dB and 1.85 dB respectively. An analysis of variance (ANOVA) was performed on the speech intelligibility scores in the conditions with similar SNR's. For the normal-hearing listeners, there was a significant main

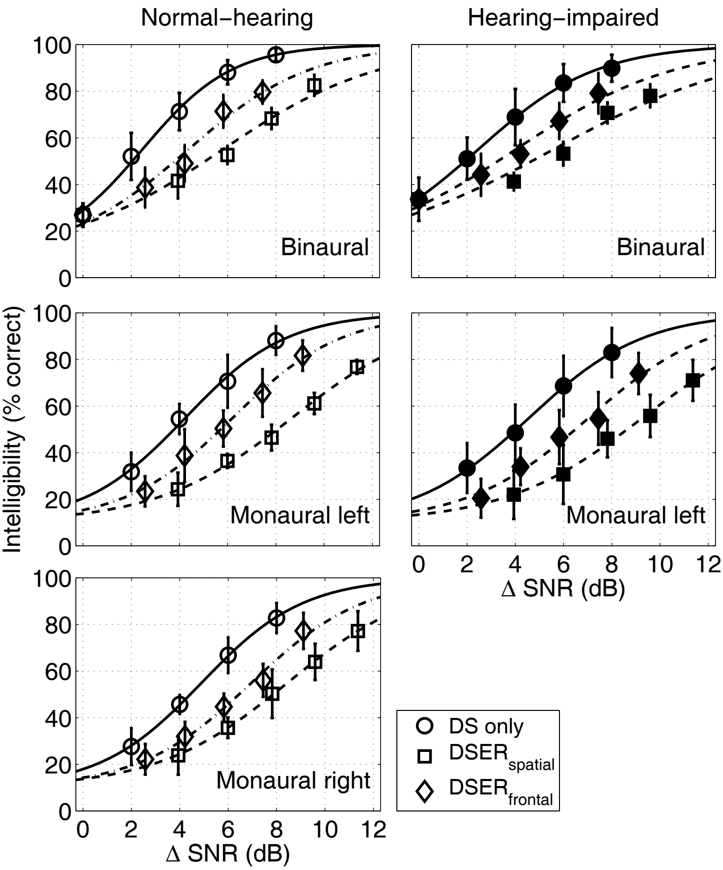


Figure 2.4: Mean speech intelligibility scores and fitted psychometric functions for binaural listening (upper row), listening with the left ear only (middle row) and listening with the right ear only (lower row). Open symbols: Normal-hearing listeners. Closed symbols: Hearing-impaired listeners. Error bars indicate ± 1 standard deviation.

effect of listening mode (binaural, left or right ear, $p < 0.01$) and condition (DS_{only}, $DSER_{spatial}$, $DSER_{frontal}$, $p < 0.01$), but no interaction effect (p = 0.77). Subsequent multiple comparisons revealed no significant difference between the speech intelligibility results for the left and the right ear, therefore the hearing-impaired listeners were only tested with the left ear. An ANOVA for the hearing-impaired listeners likewise revealed a significant main effect of listening mode and condition ($p < 0.01$), but no interaction effect (p = 0.41). For both groups of listeners, binaural speech intelligibility was significantly better than monaural speech intelligibility ($p < 0.01$). Furthermore, speech intelligibility was significantly better when the SNR was increased by adding DS energy than when it was increased by adding ER energy ($p < 0.01$) and when the ER's were presented from the front loudspeaker than when they were presented spatially distributed ($p < 0.01$). Added ER energy still improved speech intelligibility, but less efficiently than added DS energy.

2.3.2 Analysis

Efficiency of ER's for speech intelligibility

In order to quantify the benefit from ER's an efficiency factor k_{ER} was introduced:

$$k_{ER} = \frac{\Delta E^*_{ER}}{\Delta E_{ER}} \tag{2.2}$$

where ΔE_{ER} represents the total ER energy, i.e. the physical increase of the overall speech energy due to the addition of ER's and where ΔE^*_{ER} is the corresponding effective energy increase, i.e. the ER energy that is useful for speech intelligibility. The efficiency factor for the DS energy was arbitrarily set to $k_{ER} = 1$. Although it is expected that with most natural ER's an efficiency factor within $0 < k_{ER} < 1$ will occur, an efficiency factor larger than $k_{ER} = 1$ could theoretically appear when very strong room modes are present. Applying the concept of the efficiency factor k_{ER}, the total effective energy E* of the considered speech signal can be described by:

$$E^* = E_0 + k_{ER} \cdot \Delta E_{ER} \tag{2.3}$$

where E_0 is the energy of the DS alone (at the reference point, cf. Sec. 2.2.5). The physical energy increase ΔE_{ER} can be calculated from the difference of the total speech energy E (i.e. including DS and ER's) and the energy of the DS alone, i.e. $\Delta E_{ER} = E - E_0$. The effective SNR increase is then given by:

$$\Delta SNR^* = 10 \cdot log_{10}(E^*/E_0) \tag{2.4}$$

Inserting Eq. 2.3 into Eq. 2.4 results into:

$$\Delta SNR^* = 10 \cdot log_{10}(1 - k_{ER} + k_{ER} \cdot 10^{\Delta SNR/10}) \tag{2.5}$$

with $\Delta SNR = 10 \cdot log_{10}(E/E_0)$. For the case that the efficiency factor is $k_{ER} = 1$, Eq. 2.5 simplifies to $\Delta SNR^* = \Delta SNR$. The efficiency factor k_{ER} for the different ER conditions described in Sec. 3.2.1 can be derived by applying the logistic function given in Eq. 2.1 and replacing the ΔSNR by the effective ΔSNR^* (i.e., inserting Eq. 2.5 into Eq. 2.1). The modified logistic function is then first fitted to the DS alone condition (i.e., $k_{ER} = 1$) to derive the parameters SRT_{55} and s_{55}. These parameters are then used for the $DSER_{spatial}$ and $DSER_{frontal}$ conditions to derive the efficiency factor k_{ER}, which is varied to minimize the root-mean-square error between the modified logistic function and the measured data. Figure 2.5 shows the estimated k_{ER} values for the different listening conditions. Open symbols represent the normal-hearing listeners and closed symbols represent the hearing-impaired listeners. For the DS alone the efficiency is equal to one (i.e. $k_{ER} = 1$). For both groups of listeners, the efficiency of the ER's was significantly higher ($p < 0.01$) when the reflections were presented from 0° azimuth than when they were presented spatially distributed. For the frontal ER's, about half of the energy of the ER's was utilized for speech intelligibility. For the spatial ER's, only 25-34% of the reflection energy was utilized. The efficiency was similar for binaural and monaural listening in both groups.

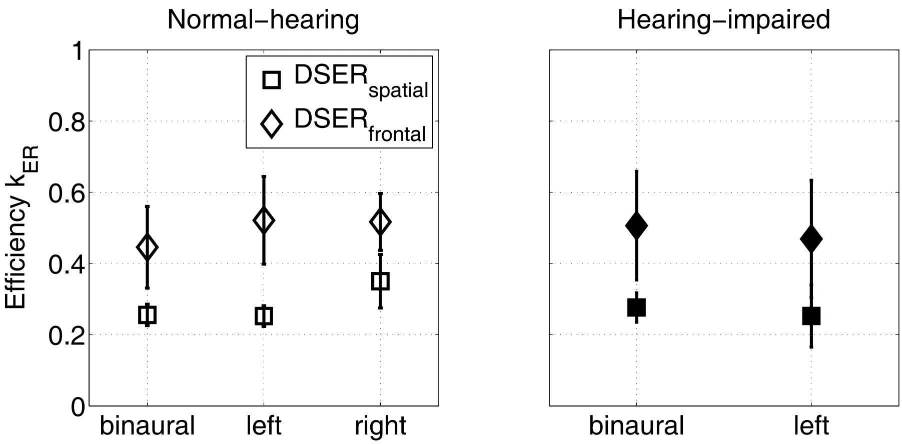


Figure 2.5: Efficiency factor k_{ER} for the different listening conditions. Open symbols: Normal-hearing listeners. Closed symbols: Hearing-impaired listeners (only the left ear was tested). Error bars indicate ± 1 standard deviation.

Binaural benefit

The binaural benefit was calculated as the difference between the binaural SNR and the monaural SNR for the left and right ear at 55% correct speech intelligibility, estimated from the fitted logistic function for each listener. Figure 2.6 shows the binaural benefit for the different listening conditions. Open symbols represent the results for the normal-hearing listeners and closed symbols represent the results for the hearing-impaired listeners. Binaural listening provided an advantage of 1.8 to 3.4 dB relative to monaural listening. Paired t-tests with Bonferroni correction were performed to test the differences in binaural benefit between the different conditions. For the normal-hearing listeners the benefit from binaural listening over right-ear-listening was similar for all conditions. For both groups the benefit from binaural listening over left-ear-listening was slightly but significantly higher for the spatial ER's than for the DS_{only} condition ($p < 0.01$). Furthermore, there was no significant difference in binaural benefit between the $DSER_{spatial}$ and the $DSER_{frontal}$ condition. For the hearing-impaired listeners the benefit from binaural listening over

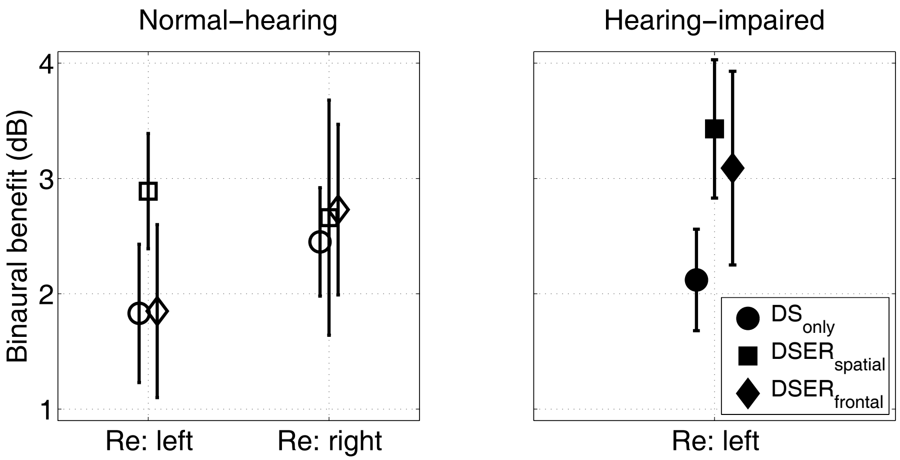


Figure 2.6: Binaural benefit expressed as the difference in SNR at 55% intelligibility between monaural (left or right ear) and binaural listening. Open symbols: Normal-hearing listeners. Closed symbols: Hearing-impaired listeners (only the left ear was tested). Error bars indicate ± 1 standard deviation.

left-ear-listening was slightly but not significantly higher than for the normal-hearing listeners.

2.4 Discussion

2.4.1 General discussion

The results in Sec. 2.3 have shown that ER energy does not improve speech intelligibility to the same extent as the same amount of DS energy in the considered simulated room. Therefore speech intelligibility can not be predicted from the total energy of the ER's, because ER energy is less efficient than DS energy. This observation is in agreement with Parizet and Polack (1992) and Soulodre *et al.* (1989) who also showed a decreased benefit from ER's, even though they used delayed and attenuated copies of the DS as ER's. The latter found a difference in speech intelligibility between the DS only condition and the added ER's condition

corresponding to a 4 dB decrease in speech level (from 7 dB to 3 dB, cf. Sec. 2.1). This is very similar to our results in Fig. 2.4 where approximately 4 dB difference can be found between the DS only and the $DSER_{spatial}$ conditions. However, Bradley *et al.* (2003) did not find this difference in speech intelligibility. They also used spectrally unfiltered ER's, so that the influence of the spectral characteristics due to wall absorptions on speech intelligibility was neglected. Furthermore, Bradley et al. used a speech intelligibility test (Fairbank's rhyme test modified by Latham; Latham, 1979) that had a very shallow slope in the dynamic range in which they measured speech intelligibility (i.e. 80-100% intelligibility). For normal-hearing listeners, an SNR increase of 10 dB was necessary to increase the speech intelligibility from 87% to 99%. This limited dynamic range might not have been sufficient to show the differences between the DS_{only} and the $DSER_{spatial}$ condition.
The discrepancies between the studies cannot be fully explained because the spectrum of the employed sound signals at the listener's location would have to be considered. However, different interference patterns of the reflections could have boosted or reduced certain frequency regions of the speech signals which in turn could have increased or decreased speech intelligibility accordingly in the different studies. For all studies, however, the benefit from ER's might have been (further) reduced when taking room absorption into account. The importance of this spectral filtering will be discussed in the next section.

In the present study the total loss of ER efficiency compared to the DS was about 50% for the frontal ER's and up to 75% for the spatial ER's, calculated by subtracting the efficiency for the ER's in Fig. 2.5 from the efficiency for the DS (i.e. $k_{ER} = 1$). In the following this efficiency loss is analyzed by looking at the combined spectrum of the ER's and the DS at the listener's location which is influenced by two main room-acoustic related effects: The filtering of the ER's due to wall absorptions in the simulated room and the comb-filtering due to the interaction of the ER's and the DS. Moreover, the direction-dependent filtering by the torso, head and pinna will be investigated.

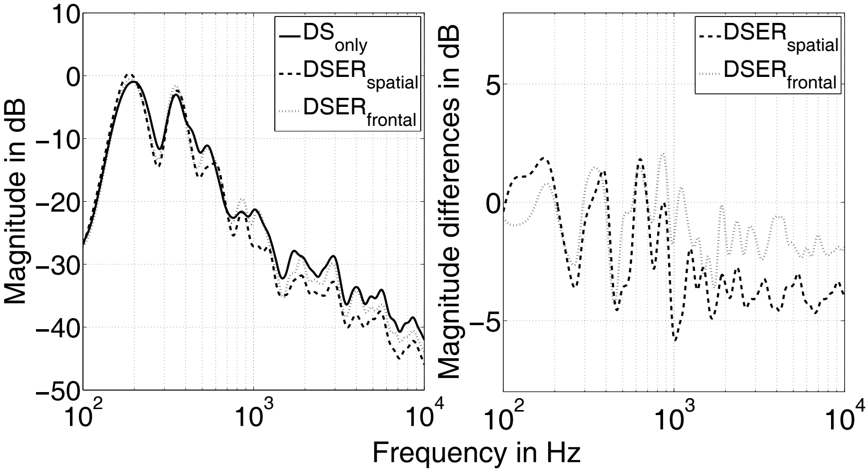


Figure 2.7: Left panel: Magnitude spectra of the DS_{only}, the $DSER_{spatial}$ and the $DSER_{frontal}$ condition. Right panel: Differences in magnitude spectrum of the $DSER_{spatial}$ and the $DSER_{frontal}$ condition relative to the DS_{only} condition (0 dB line).

2.4.2 Spectral characteristics

In the left panel of Fig. 2.7 the magnitude spectra of the speech signals for the DS_{only}, the $DSER_{spatial}$ and the $DSER_{frontal}$ condition are shown. They were measured with an omni-directional microphone (Brüel & Kjær 4943) at the listener's position and adjusted to an equal sound pressure level with the ER's contributing 6 dB to the overall speech signal. The differences between these power spectra and the spectrum corresponding to the DS_{only} condition were calculated. They are shown in the right panel of Fig. 2.7. The differences for the $DSER_{spatial}$ and the $DSER_{frontal}$ condition are shown relative to the DS_{only} condition (0 dB line). These differences between the DS_{only} condition and the conditions with ER's are influenced by the absorptions from the walls in the simulated classroom and the interference pattern (comb-filtering) of the ER's and the DS at the measurement microphone. Differences between the three conditions are furthermore introduced by the characteristics of the playback environment (weak reflections in the acoustically dampened listening room, slight differences between individual loudspeakers). The influence of the playback room can be deduced by comparing the spectra for the $DSER_{spatial}$ and the $DSER_{frontal}$ condition which, in an ideal playback room, should be identical. Figure 2.7 (right panel) shows that frequency regions important for speech intelligibility (1-4 kHz; Pavlovic, 1987) are attenuated more for the conditions with ER's compared to the DS. Due to the bandpass characteristics of speech, the total signal power is mainly determined by frequencies around 100-500 Hz (left panel Fig. 2.7). Hence the critical differences for speech intelligibility in the frequency region 1-4 kHz do not contribute much to the overall signal power. In order to investigate if these differences in spectra could have led to the difference in speech intelligibility between the conditions, the intelligibility-weighted SNR (Greenberg *et al.*, 1993) was applied to the measured data. This intelligibility-weighted SNR gives a higher weight to frequency bands important for speech intelligibility. Compared to the broadband SNR the intelligibility-weighted SNR provides the effective SNR which is a more meaningful measure with regards to speech intelligibility. The speech and noise signals were first split into 1/3 octave bands. The SNR was then calculated in each band and, before summing, each band was weighted according to its contribution to speech intelligibility. The band importance function used was taken from table 3 of the Speech Intelligibility Index (SII) standard (ANSI, 1997) and referred to average speech. The intelligibility-weighted efficiency factor was then calculated by replacing ΔSNR in Eq. 2.1 by the intelligibility-weighted ΔSNR^{*}_{iw}. The derived intelligibility-weighted efficiency factor is shown in Fig. 2.8, represented by the squares and diamonds for the spatial and frontal ER's, respectively. When taking the reduced energy in frequency regions important for speech intelligibility into account the efficiency of the spatial ER's increased significantly to about 64$\%$ (compare Fig. 2.5 and 2.8) for the binaural and monaural left condition. For the frontal ER's the efficiency increased to about 62$\%$, so that the large difference of efficiency between the spectral and frontal ER's for unweighted speech almost vanished, except for the right ear which is further discussed in Sec. 2.4.3. An increased efficiency in turn means a decreased efficiency loss. The efficiencly loss decreased by 40-50$\%$ for spatial ER's and by about 10-15$\%$ for frontal ER's. Thus, for the spatial ER's, a large part of the efficiency loss could be ascribed to the spectrum of the ER's at the listener's position. This shows the necessity to consider the realistic spectrum of the ER's when

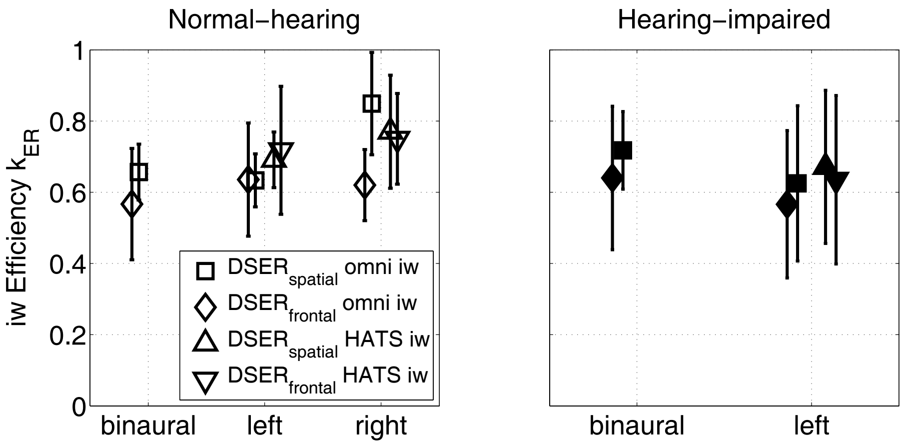


Figure 2.8: Efficiency factor k_{ER} for the different listening conditions using an intelligibility-weighted SNR. Open symbols: Normal-hearing listeners. Closed symbols: Hearing-impaired listeners (only the left ear was tested). Error bars indicate ± 1 standard deviation. Squares and diamonds: speech levels were measured with an omni-directional microphone. Triangles: speech levels were measured with HATS (head and torso simulator).

predicting their influence on speech intelligibility. If the ER's are copies of the DS, changes in the spectrum due to wall absorptions are not considered and this might overestimate the benefit from ER's. However, the considered spectral characteristics alone can not explain the total efficiency loss of the ER's. Therefore the influence of the listener's torso, head and pinna on the spectrum is discussed next.

2.4.3 Filtering by the torso, head and pinna

When considering total SNR's measured with an omni-directional microphone, speech intelligibility was better when the ER's were presented from the front than when they were spatially distributed (cf. Sec. 2.3.1). However, when considering the signals at the listener's ears, the energy contribution of the ER's depends on the direction-dependent filtering of the torso, head and pinna. This leads to changes in the spectrum of the frontal and spatial ER's. By applying the intelligibility-weighted SNR in Sec. 2.4.2, the efficiency difference between spatial and frontal ER's could be fully explained by the spectral characteristics of the ER's except for the right ear. However, this analysis did not consider the influence of the torso, head and pinna on the spectrum of the ER's. This is examined here by measuring the speech and noise levels with a Brüel & Kjær 4100 head and torso simulator (HATS) and applying the intelligibility-weighted SNR from Sec. 2.4.2. The resulting efficiency factors are shown in Fig. 2.8 by the triangles, which for both monaural conditions are about 70-75% (binaural results were not derived from the HATS measurements). When considering the signals at the listener's (HATS') ears, the spatial and frontal ER's provide equal efficiency, now also for the right ear. Hence, the difference between spatial and frontal ER's observed in Sec. 2.4.1 and 2.4.2 might be simply explained by the direction-dependent spectral changes introduced by torso, head and pinna. Compared to the efficiency factors for the omni-directional microphone in Fig. 2.8 the HATS measurements show an increase in efficiency of approximately 10%. This indicates that another 10% of the efficiency loss observed in Sec. 2.4.1 can be explained by considering the influence of the torso, head, and pinna.

The only study that has explicitly investigated the influence of the direction of an ER on speech intelligibility was a study by Nakajima and Ando (1991). They found that, when the incident angle of a single reflection was the same as for the DS ($0°$ azimuth), binaural speech intelligibility was worse than when the incident angle of the reflection was moved to $30°$, $60°$ or $90°$ azimuth. This is in contrast with our findings from Sec. 2.3.1 where frontal ER's yielded better speech intelligibility than spatial ER's. However, in their study the improved speech intelligibility for ER's from the side might have simply resulted from the direction dependent filtering of the torso, head and pinna, which leads to an increased sound pressure level for lateral ER's, particularly at angles between $45°$ and $90°$ (Goode, 2001). Using only one ER, this effect might be much more pronounced than with multiple ER's from different directions.

2.4.4 Binaural processing

Sounds originating from different directions introduce different interaural level and time differences, which can potentially be utilized by the binaural auditory system. In consequence, a binaural intelligibility benefit can often be observed for spatially separated speakers and noise sources. In the present study the condition with frontal ER's did not provide any binaural cues as all speech signal components (DS and ER's) were only presented from the front and the background noise was diffuse. Hence, solely the energy summation of the signals at the two ears, also referred to as binaural summation or binaural redundancy (Dillon, 2001; Ricketts *et al.*, 2006), could potentially provide an advantage of binaural listening over monaural listening (binaural benefit). In the $DSER_{spatial}$ condition the ER's were presented from different locations. Here, according to the precedence effect (Blauert, 1997; Litovsky *et al.*, 1999), ER's might be suppressed by the DS and thus binaural speech intelligibility could be decreased compared to the $DSER_{frontal}$ condition. The masking noise presented to one ear via insert earphones in the monaural condition might have increased the binaural benefit if central masking had occurred. Due to the relatively high signal levels on the other ear, however, it is rather unlikely that central masking could have played a role.

The binaural benefit was calculated in Sec. 2.3.2 for the different conditions and is shown in Fig. 2.6. The binaural benefit is similar for the $DSER_{spatial}$ and the $DSER_{frontal}$ condition, which indicates that the binaural advantage was only due to a binaural summation effect and that the precedence effect did not play a role for speech intelligibility. This is in agreement with, for instance, Freyman *et al.* (1998) who provided evidence in an intensity discrimination task that the precedence effect does not generally suppress reflections and that suppression might rather be limited to localization cues. Additionally the diffuse background noise might have reduced the strength of the precedence effect (Chiang and Freyman, 1998).

In Sec. 2.3.2, a slight but significant difference of about 1 dB was found between the $DSER_{spatial}$ and the DS_{only} condition for the benefit of binaural listening over left-ear listening but not over right-ear listening. This difference might be attributable to slight differences in speech intelligibility between the left and right ear at 55%

intelligibility, i.e. the point at which the binaural benefit was determined. However, the observed total binaural benefit of 1.8-3.4 dB is in good agreement with previously found binaural summation effects (Pollack, 1948; Dillon, 2001; Moore, 2007). This might suggest that at first a monaural integration of the ER's with the DS takes place at an early stage of the auditory system and that the combined signal is then processed binaurally. Due to the diffuse background noise used in the present study the binaural processing is reduced to a binaural summation effect. It needs to be investigated if different listening scenarios or a directional interferer would provide additional binaural cues that could lead to a larger binaural benefit. Moreover, pitch cues or listening-in-the-dips effects provided by speech-like interferers have to be further considered.

2.4.5 Additional factors

The influence of individual reflection delays have been investigated in several studies (Lochner and Burger, 1964; Nábělek and Robinette, 1978; Soulodre *et al.*, 1989) and were therefore not directly addressed in the present study. In general, those studies found that the contribution of an ER to speech intelligibility is larger the shorter its delay time after the DS. In the present study, only about 75% of the total efficiency loss of the ER's could be explained by their spectral characteristics at the listener's ears. A part of the missing 25% might therefore be ascribed to the time delay of the ER's compared to the DS. From Sec. 2.4.3 it can be deduced that the loss due to the temporal delays of the ER's is independent of the direction of the ER's. Although no conclusions can be made from this result about the time window in which the ER's are still integrated with the DS, it is clear from the above studies that a shorter time delay between the ER's and the DS would increase the efficiency of the ER's. In addition, other factors may have also contributed to the unaccounted efficiency loss. These may include differences between individual listeners and the HATS, the positioning of the head during the experiment or the accuracy of the applied speech-weighted SNR.

2.4.6 Hearing-impaired listeners

Apart from a slightly shallower slope, the fitted speech intelligibility curves for the hearing-impaired listeners (Fig. 2.4) were very similar to those obtained for the normal-hearing listeners in all conditions. A shallower slope is expected for hearing-impaired listeners, because the hearing loss modifies the intelligibility of the individual words of the sentence test Dantale II. This would imply different level adjustments of the words according to each individual's hearing loss in order to produce a slope as steep as for the normal-hearing listeners (Wagener and Brand, 2005). On average the SRT for the hearing-impaired listeners was 5.2 dB higher than the SRT for the normal-hearing listeners. However, when the sensitivity of each hearing-impaired listener was normalized (cf. Sec. 2.2.5), speech intelligibility resulted in the same scores as for the normal-hearing listeners. Hence, it can be concluded that hearing-impaired listeners on average have the same ability as normal-hearing listeners to integrate ER's with the DS. This is in agreement with the results of Bradley *et al.* (2003) and Nábělek and Robinette (1978) (cf. Sec. 2.1). The present study confirmed their findings for moderately and steeply sloping hearing losses combined with a realistic listening scenario with several ER's.

The hearing-impaired listeners showed a similar binaural benefit as the normal-hearing listeners. For hearing-impaired listeners an SNR improvement of 2-3 dB when listening binaurally is a significant advantage in noisy listening conditions. Hence, a bilateral hearing aid fitting seems reasonable whenever possible.

Hearing-impaired listeners' ability to understand speech is often vulnerable to late reflections. Bradley *et al.* (2003) showed that the influence of late reflections (in a room intended for speech communication) is small when the diffuse background noise is relatively more detrimental. This influence, however, depends on the background noise level and the absorption characteristics of the room and might thus change depending on the listening condition.

It is not clear how certain hearing aid algorithms like compression or directional microphones would influence ER processing. Knowing that the integration of the ER's with the DS is most likely a monaural process (cf. Sec. 2.4.4) the preservation of interaural time and level differences does not seem critical. A directional microphone on the other hand is most sensitive for sounds from the front and is less sensitive for ER's from the side or back. Hence, the positive effect of suppressing directional noises is partly reduced, because the positive contribution of spatially distributed ER's to speech intelligibility might vanish as their energy is decreased by the directional microphone.

2.5 Summary and Conclusion

This study has investigated the influence of a realistic set of ER's on speech intelligibility. In general, ER energy improved speech intelligibility, but not to the same extent as the DS. An efficiency factor k_{ER} was introduced to quantify the benefit from ER's and it could be demonstrated that a large part of the ER energy can not be utilized for speech intelligibility. The major part of this loss in efficiency was explained by the spectral characteristics of the ER's at the listener's ears and the remaining part was attributed to an auditory-internal integration window and possibly other factors. The intelligibility-weighted SNR (Greenberg *et al.*, 1993) was proposed to investigate the contribution of the spectral characteristics of ER's to the efficiency loss. It showed that the altered spectrum of (realistic) ER's compared to the DS could explain a major part of the efficiency loss.

The efficiency of the ER's should be included in models predicting speech intelligibility from the ER's energy. Only a part of the energy arriving shortly after the DS should be considered useful for speech intelligibility.

The influence of the direction of the ER's was investigated by comparing spatially distributed ER's with ER's presented from the front. Frontal ER's led to better speech intelligibility than spatial ER's. However, when using a head and torso simulator (HATS) combined with the intelligibility-weighted SNR, the difference between the two conditions vanished. Thus, the direction dependent speech intelligibility could be explained by the filtering of the torso, head and pinna.

No binaural processing of ER's other than a summation of the signals at the two ears was found. Thus it was suggested that the integration of the ER's with the DS takes place at an early monaural stage of the auditory system and that the combined signal is then processed binaurally.

Another remarkable result of this study was that hearing-impaired listeners could benefit from ER's in the same way as normal-hearing listeners, even though they had a moderately or steeply sloping hearing loss. This might be due to the observation that the integration of the ER's with the DS was found to be a monaural process, which is an advantage for hearing-impaired people with reduced binaural processing abilities.

From a room acoustic perspective the results suggest that the walls in a room should provide as little absorption as possible to ER's at frequency regions between 1-4 kHz to increase the efficiency of ER's in aiding speech intelligibility. However, this might be difficult to achieve without increasing the level of the late reflections which can be detrimental to speech intelligibility.

The benefit from ER's in this study was determined for one specific room and one specific source-receiver setup. The classroom and the source-receiver position were chosen to represent a realistic listening condition. More acoustic scenarios need to be investigated in order to generalize the present results and to resolve discrepancies between studies. However, the approach and tools used in this study could be very helpful when designing future experiments.

2.6 Acknowledgments

The authors wish to thank Torsten Dau and Sylvain Favrot for valuable comments and discussions. This research was supported by the Danish Graduate School SNAK ("Sense organs, neural networks, behavior, and communication") and Phonak AG. Parts of the work were presented at the International Symposium on Auditory and Audiological Research (ISAAR) in Helsingør, Denmark, August 26-28, 2009 and at the 36th annual convention of the German Acoustical Society (DAGA) in Berlin, Germany, March 16-19, 2010.

3

The influence of masker type on early reflection benefit and speech intelligibility *

Abstract

In a study by Arweiler and Buchholz (2011) it was recently shown that the energy of early reflections (ER's) in a room provides a benefit for speech intelligibility but cannot be considered as useful as the energy of the direct sound (DS). For the integration of the ER's with the DS, monaural auditory processing was found to be sufficient and, apart from a binaural energy summation, binaural listening did not provide an additional benefit from ER's. Arweiler and Buchholz (2011) used a diffuse speech shaped noise (SSN) in their speech intelligibility experiments, which does not provide distinct binaural cues to the auditory system. In the present study, the monaural and binaural benefit from ER's for speech intelligibility was determined in a virtual auditory environment with a diffuse SSN and three directional maskers presented from 90° azimuth: a SSN, a multi-talker babble and a reversed two-talker masker. Eight normal-hearing and eight hearing-impaired listeners participated in the experiment. For both groups and all maskers, an increased benefit from ER's was found compared to Arweiler and Buchholz (2011) which was ascribed to the different methods used to measure speech intelligibility. The directional and/or fluctuating maskers did not lead to an increased benefit from ER's compared to the diffuse SSN. Further aspects of monaural and binaural speech intelligibility for the different masker types are discussed.

* This chapter is based on Arweiler *et al.* (2011b).

3.1 Introduction

When listening to speech in a room, the sound reaching the listener's ears consists of the direct sound (DS), early reflections (ER's) and late reflections (or reverberation). ER's are usually defined to arrive at the listener within approx. 50 ms after the DS and are considered useful for speech intelligibility (Bradley *et al.*, 2003; Nábělek and Robinette, 1978). Most studies showed the benefit from ER's for a single ER with or without background noise (e.g. Lochner and Burger, 1964; Parizet and Polack, 1992) or for multiple ER's that were delayed and attenuated copies of the DS (Soulodre *et al.*, 1989; Bradley *et al.*, 2003). In such a scenario, the altered spectrum of the reflections due to wall absorption is neglected. Arweiler and Buchholz (2011) recently showed that wall absorption reduced the energy of the ER's in frequency regions important for speech intelligibility (1-4 kHz). They used a realistic ER pattern that was derived from a simulated classroom and reproduced in a virtual auditory environment with 29 loudspeakers. The altered spectrum of the ER's led to a decreased speech intelligibility of a speech signal that consisted of the DS and ER's compared to a speech signal that only consisted of the DS. Hence, the ER's were termed to be less efficient for speech intelligibility than the DS. Furthermore, no explicit binaural processing apart from a binaural summation effect was found, suggesting that the integration of the ER's with the DS reflects a monaural process. Arweiler and Buchholz (2011) argued that the diffuse noise used for the speech intelligibility measurements might have limited the binaural processing abilities, because the uncorrelated noise signals at the two ears did not provide reliable cues for the binaural auditory system. The purpose of the present study was therefore to investigate the processing of ER's for three directional maskers with and without temporal fluctuations.

The binaural advantage resulting from a masker that is spatially separated from the target speech is a summation of the advantage due to head shadow effects and the binaural benefit, which utilizes interaural time and level differences between the speech and the masker. These cues are altered when spatially distributed ER's are added to the DS and, thus, the binaural benefit might change. The binaural benefit might also be affected by temporal fluctuations in a masker. Grose and Hall (1998) found that the binaural system can utilize the relatively greater influence of the speech signal on the interaural differences during the masker dips. Based on their findings, the binaural benefit might be reduced when ER's are added to the DS because the masker dips might be partially 'filled' by the ER's. Grose and Hall (1998) used tone bursts and a detection task in their experiment. Other studies directly investigated speech intelligibility and measured the binaural benefit for speech in a fluctuating masker (e.g. Dirks and Wilson, 1969; Nábělek and Picket, 1974). However, they did not find an increased binaural benefit compared to a stationary masker.

In the present study, speech intelligibility was measured in a first condition where only the DS of the speech signal was presented over loudspeakers in an acoustically dampened room. In a second condition, spatially distributed ER's replaced part of the DS energy in the speech signal. The difference in speech intelligibility between these two conditions is a measure of the benefit from ER's. Furthermore, the difference in the binaural benefit between these two conditions reveals the effect of binaural processing as a function of the different maskers considered in this study.

The benefit from masker fluctuations (FMB) compared to a stationary masker for normal-hearing and hearing-impaired listeners has been discussed in several studies (Festen and Plomp, 1990; Peters *et al.*, 2004; Summers and Mollis, 2004; George *et al.*, 2006; Larsby *et al.*, 2008; Bernstein and Grant, 2009; Ihlefeld *et al.*, 2010). All of these studies have focused on the monaural FMB and, except for the study by Larsby *et al.* (2008), all used headphones to determine the FMB. The binaural system might be able to utilize binaural cues in the dips of the masker which might result in a larger FMB when listening binaurally compared to monaurally. In order to examine this, speech intelligibility was measured monaurally and binaurally for stationary and fluctuating maskers in a loudspeaker environment. Both normal-hearing and hearing-impaired listeners participated in the experiment. The latter group has often been characterized as having reduced binaural processing abilities (Moore, 2007).

The present study aimed at understanding: i) the effect of a directional and/or a fluctuating masker on the processing of ER's and the consequences for speech intelligibility, and ii) the role of binaural processing of ER's when the masker is directional and/or fluctuating. The results of the experiments were used to determine the benefit from masker fluctuations for monaural and binaural speech intelligibility.

3.2 Methods

3.2.1 Sound field simulations

The sound field simulations were the same as in Arweiler and Buchholz (2011). Speech intelligibility was measured in an acoustically dampened room with 29 loudspeakers arranged symmetrically around the listener (Favrot and Buchholz, 2010). This playback room had a reverberation time of $T_{30} < 100$ ms (for frequencies above 200 Hz) and ER's were attenuated by at least 10 dB at the listener's position. The room impulse response (RIR) used to create a realistic sound field was taken from a classroom modeled with the room acoustic software Odeon (version 9.1; Naylor, 1993). In this classroom, a source and a receiver were placed 3.5 m apart from each other to resemble a teacher and a student in a teaching situation. The room absorption properties of the classroom were defined in Odeon by choosing the appropriate materials for the walls, ceiling and floor. For a more detailed description of the playback room and the room model, the reader is referred to Arweiler and Buchholz (2011). The simulated RIR included the DS, the ER's and the late reflections. The DS and the ER's were used for further processing with the Loudspeaker Room Auralization System (LoRA) toolbox (Favrot and Buchholz, 2010) to produce a 29-channel RIR. Later arriving reflections were discarded to eliminate the influence of reverberation. The ER's included all reflections up to the second order, resulting in 20 ER's arriving within 55 ms after the DS at the listener. Each reflection was played from the loudspeaker that most closely matched the position of the ER in the simulated classroom. The RIR was then convolved with the speech signal and played back via the loudspeaker array. The listener was seated in the middle of the array at a distance of 1.8 m from the loudspeakers in the horizontal plane.

Speech intelligibility was measured binaurally and monaurally left for two conditions and four maskers (cf. Sec. 3.2.3). In the first condition, only the DS of the speech was presented from 0° azimuth (DS_{only}). The speech level was varied by changing the level of the DS. In the second condition, the DS of the speech was presented from 0° azimuth together with the ER's spatially distributed around the listener ($DSER_{spatial}$). The speech level was varied by changing the level of the DS and the ER together, i.e. by changing the level of the whole RIR. In this condition, the contribution of the ER's to the overall speech level was 6 dB and this remained constant when varying the speech level. In all conditions, the speech level was measured with an omni-directional microphone at the location of the center of the listener's head with the listener absent.

3.2.2 Speech material

The Danish sentence test Dantale II (Wagener *et al.*, 2003) was used to measure speech intelligibility. It is based on the Hagerman sentence test (Hagerman, 1982) where each sentence consists of five words with a fixed syntactical structure (name-verb-number-adjective-object). There are 10 alternatives for each sentence element. The listeners responded via a Matlab user interface by choosing the words they had heard from the 50 words presented on a hand-held touch screen.

3.2.3 Background noise

Four different maskers were used:

- A diffuse stationary speech-shaped noise (SSN) that was created from the Dantale II sentence material by repeatedly superimposing sentence sequences (Wagener *et al.*, 2003). The SSN was cut in uncorrelated noise segments and each of them was played from one individual loudspeaker simultaneously with all the other loudspeakers. This noise was the same as was used for the speech intelligibility measurements in Arweiler and Buchholz (2011) and thus served as a reference to this study.

- A directional stationary speech-shaped noise (SSN_{90}) which was created in the same way as the SSN, but presented from only one loudspeaker located at 90° azimuth, i.e. to the right of the listener.

- A multi-talker babble (MT_{90}) consisting of a 20-talker babble in English taken from track 3 of the compact disk CD101R3 "Auditory Tests Revised" by

AUDiTEC of St. Louis. This masker was also presented from the loudspeaker at 90° azimuth.

- A reversed two-talker masker ($revTT_{90}$) consisting of running speech from two female speakers. Two different parts of the two speakers were mixed at equal levels and time reversed. Silent gaps longer than 250 ms were removed. The female speaker was taken from track 8 of the compact disk CD B&O 101 "Music for Archimedes" by Bang & Olufsen and presented from the loudspeaker at 90° azimuth.

The temporal wave forms and the spectra for all maskers are shown in Fig. 3.1. The $revTT_{90}$ masker contains the most pronounced modulations while SSN and SSN_{90} exhibit the least pronounced modulations. In all cases, a gated-noise procedure was used, i.e. the noise started 1 s before each sentence with a 0.6 s onset ramp and ended 0.5 s after each sentence with a 0.3 s offset ramp. The masker was presented at a fixed level of 60 dB SPL for the normal-hearing listeners and at 70 dB SPL for the hearing-impaired listeners. A higher level was chosen for the hearing-impaired listeners to ensure that as much of the spectral content of the noise as possible was audible without using an uncomfortably loud presentation level. The sound pressure levels were measured with an omni-directional microphone at the location of the center of the listener's head with the listener absent and the total sound pressure level was equalized across all maskers.

3.2.4 Listeners

Eight normal-hearing and eight hearing-impaired listeners participated in the experiment. They were the same listeners as in Arweiler and Buchholz (2011), except for one normal-hearing listener who was not available for the present study. The normal-hearing listeners had thresholds of 20 dB HL (ISO, 2004) or better for both ears at octave frequencies between 0.25 and 6 kHz. Their age was between 23 and 46 years with a median age of 24 years. The thresholds for the hearing-impaired listeners are shown in Fig. 3.2. All hearing losses were of sensorineural origin and symmetric (interaural threshold differences, averaged across frequency from 0.25 to 8 kHz, were

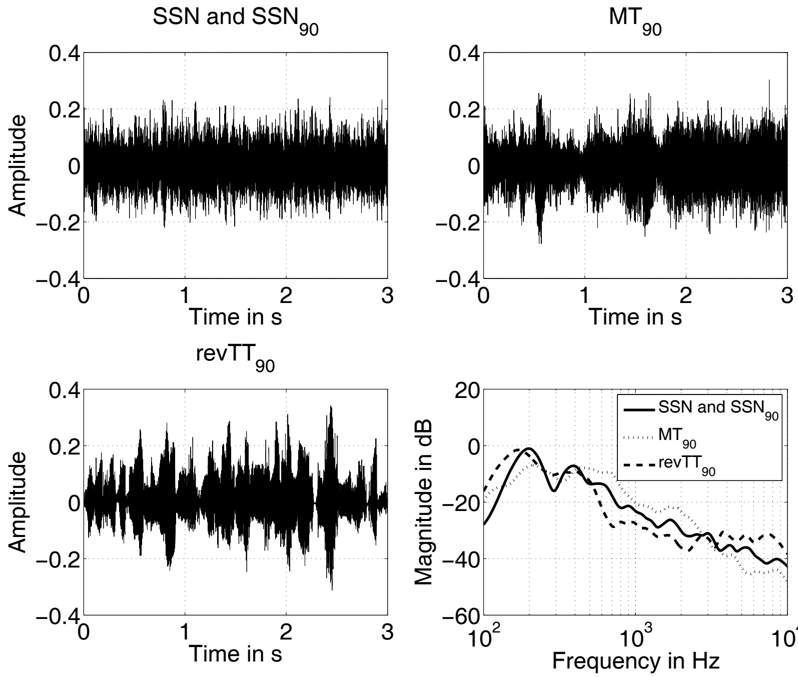


Figure 3.1: Temporal wave forms and spectra of the various maskers. The spectrum of the SSN corresponds to the average long term spectrum of the Dantale II sentences.

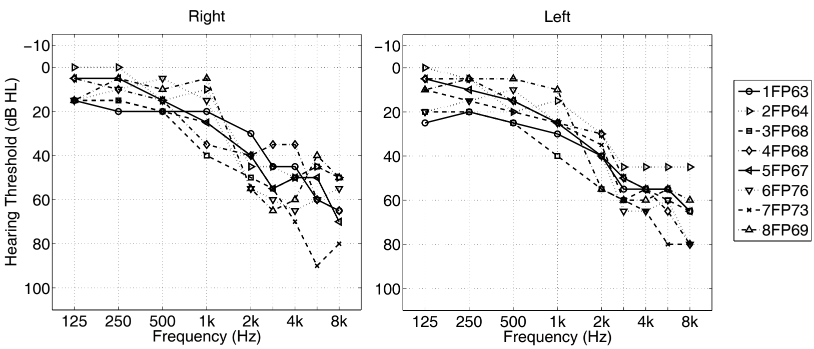


Figure 3.2: Hearing thresholds for 8 hearing-impaired listeners. The last two numbers indicate the age of the listener at the time of the listening experiment.

< 9 dB for all listeners). The hearing-impaired listeners were 63 to 76 years old with a median age of 68 years. The listeners were paid on an hourly basis for their participation.

3.2.5 Experimental procedure

Speech reception thresholds (SRT's) were measured to determine speech intelligibility. This was different from Arweiler and Buchholz (2011), where speech intelligibility was determined at fixed signal-to-noise ratios (SNR's). SRT measurements were chosen because they are less time consuming. The SRT reflects the SNR at 50% intelligibility. The speech level for the two conditions was varied adaptively with a maximum likelihood procedure (Brand and Kollmeier, 2002) depending on the number of words understood correctly. In the DS_{only} condition, the slope of the psychometric function was determined in addition to the SRT. According to Brand and Kollmeier (2002), a reliable concurrent SRT and slope estimate can be obtained with 30 sentences. Thus, in the DS_{only} condition, 30 sentences were used with the adaptive procedure, converging at 20% and 80% correct response of the psychometric function. In the $DSER_{spatial}$ condition, only the SRT was measured with 20 sentences using the adaptive procedure converging at 50% correct responses. The slope in the $DSER_{spatial}$ condition was considered to be identical to the slope of the DS_{only} condition, because the level of the DS plus ER's was varied and the ratio of the two components was kept constant. Arweiler and Buchholz (2011) showed that the slope of the psychometric function only changes when the ratio of DS and ER energy is changed, i.e. when ER energy is added or subtracted.

For the monaural speech intelligibility measurements (left ear only), an ER2 earphone (Etymotic Research) was inserted in the right ear which provided a minimum of 30 dB sound attenuation between 0.125 and 8 kHz. In addition, white noise was presented through the earphone at a level of 75 dB SPL. In this way, the non-test ear was completely masked and no headphones disturbed the sound field at the test ear. The experiments were split in two sessions, each lasting about 1.5 hours per person including breaks. Binaural speech intelligibility was measured in the

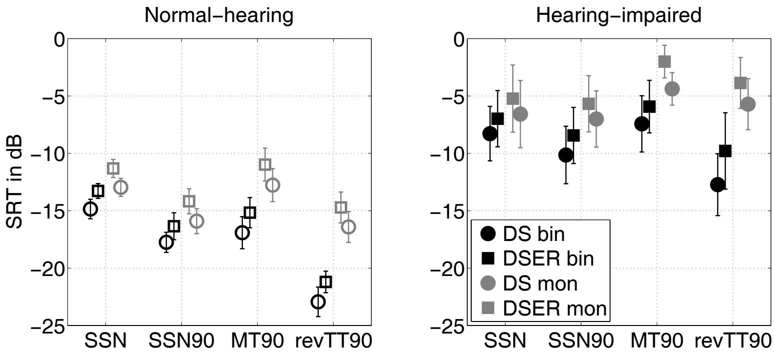


Figure 3.3: Mean monaural (grey symbols) and binaural (black symbols) SRTs for the four different maskers. Left panel: Normal-hearing listeners. Right panel: Hearing-impaired listeners. Circles represent the DS_{only} condition and squares represent the $DSER_{spatial}$ condition. Error bars indicate ± 1 standard deviation.

first session followed by a second, monaural session. The insert earphone was not removed until the monaural measurement was finished. The listeners performed a training with 20 sentences at the beginning of each session. The conditions and the order of measurements within each condition were randomized. Before the actual measurement, the listeners were instructed to look at the front loudspeaker and to hold their head upright during sentence presentation.

3.3 Results

3.3.1 Speech intelligibility data

The SRT's, averaged across the eight normal-hearing listeners, are shown in the left panel of Fig. 3.3. The corresponding data for the eight hearing-impaired listeners are shown in the right panel. Black symbols represent binaural intelligibility and gray symbols show monaural intelligibility. The circles indicate the results for the DS_{only} condition and the squares show the results for the $DSER_{spatial}$ condition. Speech intelligibility is better the more negative the SRT.

An ANOVA was performed with the listener group (normal-hearing versus hearing-impaired listeners) as between-subject factor, masker type (SSN, SSN_{90}, MT_{90} or $revTT_{90}$), listening mode (monaural or binaural) and condition (DS_{only} or $DSER_{spatial}$) as within-subject factors and SRT levels as dependent variable. There was a significant effect of listener group, masker type, listening mode and condition ($p < 0.001$). A significant interaction between listener group and masker type ($p < 0.001$) indicated that the difference in SRT between the two groups was larger for the fluctuating maskers than for the stationary maskers. Furthermore, a significant interaction between listening mode and masker type ($p < 0.001$) showed that there was a larger difference between monaural and binaural SRT's for the fluctuating maskers than for the stationary maskers.

For both groups of listeners, binaural speech intelligibility was best for the $revTT_{90}$ masker. This was most likely due to the characteristics of this masker (cf. Fig. 3.1), exhibiting the most pronounced temporal dips, the least energetic masking in frequency regions important for speech intelligibility (0.5-3 kHz) and the differences in fundamental frequency (F0) between the speech and the masker. Speech intelligibility for the MT_{90} was not better than for the stationary maskers, even though the temporal dips were more pronounced. However, the energy at frequencies between 0.5-3 kHz was higher for the MT_{90} than for the stationary maskers, which most probably led to increased energetic masking and, thus, higher SRTs.

Speech intelligibility was better for the DS_{only} condition than for the $DSER_{spatial}$ condition for all maskers and both groups of listeners. On average across maskers and listener groups the increase in speech intelligibility for the DS_{only} condition compared to the $DSER_{spatial}$ condition corresponded to an SNR difference of 1.7 dB.

Arweiler and Buchholz (2011) measured the binaural SRT for the $DSER_{spatial}$ condition with the same procedure as in the current study. Comparing this SRT between the two studies showed a very small difference of 0.4 dB for the normal-hearing and of 0.7 dB for the hearing-impaired listeners. Hence, the speech intelligibility test employed produced very reproducable results. Slope measurements in the current study resulted in about 12%/dB for the DS_{only} conditions and the

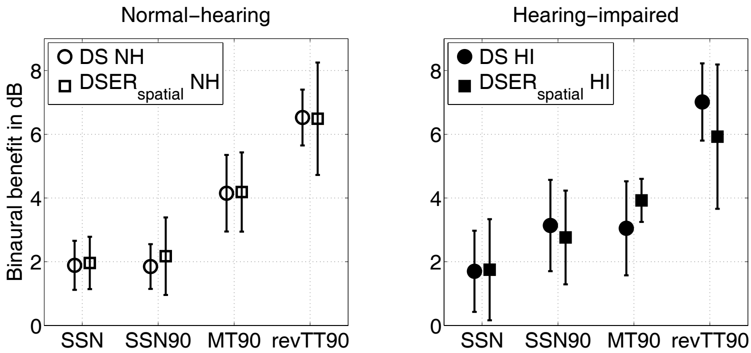


Figure 3.4: Binaural benefit expressed as the difference in SRT between monaural (left ear) and binaural listening. Open symbols: Normal-hearing listeners. Filled symbols: Hearing-impaired listeners. Error bars indicate ± 1 standard deviation.

diffuse SSN for both normal-hearing and hearing-impaired listeners which is in good agreement with Arweiler and Buchholz (2011).

3.3.2 Binaural benefit

The binaural benefit was calculated as the difference between the monaural and the binaural SRT's. Figure 3.4 shows the binaural benefit obtained in both groups of listeners for the four different maskers. Binaural speech intelligibility was always better than monaural speech intelligibility as reflected in a positive binaural benefit. For the stationary maskers (SSN and SSN_{90}), the binaural benefit corresponded to an SNR difference of about 2 dB and increased for the MT_{90} and the $revTT_{90}$. For both groups of listeners, no difference in binaural benefit was found for any masker between the DS_{only} and the $DSER_{spatial}$ conditions. For the normal-hearing listeners, an ANOVA showed a significantly larger binaural benefit for the fluctuating maskers (MT_{90} and $revTT_{90}$) than for the stationary maskers ($p < 0.001$). For example, for the $revTT_{90}$ masker, this benefit was about 4.6 dB larger than for the SSN_{90} in the DS_{only} condition (CI = 5.9151 3.4349). For hearing-impaired listeners, the benefit only increased significantly for the $revTT_{90}$ masker ($p < 0.001$) and not for the MT_{90} masker. The benefit for the $revTT_{90}$ masker was 3.9 dB larger than for the

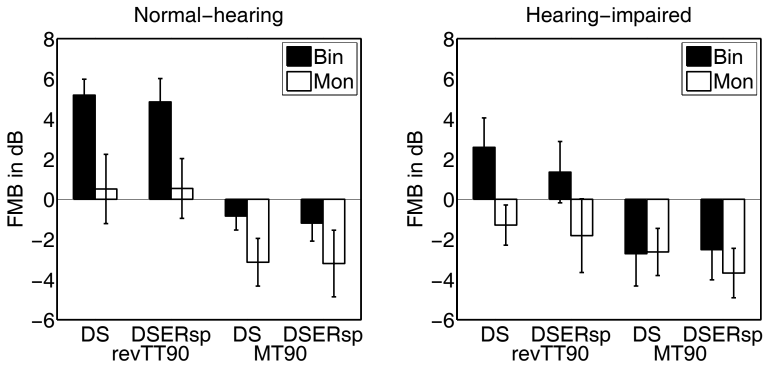


Figure 3.5: Fluctuating masker benefit expressed as the difference in SRT between the SSN_{90} condition and the MT_{90} and $revTT_{90}$ condition. Left panel: Normal-hearing listeners. Right panel: Hearing-impaired listeners. Error bars indicate ± 1 standard deviation.

SSN_{90} masker (CI = 5.7219 2.0281) in the DS_{only} condition and 3.2 dB larger (CI = 5.3493 0.9757) in the $DSER_{spatial}$ condition.

3.3.3 Fluctuating masker benefit

The FMB was calculated as the difference in SRT between the stationary masker from 90° azimuth (SSN_{90}) and the fluctuating maskers (MT_{90} and $revTT_{90}$). Figure 3.5 shows the FMB for the normal-hearing (left panel) and the hearing-impaired listeners (right panel). Both groups of listeners could benefit from the temporal fluctuations of the $revTT_{90}$ masker compared to the stationary masker (SSN_{90}), but only when listening binaurally. For the normal-hearing listeners, the SRT decreased by up to 5.2 dB when the stationary masker was replaced by the $revTT_{90}$ masker. For monaural listening, there was no significant benefit from masker fluctuations. When the stationary masker was replaced by the MT_{90} masker, no FMB could be observed. For the hearing-impaired listeners, the binaural SRT decreased by up to 2.6 dB when masker fluctuations were present. In the monaural condition, however, the fluctuations were detrimental, resulting in an increase in SRT. As for the normal-hearing listeners no FMB could be observed for the MT_{90} masker. In contrast, speech intelligibility was worse with the MT_{90} masker than with the SSN_{90} masker.

3.4 Discussion

3.4.1 Early reflection processing

One purpose of the study was to investigate the processing of ER's for a directional stationary or fluctuating masker. Arweiler and Buchholz (2011) showed that, for a diffuse SSN masker, increased DS energy leads to better speech intelligibility than increased ER energy. The same result was found in the present study where speech intelligibility was better in the DS_{only} condition than in the $DSER_{spatial}$ condition both for a diffuse masker and a directional masker (Fig. 3.3). However, Arweiler and Buchholz (2011) found a decrease in intelligibility for the $DSER_{spatial}$ condition compared to the DS_{only} condition corresponding to a difference in SNR of about 4 dB. For the diffuse SSN in the present study only a decrease corresponding to an SNR difference of about 1.5 dB SNR was found. Thus, in the present study, the benefit from ER's was larger than in Arweiler and Buchholz (2011). An explanation for this difference might be the different methods used to estimate the SRT, i.e. the fixed SNR's with varying ER energy versus the adaptive SRT estimation with fixed ER energy. Both methods led to very similar slopes for the DS_{only} condition but the latter method produced much steeper slopes in the $DSER_{spatial}$ condition, because the slopes for DS and $DSER_{spatial}$ were considered identical (cf. Sec. 3.2.5). This is in contrast to results from a study by Bradley *et al.* (2003). They used a speech intelligibility test with a very shallow slope (an SNR increase of 10 dB was necessary to increase the speech intelligibility from 87% to 99%) and found an even larger benefit from ER's. However, the increased benefit might rather be ascribed to the limited dynamic range in which they measured speech intelligibility (80-100%) than to the shallow slope. Thus, the method with which speech intelligibility is measured might play an important role for the observed ER benefit.

For the $directional$ maskers, the DS_{only} condition resulted on average in an improved speech intelligibility corresponding to a 1.8 dB SNR increase compared to the $DSER_{spatial}$ condition, which is similar to the 1.5 dB increase for the diffuse SSN. Thus, a directional masker did neither facilitate nor disturb the integration of the ER's with the DS compared to a diffuse masker. Furthermore, in the $DSER_{spatial}$

condition, the ability to suppress the noise masker was not improved for a directional masker compared to the diffuse masker. This could have been assumed if the directional and the diffuse masker had influenced the correlation of the ER's at the listener's ears differently (with the directional masker introducing less decorrelation of the speech signal and, thus, facilitating binaural unmasking).

3.4.2 Binaural benefit

In Fig. 3.4, the binaural benefit is shown for the different noise maskers. Arweiler and Buchholz (2011) found a 1.8-3.4 dB binaural benefit (depending on the ear and the group of listeners) for a diffuse SSN and no difference between the $DSER_{spatial}$ condition and a condition where both, the reflections and the DS, were presented from $0°$ azimuth. They suggested that only a binaural summation effect led to improved speech intelligibility when listening binaurally and that the precedence effect did not influence this benefit. The present study found a 2 dB binaural benefit for the diffuse SSN, which is in good agreement with the binaural summation effect found in Arweiler and Buchholz (2011) and in other studies (Pollack, 1948; Dillon, 2001; Moore, 2007). When a directional masker is used, binaural unmasking effects occur due to interaural differences of the speech and the masker. These difference cues, together with head shadow effects, are reflected in the binaural SRT measurements in the present study. With the monaural SRT measurement, only head shadow effects can provide an advantage for speech intelligibility so that the difference between monaural and binaural SRT's solely represents binaural unmasking based on difference cues. This binaural benefit amounts to 2-3 dB for the SSN_{90} condition both for normal-hearing and hearing-impaired listeners. Hence, the binaural benefit is not influenced by the direction of the SSN masker (diffuse vs. directional) even though the underlying mechanisms for the benefit are assumed to be different (binaural summation vs. binaural unmasking). Furthermore, no difference was found for the binaural benefit between the DS_{only} and the $DSER_{spatial}$ condition. This suggests that binaural processing is not involved in the integration of the ER's with the DS. Hence, it is assumed that this reflects a monaural process for both the directional and the diffuse masker.

An increased binaural benefit was observed when a speech masker was used instead of the noise masker (Fig. 3.4). Hawley *et al.* (2004) found a larger binaural benefit for two speech or reversed speech maskers presented from 90° azimuth than for two stationary or modulated noise maskers presented from the same location. They concluded that "with more than one interferer the binaural system is more effective at alleviating interference from a speech or reversed speech source than noise or modulated noise". This is in agreement with the results of the present study even though the underlying mechanisms for this effect are not completely clear. Hawley *et al.* (2004) assumed that a part of the increased binaural benefit for two speech type maskers compared to noise maskers might be due to an increased release from "informational" masking when the speech and masker are spatially separated. However, with the maskers used in the present study, a detailed analysis is difficult, because they differed in various properties (e.g., spectrum, fundamental frequency, modulation depths). Another study by Dirks and Wilson (1969) resulted in a slightly decreased binaural benefit for a speech masker compared to a noise masker when the target speech and the masker were spatially separated. A third study by Goverts *et al.* (2007) also found less binaural benefit for an interfering female talker than for a stationary noise. In their study, using headphones, the "monaural" condition was a diotic presentation of the speech and the noise (N_0S_0) and the "binaural" condition was a dichotic presentation (N_0S_π). They concluded that the speech masker was a less effective masker in the diotic condition compared to the stationary masker and, thus, binaural unmasking was reduced for the speech masker. This is in contrast to the results from the present study where the SSN_{90} and the $revTT_{90}$ were similarly effective in the monaural conditions (cf. Fig. 3.3). However, using phase reversed speech presented over headphones might not provide a perceptual spatial separation of the speech and the noise signal as clear as in the present study and, thus, the results should only be compared with caution.

A study by Carlile and Wolfe (2005) examined the influence of externalization on masking release. Two single talkers at ±30° azimuth were used as speech maskers. They found that, when binaural cues were present, the externalization of the speech and masker signals (i.e. the use of individual head related transfer functions (HRTF) for headphone presentation) provided a 5 dB intelligibility advantage compared to the speech and masker being presented dichotically without HRTF's. When the sound signals were low-pass filtered, however, the intelligibility advantage reduced to 1 dB. Thus, in the present study, the use of loudspeakers (rather than headphones) might have slightly increased the binaural release from masking for the fluctuating maskers.

As described by Arweiler and Buchholz (2011) central masking could have played a role in the monaural condition where a white noise was presented to the right ear with an insert earphone. Central masking refers to threshold elevation in one ear due to a masker in the contralateral ear that is not due to cross-talk, but takes place within the central auditory nervous system (Roeser and Clark, 2007). If central masking had occurred in the present study, increased monaural SRT's would have been obtained and the binaural benefit would have been increased. An increased binaural benefit (in addition to the 2 dB binaural benefit due to binaural summation/unmasking) could only be observed for the fluctuating maskers compared to the stationary maskers. Central masking might have occurred in the temporal dips of the fluctuating masker, such that the dips of the masker were not as useful for speech intelligibility in the monaural condition as in the binaural condition, because the white noise "centrally" filled these dips. To test this hypothesis, a headphone experiment was performed with four normal-hearing listeners. Speech intelligibility was measured with Dantale II sentences for the SSN_{90} and the $revTT_{90}$ masker. The speech and noise stimuli were filtered with impulse responses recorded on a Brüel & Kjær 4100 head and torso simulator (HATS) in the playback room for $0°$ and $90°$ azimuth, respectively. In the first (monotic) condition, SRT's were measured with the speech and masker presented simultaneously on the left ear. In the second (dichotic) condition, the filtered speech and masker were again presented to the left ear and, in addition, a white noise was presented to the right ear. The SSN_{90} and the $revTT_{90}$ masker had a level of 60 dB SPL measured at the centre of the listener's head. The white noise was presented at a constant level of 75 dB SPL. The speech was varied adaptively. Surprisingly, for the SSN_{90} masker, SRT's were on average 0.75 dB better for the dichotic condition, i.e. when an additional noise was presented to the right ear. For the $revTT_{90}$ masker, SRT's were slightly better in the monotic condition than in the dichotic condition, indicating that central masking effects might have influenced the monaural speech intelligibility results. However, the effect was very small, such that only a part (about 1 dB) of the increased binaural benefit found for the fluctuating maskers (Sec. 3.3.2) might have been due to central masking. Furthermore, hearing-impaired listeners should have received less central masking because the presentation level of the stationary and fluctuating maskers was higher (70 dB SPL compared to 60 dB SPL for the normal-hearing listeners), but the binaural benefit was nevertheless very similar to the binaural benefit for the normal-hearing listeners.

3.4.3 Fluctuating masker benefit

The FMB was determined monaurally and binaurally for two temporally fluctuating maskers. A FMB was observed only for the $revTT_{90}$ masker and not for the MT_{90} masker (Fig. 3.5). The FMB was larger when listening binaurally compared to almost no or even negative FMB when listening monaurally. Also here, central masking could have resulted in a larger binaural than monaural FMB (cf. Sec. 3.4.2). However, even if minor (1 dB) central masking effects played a role, the binaural FMB will still be larger than the monaural FMB (Fig. 3.5). Furthermore, Festen and Plomp (1990) only found a very small monaural FMB for normal hearing-listeners and a negative FMB for hearing-impaired listeners when the target were sentences spoken by a female and the masker was a reversed talker with the same sex. Larsby *et al.* (2008) also found a negative FMB for older hearing-impaired listeners in a diotic listening condition with reversed female discourse as masker. Other studies used an opposite-sex or modulated noise masker which resulted in a much larger FMB of up to 10 dB (Peters *et al.*, 2004; Summers and Mollis, 2004; George *et al.*, 2006; Bernstein and Grant, 2009; Ihlefeld *et al.*, 2010). It seems that, for monaural listening, the similarity of target and masker reduces the FMB. On the other hand, when listening binaurally, there is a large FMB, despite target and masker similarity. Thus, the FMB depends on the properties of the speech and masker especially in a monaural listening condition.

It has been discussed in the literature whether the FMB depends on the SNR at which speech intelligibility is measured. Larsby *et al.* (2008) found that, in difficult listening conditions, i.e. at low SNR's, speech intelligibility is better with a fluctuating masker than a stationary masker. Thus, the FMB is large. In an easy listening condition, i.e. at high SNR's, speech intelligibility is better for a stationary masker than for a fluctuating masker. This can be transferred to the current results when looking at the SNR's used for monaural and binaural speech intelligibility measurements in the $revTT_{90}$ condition (Fig. 3.3). At higher SNR's (monaural measurement), the distraction effect from the fluctuations in the masker is dominant, therefore speech intelligibility is similar to that in the SSN_{90} condition. At lower SNR's (binaural measurement), the dips in which potentially top-down processing can be activated are the dominant factor. Therefore, speech intelligibility becomes significantly better for the $revTT_{90}$ masker. For the MT_{90} masker the fluctuations, were probably not pronounced enough to utilize top-down processing, but distracting enough to decrease monaural speech intelligibility.

Several studies found that hearing-impaired listeners did not receive the same benefit from masker fluctuations as normal-hearing listeners (Festen and Plomp, 1990; Peters *et al.*, 2004; Summers and Mollis, 2004; George *et al.*, 2006; Larsby *et al.*, 2008) and attributed this to reduced audibility, reduced temporal and/or spectral resolution, or age. Bernstein and Grant (2009), however, found that the reduced FMB in hearing-impaired listeners can be explained by the differences in SNR at which normal-hearing and hearing-impaired listeners are tested, the latter usually requiring a more favorable SNR. In the present study, speech intelligibility for both groups of listeners was inherently measured at different SNR's. Figure 3.6 shows the FMB as a function of the SRT for the SSN_{90} condition. Open symbols represent the normal-hearing listeners and filled symbols indicate the hearing-impaired listeners. For each individual group of listeners, the FMB decreases with increasing SNR, which is in agreement with Bernstein and Grant (2009). However, some hearing-impaired listeners received the same FMB as the normal-hearing listeners, even though their speech intelligibility was measured at a lower SNR. This is primarily seen, when comparing the binaural FMB of the hearing-impaired listeners with the monaural FMB of the normal-hearing listeners. Thus, it might be concluded that the FMB does not solely depend on the SNR but also on the listening mode (monaural or binaural). The reason for the reduced FMB for the hearing-impaired listeners compared to the normal-hearing listeners in the present study remains open.

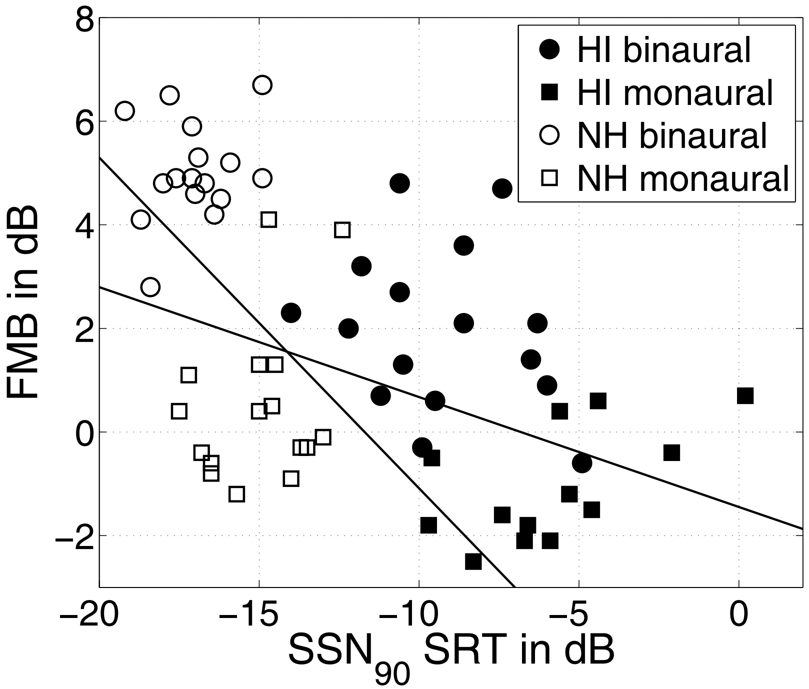


Figure 3.6: Fluctuating masker benefit for each individual listener and as a function of the SSN_{90} SRT for the DS_{only} and the $DSER_{spatial}$ condition. Open symbols: Normal-hearing listeners. Filled symbols: Hearing-impaired listeners. Lines are least-square fits to the individual normal-hearing and hearing-impaired data, respectively.

3.5 Summary and Conclusion

The present study investigated the effect of a directional and/or fluctuating masker on ER processing and the consequences for monaural and binaural speech intelligibility. Speech intelligibility was consistently better when the speech signal consisted of the DS only than when ER's were included. The difference in speech intelligibility between these two conditions was similar across the different masker types. However, the absolute benefit from ER's depends on the specific measurement method used to determine the speech intelligibility (fixed SNR versus SRT measurement).

A binaural benefit of 2-3 dB was observed for the diffuse SSN as well as for the directional SSN_{90}, indicating that the direction of the masker does not play a role for the processing of the ER's. The binaural benefit was found to be increased for the fluctuating maskers, but this increase was similar for the DS_{only} and the $DSER_{spatial}$ conditions. Thus, the binaural cues that are available in a directional and/or fluctuating masker condition do not affect the integration of the ER's and the DS in a different way than in the case of a diffuse masker. The integration of the ER's and the DS can therefore be assumed to result from monaural auditory processing for both masker types.

Monaural and binaural speech intelligibility were compared for the temporally fluctuating maskers ($revTT_{90}$ and MT_{90}) and the SSN_{90} masker. A fluctuating masker benefit (FMB) was observed only for the $revTT_{90}$ masker. Binaural speech intelligibility resulted in an increased FMB compared to monaural speech intelligibility. The reduced monaural FMB could be a result of the monaural auditory system being more susceptible to target-masker similarity, e.g. when target and masker have the same sex. In agreement with Bernstein and Grant (2009), the FMB for each listener group was found to depend on the stationary noise SNR. However, in addition, it has to be taken into account if the speech intelligibility was measured monaurally or binaurally when comparing both groups of listeners.

The influence of a directional and/or fluctuating masker on ER processing was similar for the hearing-impaired listeners and the normal-hearing listeners. The direction and the temporal characteristics of the masker did not have an effect on the binaural processing of the ER's. Thus, also in difficult listening conditions with a speech-like masker, hearing-impaired listeners receive the same benefit from ER's as normal-hearing listeners.

4

Speech intelligibility with binaurally linked hearing aids *

Abstract

Conventional compression algorithms in bilateral hearing-aid fittings distort the interaural level differences (ILD's) due to independent gain characteristics at the two ears. By transmitting signals between hearing aids, the compression can be coordinated and the same gain can be applied to both ears, thus preserving the ILD's. The present study investigated the influence of such "binaurally linked" hearing aids on speech intelligibility. Seven hearing-impaired listeners with a symmetric hearing loss were fitted with hearing aids connected to a hearing aid research platform (HARP). A fast and a slow compression algorithm were implemented on the platform. For the binaural link, the average gain of the left and right hearing aid was applied to both hearing aids. Speech reception thresholds (SRT's) were measured in a loudspeaker setup with the target speech and the masker spatially separated. Slightly, but not significantly, lower SRT's were achieved for the binaurally linked processing than for the unlinked processing using the same compression speed. The difference between monaural and binaural speech intelligibility was independent of the hearing aid algorithm. Thus, the preservation of the exact ILD information does not seem to be critical for binaural processing and speech intelligibility in the considered conditions.

* This chapter is based on Arweiler *et al.* (2011c).

4.1 Introduction

Wireless technology has become an integral part of modern hearing aids. In a bilateral fitting it has become possible to link the hearing aids on the left and right ear wirelessly and to exchange information between them. The transmission of either data parameters or the raw audio data is possible. Transferring low bit-rate data parameters is used to coordinate (or synchronize) settings in the two hearing aids, e.g. volume, program or microphone mode settings. While this coordination is useful in most listening conditions, it is not desirable in asymmetric listening conditions, e.g. in a car, when speech is present only on one side. In such a condition, the transmission of the raw signal data has proven useful. This sample-by-sample transmission requires a larger bandwidth and higher processing speed than parameter transmission and is, for example, realized with near-field magnetic induction (NFMI). Richards *et al.* (2006) used such a binaural link with audio data transmission in a simulated car situation. Assuming that the car noise is the dominant signal on one ear in such a condition, the noise estimate at that ear was subtracted from the speech+noise signal at the other ear, such that the SNR at the better ear was increased. The binaural link led to improved speech intelligibility in this specific condition for hearing-impaired listeners. Furthermore, audio data transfer has been used to restore binaural sound cues (Behrens, 2008). In particular, interaural level differences (ILD's), which would be altered by the compression algorithms of two independent hearing aids (Moore *et al.*, 1992), can be preserved by applying linked amplification. With independent compression and thus independent gain of the hearing aid signals at the two ears, more gain is applied to the ear receiving the softer signal and less gain is applied to the ear receiving the louder signal leading to a decrease in ILD's. This decrease is proportional to the compression ratio (Keidser *et al.*, 2006; Behrens *et al.*, 2009). Hence, it seems beneficial to apply the same gain on both ears, provided that the hearing-impaired person has a symmetric hearing loss. The preservation of spatial cues with binaurally linked hearing aids has been studied in terms of localization and sound quality in one study by Sockalingam *et al.* (2009). They found that fewer localization errors were made and that the sound was rated as more natural in a café or street environment when the binaural communication link was turned on.

While spatial cues are obviously useful for the localization of sound they can also provide an advantage for speech intelligibility. When a masker signal is spatially separated from a speech signal, the difference in the perceived location of the speech and the masker can improve speech intelligibility, particularly if the masker is a speech signal instead of a noise signal (Shinn-Cunningham, 2003). No studies have so far been published on the effect of binaurally linked hearing aids on speech intelligibility. The goal of the present study was therefore to investigate if the preservation of spatial cues with binaurally linked hearing aids could provide a benefit for speech intelligibility.

When speech intelligibility is measured in a condition where the speech and the masker are spatially separated, part of the spatial release from masking is due to the head shadow effect where the ear further away from the masker is provided with a more favorable SNR. However, an additional “true” binaural benefit is commonly observed which results from the binaural auditory system being able to utilize the spatial cues available in such a scenario (Dillon, 2001). This binaural benefit can be determined by comparing monaural and binaural intelligibility in the same listening condition. If binaurally linked hearing aids preserve the spatial cues, such that they can be used by the binaural system then the binaural benefit should be larger for linked hearing aids than for unlinked hearing aids.

Independent of a linked or unlinked processing strategy, the time constants of the compression algorithm play a role for speech intelligibility. When the masker is temporally fluctuating (e.g. a speech-like masker), fast compression can improve "listening in the dips" by amplifying the target speech signal in the short pauses of the masker (Moore, 1999). However, when the target speech and the speech-like masker are processed together by a fast compressor they may acquire a common modulation which makes them more difficult to separate (Stone and Moore, 2007; Plomp, 1988). A slow compression algorithm has the advantage of preserving the temporal envelope of the target speech which is important for speech intelligibility (Drullman *et al.*, 1994). However, slow compression does typically not provide any benefit in terms of "dip listening". Thus, it seems difficult to predict if the highest speech intelligibility scores can be achieved when a binaural link is combined with slow compression or when it is combined with fast compression.

The above mentioned studies that investigated the advantage of a binaural link have used commercial hearing aids in their experiments. While commercial hearing aids have the advantage that they can provide realistic signal processing, i.e. hearing aid settings that the hearing aid users would use in their daily routine, the interaction of different signal processing algorithms is often difficult to control for the experimenter. Furthermore, different hearing aid manufacturers use different signal processing approaches, so that experiments carried out with a hearing aid from one manufacturer typically represent only one signal processing approach. A standardized platform which provides full control over independent signal processing algorithms is therefore desirable. Such a hearing aid research platform (HARP) was used in the present study with a binaural link algorithm, a compression algorithm and a linear amplification algorithm implemented. All stimuli were presented to the listener in a 3D loudspeaker setup. The HARP, together with such a virtual auditory environment, represent an advanced approach for testing hearing aid algorithms.

4.2 Methods

4.2.1 Listeners

Seven hearing-impaired listeners participated in the study. They were the same listeners as in Arweiler and Buchholz (2011) and Arweiler *et al.* (2011b), except that one listener was not available for the present experiment. Their age was between 64 and 77 years with a median age of 68. Median hearing thresholds are shown in Fig. 4.1. All listeners had a sensorineural and symmetric hearing loss (interaural threshold differences averaged across audiometric frequencies were $\leq$ 5 dB). Hearing aid usage and experience differed between listeners. One listener had never worn a hearing aid before, two listeners had hearing aids, but did not use them, two listeners occasionally used their hearing aids and two listeners always used their hearing aids (i.e. all day long). Four out of the six listeners with hearing aids were bilaterally

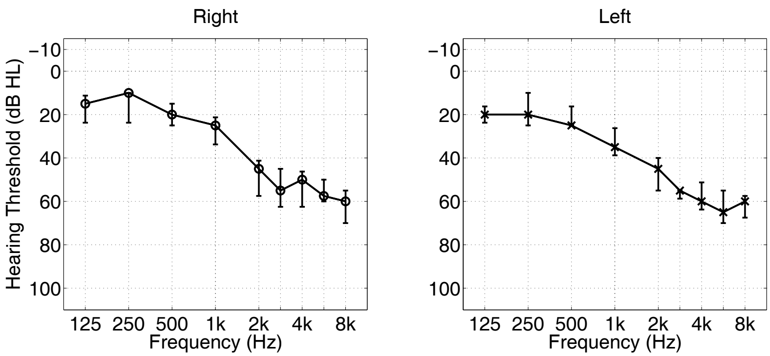


Figure 4.1: Median hearing thresholds for the 7 hearing-impaired listeners. Error bars indicate the interquartile range.

fitted with a behind-the-ear (BTE) hearing aid and an open canal fitting (slim tube and dome). The remaining two listeners were fitted with a completely-in-the-canal (CIC) hearing aid on the left ear only and with bilateral in-the-ear (ITE) hearing aids, respectively. All listeners with hearing aids had owned them for at least 2.5 years.

4.2.2 Stimuli

The speech stimuli were sentences from the Danish sentence test Dantale II (Wagener *et al.*, 2003). Each sentence consists of five words with a fixed syntactical structure (name-verb-number-adjective-object) spoken by a woman. For each sentence element 10 alternatives are available. After each sentence presentation the listener had to chose the words they had heard from the 50 words presented via a Matlab user interface on a hand-held touch screen. The level of the sentences was varied adaptively for speech reception threshold (SRT) estimation (cf. Sec. 4.2.4). All sentences were presented from a loudspeaker at $0°$ azimuth.

The masker stimuli were two of the stimuli used in Arweiler *et al.* (2011b), a stationary speech shaped noise masker (SSN_{90}) and a reversed two-talker masker ($revTT_{90}$). The SSN_{90} was created by repeatedly superimposing sequences of the Dantale II sentences (Wagener *et al.*, 2003). Consequently, the masker had the same average long-term spectrum as the sentences. The $revTT_{90}$ consisted of running speech from two female speakers. One speaker was time delayed with respect to the other before the two speakers were mixed at an equal level and time reversed with silent gaps longer than 250 ms removed. The female speaker was taken from track 8 on the compact disk CD B&O 101 "Music for Archimedes" by Bang & Olufsen. Both maskers were presented from a loudspeaker at 90° azimuth, hence the indication 90 (cf. Sec. 4.2.4). The presentation of the maskers was gated, i.e. the noise started 1 s before each sentence, with a 0.6 s onset ramp, and ended 0.5 s after each sentence, with a 0.3 s offset ramp. The broadband root-mean-square sound pressure was equalized across the two maskers and levels were measured with an omni-directional microphone at the location of the center of the listener's head with the listener absent. The maskers were presented at a constant level of 70 dB SPL.

4.2.3 Hearing aid research platform (HARP)

Setup

The HARP consists of a real-time target machine from Speedgoat (www.speedgoat.ch) which is programmed, controlled and monitored by a host computer. The target machine is a dual-core PC with two PCI I/O boards, the input board with 24 Bit A/D converters and the output board with 16 Bit D/A converters. The host computer runs Matlab and SimuLink software to develop hearing aid signal processing models as well as to control the real-time processes on the target machine. The developed SimuLink models are automatically translated into C-code with Real Time Workshop, xPC Target, and Visual Studio. Any hearing aid signal processing algorithm was first developed and tested in simulated real-time on the host computer using Simulink. The successful hearing aid signal processing algorithm was then compiled and transferred onto the target machine, where finally the hearing aid signal processing algorithm was run in real-time. The host computer was then used to monitor the target machine processing.

Two BTE hearing aid satellites from Phonak were connected to the target PC via a purpose build battery-powered four-channel microphone preamplifier from Phonak and a Behringer Powerplay Pro-8 HA8000 headphone amplifier. Each satellite had one receiver and two microphones, a front microphone and a back microphone, but only the front microphone was used.

The calibration of the different hardware components of the HARP, including the hearing aid satellites, was done with a PC using a high quality RME DIGI 96/8 sound card with internal 8-channel input and output board extension and running Matlab. Where applicable, a hearing aid test box (Affinity 2.0 system) was used which was controlled via the hearing aid test (HIT) software from Interacoustics. The test box employed a reference microphone to control the sound pressure level produced by a loudspeaker inside an acoustically sealed box. Moreover, the test box contained a 2cc coupler (with a measurement microphone) to measure the sound pressure level produced by the hearing aid satellite receiver. Equalization filters were realized by IIR filters and fitted to the inverse absolute spectrum of the measured transfer functions using Matlab.

Hearing aid signal processing algorithms

Three signal processing algorithms were implemented on the research platform: (i) a linear signal processing algorithm, (ii) a wide dynamic range compression (WDRC) algorithm and (iii) a binaural link algorithm. For all algorithms, the input waveform was first analysed by a 4th order Gammatone bandpass filterbank with 24 auditory filters. Afterwards, a resynthesis stage combined all frequency channels to form the output waveform (Hohmann, 2002). In the linear algorithm, all sound signals were processed linearly, i.e. the same gain was applied for all input levels (cf. Sec. 4.2.4). In the WDRC algorithm the input signal was first processed by a compressor using a peak level detector. Afterwards, spectral smoothing was applied to limit gain variations between adjacent frequency bands. In this way, the decrease in spectral contrast that is generally introduced by the compressor (i.e. spectral peaks are more compressed than spectral dips) was reduced as well as the uncontrolled reinsertion of energy from adjacent frequency bands after compression (due to overlapping filters in the filterbank) was limited. In the binaural link algorithm, the compressors of the left and right hearing aid were linked, i.e. the average gain of both compressors was applied to each hearing aid, thus retaining interaural level differences. An ideal binaural link was simulated through sample-by-sample transmission in all 24 channels introducing no additional time delay.

The change in ILD's that occurs when no binaural link is applied to the signal processing is shown in Fig. 4.2 for the $revTT_{90}$ and the SSN_{90} masker. The solid line represents the ILD's before signal processing and the dashed lines after signal processing (for fast and slow compression, respectively). The ILD's were calculated by using an offline model of the HARP. The input target and masker signals (as recorded at the microphone of the hearing aid fitted to a head and torso simulator (HATS from Brüel & Kjær)) were compared in each frequency channel to the output target and masker signals after the mixture of the target and masker had been processed by the compressor. The insertion gains (IG's) used for the compressor algorithm were derived from the left ear of the median audiogram of the hearing-impaired listeners. The independent compression on both hearing aids introduced a reduction in ILD's for both the target (presented from 0° azimuth) and the masker (presented from 90° azimuth) of approximately 4 dB and 5.1 dB, respectively (averaged over the 24 frequency channels and the two masker types). This will be addressed again in Sec. 4.4.

For all algorithms, an expander at low input levels ensured the suppression of microphone noise. The expander threshold was 12 dB above the microphone noise in each frequency channel. Furthermore, a soft limiter was activated when a maximum total (frequency independent) output of 100 dB SPL was reached. The total input/output delay of the HARP was 6.3 ms.

4.2.4 Experimental procedure

Hearing aid fitting

The HARP was first calibrated such that, at the output of a 2cc coupler, a flat frequency response was achieved. However, when the hearing aid is fitted to the listener's ear and the sound is presented via loudspeakers, the sound pressure level

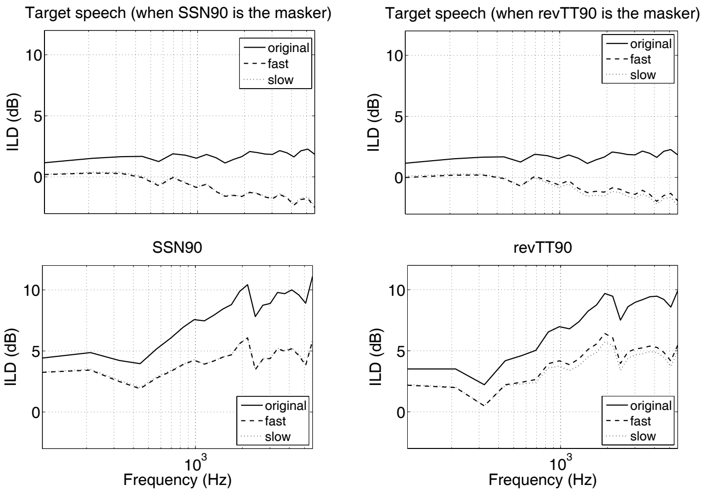


Figure 4.2: Interaural level differences (ILD's) before and after signal processing for the target, the SSN_{90} masker and the $revTT_{90}$ masker. Solid line: ILD's for target and masker before signal processing. Dashed and dotted line: ILD's for target and masker after signal processing for fast and slow compression respectively.

has to be controlled at each individual's ear drum. Thus, the free field to BTE microphone transfer function and the difference between the 2cc coupler and the real ear (RECD) have to be equalized. In this way, it is made sure that the (aided) sound pressure at the ear drum is equal to the sound pressure at the ear drum without any hearing aid present. In consequence, the gain provided by the HARP is directly equal to the real-ear insertion gain (REIG, Dillon, 2001) which can be derived from a generic fitting rule. For the equalization of the free field to BTE microphone transfer function, the transfer function from Bentler and Pavlovic, 1989 (Table 1, column B) was used. For the equalization of the RECD, the REIG was determined for each individual ear with the Affinity 2.0 system. First, the real ear unaided gain (REUG) was measured with a probe microphone. Next, the real ear aided gain (REAG) was measured with the hearing aid satellites fitted via closed (no venting) custom made earmolds to the listener's ears and with the HARP providing a 15 dB frequency independent linear gain. A warble tone with 70 dB SPL was used as input signal and presented via an external loudspeaker connected to the Affinity system. The

REIG was automatically calculated by subtracting the REUG from the REAG. An IIR filter was then derived from the inverse of the REIG. The accuracy of the equalization was verified by applying the free field to BTE and the REIG equalization filter to the signal processing on the HARP and repeating the REAG measurement. The resulting frequency response showed a relatively flat 15 dB gain at frequencies between 500 and 5000 Hz (see Appendix Insertion gain measures).

The "equalization" procedure provided with the CamFit software (version 1.0; Moore *et al.*, 1999) was used to derive individual IG's and CR's in 12 frequency channels for each listener based on their audiogram. The CamFit frequency channels were defined by the 24 HARP frequency channels so that one CamFit channel covered two HARP channels over a bandwidth of approximately 6000 Hz. The derived IG's (for tonal inputs) and CR's were interpolated between the center frequencies of the frequency channels. The compression threshold (CT) in each frequency channel was defined as 25 dB below the corresponding average speech level (65 dB SPL). That means, with the speech presented at a level of 10 dB below the average speech level (which was assumed to correspond to the lowest SRT) and the dynamic range of speech being $\pm$ 15 dB, all components of the speech signal would still be processed in the compressive region. Two compression strategies were employed: a fast and a slow one. The fast strategy used a compressor attack time (AT) of 5 ms and a release time (RT) of 50 ms. For the slow strategy the AT was 5 ms and the RT was 1000 ms.

For the linear algorithm, IG's were derived from the fast unlinked compressor algorithm. Linear IG's corresponded to the IG's of the compressor algorithm for each individual listener when the input was a SSN with 70 dB SPL. This input signal had the same properties and level as the SSN_{90} used in the speech intelligibility measurement.

Overall, five different conditions were tested: Binaurally unlinked combined with fast compression, slow compression and linear amplification and binaurally linked combined with fast and slow compression. The binaural link was not combined with linear amplification because the linear gains to the left and right ear were almost identical due to the symmetric hearing loss.

Speech intelligibility measurements

The listener was seated in the middle of a 29-loudspeaker array. The speech signal was always presented from the front loudspeaker at 0° azimuth and the masker was presented from the loudspeaker at 90° azimuth (to the right of the listener). All signals were calibrated with an omni-directional microphone at the location of the center of the listener's head with the listener absent. Speech reception thresholds (SRT's) were measured monaurally and binaurally with 20 sentences for each listener and condition. The level of the speech signal was varied adaptively with a maximum likelihood procedure (Brand and Kollmeier, 2002) depending on the number of words understood correctly. In the monaural condition, the right earmold was removed and the ear was plugged with an insert earphone (ER2 from Etymotic Research) delivering a white noise with a level of 75 dB SPL. The earphone itself provided a minimum of 30 dB sound attenuation between 0.125 and 8 kHz. For the binaural link condition, it was necessary to keep both hearing aids in place behind the ears, also in the monaural measurement to pick up the signals at both ears. In order not to bias the listeners, both hearing aids were kept behind the ear for all monaural measurements.

The hearing-impaired listeners attended 3 sessions. In the first session (1h), after ear inspection, ear impressions were taken for the manufacturing of the closed skeleton earmolds and hearing thresholds were measured. In the second session (3h), the hearing aids were fitted as described in Sec. 4.2.4 and speech intelligibility measurements were started. Four listeners started with binaural measurements and 3 listeners started with monaural measurements. In the third session (3h), measurements from the second session were finished and the remaining measurements were performed. Altogether, 20 SRT measurements were performed (5 hearing aid algorithms x 2 maskers x monaural/binaural). The masker conditions and the order of measurements for each masker were randomized. Before the first measurement, listeners performed a training with 20 sentences and were instructed to look at the front loudspeaker and to hold their head upright during sentence presentation. Listeners were paid on an hourly basis for their participation.

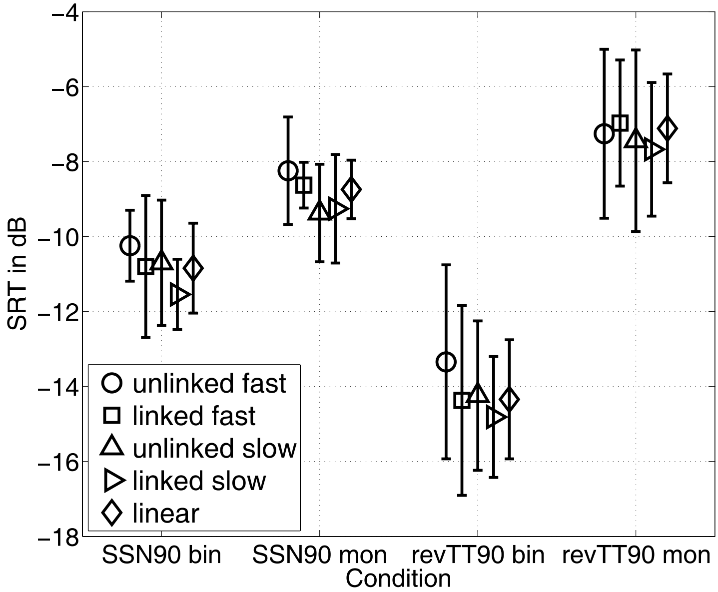


Figure 4.3: Mean SRT's for the different hearing aid processing algorithms and conditions. Error bars indicate $\pm$ 1 standard deviation.

4.3 Results

4.3.1 Speech intelligibility data

Figure 4.3 shows the mean SRT's averaged across the seven hearing-impaired listeners. In general, speech intelligibility results varied across the four different conditions. Binaural speech intelligibility was better for the $revTT_{90}$ than for the SSN_{90} masker and monaural speech intelligibility was slightly better for the SSN_{90} than for the $revTT_{90}$.The linked hearing aid conditions tended to yield lower SRT's than the unlinked conditions when combined with the same compressor speed (fast or slow). Slow compression yielded lower SRT's than fast compression. Five paired t-tests with Bonferroni correction for multiple comparisons were performed on the results. T-tests were used instead of an analysis of variance to eliminate inter-subject variability. The binaural SRT's were analyzed for the unlinked versus the linked conditions. In the case of fast compression, the binaural link did not provide significantly lower SRT's than the unlinked condition. In the case of slow compression, a significant difference between linked and unlinked SRT's was found (p = 0.024, CI = 0.1073 1.3069), but the difference did not remain significant after Bonferroni correction. For the $revTT_{90}$there was no significant difference between the linked algorithm combined with slow versus fast compression, indicating that the speech intelligibility with the fluctuating masker was independent of the compressor speed. Finally, the SRT's obtained with the linear processing were not significantly different from the SRT's obtained with fast or slow compression. Figure 4.4 shows the individual SRT results for the unlinked condition (ordinate) against the corresponding individual SRT's for the linked condition (abscissa). The filled symbols indicate slow compression and the open symbols indicate fast compression. Data points below the diagonal indicate lower SRT's for the linked condition. Even though there was no significant effect of the binaural link on speech intelligibility when averaged across listeners, some listeners showed a consistent benefit from the binaural link. For example, the listeners indicated by the circle and the diamond had a lower SRT with the linked hearing aid in almost all experimental conditions.

4.3.2 Binaural benefit

The difference between monaural and binaural SRT's is shown as the binaural benefit in Fig. 4.5. Squares indicate the binaural benefit for the SSN_{90} and circles represent the binaural benefit for the $revTT_{90}$. For both maskers, an analysis of variance (ANOVA) did not show a significant difference in binaural benefit between hearing aid algorithms. The average binaural benefit was 2 dB for the SSN_{90} and 6.9 dB for the $revTT_{90}$.

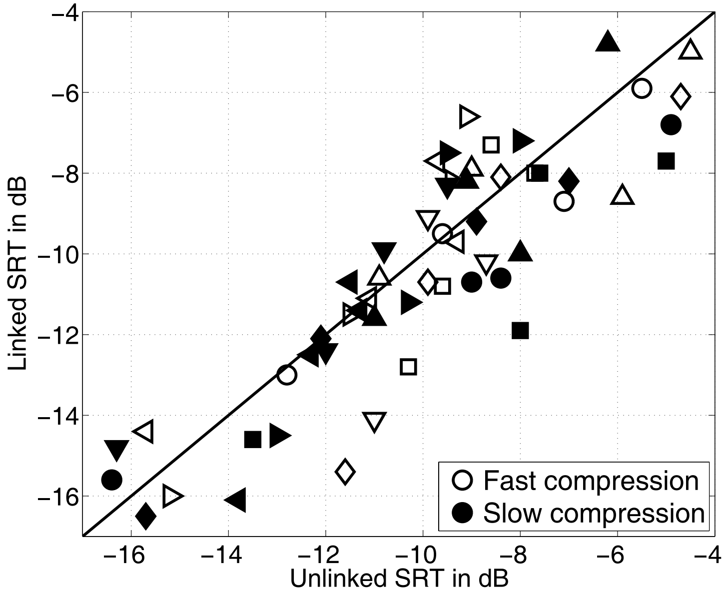


Figure 4.4: Individual SRT's for the unlinked condition as a function of the linked condition. Filled symbols: slow compression. Unfilled symbols: fast compression. Each symbol type represents one hearing-impaired listener. Data points below the diagonal indicate better SRT's for the linked condition.

.

4.4 Discussion

The aim of this study was to investigate if the preservation of ILD cues, as realized in binaurally linked hearing aids, can improve speech intelligibility compared to unlinked hearing aids. In Fig. 4.2 the expected decrease in ILD's due to unlinked compression was shown. However, the ILD's are not only decreased because of the asymmetric listening condition in which the ear closer to the masker receives relatively less gain and the ear further away from the masker receives relatively more gain. Slight differences in hearing threshold between the two ears also lead to an imbalance in gain. However, hearing-impaired listeners seem to be able to adapt to and learn the change

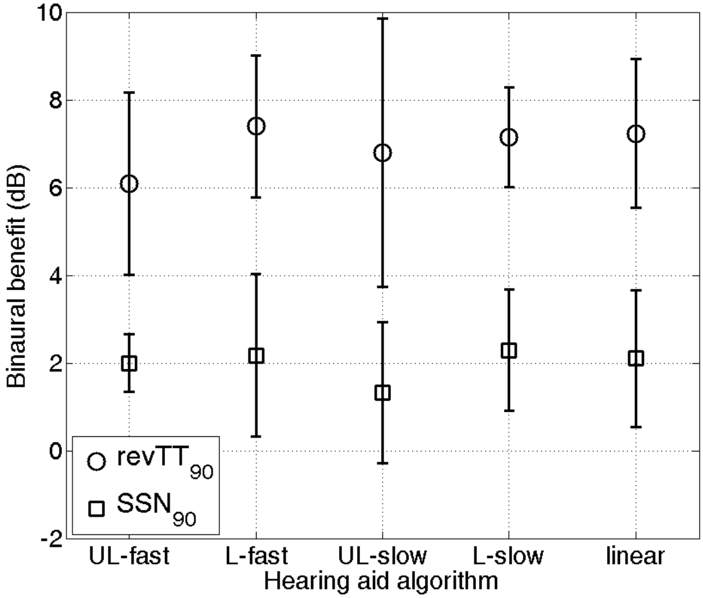


Figure 4.5: Mean binaural benefit for the different hearing aid processing algorithms and maskers, calculated as the the difference between the mean monaural and binaural SRT's. UL: unlinked. L: linked. Error bars indicate $\pm$ 1 standard deviation

.

in ILD's resulting from sensitivity differences between the two ears (Blauert, 1997). This acquired adaptation is again disturbed when IG's are applied that compensate for the sensitivity differences. Furthermore, the attack time (AT) might have influenced the perceived ILD. Musa-Shufani *et al.* (2006) showed that just noticeable differences for ILD's increased with decreasing AT. A longer AT gives the listener the possibility to analyze the ILD before it is compressed. Thus, the ILD's shown in Fig. 4.2 only represent the ILD's due to compression, because a symmetric audiogram based on the left ear median hearing thresholds was used for calculation.

Gain differences at the two ears resulted in ILD's that were on average about half of the original ILD's. This was expected considering that the compression ratios derived from the CamFit software for the median audiogram were between 1.0 and 2.8 in the different frequency channels. Compared to the original ILD's of the masker the compressed ILD's corresponded to a shift of the masker towards the location of the target. However, the reduction of ILD's for the speech signal introduced a shift of the target away from the masker. The shift of the target and the masker in the same direction (by slightly different amounts) thus only slightly decreased the spatial separation. Hence, the preservation of the ILD's by the binaural link did not lead to an improved spatial release from masking compared to the unlinked condition.

The linked hearing aid algorithm could restore the ILD's that were reduced due to compression. However, regarding the speech intelligibility data only a small difference in SRT was found between the linked and unlinked algorithm. The linked gain for the compression system was mainly determined by the louder masker signal. Thus, no improvement in SNR could be expected for the used target/masker configuration. A binaural link can suppress the masker only in very specific conditions (Kates, 2008). When the speech signal is, for example, presented from one side of the listener with a higher sound level than the masker presented from the opposite side, then the louder speech signal at one ear reduces the gain of the masker signal at the opposite ear. The effect is small, however, and in the order of 1 dB SNR improvement. Furthermore, problems of hearing-impaired listeners to understand speech in noise most often occur at negative SNR's.

Even though unlinked compression reduces the ILD's, this might not influence the localization abilities of hearing-impaired listeners and hence, speech intelligibility might not be worse than for linked compression. Keidser *et al.* (2006) measured the localization error with WDRC hearing aids which almost halved the ILD's compared to a linear system. Surprisingly, the reduction of ILD's had no significant effect on localization performance. Musa-Shufani *et al.* (2006) measured the influence of compression on just-noticeable-differences for ILD's with hearing-impaired listeners and compared the results to a localization task based on ILD's and ITD's. They concluded that hearing-impaired listeners predominantly rely on ILD's for the localization of high-frequency stimuli and, even though they were reduced due to the compression, the localization accuracy was only decreased by 5° compared to normal-hearing listeners.

The relatively small difference in speech intelligibility between the linked and unlinked condition might also result from a limited correlation between localization abilities and speech understanding in noise. A study by Noble *et al.* (1997) investigated the relationship between localization, detection of spatial separateness and speech intelligibility in noise for hearing-impaired listeners. For the group of hearing-impaired listeners with sensorineural hearing loss, they did not find consistent relationships between the localization task, the spatial separateness task, and the speech task.

The binaural benefit determined from the monaural and binaural SRT's (see Fig. 4.5) was not significantly different across hearing aid algorithms, i.e. there was no benefit from linking the hearing aids in terms of improved binaural speech intelligibility. Moore *et al.* (1992) studied the influence of a compression hearing aid and a linear hearing aid on the binaural benefit for speech intelligibility. Averaged across hearing-impaired listeners they found a binaural benefit of approx. 2.5 dB for both processing strategies. They presented a babble noise from $\pm 90^\circ$ and speech from 0° azimuth. The monaural condition was realized by turning off the hearing aid at the non-test ear. The binaural benefit is in good agreement with the binaural benefit for the SSN_{90} in the present study, even though the non-test ear in Moore *et al.* (1992) was not completely masked in the monaural condition. This might have resulted in a reduced binaural benefit. Moore *et al.* (1992) did not determine to what extend the ILD's were reduced by the compression but simply concluded that the independent compression at the two ears did not adversely affect the binaural cues. The present study has shown that *despite* the reduction of ILD's introduced by independent compression the binaural benefit is not decreased compared to other hearing aid processing algorithms. It remains unclear why these reduced ILD's do not reduce localization and/or speech intelligibility in hearing-impaired listeners. Edwards (2010) suggested that less compression in the hearing-impaired cochlea leads to a larger perceived ILD than the non-linear behavior of the cochlea in normal-

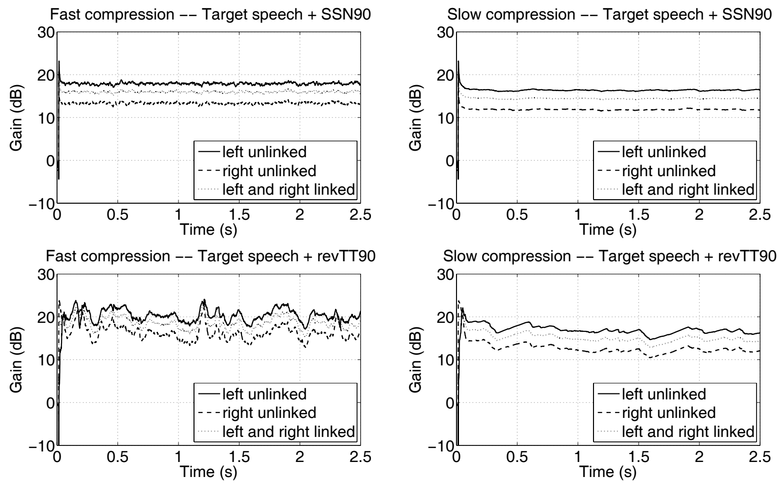


Figure 4.6: Linked and unlinked (left and right) gains applied for fast and slow compression calculated for one exemplary listener and averaged over all frequency bands. Upper panels: gains with target speech and SSN_{90} masker, lower panels: gains with target speech and $revTT_{90}$ masker. Left panels: fast compression, right panels: slow compression.

hearing listeners. This might counteract the reduction of ILD's through hearing aid compression. More evidence is needed for this assumption.

On average, the hearing-impaired listeners showed consistently better speech intelligibility for the slow compression than for the fast compression algorithm (Fig. 4.3). This is surprising considering the gains applied to the speech and the masker with slow and fast compression. These are shown in Fig. 4.6 for the SSN_{90} masker (upper panels) and the $revTT_{90}$ masker (lower panels) and for fast (left panels) and slow compression (right panels). The upper and lower curve in each panel show the gains for the left and right ear when independent compression is used. The middle curve shows the (identical) gain for both ears when the hearing aids are linked. The gains were calculated for one specific hearing-impaired listener, with the mixture of speech and masker as the input signal and averaged over frequency. Less gain is applied to the input signal when slow compression is used than when fast compression is used. This is due to the behavior of the compressor. For a long RT (slow compression), the gain response to a decrease in input level is slow such that less gain is applied than for a fast RT. For the unlinked condition less gain is applied to both ears with slow compression than with fast compression. In the linked condition (compared to unlinked) the gain is reduced on the better (left) ear and increased on the noisy (right) ear (Fig 4.6). Even though the gain differences between slow and fast compression as well as between linked and unlinked are not small, they did not have a large effect on speech intelligibility.

It should be noted that the hearing-impaired listeners were not acclimatized to the hearing aids and some of them never or hardly wore their own hearing aids. Moore *et al.* (1992) pointed out, that "this is important because if a person has had a hearing impairment for many years, it may take some time for them to learn to use the new cues provided by a compression system." Thus, even though the linked hearing aids provided measurable benefit in terms of preserved ILD's, the hearing-impaired listeners might have simply not been able take advantage of it.

Despite the time it may take to get acclimatized to a hearing system, four hearing-impaired listeners received an instantaneous benefit from the hearing aids. The aided SRT's measured in the present study were compared with the unaided SRT's from the study by Arweiler *et al.* (2011b) for the conditions that used the same stimuli. The same hearing-impaired listeners participated in both studies, except for one who was not available for the present study. On average across hearing aid processing algorithms, these four listeners showed an improvement in SRT of 1.62 dB. This improvement corresponds to an increase in speech intelligibility of almost 20%, assuming a slope of the psychometric function of 12%/dB as measured in Arweiler *et al.* (2011b). Three of the hearing-impaired listeners did not show an instantaneous benefit in the aided conditions and the SRT was on average 0.33 dB worse than without hearing aids.

An improvement in speech intelligibility was not necessarily expected with the algorithms used in the HARP. None of the algorithms was designed to improve the SNR, even though in some cases compression can increase the SNR (Naylor and Johannesson, 2009). Furthermore, with the speech presented from a location in front of the listener and the masker presented from the side of the listener, a BTE hearing aid can lead to a decreased head shadow effect (Festen and Plomp, 1986). Due to the microphone location behind the ear, the masker noise at the contralateral ear is less attenuated than would be the case at the ear canal entrance. This can result in decreased speech intelligibility. On the other hand, the improved sensitivity of the hearing-impaired listeners in high frequencies, due to frequency dependent amplification in the aided conditions, might have provided an increased benefit from the head shadow compared to the unaided conditions.

4.5 Summary and Conclusion

Binaural hearing aids, as opposed to bilateral hearing aids, can preserve ILD's which are assumed to play a role for understanding speech in noise. In the present study, a hearing aid research platform (HARP) was used to investigate the benefit from binaurally linked hearing aids for speech intelligibility in hearing-impaired listeners.

In a loudspeaker setup, where the target speech and the masker were spatially separated, the binaural link resulted in slightly, but not significantly better speech intelligibility than the unlinked hearing aid algorithm (for the same compression speed). The best speech intelligibility results were achieved when the linked hearing aid was combined with slow compression. Despite the altered ILD's that were introduced by the unlinked WDRC, binaural processing was not adversely affected, resulting in the same binaural benefit for unlinked and linked hearing aid signal processing.

Thus, the results from the present study do not support the hypothesis that binaurally linked hearing aids improve speech intelligibility. However, the binaural link might still be useful to improve sound quality in terms of naturalness and spatial awareness, as was shown in previous studies (Sockalingam *et al.*, 2009; Behrens, 2008). Furthermore, a binaural link can provide other advantages than 'only' preserving ILD's. Active binaural noise reduction systems are, for example, being developed based on interaural cross-correlation (Kates, 2008). Other binaural systems have used the individual signals at the two ears to determine the location of the sound signal, such that a source in front of the listener will receive more gain than a source at other locations (Hohmann *et al.*, 2002; Wittkop and Hohmann, 2003). However, these algorithms have so far either not been suitable for the implementation in binaural hearing aids because of the limited data transfer or the benefit for speech intelligibility could not be demonstrated. More research is therefore needed to find the link between localization, spatial awareness, binaural processing and speech intelligibility in noise. The independent hearing aid research platform in combination with a virtual auditory environment as used in the present study may represent versatile tools to further develop outcome measures that show the benefit of hearing aids for hearing-impaired listeners.

5

Effect of hearing aid signal processing on early reflections and the consequences for speech intelligibility

5.1 Introduction

In Chapters 2 and 3 it was shown that hearing-impaired listeners can benefit from early reflections (ER's) in the same way as normal-hearing listeners. Speech intelligibility improved when ER energy was added to the speech signal. The question was raised if certain hearing aid algorithms, like a directional microphone or compression, could decrease the ER benefit. In Chapter 4, the influence of different hearing aid algorithms on speech intelligibility was investigated with anechoic speech as target signal. In the present study, the anechoic speech signal was replaced by a speech signal that included ER's ($DSER_{spatial}$ condition, as used in Chapter 2 and 3) and speech intelligibility was measured with the same hearing aid algorithms as in Chapter 4. Thus, the effects of compression and hearing aid microphone placement on the ER benefit are considered here.

The benefit from the ER's was mainly determined by their spectrum compared to the DS (Arweiler and Buchholz, 2011). Due to wall absorption, the energy of the ER's was decreased at high frequencies which in turn led to a decreased speech intelligibility. With a hearing aid, the spectrum might furthermore change due to the microphone location when a BTE hearing aid is used. The influence of the pinna on the spectrum of the signal is different for sound signals arriving behind the pinna than for signals arriving at the ear canal entrance. The incident angle of the incoming sound also plays a role, such that the spectrum might be altered differently for the ER's, which are spatially separated, and the DS, which is presented from the front direction. The location of the hearing aid microphone also decreases the head shadow (Festen and Plomp, 1986) which might be advantageous for the integration of the ER's with the DS. The spectrum of the sound signals is also altered by multi-channel frequency compression. However, it is unclear if this has an effect on the integration of the ER's with the DS. Attack (AT) and release times (RT) of a compressor system could also play a role.

In this study, the same methods as in Chapter 3 and 4 were used and are described briefly in the following section. The effect of the different hearing aid algorithms on ER processing will be evaluated by comparing the results from the present study to the results from Chapters 3 and 4. More specifically, the difference in speech intelligibility between the condition with DS only and the condition with added ER's is expected to be larger with hearing aids than without hearing aids, if hearing aid signal processing interferes with the processing of ER's.

5.2 Methods

Monaural and binaural speech intelligibility tests were performed with the same hearing aid algorithms as in Chapter 4: Unlinked with fast compression, unlinked with slow compression, linked with fast compression, linked with slow compression and linear signal processing. The same maskers were used: SSN_{90} and $revTT_{90}$. They were again presented from 90° azimuth in a loudspeaker setup of 29 loudspeakers. The DS_{only} speech signal corresponded to the one used in Chapter 3 and 4 and the speech signal including the ER's was identical to the $DSER_{spatial}$ signal in Chapter 3. In the $DSER_{spatial}$ condition the contribution of the ER's to the overall speech level was 6 dB. The DS was presented from the loudspeaker at 0° azimuth and the ER's were again presented spatially distributed from different loudspeakers around the listener. The level of the maskers was 70 dB SPL, measured at the center of the listener's head with the listener absent. The level of the speech (Dantale II) was varied adaptively to determine the speech reception threshold (SRT). The same seven hearing-impaired listeners as in Chapter 4 participated in this study.

5.3 Results

5.3.1 Speech intelligibility data

The SRT's for the $DSER_{spatial}$ condition, averaged across the hearing-impaired listeners, are shown as black symbols in Fig. 5.1. The gray symbols represent the mean SRT's for the DS_{only} condition from Chapter 4 (Fig. 4.3) and the stars indicate the mean SRT's for the unaided $DSER_{spatial}$ and DS_{only} condition from Chapter 3 (Fig. 3.3). The error bars are omitted for better readability. A t-tests for each of the four conditions (SSN_{90} binaural, SSN_{90} monaural, $revTT_{90}$ binaural and $revTT_{90}$ monaural) was performed on the differences between the $DSER_{spatial}$ and the DS_{only} condition. For all four conditions, speech intelligibility was significantly better in the DS_{only} condition than in the $DSER_{spatial}$ condition ($p < 0.001$, after Bonferroni correction for multiple testing). Averaged across hearing aid algorithms, the difference in speech intelligibility between the DS_{only} condition and to the $DSER_{spatial}$ condition corresponded to an SNR difference of 1.8 dB. This is very similar to the unaided difference between the DS_{only} and the $DSER_{spatial}$ condition of 1.7 dB (Chapter 3, p. 40). Thus, from the 6 dB energy contribution of the ER's, 1.8 dB could not be used for speech intelligibility.

5.3.2 Binaural benefit

The binaural benefit was calculated as the difference between the monaural SRT's and the binaural SRT's. In order to investigate the influence of the hearing aid algorithms on the binaural benefit, the unaided binaural benefit from Chapter 3, the aided binaural benefit for the DS_{only} condition from chapter 4 and the aided binaural benefit for the $DSER_{spatial}$ condition from the present study were compared. The aided binaural benefit was subtracted from the unaided binaural benefit for each signal

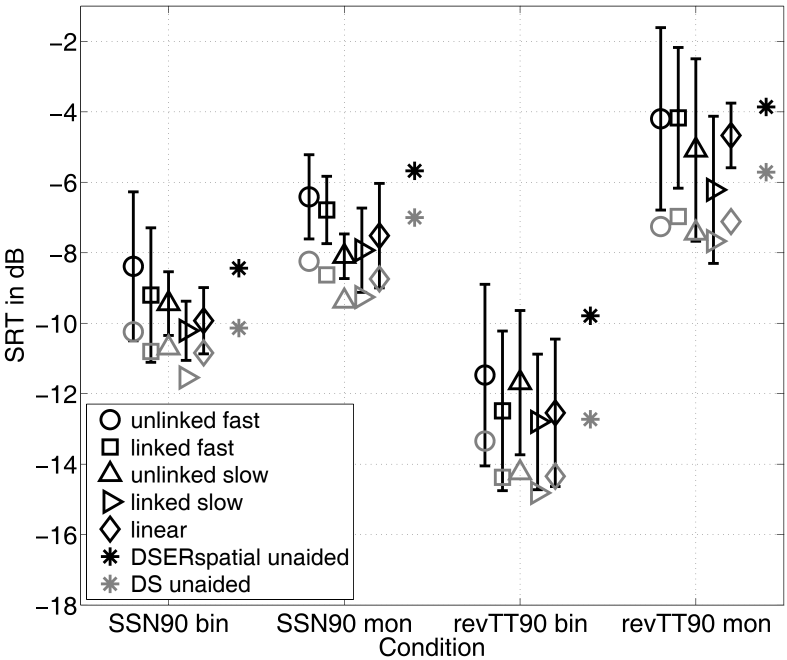


Figure 5.1: Mean SRT's for the different hearing aid processing algorithms and conditions. Black symbols: $DSER_{spatial}$. Error bars indicate $\pm$ 1 standard deviation. Gray symbols: DS_{only}, replotted from Chapter 4, Fig. 4.3. Error bars are omitted for better readability. The black and gray star are replotted from Chapter 3, Fig. 3.3, representing the mean SRT's for the unaided $DSER_{spatial}$ condition and the unaided DS_{only} condition, respectively. Error bars are omitted for better readability.

processing algorithm. The difference is shown in Fig. 5.2 for the two maskers and the five algorithms. Circles indicate the DS_{only} condition and squares represent the $DSER_{spatial}$ condition. Open and closed symbols show the SSN_{90} and the $revTT_{90}$, respectively. A 2-way analysis of variance (ANOVA) was performed for each masker with hearing aid algorithms and conditions (DS_{only} and $DSER_{spatial}$) as main factors and the difference in binaural benefit as dependent variable. For the SSN_{90} masker, there was no significant effect of hearing aid algorithm or condition. For the $revTT_{90}$ masker, there was no significant effect of hearing aid algorithm, but a

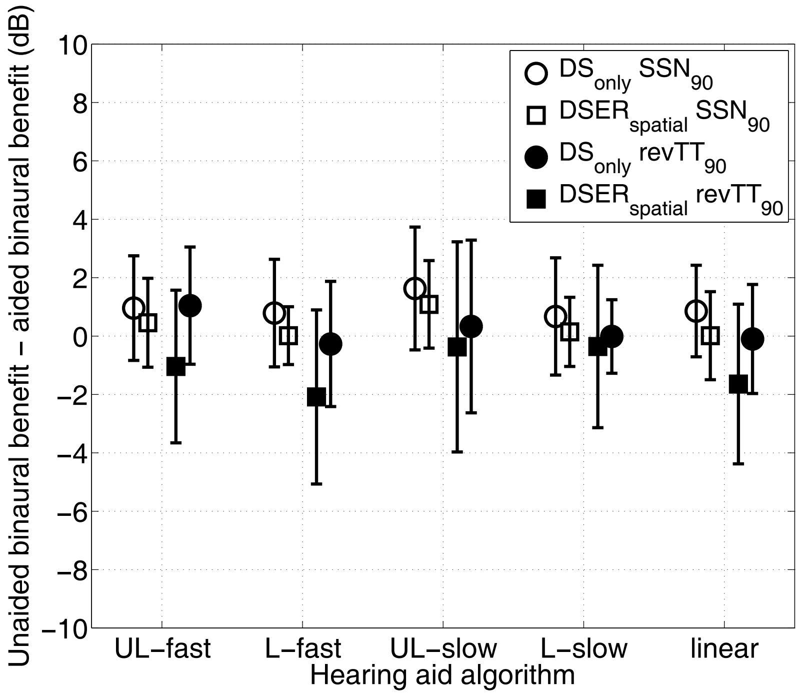


Figure 5.2: Mean difference across listeners between the unaided and aided binaural benefit. Positive data points represent a larger binaural benefit for unaided listening. Circles indicate the DS_{only} condition and squares represent the $DSER_{spatial}$ condition. Open symbols: SSN_{90}. Filled symbols: $revTT_{90}$. UL: unlinked. L: linked. Error bars indicate ±1 standard deviation.

significant effect of condition (p = 0.04). However, subsequent multiple comparisons did not reveal a significant difference between the DS_{only} and the $DSER_{spatial}$ condition for any of the algorithms.

5.4 Discussion and Conclusion

The hearing aid algorithms used in the present study did not have a detrimental effect on the benefit from ER's for speech intelligibility. On average, the same difference between the DS_{only} and the $DSER_{spatial}$ condition was found for speech intelligibility with and without hearing aids (1.7 versus 1.8 dB). The location of the hearing aid microphone did not play a critical role, i.e. the different spectral changes introduced by the pinna for the DS_{only} and the $DSER_{spatial}$ condition might not have been large enough to be reflected in the speech intelligibility results. Furthermore, the hearing aids did not influence the binaural benefit. The independent compression for the unlinked hearing aids reduced the ILD's but did not affect binaural processing. Neither did the preservation of the ILD's by the binaural link improve binaural processing. In both, the DS_{only} and the $DSER_{spatial}$ condition, the binaural benefit was not affected by the hearing aid. The integration of ER's with the DS is considered to be a monaural process (Arweiler and Buchholz, 2011). Hence, the binaural cues provided by the ILD's did not play a role for the integration process.

ER's are important for speech intelligibility. The present study showed that hearing aids with 'simple' signal processing algorithms maintain the ER benefit. The signal processing in modern hearing aids is based on compression algorithms like in the present study. In addition, more advanced algorithms are typically used with the focus on increasing speech intelligibility in noise. Spectral subtraction-based signal enhancement methods might affect the DS and the ER's equally and might thus not have an effect on the ER benefit. Multi-microphone arrays, on the other hand, might decrease the benefit from ER's. Depending on the directionality of such an array, part of the ER energy might be suppressed and the positive effect of the directional microphone on the SNR might be significantly reduced.

No reverberation was considered in this study. If the full room information is used for both the speech and the masker signal the useful effects of the ER's and the detrimental effects of the reverberation will interact. Future studies might investigate this interaction and the consequences for the ER benefit.

The hearing aid signal processing algorithms implemented on a research platform and the virtual auditory environment used in the present study provide a basis for realistic hearing aid testing. In the future, more complex algorithms can be implemented and more complex listening scenarios can be created to investigate speech intelligibility in hearing-impaired listeners.

6 Overall summary and discussion

In this thesis, different aspects of spatial hearing were investigated and discussed. Chapter 2 focused on the benefit from ER's for speech intelligibility. Speech intelligibility was measured as a function of the DS energy and the ER energy comprised in the speech signal. An efficiency factor was derived from the intelligibility results which showed that ER energy was less useful for speech intelligibility than DS energy. This was in contrast to a study by Bradley *et al.* (2003) who observed that ER energy contributed to speech intelligibility in a similar way than DS energy. They used delayed and attenuated copies of the DS to simulate the ER's and did not consider the spectral characteristics of the ER's. Hence, the speech intelligibility results in Chapter 2 were analyzed based on the spectrum of the ER's. The SII-based weighting of the efficiency factor according to the importance of different frequency regions for speech intelligibility revealed that the altered spectrum of the ER's compared to the DS caused the decrease in speech intelligibility. Thus, it was concluded that absorptions from the walls, the ceiling or the floor of a room need to be taken into account when determining the importance of ER's for speech intelligibility. Bradley *et al.* (2003) and other studies did not consider this change in the spectrum for their simulated ER's and thus might have overestimated the benefit from ER's. The intelligibility-weighted efficiency factor provides a tool to determine how useful ER's are in a given room. Only one room with one source-receiver setup was used in this experiment and it can be argued that the benefit from ER's changes for different rooms and different locations of the source and the receiver. Furthermore, no late reflections (reverberation) were added to the speech signal and the considered masker was anechoic, i.e. contained no reflections at all. There are numerous possibilities to extend and generalize this experiment. Speech intelligibility measures, however, are time consuming and speech intelligibility models might be able to predict speech intelligibility in various scenarios, with and without reflections. The results from the present study might provide constraints for such models of speech intelligibility.

The work presented in Chapter 2 focused furthermore on the auditory processing of ER's. In particular, it was examined if the precedence effect had an influence on speech intelligibility when the speech signal included spatially distributed ER's. The direction of the ER's is known to be suppressed by the precedence effect but it was unclear if parts of the energy of the ER's would also be suppressed. Speech intelligibility was thus measured for frontal ER's and for spatial ER's and better results were obtained for the frontal ER's. However, this difference could be explained by the influence of the torso, head and pinna on the direction of the ER's and the influence of the playback room. Thus, a suppression of ER energy due to the precedence effect was not found.

The processing of ER's had not been investigated in earlier studies in terms of monaural versus binaural listening. It was assumed in the present study that, if binaural processing was necessary for the integration of ER's with the DS, then hearing-impaired listeners with reduced binaural processing abilities might show a reduced benefit from ER's compared to normal-hearing listeners. All listeners therefore performed the speech intelligibility measurements monaurally and binaurally. The difference between the monaural and the binaural intelligibility results was referred to as the binaural benefit. For all tested conditions (DS_{only}, $DSER_{spatial}$ and $DSER_{frontal}$), the binaural benefit was similar and corresponded to a binaural summation effect. Hence, binaural listening did not provide an advantage for speech intelligibility beyond a binaural energy summation, neither for the DS alone, nor for spatial or frontal ER's. It was concluded, that monaural processing is sufficient in terms of the integration of ER's with the DS, and that binaural processing does not affect the benefit from ER's. It was thus not surprising that hearing-impaired listeners showed the same benefit from ER's for speech intelligibility. However, the question was raised, wether a directional masker could have provided more distinct binaural cues than a diffuse masker, such that a difference in binaural benefit between the DS_{only} and the $DSER_{spatial}$ condition would occur.

The experiments from Chapter 2 were therefore repeated with a directional and/or a temporally fluctuating masker. Chapter 3 showed that, again, the ER benefit was very similar for monaural and binaural listening in the DS_{only} and the $DSER_{spatial}$ condition. It was concluded that, even with a directional masker, ER's are first integrated monaurally with the DS and that the combined signal is then processed further. The masker type did not play a role, presumably because noise suppression takes place at a binaural stage after the processing of the ER's.

Apart from investigating ER processing, monaural and binaural speech intelligibility was compared in Chapter 3 for the different maskers. An increased binaural benefit was found for the temporally fluctuating maskers. Only a few studies have investigated the binaural benefit for fluctuating maskers and the results were therefore difficult to compare and to explain. However, the similarity of the target and the masker might play a role, whereby the binaural system might be more efficient to suppress other speech maskers than noise maskers when they are spatially separated from the speech signal. The presentation of the signals over loudspeakers and thereby preserving individual HRTF cues might also have increased the binaural benefit compared to studies that have used headphone presentation.

The increased binaural FMB compared to the monaural FMB was explained in a similar way, such that the target-masker similarity decreased monaural speech intelligibility for a speech-like masker relatively more than for a stationary masker. Bernstein and Grant (2009) explained the difference in FMB between normal-hearing and hearing-impaired listeners with the difference in the stationary-noise SNR at which speech intelligibility is measured, with the hearing-impaired listeners usually requiring a more favorable SNR. The reason for the reduced FMB for the hearing-impaired listeners in the present study remains unclear.

When the fluctuating masker was used instead of the stationary masker, central masking effects needed to be considered. An additional experiment was performed that could demonstrate that such effects did not influence the results. By presenting a spectrally shaped noise instead of white noise through the insert earphone, these effects could have been minimized beforehand. Low-pass filtering according to the speech spectrum could have reduced the masking effect in higher frequency regions. In the experiments using hearing aids central masking effects were even more unlikely. The amplification of the loudspeaker signals by the hearing aid increased the level difference to the earphone signal which in turn should have decreased central masking. Central masking effects should have become apparent in the binaural benefit, which was very similar for the condition with and without hearing aids though.

The influence of hearing aid signal processing on spatial cues and the consequences for speech intelligibility were investigated in Chapters 4 and 5. Binaurally linked hearing aids were used with fast and slow compression. In contrast to unlinked hearing aids which distort ILD cues, the binaural link preserves the natural ILD's. ILD's were considered important for the benefit from a spatial separation of the speech and the masker signal due to the head shadow effect and binaural processing abilities. ILD cues are most effective for frequency regions above 1.5 kHz. For the hearing-impaired listeners, frequency-dependent amplification was applied according to each individual's hearing loss, such that they could use the ILD information available at higher frequencies. For binaural listening, however, no difference in speech intelligibility was found between the linked and the unlinked condition. Apparently, the (precise) preservation of ILD's was not necessary for improved speech intelligibility in the considered conditions. One possible explanation might be that the hearing-impaired listeners were not able to instantaneously use the ILD cues, because they were not acclimatized to the hearing aids. Furthermore, they might have relied more on low-frequency ITD cues. This might also be reflected in the constant binaural benefit across hearing aid algorithms. As described in Chapter 3, the binaural benefit for the SSN_{90} masker results from binaural unmasking due to interaural difference cues. This binaural unmasking is mostly determined by the ITD's (Bronkhorst and Plomp, 1988), which are most useful at low frequencies. Thus, the binaural benefit might be unaffected by the linked signal processing which preserves the high-frequency ILD's. For the same reason, it can be assumed that the binaural benefit was similar in the aided and unaided conditions. Without hearing aids, the hearing-impaired listeners still had access to ITD cues because of their almost normal hearing thresholds in the low frequencies. It can be speculated then, that the increased unaided and aided binaural benefit for the $revTT_{90}$ masker might result from an improved access to ITD's, presumably in the dips of the masker. This could be clarified by a repetition of the experiments of Bronkhorst and Plomp (1988) with a speech-like masker.

The speech intelligibility results were not influenced by the type of compression used (fast versus slow). Stone and Moore (1999) compared speech intelligibility for different compression systems, including fast and slow compression. They found that, even though the compression algorithms were markedly different, the resulting speech intelligibility was similar for all algorithms. As was found in the present study, the precise form of compression did not affect their speech intelligibility results.

Early reflection processing was likewise not affected by any of the hearing aid algorithms. Thus, the benefit from ER's was maintained. It still needs to be investigated if directional microphones might suppress part of the ER energy and, hence, reduce speech intelligibility. This might partly reduce the positive effect of a directional microphone with regard to noise suppression.

Overall, the results and methods of this thesis might be encouraging to further investigate spatial hearing and speech intelligibility in hearing-impaired listeners with and without hearing aids.

Bibliography

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

Arweiler, I. and Buchholz, J. M. (**2011**). “The influence of spectral and spatial characteristics of early reflections on speech intelligibility”, J. Acoust. Soc. Am., **130**, 996–1005.

Arweiler, I., Buchholz, J. M. and Dau, T. (**2011b**). “The influence of masker type on early reflection processing and speech intelligibility (L)”, J. Acoust. Soc. Am. (under review).

Arweiler, I., Buchholz, J. M. and Dau, T. (**2011c**). “Speech intelligibility with binaurally linked hearing aids”, Int. J. Audiol. (under review).

Behrens, T. (**2008**). “Spatial hearing in complex sound environments”, Hear. Rev. **15**, 94–102.

Behrens, T., Maas, P., and Neher, T. (**2009**). “A method for quantifying the effects of non-linear hearing aid signal processing on interaural difference cues in conditions with multiple sources”, in: J. M. Buchholz, T. Dau, J. C. Dalsgaard, T. Poulsen, eds. *Binaural Processing and Spatial Hearing, Proceedings of the Second International Symposium on Auditory and Audiological Research (ISAAR 2009)*, Helsingør, Denmark, Centertryk A/S.

Bentler, R. A. and Pavlovic, C. V. (**1989**). “Transfer functions and correction factors used in hearing aid evaluation and research”, Ear Hear. **10**, 58–63.

Bernstein, J. G. W. and Grant, K. W. (**2009**). “Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners”, J. Acoust. Soc. Am. **125**, 3358–3372.

Blauert, J. (**1997**). *Spatial hearing* (The MIT Press, Cambridge, MA), p. 279.

Bradley, J. S., Sato, H., and Picard, M. (**2003**). "On the importance of early reflections for speech in rooms", J. Acoust. Soc. Am. **113**, 3233–3244.

Brand, T. and Kollmeier, B. (**2002**). "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests", J. Acoust. Soc. Am. **111**, 2801–2810.

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

Carlile, S. and Wolfe, J. (**2005**). "High frequencies and spatial location in speaker segregation", Poster session presented at the annual meeting of the American Audiological Association (AAA), Washington DC.

Chiang, Y. C. and Freyman, R. L. (**1998**). "The influence of broadband noise on the precedence effect", J. Acoust. Soc. Am. **104**, 3039–3047.

Dillon, H. (**2001**). *Hearing aids* (Thieme, New York), pp. 379–380.

Dirks, D. D. and Wilson, R. H. (**2002**). "The effect of spatially separated sound sources on speech intelligibility", J. Speech Hear. Res. **12**, 5–38.

Drullman, R., Festen, J. M. and Plomp, R. (**1994**). "Effect of temporal envelope smearing on speech reception", J. Acoust. Soc. Am. **95**, 1053–1064.

Edwards, B. (**2010**). "Aided psychoacoustics and binaural perception with hearing aids", Presentation at the International Hearing Aid Conference (IHCON), Lake Tahoe, USA.

Favrot, S. and Buchholz, J. M. (**2010**). "Lora: A loudspeaker-based room auralization system", Acta Acust. United Ac. **96**, 364–375.

Festen, J. M. and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing", J. Acoust. Soc. Am. **88**, 1725–1736.

Festen, J. M. and Plomp, R. (**1986**). "Speech-reception threshold in noise with one and two hearing aids", J. Acoust. Soc. Am. **79**, 465–471.

Freyman, R. L., McCall, D. D., and Clifton, R. K. (**1998**). "Intensity discrimination for precedence effect stimuli", J. Acoust. Soc. Am. **103**, 2031–2041.

George, E. L. J., Festen, J. M., and Houtgast, T. (**2006**). "Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners", J. Acoust. Soc. Am. **120**, 2295–2311.

Goode R.L. (**2001**). “Auditory physiology of the external ear”, in: *Physiology of the ear*, edited by A.F. Jahn and J. Santos-Sacchi (Singular, San Diego, CA), p. 151.

Goverts, S. T., Derleux, M., Festen, J.M. and Houtgast, T. (**2007**). “The influence of masker type on the Binaural Intelligibility Level Difference”, in: *Hearing – From Sensory processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp and J. Verhey (Springer, Berlin Heidelberg), p. 551–558.

Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (**1993**). “Intelligibility-weighted measures of speech-to-interference ratio and speech system performance”, J. Acoust. Soc. Am. **94**, 3009–3010.

Grose, J. H. and Hall, J. W.. (**1998**). “Masker fluctuation and the masking-level difference”, J. Acoust. Soc. Am. **103**, 2590–2594.

Haas, H. (**1951**). “Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache [On the influence of a single echo on the intelligibility of speech]”, Acustica **1**, 49–58.

Hagerman, B. (**1982**). “Sentences for testing speech-intelligibility in noise”, Scand. Audiol. **11**, 79–87.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (**2004**). “The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer”, J. Acoust. Soc. Am. **115**, 833–843.

Hohmann, V. (**2002**). “Frequency analysis and synthesis using a Gammatone filterbank”, Acta Acust. United Ac. **88**, 433–442.

Hohmann, V., Nix, J., Grimm, G., and Wittkop, T. (**2002**). “Binaural noise reduction for hearing aids”, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Orlando, USA.

Ihlefeld, A., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (**2010**). “Simulations of cochlear-implant speech perception in modulated and unmodulated noise”, J. Acoust. Soc. Am. **128**, 870–880.

ISO 389-8 (**2004**). “Acoustics—Reference zero for the calibration of audiometric equipment—Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones”, International Organization for Standardization (Geneva).

Kates, J. M. (**2008**). *Digital Hearing Aids* (Plural Publishing, San Diego, CA), pp. 411–416.

Keidser, G., Rohrseitz, K., 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”, Int. J. Audiol. **45**, 563–579.

Larsby, B., Hälgren, M., and Lyxell, B. (**2008**). “The interference of different background noises on speech processing in elderly hearing impaired subjects”, Int. J. Audiol. **47**, 83–90.

Latham, H. G. (**1979**). “Signal-to-noise ratio for speech-intelligibility - auditorium acoustics design index”, Appl. Acoust. **12**, 253–320.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (**1999**). “The precedence effect”, J. Acoust. Soc. Am. **106**, 1633–1654.

Lochner, J. P. A. and Burger, J. F. (**1964**). “Influence of reflections on auditorium acoustics”, J. Sound Vib. **1**, 426–454.

Lochner, J. P. A. and Burger, J. F. (**1958**). “The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech”, Acustica **8**, 1–10.

Moore, B. C. J. (**2007**). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*, 2nd ed. (Wiley, Chichester), chapter 7.

Moore, B. C. J., Johnson, J. S., Clark, T. M., and Pluvinage, V. (**1992**). “Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss”, Ear Hear. **13**, 349–370.

Moore, B. C. J., Galsberg. B. R., and Stone, M. A. (**1999**). “Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression”, Brit. J. Aud. **33**, 241–258.

Musa-Shufani, S., Walger, M., von Wedel, H., and Meister, H. (**2006**). “Influence of dynamic compression on directional hearing in the horizontal plane”, Ear Hear. **27**, 279–285.

NáběIek, A. K. and Robinette, L. (**1978**). “Influence of precedence effect on word identification by normally hearing and hearing-impaired subjects”, J. Acoust. Soc. Am. **63**, 187–194.

NáběIek, A. K. and Picket, J. M. (**1974**). “Reception of consonants in a classroom affected by monaural and binaural listening, noise, reverberation, and hearing aids”, J. Acoust. Soc. Am. **56**, 628–639.

Nakajima, T. and Ando, Y. (**1991**). “Effects of a single reflection with varied horizontal angle and time-delay on speech-intelligibility”, J. Acoust. Soc. Am. **90**, 3173–3179.

Naylor, G. M. (**1993**). “Odeon - another hybrid room acoustical model”, Appl. Acoust. **38**, 131–143.

Naylor, G. and Johannesson, R. B. (**2010**). “Long-Term Signal-to-Noise Ratio at the Input and Output of Amplitude-Compression Systems”, J. Am. Acad. Audiol. **20**, 161–171.

Noble, W., Byrne, D., and Ter-Horst, K. (**1997**). “Auditory localization, detection of spatial separateness and speech hearing in noise by hearing impaired listeners”, J. Acoust. Soc. Am. **102**, 2343–2352.

Parizet, E. and Polack, J. D. (**1992**). “The influence of an early reflection upon speech-intelligibility in the presence of a background-noise”, Acustica **77**, 21–30.

Pavlovic, C. V. (**1987**). “Derivation of primary parameters and procedures for use in speech-intelligibility predictions”, J. Acoust. Soc. Am. **82**, 413–422.

Peters, R. W., Moore, B. C. J., and Baer, T. (**1997**). “Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people”, J. Acoust. Soc. Am. **103**, 577–587.

Plomp, R. (**1988**). “The negative effect of amplitude compression in multi- channel hearing aids in the light of the modulation-transfer function”, J. Acoust. Soc. Am. **53**, 1646–1657.

Pollack, I. (**1948**). “Monaural and binaural threshold sensitivity for tones and for white noise”, J. Acoust. Soc. Am. **20**, 52–57.

Richards, V. M., Moore, B. C. J., and Launer, S. (**2006**). “Potential benefits of across-aid communication for bilaterally aided people: Listening in a car”, Int. J. Audiol. **45**, 182–189.

Ricketts, T. A., Grantham, D. W., Ashmead, D. H., Haynes, D. S., and Labadie, R. F. (**2006**). “Speech recognition for unilateral and bilateral cochlear implant modes in the presence of uncorrelated noise sources”, Ear And Hearing **27**, 763–773.

Roeser R.J. and Clark J.L. (**2007**). “Clinical masking”, in: *Audiology: Diagnosis*, 2nd ed., edited by R.J. Roeser, M. Valente and H. Hosford-Dunn (Thieme, New York, NY), p. 263.

Shinn-Cunningham, B. G. (**2003**). "Spatial hearing advantages in everyday environments", Proceedings of the ONR workshop on Attention, Perception, and Modeling for Complex Displays, Troy, NY.

Sockalingam, R., Holmberg, M., Eneroth, K., Schulte, M., (**2009**). "Binaural hearing aid communication shown to improve sound quality and localization", Hear. J. **62**, 46–47.

Soulodre, G. A., Popplewell, N., and Bradley, J. S. (**1989**). "Combined effects of early reflections and background-noise on speech-intelligibility", J. Sound Vib. **135**, 123–133.

Stone, M. A. and Moore, B. C. J. (**1999**). "Tolerable hearing aid delays. I. Estimation of limits imposed by the auditory path alone using simulated hearing losses", Ear Hear. **20**, 182–192.

Stone, M. A. and Moore, B. C. J. (**2007**). "Quantifying the effects of fast-acting compression on the envelope of speech", J. Acoust. Soc. Am. **121**, 1654–1664.

Stone, M. A., Moore, B. C. J., Alcï£¡ntara, J. I. and Glasberg, B. R. (**1999**). "Comparison of different forms of compression using wearable digital hearing aids", J. Acoust. Soc. Am. **106**, 3603–3619.

Summers, V. and Molis, M. R. (**2002**). "Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level", J. Speech Lang. Hear. Res. **47**, 245–256.

Wagener, K., Josvassen, J. L., and Ardenkjaer, R. (**2003**). "Design, optimization and evaluation of a danish sentence test in noise", Int. J. Audiol. **42**, 10–17.

Wagener, K. C. and Brand, T. (**2005**). "Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters", Int. J. Audiol. **44**, 144–156.

Wittkop, T. and Hohmann, V. (**2003**). "Strategy-selective noise reduction for binaural digital hearing aids", Speech Comm. **39**, 111–138.

Contributions to Hearing Research

Vol. 1: *Gilles Pigasse*, Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials, Dec. 2008.

Vol. 2: *Olaf Strelcyk*, Peripheral auditory processing and speech reception in impaired hearing, Jun. 2009.

Vol. 3: *Eric R. Thompson*, Characterizing binaural processing of amplitude-modulated sounds, Aug. 2009.

Vol. 4: *Tobias Piechowiak*, Spectro-temporal analysis of complex sounds in the human auditory system, Sep. 2009.

Vol. 5: *Jens Bo Nielsen*, Assessment of speech intelligibility in background noise and reverberation, Dec. 2009.

Vol. 6: *Helen Connor*, Hearing aid amplification at soft input levels, Jan. 2010.

Vol. 7: *Morten Løve Jepsen*, Modeling auditory processing and speech perception in hearing-impaired listeners, May 2010.

Vol. 8: *Sarah Verhulst*, Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions, Jun. 2010.

Vol. 9: *Sylvain Favrot*, A loudspeaker-based room auralization system for auditory research, Jun. 2010.

Vol. 10: *Sébastien Santurette*, Neural coding and perception of pitch in the normal and impaired human auditory system, Jun. 2011.

Vol. 11: *Iris Arweiler*, Processing of spatial sounds in the impaired auditory system, Jul. 2011.

Insertion gain measures

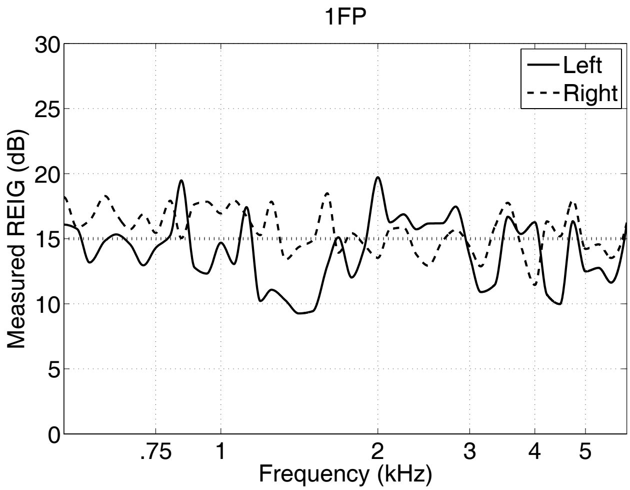


Figure 1: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 1FP.

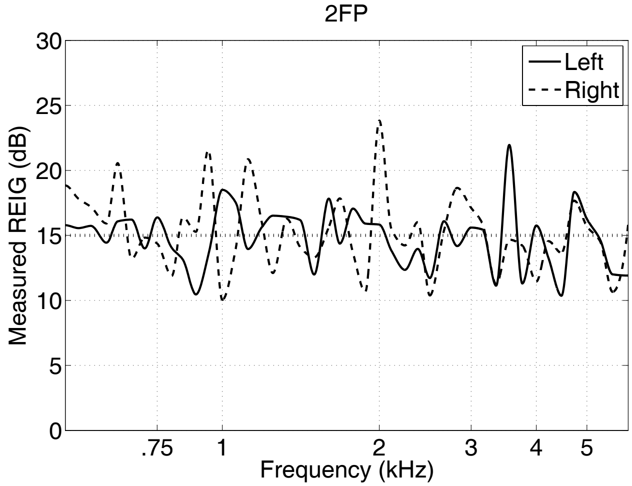


Figure 2: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 2FP.

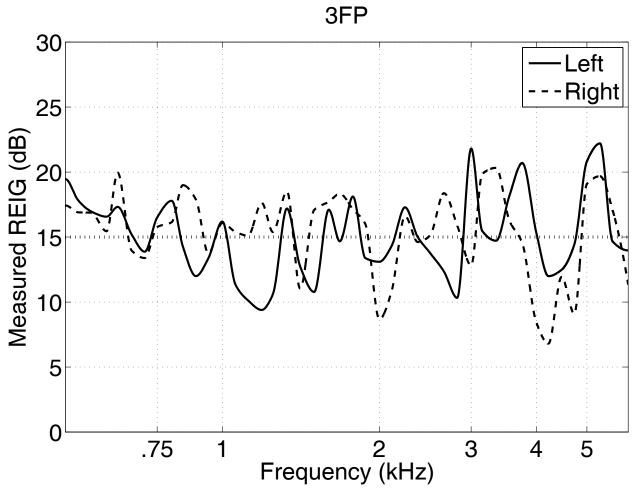


Figure 3: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 3FP.

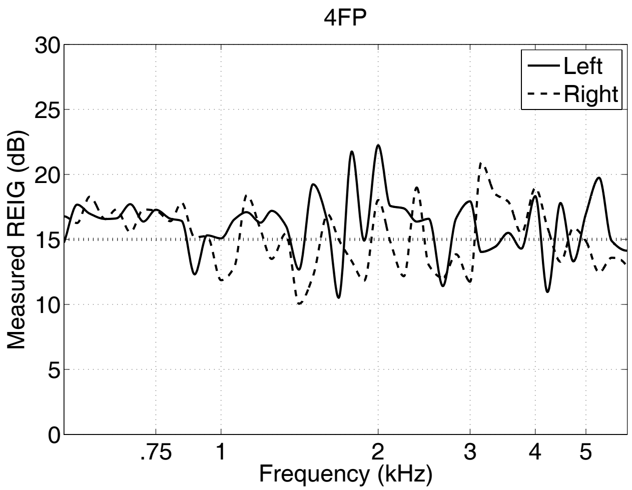


Figure 4: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 4FP.

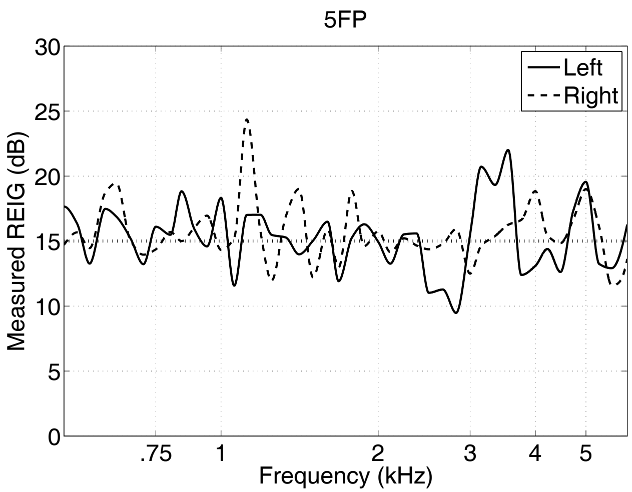


Figure 5: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 5FP.

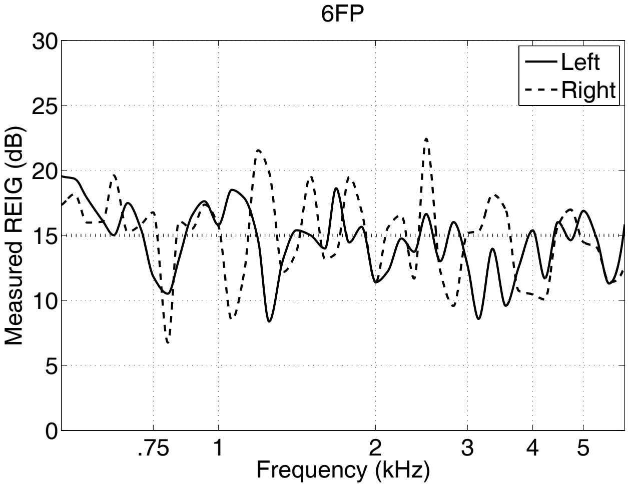


Figure 6: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 6FP.

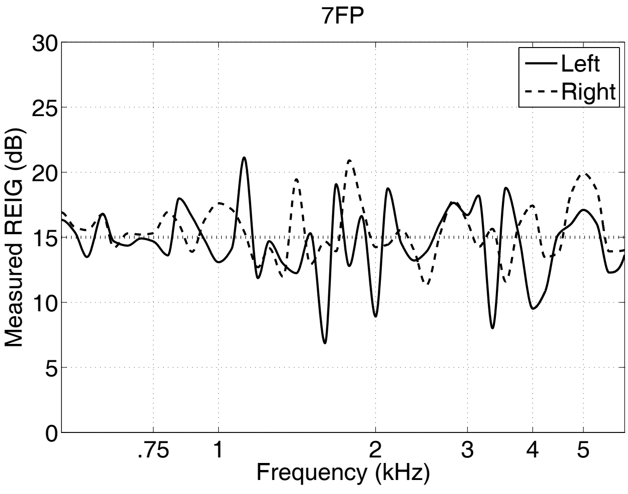


Figure 7: Insertion gain measured for left and right ear at an input level of 15 dB SPL. Listener 7FP.

Understanding speech in complex acoustic environments presents a challenge for most hearing-impaired listeners. In conditions where normal-hearing listeners effortlessly utilize spatial cues to improve speech intelligibility, hearing-impaired listeners often struggle. Here, the influence of two such spatial cues – early reflections and interaural level differences – on speech intelligibility was investigated in listeners with normal and impaired hearing using a virtual auditory environment and an independent hearing aid research platform. This work contributes to the understanding of early reflection processing in the auditory system and may have implications for speech perception models and the development of compensation strategies in future generations of hearing instruments.

Centre for Applied Hearing Research
Department of Electrical Engineering

Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
Fax: (+45) 45 93 16 34
www.elektro.dtu.dk

ISBN 978-87-92465-84-9