Effects of reverberation on speech intelligibility in normal-hearing and hearing-impaired listeners

ERH 2011-27-E

Author: Claudia Breitsprecher

Date: 12. December 2011

Classification: External

E R I K S H O L M

OTICON A/S

RESEARCH CENTRE ERIKSHOLM

243, Kongevejen

DK-3070 Snekkersten

Denmark

EFFECTS OF REVERBERATION ON SPEECH INTELLIGIBILITY IN NORMAL-HEARING AND HEARING-IMPAIRED LISTENERS

M.SC. THESIS OF

CLAUDIA BREITSPRECHER

SUPERVISED BY

TORSTEN DAU
CENTER OF APPLIED HEARING RESEARCH

SØREN LAUGESEN
ERIKSHOLM
OTICON RESEARCH CENTER

EFFECTS OF REVERBERATION ON SPEECH INTELLIGIBILITY IN NORMAL-HEARING AND HEARING-IMPAIRED LISTENERS

M.SC. THESIS PREPARED BY

CLAUDIA BREITSPRECHER

SUPERVISORS

TORSTEN DAU
Department of Electrical Engineering
Centre for Applied Hearing Research (CAHR)
Technical University of Denmark
Ørsteds plads, Building 352
DK - 2800 Kgs. Lyngby
Denmark

SØREN LAUGESEN
Eriksholm
Research Center Oticon A/S
Kongevejen 241
DK - 3070 Snekkersten
Denmark

Project period:	01-03-2011 – 30-09-2011
ECTS points:	35 points
Edition:	First
Comments:	This report is part of the requirements to achieve the Master of Science (M.Sc.) in Acoustical Engineering at the Technical University of Denmark.
Rights:	©Claudia Breitsprecher, 2011

ABSTRACT

The effect of reverberation on speech intelligibility in normal hearing (NH) and hearing impaired (HI) listeners was studied in two experiments. The first experiment examined the sentence intelligibility in the presence of different noises for monaural and binaural listening. The performance of both listening groups was negatively affected by reverberation. The obtained binaural benefit was bigger for the speech masker than for the non-speech masker. In addition, when expressed in terms of difference in SRT between the anechoic and the reverberant condition, the effect of reverberation was of the same magnitude for NH and HI listeners. The second experiment studied the perception of consonants in the absence of noise. Due to a ceiling effect for the NH listeners only the results for the HI listeners could be analysed. For those listeners, performance worsened in the reverberant condition compared to the anechoic condition and improved for binaural as compared to monaural listening.

The patterns arisen from the relationships of reverberation, listening mode (monaural vs. binaural) and the perception of speech were similar in both experiments. If the underlying mechanisms are similar for both sentences and consonants was, however, not obvious from the obtained results.

Keywords: Speech intelligibility, reverberation, hearing impairment, binaural listening

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisors, Søren and Torsten, who supported my throughout the entire project and who, in their own ways, brought the best out of me. This thesis would not be the same if it wasn't for our discussions and their review of earlier versions of this report.

A big thank you to all my colleagues at Eriksholm for making the last couple of months such a pleasant experience. In particular I would like to thank Peter Djørup, who never got tired helping me with *all* kinds of technical issues. In addition, I would like to thank Renskje Hietkamp for helping me out the recruitment of the test subjects and with the paper work that comes along with it.

I would also like to thank Morten Løve Jepsen for providing the consonant test software and for giving me the right direction in analysing the results. Thank you to Lise Bruun Hansen for sharing her knowledge about Danish consonants with me.

Further, I would like to thank Björn Ohl who provided the software for the room acoustic measurements and who made sure I used it correctly.

A very big thank you to all my test subjects for their time and patience.

I would like to thank Alexandre Chabot-Leclerc for letting me use his computer and his Latex skills to beautify my thesis. Also, a very big thank you to Richard McWalter and Edd Collett for native proofreading of the report.

I would also like to mention my fellow acoustic students who made the last two years so much more than just an academic experience and who I am sharing so many great memories with now.

In addition, I would like to express my gratitude to my parents who always encouraged me to go my own way and who never doubted that I would.

Last but certainly not least, I would like to thank Theo Harpel for putting up with me over the last months.

CONTENTS

1 INTRODUCTION

Listening to speech is an essential component of our everyday life. If the perception of speech is distorted, e.g. due to adverse listening conditions or hearing impairment, our communication ability is negatively affected. In everyday listening conditions a target speech signal is often partly masked by noise or by one or more simultaneous talkers that in most cases originate from different locations and distances than the target speech signal. In addition, the speech signal may also be distorted by adverse listening conditions such as room reverberation.

The extent to which the perception of speech is affected by adverse listening conditions and simultaneous talkers is strongly dependent on the hearing status of the listeners. Persons with mild to moderate hearing losses usually understand speech reasonably well in quiet environments when talking to only one person. Difficulties occur when more than one person is talking at once or when background noise or reverberation are present. People with severe or profound hearing losses usually have difficulties when talking to only one person in a quiet room and suffer even more when noise or reverberation are present (cf. Moore, 2007, Ch. 8). Normal hearing (NH) people also report to having difficulties understanding speech in reverberant listening conditions, although their difficulties do not seem to be as severe as for hearing impaired (HI) people.

The reasons why HI listeners have difficulties understanding speech are manifold. As a result of elevated hearing thresholds the suprathreshold spectral speech energy for unaided HI listeners is less compared to NH listeners. In addition, hearing losses are usually frequency dependent meaning that even if in some frequency regions the signal is well above threshold in others it might not be. Difficulties in understanding speech might therefore occur due to the fact that the speech signal is simply not audible, even if only in some frequency regions (Humes, 2007). In addition, it has been argued that problems in understanding speech occur due to a reduced ability to discriminate sounds that are well above threshold (Glasberg and Moore, 1988-1989). That is, even if the speech is audible, HI persons will still have problems segregating the speech from interfering sources and, hence understanding the speech.

While the effect of noise on speech intelligibility has been investigated quite well (for review see e.g. Moore, 2007, Ch. 8.II), the perception of speech in reverberant listening environments has not been explored well yet. Therefore, the focus of the present study was to quantify the difficulties of NH and HI listeners when listening to speech in a reverberant environment. Of particular interest is whether or not the two groups of listeners are affected differently when reverberation is present. The first experiment that is presented in Ch. 2 examined the sentence intelligibility in noise and reverberation. A speech signal, however, may be divided into different units. Starting with sentences the next smaller unit would be words. These can again be divided into smaller units, syllables. Syllables usually consist of a combination of consonants and vowels. The perception of consonants in reverberation was studied in a second experiment as described in Ch. 3. In both experiments monaural and binaural listening were compared. Before presenting methods and results of the named experiments useful background information is given and preceding studies related to the topic are reviewed in the respective chapters. After discussing the results of each experiment individually a general discussion is given in Ch. 4. Here the findings of the two experiments are compared with regard to the similarities of the relationship of reverberation, listening mode (monaural vs. binaural) and the different units of speech, respectively. Finally the entire project will be summarised in Ch. 5.

2

MONAURAL AND BINAURAL SENTENCE INTELLIGIBILITY IN NOISE AND REVERBERATION

Understanding speech in everyday listening situations can be a challenging task. Usually, the listener tries to focus his/her attention to a target talker while maintaining awareness of the entire auditory scene. Here, the perception of a target talker can be negatively affected by other competing talkers, other noise sources or reverberation. The success of communication in such situations is strongly affected by the listener's ability to perceptually segregate the target speech from interfering sources. There are a number of acoustic cues, such as difference in fundamental frequency, onsets and offsets, prosody, intensity levels and spatial location, that can help segregating the different sound sources. Spatially separating the sources, for instance, gives rise to different cues that help the listener to detect and understand the target talker. The spatial separation of target and the masker gives rise to different target-to-masker ratios (TMR) across the ears. This happens because the head and body of the listener interact with the incoming sound. This interaction is dependent on the sound wavelength. Hence, TMR changes with spatial location can mostly be found at high frequencies. Whenever the TMR differs across the two ears, the ear with the better TMR can directly improve the performance in speech recognition. This effect is usually referred to as better-ear-listening. Furthermore, differences in spatial location provide binaural processing cues such as interaural time differences (ITD) and interaural level differences (ILD) that help segregating the sources in addition to the benefit of increased intelligibility due to increased TMR. These binaural cues are especially helpful for understanding speech near threshold (Bronkhorst, 1999).

Whenever a listener finds himself/herself in a listening situation with various competing talkers he/she can take advantage of the stated cues arising from spatial separation to selectively attend to a sound source from a particular location and ignore the competing sound sources from other directions. However, the benefit of spatial separation is affected by a number of factors. For instance, it has been shown that a greater *spatial release from masking (SRM)* can be found for speech interferers rather than for noise interferers (see e.g. Noble and Perrett, 2002; Freyman *et al.*, 1999). This can be explained by different components contributing to masking. A target signal can be masked energetically by a competing signal when they occur in the same frequency region at the same time due to the overlap of their representations in the peripheral auditory system. Ongoing speech, however, fluctuates in amplitude and frequency. Thus, the amount of this so-called *energetic masking* on an interfering speech signal varies over time. Still, it is more difficult to segregate a target talker from an interfering talker compared to interfering noise. The reason for this is another component of masking that cannot be accounted for by peripheral masking. In this type of masking, both the target and the masker are audible but the listener is not able to perceptually segregate the signals and successfully draw attention to the target signal. This component of masking is usually referred to as *informational masking* (see e.g. Brungart, 2005; Marrone *et al.*, 2008a) and is more effective when the target and interfering signals are more similar (Brungart, 2001; Noble and Perrett, 2002). However, the mechanisms relevant for this kind of masking are not as clearly understood as those relevant for the purely peripheral energetic masking.

The amount of informational masking produced by an interfering signal is not easy to quantify. Therefore, Arbogast *et al.* (2005) and Arbogast *et al.* (2002) decided to use processed speech as interfering signals in order to control the amount of energetic and informational masking. To do so, three types of maskers, varying in the amount of informational and energetic masking, were designed. NH and HI subjects were tested in conditions where the target and interfering signals were either collocated or spatially separated. The mean SRM was found to be bigger for the interferer causing mainly informational masking compared to those causing purely energetic or both informational and energetic masking. In addition, the SRM was only significantly different between NH and HI listener for the interferer causing mainly informational masking and not for the other two. However, even though the amount of SRM in HI listeners was smaller than in NH listeners, the results implied that NH as well as HI listeners are able to use the spatial separation of target and interferer to decrease informational masking. As pointed out by Arbogast *et al.* the spatial separation of the signals enables the listener to label each signal accordingly as target and interferer resulting in less confusion between the two, thus reducing informational masking.

However, since the spatial separation was achieved by presenting the interferer from only one side of the listener it remained unclear to what extent the head-shadow effect and binaural interaction contribute to the advantage of spatial separation. In a symmetrical setup, however, the head-shadow effect on the one side of the listener will be counterbalanced with that on the other side (Bronkhorst and Plomp, 1992) and thus, better-ear-listening will be minimized. Noble and Perrett (2002) conducted a number of experiments using a setup where the competing signals were displaced symmetrically from the target speech signal at 0°. As well as for the studies using asymmetrical setups a SRM was found in this study. Again, the biggest benefit of spatial separation was found for the speech interferer compared to the noise interferer. Another interesting finding is that the SRM is bigger when the interfering talker is of the same gender as the target talker. This can be related to findings by Brungart (2001) that showed that the performance in speech recognition was worst when the target and interfering talkers were of the same gender and to the general assumption that the SRM is primarily due to the release from informational masking (Freyman *et al.*, 1999; Arbogast *et al.*, 2005).

To what extent the acoustic properties of a room affect the speech intelligibility of NH and HI listeners has not been the focus of many studies to date. It is, however, generally assumed that reverberation reduces the fast temporal modulations of the speech signal by filling minima in the envelope and, hence smearing the temporal envelope of the speech signal. This reduction in the signal's temporal fluctuation reduces the intelligibility of sentences (see e.g. Duquesnoy and Plomp, 1980). In addition, the efficiency of the cues arising from spatial separation is likely to be reduced. As reverberation and echoes interact with the signals reaching the ears, the TMR at the ears are modified compared to non-reverberant room conditions (see e.g. Shinn-Cunningham, 2003; Breitsprecher, 2009). Thus, the advantage of improved TMR arising from spatial separation is reduced. In addition, reverberation decorrelates the signals at the two ears which in turn reduces the efficiency of binaural processing cues. As a result, the segregation of target and interfering signal is not as obvious as in anechoic conditions. To what extent this affects the SRM has not yet been examined in detail. In addition, it is not obvious if NH and HI listeners are affected in the same way.

Another factor to affect the segregation of two competing signals is the listening mode, i.e. whether a person is listening monaurally or binaurally. A study by Braasch and Hartung (2002), for instance, examined the detection of a broadband noise in the presence of a second broadband noise for monaural and binaural listening. It was found that the difference between the listening modes was only minimal when the competing signals were collocated. When the signals were spatially separated, however, binaural listening was superior to monaural listening. Thus, as pointed out by the authors, it can be assumed that binaural cues are important when segregating two concurrent sounds that are separated spatially in the horizontal plane. When listening monaurally, however, the lack of these cues impairs the ability to discriminate a target from an interfering sound and other cues are necessary to segregate the sounds. In addition, in their study, the performance was poorer in reverberant compared to anechoic listening conditions. This indicates that sound segregation is affected negatively in reflective environments. It has been shown that monaural listening nearly eliminates the benefit of spatial separation when listening to speech in the presence of competing talkers. In a study by

Marrone *et al.* (2008b) the SRM was examined for a three-talker interferer. The interferer was either presented collocated with the target or symmetrical positioned around the target speech signal. This spatial separation resulted in a release from masking. The benefit from spatial separation, however, nearly vanished for monaural listening. Marrone *et al.* (2008b) showed that, when the interferers were presented at $\pm 90°$ the masking thresholds were approximately equivalent to the collocated thresholds. A control measure, where both ears were closed but the target level was increased to restore audibility, the SRM was restored. The conclusion was that, since the loss of normal pinna cues was not the reason for the eliminated release from masking, the advantages found in the study depended on binaural listening. Another interesting finding from that study is that the amount of reverberation in the listening environment influences the magnitude of the SRM. Increasing the reverberation reduced the spatial release. However, all listeners tested in this study were NH. It remains unclear if the findings can be applied to HI listeners.

The main goal of the present study is to examine the relationship between sentence intelligibility and reverberation for NH and HI listeners and, if the listening groups are affected differently. A comparison of speech intelligibility of NH and HI listeners in reverberation was not available in any of the cited studies. It is, however, hypothesized that reverberation reduces the speech intelligibility for both listening groups. Based on the difficulties reported by HI listeners in reverberant environments it can be assumed that HI are more affected by reverberation than NH listeners. It is also of interest if different kinds of interferers are able to mask target speech differently. Due to the higher similarity to target speech, a speech interferer should be able to mask the target speech more effectively than a noise interferer. In addition, binaural and monaural listening are to be compared. Given the results from e.g. Marrone *et al.* (2008b) and the general assumption that binaural listening provides additional cues to separate competing signals, binaural listening is hypothesized to be superior to monaural listening.

2.1 METHODS

2.1.1 *Rooms*

The experiments were conducted in two different rooms. One of them is the anechoic chamber of the research center Eriksholm. The chamber has a inner volume of 4.3 x 3.4 x 2.7 m^3 and has been shown to be anechoic down to a frequency of 100 Hz. The second room is a rather big room (V = 8.95 x 5.68 x 4.67 m^3) with walls that are made of acoustically hard materials. In order to estimate the reverberation time of the room, the room impulse response (RIR) was measured for a number of loudspeaker and microphone locations. This was done using a custom-made software package based on MatLab®. More information regarding the software package to measure the RIRs is available in Ohl (2009). From the measured RIRs, however, decay curves were calculated using the backward Schroeder integration (Schroeder, 1965) from which the reverberation time can be determined. The calculations to obtain the decay curves were integrated in the same software package mentioned before. Since the level of the background noise was rather high, it was only possible to determine the reverberation time in terms of T_{20}. T_{20} is determined from a decay of 20 dB, starting at -5 dB. This decay time is then extrapolated to a decay of 60 dB assuming that the part of the decay curve is representative for the entire decay. The resulting reverberation time, as shown in the left panel of Fig. 2.1, was determined to be approximately 0.8-0.9 s in the medium and high frequencies. However, due to a high reverberation level the room is perceived as more reverberant. Having this moderate reverberation time this room offers the ability to examine the effect of reverberation whilst keeping the room acoustics comparable to everyday listening conditions. In addition, the room was also large enough to accommodate the setup used for the listening experiments as shown in the right panel of Fig. 2.1. A factor determining if the room was large enough was whether the distance between loudspeaker and listener could be beyond the critical distance d_{crit} of the room. The critical distance was estimated with the help of a level meter. While presenting a stationary noise signal from one of the loudspeakers it was determined at which distance the level did not decrease anymore when moving away from the source. d_{crit} is defined as the point where the levels of direct and reverberant sound field are equal. In the given room, d_{crit} was estimated as approximately 2.5 m.

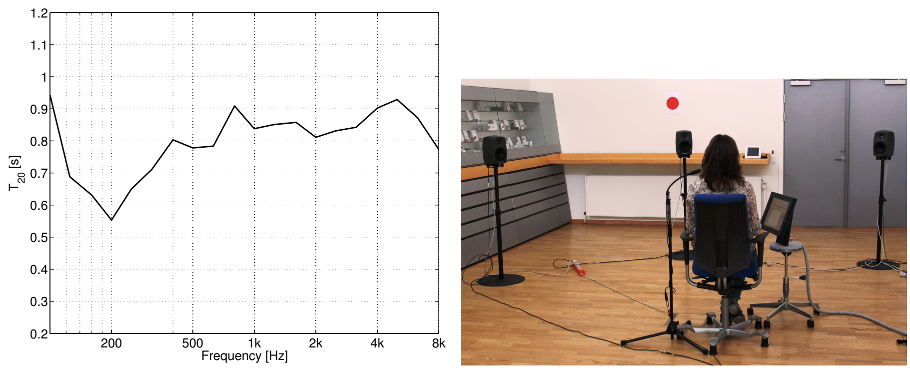

Figure 2.1: Left side: Frequency dependent reverberation time measured in the reverberant room used for the listening experiments.
Right side: The setup used in this experiment as installed in the reverberant room. For further explanations about the setup, see Sec. 2.1.2.

2.1.2 *Listening experiments*

Sentence intelligibility was tested using the Danish version of the hearing in noise test (HINT) (Nielsen and Dau, 2011). The HINT measures the listener's ability to understand speech in the presence of noise. In the present study two different types of noise were used. As pointed out before, reverberation modifies the temporal modulations of a signal and, thus the interferers used in this study ought to have modulations in the temporal domain. In contrast, it is expected that the use of stationary noise would not be able to show the effect of reverberation. In order to minimize differences in segregation cues arising from the temporal modulations of the maskers it is considered advantageous that the modulations of the two maskers are as similar as possible. As also mentioned before, the similarity between target and interfering signal has an high impact on how well the target signal can be segregated from the interfering signal (Brungart, 2001). Therefore both interfering signals are to be of the same gender as the HINT talker, i.e. male. Thus, one of the two interferes is a two-talker babble obtained from running speech of two male talkers reading H.C. Andersen's 'Nattergalen' (The Nightingale). In order to avoid short-term better-ear advantages due to long talking pauses, these were removed for both talkers using Adobe Auditon CS 5.5. In the further course of this report, this interferer will be referred to as *'speech masker'*. The second masking signal is the one-speaker male ICRA noise (Dreschler *et al.*, 2001), i.e. a modulated noise having long-term average frequency spectrum and modulation properties equivalent to a single male talker. As well as for the speech masker, long pauses in the signal were removed. Throughout this report this masking signal will be referred to as *'Non-speech masker'*. In order to avoid different spectral cues among both maskers that might result in that one of the masker is easier to segregate from the target masker, the long-term average frequency spectra of both interferers were matched to that of the HINT sentences. The matching is illustrated in Fig. 2.2. A correction filter was designed from the frequency-dependent differences between the spectrum of the HINT sentences and the signal to be matched. The filter was then applied to the respective signal.

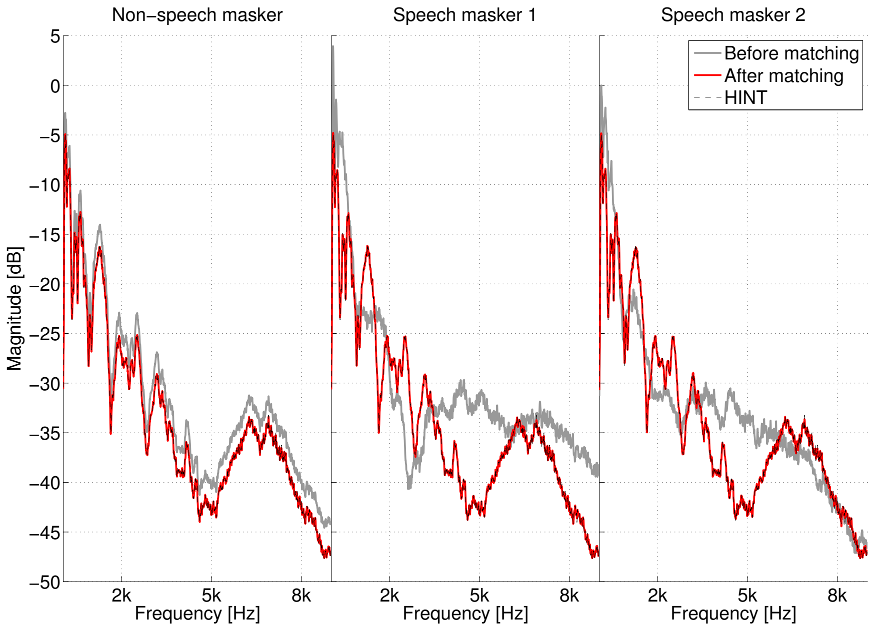

Figure 2.2: The three signals used to mask the target speech signal, i.e. HINT sentences, before and after applying a correction filter to match the long-term average spectrum to that of the target speech. The most left panel shows the non-speech masker and the middle and right panel show the two male talkers the speech masker is constructed of.

The setup used for the listening experiments is shown in Fig. 2.3. Three loudspeakers were positioned at different horizontal angles relative to the listener, that is $0°$ and $\pm45°$. The target signal was presented from the front loudspeaker while the interfering signals were always presented from the loudspeakers at $\pm45°$. In case of the speech masker, one of the talkers was presented from each loudspeaker, respectively. The non-speech masker was presented from each of the loudspeakers. In both cases, the signal presented from one loudspeaker was delayed relative to the other loudspeaker. The symmetrical separation of the interfering signals was chosen in order to minimize better-ear listening due to the head shadow effects described earlier. During the experiment the test subject was seated in the middle of the setup and was instructed to look straight ahead to the loudspeaker positioned at $0°$. The distance between loudspeaker and listener was approximately 2.7 m in the reverberant room which is beyond the critical distance of the room.

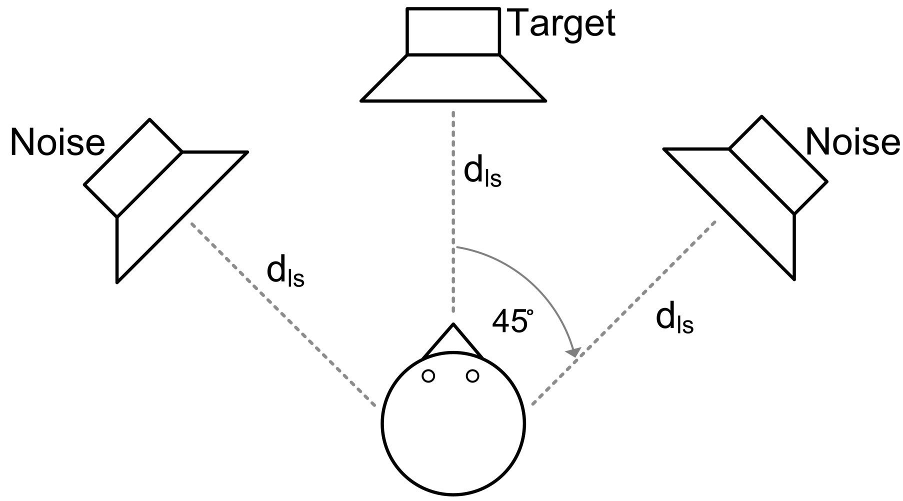

Figure 2.3: The setup used in the listening experiment. The target signal was presented from the front loudspeaker and the interfering signals were presented from two loudspeaker that were displaced by ±45° from the front loudspeaker. The distance between the listener and the loudspeaker d_{ls} was 2.7 m in the reverberant room which was beyond the critical distance.

During the test, the test subject was asked to repeat the sentences coming from the front loudspeaker. Based on the sentence score of the subject the level of both the speech and the target signal was then varied adaptively. With the variation of the level of both target and masking signal the TMR was changed whilst keeping the overall level constant. This TMR adjustment was chosen in order to avoid very high output levels when testing with HI test subjects (see Ch. 2.1.4). For every tested condition, the speech reception threshold (SRT), i.e. the TMR for understanding 50% of the speech correctly, was determined after 20 sentences (cf. Biologic Systems Corp, 2005, Appendix B).

Speech intelligibility was tested both monaurally and binaurally. The importance of closing the ear completely when comparing monaural and binaural listening in reverberant conditions was demonstrated by MacKeith and Coles (1971). Therefore, when testing monaurally one ear was closed using insert foam phones, which then again were closed with wax. In addition, the subject was wearing an ear muff on the side that ought to be closed. Whether or not closing one ear like this results in strictly monaurally listening has not been verified. It is, however, expected that it is closed well enough to assume monaural listening over a wide frequency range. For the NH subjects the subject's least favourable ear was closed for the monaural condition. If the subject did not have a more favourable ear, the right ear was closed. For the hearing impaired listeners, the ear with the poorer hearing threshold was closed and, if the thresholds were very similar, i.e. variations of $\leqslant$ 5dB across ears, the right ear was closed.

The listening experiments were conducted in two sessions. During the first session all conditions in the reverberant room were measured. In the second session, which was scheduled 4–6 weeks after the first one, the listening experiments were repeated in the anechoic chamber. In both sessions the four experimental conditions arising from the two types of interferer and the two listening conditions were tested in a balanced order according to the Latin-square design in order to avoid confounding effects. Balancing across the test rooms was not possible due to construction works near the reverberant room during the time period of the experiments. A high background level would have disturbed the measurements in the reverberant room and, hence the measurements in this room needed to be finished before the construction works began.

2.1.3 *Listeners*

Five NH and 12 HI listeners participated in the test. All NH participants were employees at Oticon, their average age was 31.8 and ranged from 25 to 38 years. The HI listeners were recruited from Erikholm's client database if they had a mild to moderate sloping hearing loss that did not exceed 75 dB HL and did not differ more than 15 dB across ears in any of the audiometric frequencies. They were not paid for their participation but reimbursed for their travel expenses. The tone audiograms of the HI participants, measured with insert ear phones, are shown in Fig. 2.4. The average age of the HI subjects was 64.5 years and ranged from 50 to 75.

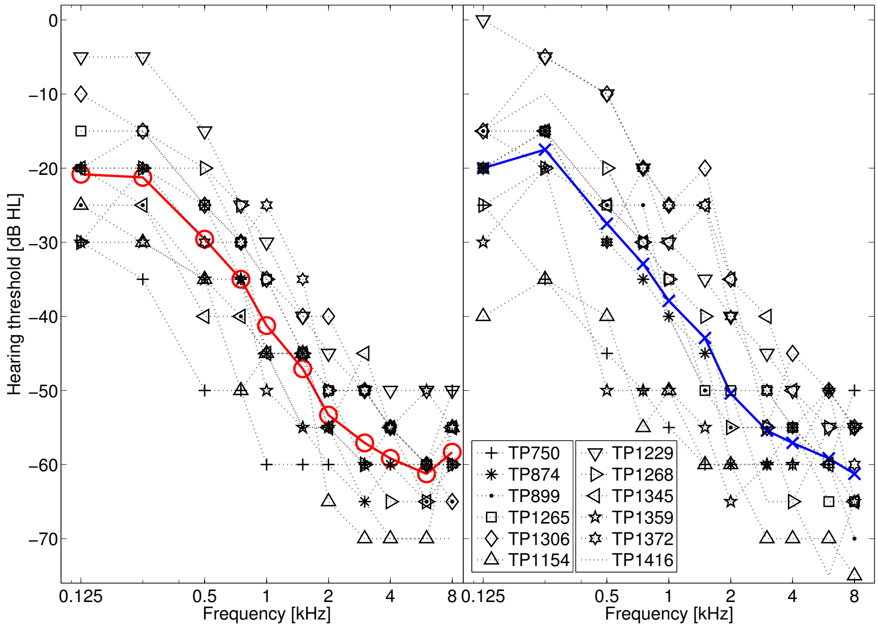

Figure 2.4: Tone audiograms measured with insert ear phones for the right and left side of the hearing impaired participants. The average hearing loss across all subjects is indicated by the straight lines where the circular markers (o) indicate the right side and the crosses (x) the left side.

2.1.4 *Hearing loss compensation*

The microphone location in the hearing aid affects how the incoming sounds are filtered due to the head and pinnae and this has an impact on the resulting signal-to-noise ratios (SNR) at the listener's ears (see e.g. Breitsprecher, 2009). Hence, if the HI subjects were fitted with hearing instruments during the listening experiments, the SNR at the ears would be different compared to those the normal hearing subjects are exposed to. In order to avoid differences across the listening groups in the way the incoming sound is filtered, the HI subjects were tested without hearing aids. Another reason not to use hearing aids in this experiment was to avoid bandwidth and gain limitations caused by hearing instruments. However, to assure audibility the signals presented to the HI subjects needed to be amplified. As can be seen in Fig. 2.4 the HI test subjects suffered from high-frequency hearing loss. Hence, simply presenting signals at high presentation levels would not make some of the high-frequency signal energy audible and thus would not eliminate the problem of inaudibility. Therefore all signals presented to the HI subjects were spectrally shaped depending on the individual hearing loss of the subject. This shaping ensures that the signals were sufficiently audible for the test subject. According to Humes (2007) a signal is sufficiently audible for the listener if the rms spectrum of the stimulus is 10–15 dB above the hearing threshold in that same frequency region.

Hence, in order to assure audibility for every HI test subject a filter was designed from a frequency dependent gain based on the individual hearing threshold of the subject. The gain to assure sufficient audibility was determined for the spectrum of a signal constructed from all the HINT sentences and the filter designed from this was then applied to all the signals presented to the test subject. This ensured that all signals were amplified in the same way. To determine the gain that is necessary to make the signals sufficiently audible the spectrum constructed from the HINT sentences was first summed in third-octave bands. Each third-octave band rms was then scaled so that the overall free-field sound pressure was the same that would be presented to a NH subject, which in the present study was 65 dB SPL. Finally, for every third-octave band it was determined if the signal is sufficiently loud by checking if the signal level in the given band was above the sum of the test subject's minimum audible field (MAF) and a value determining how much above the hearing threshold the signal should be (see Fig. 2.5). The calculation of the MAF was based on data from Moore *et al.* (2008). If the signal was not loud enough a gain value was be determined assuring that the rms in the given frequency band exceeds the test person's MAF level by a certain amount. This amount was 15 dB in the frequencies bands with centre frequency up to 2.5 kHz and then down to 4 dB at the band with f_c = 8 kHz. This was necessary to avoid too high output levels. Likewise, the gain was not increased below 250 Hz and above 8 kHz.

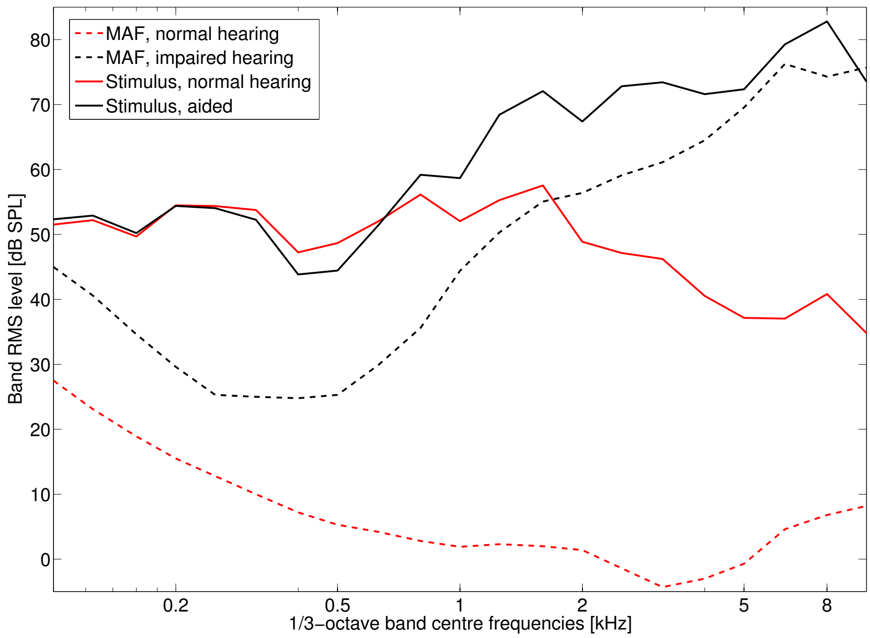

Figure 2.5: Illustration of the hearing loss compensation using the example of one HINT sentence: Minimum Audible Field (MAF) for normal and impaired hearing (assumed as right-left average), and spectrum of the stimulus for normal and impaired hearing (aided). The free-field level of the stimulus presented to the NH listeners was set to 65 dB SPL and the rms level of the stimulus presented to the HI listener was raised above threshold by 15 dB from 0.2-2.5 kHz and down to 4 dB at 8 kHz.

2.2 RESULTS

The sentence intelligibility, expressed in terms of the SRT in dB TMR, for all masking conditions in the two different rooms is summarised in Fig. 2.6. Overall, the NH subjects performed better than the HI subjects, as indicated by their lower SRT. In addition, both listening groups performed worse in reverberation than in anechoic conditions. Furthermore, binaural listening was superior to monaural listening for both rooms and for both types of interferers. The benefit of binaural listening was, however, bigger for the speech masker compared to the non-speech masker. This is true for both listening groups and listening environments.

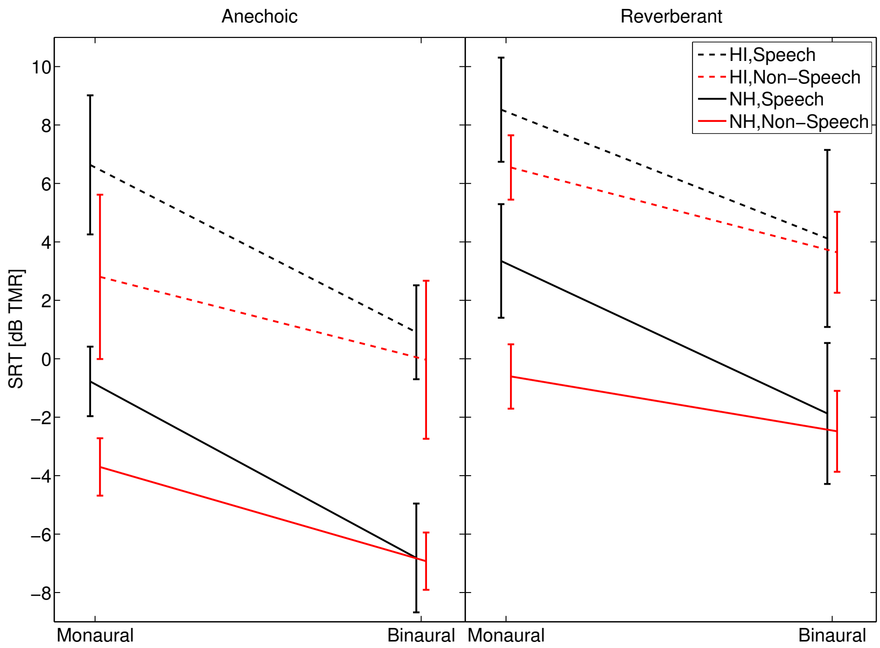

Figure 2.6: Mean SRT in dB TMR in anechoic and reverberant conditions averaged across 5 NH and 12 HI listeners. The SRT was measured in the presence of two types of noise, i.e. a speech masker and a non-speech masker, for monaural and binaural listening conditions. The error bars represent one standard deviation in each condition.

A repeated measures analysis of variance (ANOVA) with hearing status as predictor and room, masker and listening mode (monaural or binaural) as repeated measures was used in order to identify factors that significantly affect the sentence intelligibility in the given experiment. All main effects were proven to have a significant effect on the intelligibility on a significance level of $\alpha = 0.01$ (hearing status: $p < 0.001$, room: $p < 0.001$, masker: $p < 0.001$, listening mode: $p < 0.001$, see also Tab. A.1). In addition, the interaction of masker and listening mode was found significant on a level of $\alpha = 0.01$ ($p < 0.001$). No other interaction was found to be significant.

In order to examine if the two listening groups are affected in the same way by reverberation, the difference between the SRT obtained in the anechoic condition was subtracted from that in the reverberant condition. The mean differences can be seen in Fig. 2.7.

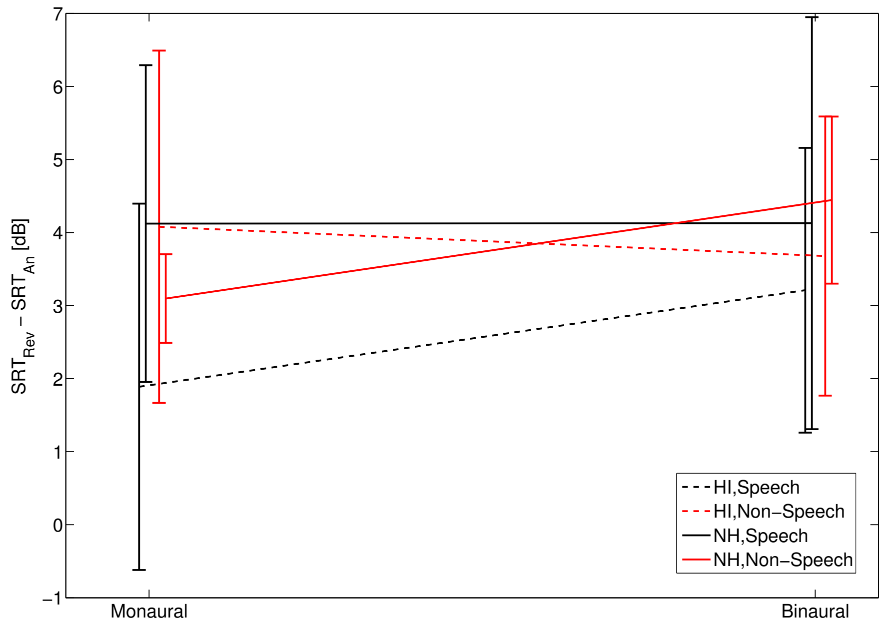

Figure 2.7: Mean differences of the SRT in the reverberant and the SRT in the anechoic condition for speech and non-speech masker for the NH and HI listeners. The positive values indicate a higher SRT in the reverberant environment. The error bars represent one standard deviation in each condition.

A repeated measures ANOVA with hearing status as predictor and masker and listening mode as repeated measures was used to identify factors that might affect the difference in performance between anechoic and reverberant listening condition. None of the named main effects was significant on a level of $\alpha = 0.01$ (see Tab. A.2). This is in line with the mean difference in performance between reverberant and anechoic listening environment presented in Fig. 2.7 that do not show a clear difference between the listening groups, the listening mode or the type of masker.

2.3 DISCUSSION

The main question of this experiment was whether or not NH and HI listeners are affected similarly by reverberation when listening to speech in the presence of interferers. As pointed out in Sec. 2.2 reverberation significantly affected the performance in the experiment. Interestingly, the interaction of room and hearing status was non-significant ($p = 0.26$). In addition, the difference between the performance in the reverberant and the anechoic environment, as presented in Fig. 2.7, was not affected significantly by the hearing status of the listener. This indicates that both listening groups are affected similarly by reverberation which is in contrast to the hypothesis constructed earlier. However, even though both listening groups were affected similarly when reverberation was added to a listening situation on an objective scale, the added reverberation might have been perceived as more disturbing by the HI listeners. Considering a listening situation where the task is to understand a target talker in the presence of competing talkers that talk at the same level, the TMR is approximately 0 dB if not higher. As shown in the previous section the negative TMR for the NH listeners indicates that it is rather easy for them to understand the target speech in anechoic conditions. When reverberation is added, however, the task becomes somewhat more difficult for them. For the HI listeners, in contrast, the task of understanding the target signal in the presence of competing signals is already difficult in the anechoic condition, as indicated by their higher SRTs. The added reverberation may therefore be perceived as even more disturbing than for the NH listeners.

In general, the HI listeners performed worse than the NH listeners for the presented experiment. As pointed out above, this poorer performance is not affected by the room. Interestingly, it is neither affected by the type of interferer and the listening mode as indicated by non-significant interaction of hearing status and any of the other main effects ($p = 0.73$ and $p = 0.44$, respectively).

Binaural listening was always superior to monaural listening and, disregarding the differences in the listening modes, the non-speech interferer was less effective than the speech masker. This is in line with other studies that compared the the speech intelligibility when the target speech is competing with either a speech or a noise masker (see e.g. Freyman *et al.*, 1999; Noble and Perrett, 2002). Interestingly, the difference between the obtained results for the two maskers was bigger for the monaural listening mode. In this case, the speech reception was worse when the interferer was speech compared to when it was non-speech. For binaural listening, however, the differences between the maskers nearly vanished. This effect may be compared to results obtained in experiments where the SRM was found for different kind of maskers. As pointed out by Freyman *et al.* (1999) the SRM can also be found when the spatial separation is only perceived. By utilizing the precedence effect Freyman *et al.* (1999) created an illusion of target and interfering signal and found the same effects that can be seen for an actual separation of the signal. That is, the SRM is bigger for speech than for non-speech maskers. In the present experiment the spatial separation of the competing signals was only perceived when listening binaurally. In the monaural case, however, the listeners were no longer able to distinguish the source locations (cf. Braasch and Hartung, 2002). This indicates that the cues used to segregate the competing signals were different in the monaural listening mode and in the binaural listening mode. When listening binaurally the interaction cues arising from the spatial separation of the signal seem to be sufficient to perceptually segregate the competing signals and correctly identify the target signal. These binaural cues are not available when listening monaurally and hence other cues are necessary to segregate the signals. All interfering signals were chosen to represent the same gender as the target signal and their long-term average spectra were matched to the target. Therefore, the cues to segregate the signal are unlikely to be found in the spectral domain. In addition, the selected interfering signal differ only little in their temporal envelope. Differences in their temporal fine-structure might have contributed to the effectiveness of the maskers. The cues arising from that are, of course, also available when listening binaurally but the cue of spatial location seems to be predominant in that case. The insignificance of the interaction of hearing status and listening mode and interferer, respectively, indicates that the cues used to segregate the competing signals are available and used to the same extent by the NH and HI listeners.

Still, as pointed out before, the HI listeners generally performed worse than the NH listeners. Since audibility was ensured by spectrally shaping the stimuli for the HI listeners, the reason for that offset might be due to some suprathreshold mechanisms.

3 MONAURAL AND BINAURAL CONSONANT PERCEPTION IN REVERBERATION

The findings from the first experiment are valid for the reception of sentences in adverse listening situations. To what extent reverberation and the listening mode affect the reception of smaller units of speech, i.e. consonants, in NH and HI listeners was studied in a second experiment. Again, it is of interest if NH and HI listeners are affected similarly. In addition, it is of interest whether the same dependencies on reverberation and listening mode found in the first experiment can be observed for consonants as well.

While the effect of different noises on the reception of consonants has been studied quite well (see e.g. Phatak and Allen, 2007; Phatak *et al.*, 2008), the effect of reverberation and listening mode has been the interest of only few studies. Nabelek and Pickett (1974), for example, examined the effect of reverberation and listening mode on the perception of consonants. In this experiment, the consonants were presented in the presence of both steady and impulsive noise to NH listeners. The obtained results were analyzed both in terms of overall percent correct reception and in more specific ways to examine the effect of reverberation on different phonetic features of the consonants. The more specific analyses might be able to discover any differential effects of the experimental conditions on different groups of consonants. One of the findings from that study was that both listening mode and reverberation affect the overall percent correct perception of the consonants. The performance of the listeners was better when listening binaurally and in listening environments with less reverberation. In addition, the advantage of binaural hearing was reduced with prolongation of the reverberation time. With respect to the binaural gain the authors concluded that the benefit was due to the additional information available at the second ear. In combination with the information from the other ear, the consonant reception is improved. Regarding the poorer performance with increasing reverberation the authors concluded that in addition to the direct sound, more energy is reflected from the walls and ceiling. As a result the temporal structure of the speech is smeared. Not only does it impair the reception of the consonants, it also reduces the efficiency of the binaural cues. When analysing the effect of listening mode and reverberation on different phonetic features the binaural gain was found to be bigger for the manner and voicing feature than for the place feature. The authors concluded that this was due to the fact that binaural listening provides low-frequent information which then again contains more voicing and manner cues.

In terms of reverberation, the place feature suffered more than the manner and voicing feature. It should, however, be noted that all of these findings were obtained in the presence of noise. The effects of reverberation alone or noise alone on the perception of consonants was not studied.

In contrast, Helfer (1994) tested the perception of consonants also in the absence of noise. In this way, it could be studied how reverberation alone and noise alone affect the reception of consonants, respectively. One of the findings was that consonant error patterns produced in reverberation and noise are indeed relatively independent. This suggests that reverberation and noise interfere with the phonetic features of the consonants in different ways. When analysing the consonant perception in terms of percent-correct scores Helfer (1994) found that binaural listening was superior to both monaural and diotic listening. The difference in performance between monaural and diotic listening, however, was insignificant. Based on this it was concluded that the binaural advantage was not due to a simple increase of intensity resulting from binaural summation. The biggest binaural benefit appeared for the consonant group of back plosives (cf. Tab. 3.1). It was suggested that this was due to the frequency dependence of the binaural benefit. An advantage of binaural listening usually relies on low-frequency information which is contained to a substantial amount in back plosives.

In the present experiment the effect of reverberation and listening mode on the perception of consonants was studied for NH and HI listeners. In accordance with the findings from the first experiment and the studies mentioned above it is hypothesized that reverberation affects the consonant perception negatively. Again, a comparison of the performance of the two listening groups was not available in any of the cited studies. Just like in the first experiment it is, however, assumed that the consonant perception in reverberation is better for NH than for HI listeners. Binaural listening is hypothesized to increase performance compared to monaural listening. Since binaural listening reduces the perception of reverberation (Koenig, 1950) the benefit should be even bigger in reverberant listening conditions. In addition, the amount of the binaural benefit should be dependent on the frequency content of the consonants. A bigger binaural benefit should be obtained for those groups of consonants providing more low-frequeny information such as back plosives.

3.1 METHODS

The listening experiment to study the perception of consonants was conducted in the same two rooms described in Sec. 2.1.1. In this way, the effect of reverberation on the consonant perception could be examined. During the experiment the listener was seated in the same setup as in the previous experiment, see Fig. 2.3.

To test the consonant perception a consonant-vowel (CV) test for Danish consonants was used. Details regarding the development of the test can be found in Jepsen and Dau (2011). The test, as used here, consisted of 15 Danish consonants that were followed by the vowel /a/. In general, consonants may be divided into different groups depending on the place and manner of their production. The grouping of the consonants used in this experiment can be found in Tab.3.1. Please note, that this grouping is only valid for Danish consonants that are placed initially like in the used CV test. Consonants of other languages may be produced differently.

	Place of Articulation				
	Front			*Back*	
Manner of Articulation	Labial	Dental/ Alveolar	Palato alveolar	Palatal/ velar	glottal
Plosives	b p	d t		g k	
Nasals	m	n			
Fricatives	v f	s	sj		h
Lateral		l			
Approximant				j	

Table 3.1: Grouping of the Danish consonants used in the experiment in terms of place and manner of articulation when placed initially as in the used CV material. Grouping in accordance with Grø nnum (1998).

As pointed out by Helfer (1994), consonant perception is affected differently by reverberation and noise. Since the focus in the present study is the effect of reverberation the consonant perception in this experiment was tested in the absence of noise. In addition, using a modulated noise which is more likely to be affected by reverberation than stationary noise is not trivial. Since the stimuli in this test are very short the detectability would vary over time with a noise that varies in time. The CV-sounds were presented from the loudspeaker straight ahead of the listener. The listener's task was to indicate the perceived consonant on a touchscreen that displayed all of the consonants available in the test, cf. Fig. 3.1. If in doubt, the listener was allowed to repeat the presentation of the CV-sound once. In addition, when sure that the given response was incorrect, the test subject was allowed to regret the last response in which case the last CV-sound was presented again. Feedback about the correctness of the response, however, was not provided.

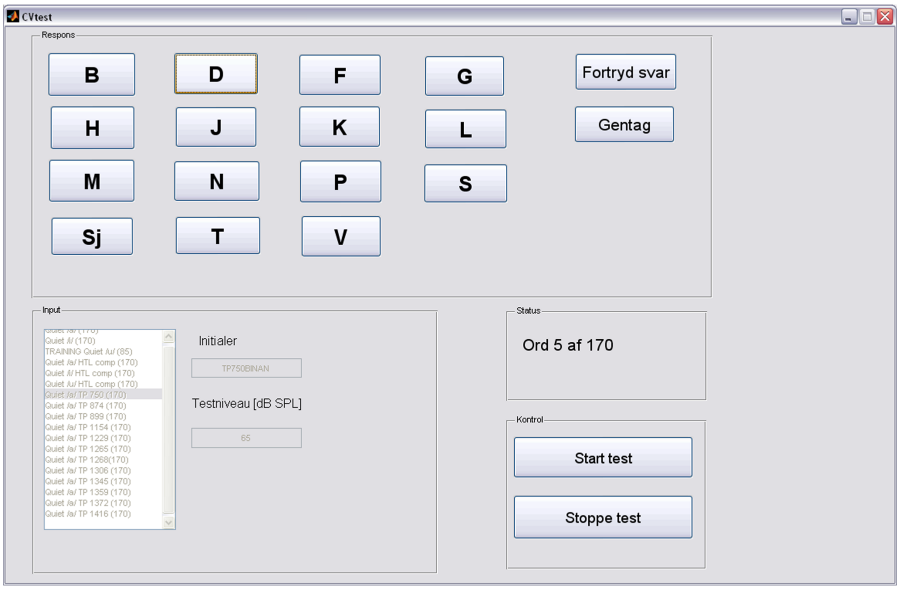

Figure 3.1: Graphical user interface used in the CV test. The test subject indicated the perceived consonant after listening to the stimulus. The listener was allowed to repeat the presentation and to regret the given answer.

The perception of the consonants was tested both monaurally and binaurally. The monaural listening condition was achieved in the same way as in the previous experiment, see Sec. 2.1.2.

The listeners who participated in this test were the same as in the previous experiment, see Sec. 2.1.3. Again, modified filtering of the incoming sounds due to hearing aid microphone location was not desired. Hence, the HI listeners were tested without hearing aids. But, just like in the previous experiment, in order to ensure audibility the stimuli were spectrally shaped depending on the individual hearing loss. The gain filter applied to the stimuli was the same as in the previous experiment. Therefore, the shaping as such was the same as described in Sec. 2.1.4. The scaling of the signals before filtering, however, was different. Like in the previous experiment the scaling was done relative to a signal constructed from all HINT sentences since the gain filter was originally designed for this signal. However, the ratio of silence and actual stimuli is much higher for the CV signals than for the signal constructed of the HINT sentences. Scaling to the overall free-field sound pressure level as presented to the NH listeners would therefore result in very high output levels for the CV stimuli. In addition, the spectrum of the signal constructed from all HINT sentences differs from the spectrum of the signal constructed from all CV sounds (see Fig.3.2). If the entire spectra were matched the resulting output levels after filtering would again be very high for the CV signals. Therefore, the scaling of the CV sounds to the HINT sentences was done such that the rms-level of the CV stimuli was the same as for the HINT sentences for frequencies above 2.5 kHz after all silence was removed from both signals. As a result a scaling factor was found which was applied to all the CV sounds before filtering.

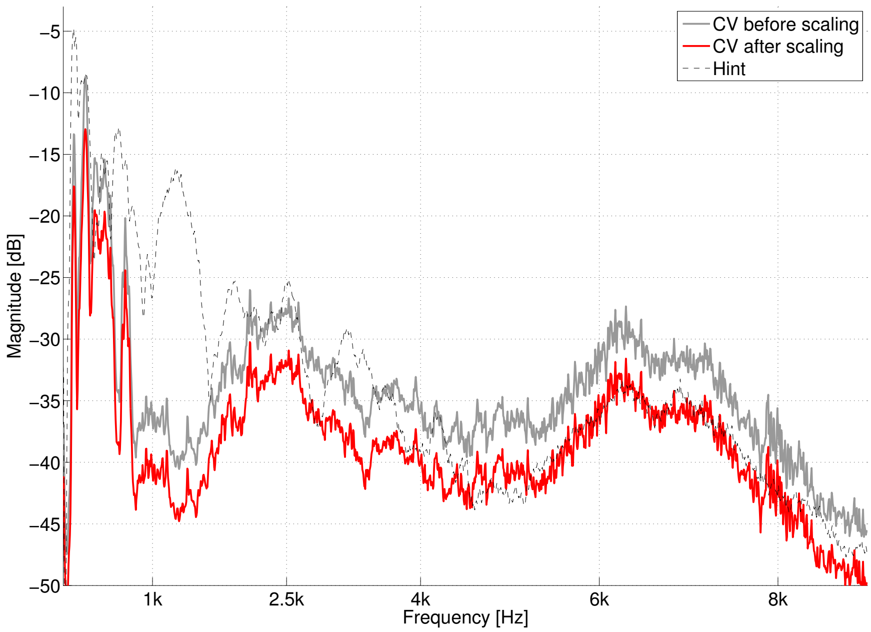

Figure 3.2: Scaling of the CV signals before applying the gain filter to ensure audibility for the HI listeners. The scaling was done such that the rms-level of the signal constructed from all CV sounds was the same as for the HINT sentences for frequencies above 2.5 kHz.

3.2 RESULTS

The results obtained from the consonant test are first analysed in terms of overall percent correct, see Sec. 3.2.1, and then in terms of correct scores for different groups of consonants, see Sec. 3.2.2.

3.2.1 *Percent correct scores*

The perception of consonants in terms of overall percent correct is summarised in Fig. 3.3. Overall the NH listeners performed better than the HI listeners, as indicated by their higher percent-correct scores. The NH listeners' performance was very good throughout all experimental conditions. This indicates a ceiling affect which will be discussed in detail in Sec. 3.3. The HI listeners, however, performed better in anechoic than in reverberant conditions, and when listening binaurally.

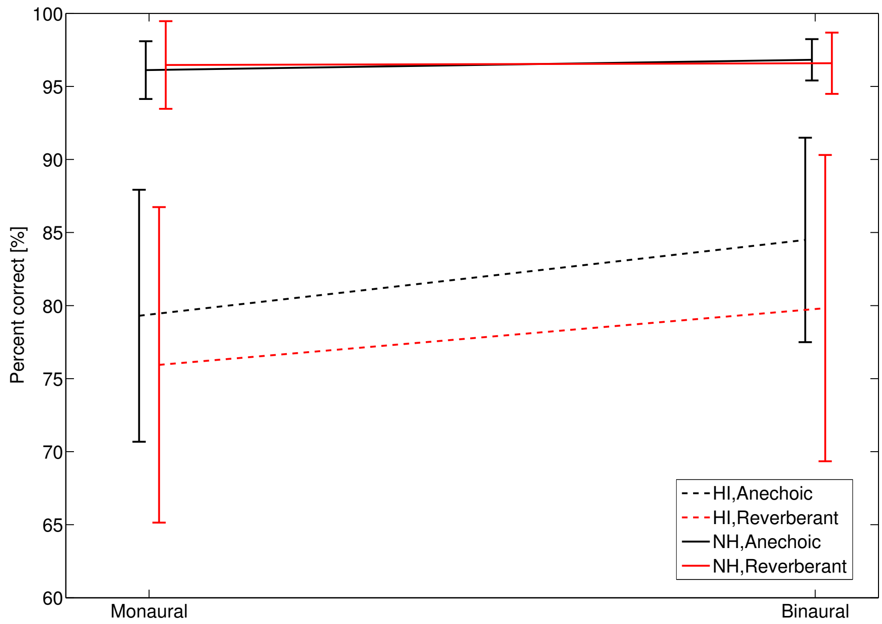

Figure 3.3: Percent correct scores for consonant perception in anechoic and reverberant listening conditions averaged across 5 NH and 12 HI listeners. Consonant perception was tested monaural and binaural. The error bars represent one standard deviation in each condition.

A repeated measures ANOVA with room and listening mode as repeated measures was used in order to identify factors that significantly affect the consonant perception. Due to the ceiling effect found for the NH listeners that are likely to bias the outcome of this analysis the results for the NH were excluded. The analysis is therefore only valid for the HI listeners. The HI listeners' performance in terms of percent correct scores was significantly affected by the room ($p = 0.003$) and the listening mode ($p = 0.004$) on a statistical level of $\alpha = 0.01$ (see also Tab. A.3). The interaction of room and listening mode, however, was insignificant ($p = 0.91$).

3.2.2 *Correct responses for individual consonant groups*

A more detailed analysis on different groups of consonants was undertaken to examine if reverberation and listening mode affect those groups differently. Again, this detailed analysis was only done for the HI listeners due to the ceiling effect for the NH listeners. The analysis was done in terms of correct responses (CR) for individual consonants that were later grouped according to Tab.3.1.

First, individual confusion matrices (CM) for each of the HI listeners were obtained. CM were found for the four different experimental conditions arising from listening environment and listening mode. For a given condition, the correct responses of consonants that are articulated similarly were averaged for each listener, respectively. Here, the place of articulation was distinguished in terms of back and front. Back and front consonants were then divided into smaller groups depending on the manner of articulation, see Tab.3.1. As a result, for each listener four values given by the experimental condition were assigned to seven groups of consonants. In each group the CR were averaged across all HI listeners as represented in Tab.3.2.

A Wilcoxon matched paired test was conducted for each of the seven consonants groups with respect to the listening environment (see Tab. A.5) and listening mode (see Tab. A.4), respectively. In this way, it could be tested whether the experimental conditions affect the reception of the different groups significantly.

		Anechoic		Reverberant	
		Monaural	Binaural	Monaural	Binaural
Front	Plosives	8.60 (1.09)	9.04 (1.28)	8.52 (1.05)	8.30 (1.28)
	Nasals	8.58 (2.11)	9.00 (1.33)	**7.50** (1.92)	**8.45** (1.97)
	Fricatives	8.69 (1.52)	9.35 (1.37)	8.55 (1.35)	8.84 (1.72)
	Lateral	9.33 (2.39)	10.00 (1.71)	10.64 (3.44)	10.64 (1.43)
Back	Plosives	9.75 (1.88)	9.96 (1.86)	**9.09** (1.66)	**9.95** (2.05)
	Fricatives	9.25 (2.42)	9.67 (2.35)	8.09 (2.55)	9.63 (2.50)
	Approximant	10.67 (1.15)	11.50 (0.90)	**9.36** (2.55)	**10.45** (2.29)

Table 3.2: Number of correct responses for different groups of consonants averaged across all HI subjects. The grouping was realised in terms of place and manner of articulation in anechoic and reverberant conditions for monaural and binaural listening. The standard deviation for each condition is given in brackets. A significant effect on a level of α=0.05 of reverberation in a given group is indicated by underlining and a significant effect of listening mode is indicated in **bold**.

The performance in terms of CR was better for the back consonants than for the front consonants for all experimental conditions when the manner of articulation was disregarded. When looking only at manner of articulation, the best performance in terms of CR was obtained for the approximant in the anechoic environment and for the lateral consonant in the reverberant environment. The worst performance was found for the nasals, independent of the listening environment.

Overall, consonant perception in terms of CR was better in anechoic conditions compared to reverberant conditions as indicated by the higher number of CR. Here, the mean difference of performance between anechoic and reverberant conditions when disregarding the listening mode was bigger for the back consonants compared to the front consonants. A significant effect of reverberation was, however, only found for the front plosives when listening binaurally ($p = 0.01$). In general, binaural listening was superior to monaural listening, both in anechoic and reverberant listening conditions. In the anechoic listening environment the binaural benefit was on average the same for both front and back consonants. In reverberation, however, the difference between monaural and binaural listening was bigger for the back compared to the front consonants. If present, a significant binaural benefit was only found in reverberant listening conditions. A significant binaural benefit in the reverberant condition was found for front nasals ($p = 0.01$), back plosives ($p = 0.04$) and the back approximant ($p = 0.05$). For the back fricatives the significance of the difference between monaural and binaural listening in reverberation was just missed ($p = 0.06$).

3.3 DISCUSSION

One of the main questions was whether the same effects of reverberation and listening mode found in Ch. 2 can be found for the perception of consonants. In general, NH listeners performed better than HI listeners which is in agreement with the first experiment as well as with the hypothesis constructed earlier. However, the performance of the NH listeners was essentially independent of both reverberation and listening mode. The performance as such was nearly perfect across all experimental conditions. At that point, variations of reverberation or listening mode had no further impact on the performance of the NH listeners. This ceiling effect could be avoided by modifying the used consonant test. A noticeable effect of both reverberation and listening mode for NH listeners was, for example, found by Helfer (1992). The reverberation time of the room used in the cited experiment was relatively flat across frequency and averaged 1.0 sec between 0.5 and 2 kHz which is only slightly above the average across the same frequencies for the reverberation time of that room used in the presented experiment, i.e. 0.8 sec. However, as studied by Helfer (1992, 1994) initial consonants are less affected by reverberation than final consonants. Using a vowel-consonant (VC) instead of a CV test might therefore be able to reveal the effect of reverberation even for NH listeners. During the time of the experiment, however, a VC for Danish consonants was not available. As pointed out by Helfer (1992) when using the vowel /a/ performance was better compared to when the vowels /i/ and /u/ were used. This was true for both for monaural and binaural listening. Hence, another way to avoid the ceiling effect would be to use a vowel different than /a/.

The performance of the HI listeners, however, was affected by both reverberation and listening mode. The performance was better in anechoic compared to reverberant conditions which is in agreement both with findings from other studies (see Nabelek and Pickett, 1974; Helfer, 1992, 1994) and with the findings from the sentence intelligibility test of the first experiment. The duration of the stimuli presented in the CV test was much shorter than that in the sentence intelligibility test. In particular, the duration is shorter than the reverberation time of the room. Therefore, it can only be speculated that the modifications of the signals due to reverberation, i.e. temporal smearing, are the same for the sentences and the CV sounds.

In addition, binaural listening was superior to monaural listening. Again, this is in agreement with the findings for the sentence intelligibility test as well as with the studies mentioned above. Even though a binaural benefit was observable in both anechoic and reverberant conditions a significant binaural advantage could only be found for a number of consonant groups in reverberation. Combining information from both ears suppresses the perception of echos (Zurek, 1979) that are, of course, only present in reflective listening conditions. In addition, there are more cues available in reflective environments that the binaural system can take advantage of compared to when listening in anechoic environments. Therefore, it is not surprising that the binaural benefit was found significant mostly in reverberant rather than in anechoic listening conditions. However, performance improved when listening binaurally compared to listening monaurally in the anechoic enviroment as well. As found by e.g. Helfer (1994), this effect cannot be attributed to increased intensity from binaural summation. In addition, binaural loudness summation is significantly less when stimuli are presented through loudspeakers as compared to when presented via earphones (see Epstein and Florentine, 2009; Florentine *et al.*, 2011, Ch. 8.5.2). A binaural benefit in the perception of consonants in quiet, i.e. in the absence of noise and reverberation, was also found by Helfer (1992). The underlying mechanisms for this benefit were, however, not clear as pointed out by the author. However, the fact that the binaural gain is also present in the anechoic condition indicates that there is an, not yet understood, effect that can be attributed to the cues arising from binaural listening.

The binaural benefit was bigger for back consonants compared to front consonants, especially in reverberant conditions. This can be explained by the frequency dependency of the binaural benefit. As pointed out by other studies the benefit of binaural listening relies on low-frequency information that is provided to a bigger extent from back rather than front consonants (see e.g. Grø nnum, 1998).

4 OVERALL DISCUSSION: COMPARISON OF SENTENCE AND CONSONANT PERCEPTION

The experiments presented in Ch. 2 and 3 examined the effect of reverberation and listening mode on the perception of two different units of speech, respectively. While the intelligibility of macroscopic speech, i.e. sentences, was studied in the first experiment, the perception of speech on a microscopic level, i.e. consonants, was examined in the second experiment. It is, of course, of interest, if the same relationships between reverberation, listening mode and speech could be found for the different units, respectively. It should, however, be noted that the experiments on sentence intelligibility were conducted in the presence of noise while the consonant perception was tested in the absence of noise. Therefore, even though similar trends may be observable, the comparison of the results is not straightforward.

The perception of speech is affected negatively by reverberation. Even though this is true for both macro- and microscopic speech, the underlying processes may be different. The negative effect of reverberation on the intelligibility of sentences can primarily be explained by smearing effects on the temporal envelope of the signal which results in the target signal being less intelligible and more difficult to be segregated from an interfering signal. Since the consonants' duration is shorter than the reverberation time of the room used in this experiment it is highly speculative to assume the same smearing effects at a microscopic speech level as well. However, since the negative effect of reverberation was observed for consonants it can be assumed that reverberation modifies the microscopic units of speech in a way that they become less intelligible.

Furthermore, performance improved when listening binaurally compared to when listening monaurally. This is true for both for the perception of sentences and consonants. Listening binaurally improved the performance in the sentence intelligibility test because of giving rise to cues that help segregating the spatially separated target and interfering signals. The perception of the consonants in the reverberant condition could be improved when listening binaurally since the arising cues help suppressing the reverberation present in the room. However, the fact that the binaural benefit was observable also in the anechoic conditions both for consonant perception and sentence intelligibility indicates the presence of an additional binaural effect that is not yet understood.

5 SUMMARY

In the present study the effect of reverberation on speech intelligibility of NH and HI listeners was examined in two experiments. In both experiments the performance was tested for monaural and binaural listening.

The intelligibility of sentences in the presence of a speech masker and a non-speech masker was tested in the first experiment. Here, a negative effect of reverberation could be found for both listening groups. In addition, the listeners performed better when listening binaurally than when listening monaurally. The binaural benefit was bigger for the speech masker than for the non-speech masker. This can be explained by different mechanisms to segregate the competing sound sources and to focus on the target speech signal. The cues to perceive the spatial separation of the signals were only available for binaural listening. In the monaural case the target speech was more difficult to segregate from the speech masker than from the non-speech masker. As a result, the performance was worse for the speech masker than for the non-speech masker when listening monaurally. The effect of reverberation was the same for the NH and the HI listeners when expressed in terms of difference in SRT for the anechoic condition and the reverberant condition. Considering that the HI listeners already needed rather high TMR to understand the speech in the anechoic condition it can, however, be argued that the added reverberation is perceived more severe by the HI listeners than by the NH listeners.

The effect of reverberation on the perception of consonants was studied in the absence of noise. The obtained results could only be analysed for the HI listeners due to a ceiling effect for the NH listeners. The ceiling effect can be avoided in future studies by modifying the consonant test in terms of position of the consonant and choice of the vowel. The performance of the HI listeners, however, was negatively affected by reverberation. In addition, the consonant perception was better for binaural listening than for monaural listening. While the binaural gain in the reverberant condition can be attributed to the echo suppression due to binaural listening (Zurek, 1979), the fact that the binaural gain was also present in the anechoic conditions indicates an additional effect that can be attributed to binaural listening which is not understood to date.

Comparing the patterns arising from the effect of reverberation, listening mode on the perception of the two different units of speech, similar relationships could be found. Both sentence intelligibility and consonant perception were negatively affected by reverberation and improved for binaural listening compared to monaural listening. If the underlying processes are similar is, however, not obvious from the present data.

A

APPENDIX: OUTCOMES OF THE STATISTICAL ANALYSES

A.1 EFFECT OF HEARING STATUS, ROOM, TYPE OF MASKER AND LISTENING MODE ON THE SENTENCE INTELLIGIBILITY

Effect	p-value
(1) **Hearing status**	**0.000009**
(2) **Room**	**0.000000**
(3) **Masker**	**0.000005**
(4)**Monaural/Binaural**	**0.000000**
Room*Hearing status	0.255811
Masker*Hearing status	0.725695
Monaural/Binaural*Hearing status	0.443216
Room*Masker	0.474996
Room*Monaural/Binaural	0.381120
Masker*Monaural/Binaural	**0.000045**
Room*Masker*Hearing status	0.216881
Room*Monaural/Binaural*Hearing status	0.216881
Masker*Monaural/Binaural*Hearing status	0.269185
Room*Masker*Monaural/Binaural	0.878881
(1)*(2)*(3)*(4)	0.156787

Table A.1: Repeated measures analysis of variance for the SRT-values obtained from the sentence intelligibility test. The effects marked in **bold** are significant at $\alpha = 0.01$.

Effect	p-value
Hearing status	0.267977
Masker	0.389428
Monaural/Binaural	0.462156
Masker*Hearing status	0.868139
Monaural/Binaural*Hearing status	0.212322
Masker*Monaural/Binaural	0.854552
Masker*Monaural/Binaural*Hearing status	0.155031

Table A.2: Repeated measures analysis of variance for the difference in SRT-values between the anechoic condition and the reverberant condition in the sentence intelligibility test. None of the effects is significant at $\alpha = 0.01$.

A.2 EFFECT OF ROOM AND LISTENING MODE ON THE CONSONANT PERCEPTION

A.2.1 *Overall percent correct scores*

Effect	p-value
Room	**0.003268**
Monaural/Binaural	**0.004297**
Room*Monaural/Binaural	0.907300

Table A.3: Repeated measures analysis of variance for overall percent scores obtained in the consonant perception test. The effects marked in **bold** are significant at $\alpha = 0.01$.

A.2.2 *Correct responses for individual consonant groups*

		Anechoic	Reverberant
Front	Plosives	0.168	0.415
	Nasals	0.266	**0.012**
	Fricatives	0.055	0.398
	Lateral	0.241	0.790
Back	**Plosives**	0.624	**0.041**
	Fricatives	0.327	0.059
	Approximant	0.114	**0.050**

Table A.4: Wilcoxon matched pairs to test for significant binaural benefit in different groups of consonants. For the groups marked in **bold** the binaural benefit was significant at $\alpha = 0.05$.

		Monaural	Binaural
Front	**Plosives**	0.441	**0.009**
	Nasals	0.286	0.183
	Fricatives	0.445	0.056
	Lateral	0.100	0.285
Back	Plosives	0.265	0.813
	Fricatives	0.086	0.735
	Approximant	0.126	0.161

Table A.5: Wilcoxon Matched pairs to test for significant binaural benefit in different groups of consonants. For the groups marked in **bold** the binaural benefit was significant at $\alpha = 0.05$.

BIBLIOGRAPHY

Arbogast, T., Mason, C., and Kidd, G. (**November 2002**), "The effect of spatial separation on informational and energetic masking of speech." Journal of the Acoustical Society of America **112**(5), 2086–2098.

Arbogast, T., Mason, C., and Kidd, G. (**April 2005**), "The effect of spatial separation on informational masking of speech in normal-hearing and hearing impaired listeners." Journal of the Acoustical Society of America **117**(4), 2169–2180.

Biologic Systems Corp (**2005**), *HINT PRO: Hearing in Noise Test User's and Service Manual*, Bio-logic Systems Corp, Mundelein, Illinois.

Braasch, J. and Hartung, K. (**2002**), "Localization in the Presence of a Distracter and Reverberation in the Frontal Horizontal Plane. I.Psychoacoustical Data," Acta Acoustisca United with Acoustica **88**, 942–955.

Breitsprecher, C. (**2009**), "The Influence of Microphone Location and Reverberation on Better-Ear Effects," B.Eng. thesis, University of Applied Sciences Oldenburg.

Bronkhorst, A. (**1999**), "The cocktail party effect: Research and applications," Journal of the Acoustical Society of America **105**, 1150.

Bronkhorst, A. and Plomp, R. (**1992**), "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," Journal of the Acoustical Society of America **92**(6), 3132–3139.

Brungart, D. (**March 2001**), "Informational and energetic masking effects in the perception of two simultaneous talkers," Journal of the Acoustical Society of America **109**(3), 1101–1109.

Brungart, D. (**2005**), "Informational and Energetic Masking Effects in Multitalker Speech Perception," in *Speech Separation by humans and machines*, edited by P. Divenyi (Kluwer Academic Publishers), pp. 261–267.

Dreschler, W., Verschuure, H., Ludvigsen, C., and Westermann, S. (**2001**), "ICRA Noises: Artificial Noise Signals with Speech-like Spectral and Temporal Properties for Hearing Instrument Assessment," International Journal of Audiology **40**, 148–157.

Duquesnoy, A. and Plomp, R. (**1980**), "Effect of reverberation and noise on the intelligibility of sentences in case of presbyacusis," Journal of the Acoustical Society of America **68**, 537–544.

Epstein, M. and Florentine, M. (**2009**), "Binaural loudness summation for speech and tones presented via earphones and loudspeakers," Ear Hear **30**, 234–237.

Florentine, M., Popper, A., and Fay, R. (**2011**), *Loudness* (Springer Handbook of Auditory Research).

Freyman, R., Helfer, K., McCall, D., and Clifton, R. (**December 1999**), "The role of perceived spatial separation in the unmasking of speech," Journal of the Acoustical Society of America **106**(6), 3578–3588.

Glasberg, B. and Moore, B. (**1988-1989**), "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech," Scandinavian Audiology **Suppl 32**.

Grø nnum, N. (**1998**), *Fonetik og Fonologi - almen og dansk* (Akademisk Forlag A/S).

Helfer, K. (**December 1992**), "Aging and the Binaural Advantage in Reverberation and Noise," Journal of Speech and Hearing Research **35**, 1394–1401.

Helfer, K. (**April 1994**), "Binaural Cues and Consonant Perception in Reverberation and Noise," Journal of Speech and Hearing Research **37**, 429–438.

Humes, L. (**2007**), "The Contributions of Audibility and Cognitive Factors to the Benefit Provided by Amplified Speech to Older Adults," Journal of the American Academy of Audiology **18**, 590–603.

Jepsen, M. and Dau, T. (**2011**), "Confusion of Danish consonants in white noise," in *"Speech Perception and Auditory Disorders", proc. of the International Symposium on Auditory and Audiological Research*, submitted.

Koenig, W. (**January 1950**), "Subjective Effects in Binaural Hearing," Journal of the Acoustical Society of America **22**(1), 61–62.

MacKeith, N. and Coles, R. (**1971**), "Binaural advantages in hearing of speech," The Journal of Laryngology and Otology **85**, 213–232.

Marrone, N., Mason, C., and Kidd, G. (**2008**a), "Evaluating the Benefit of Hearing Aids in Solving the Cocktail Party Problem," Trends in Amplification **12**, 300–315.

Marrone, N., Mason, C., and Kidd, G. (**2008**b), "Tuning in the spatial dimension: Evidence from a masked speech identification task," Journal of the Acoustic Society of America **124**, 1146–1158.

Moore, B. (**2007**), *Cochlear hearing loss – Physiological, Psychological and Technical Issues.* (John Wiley and Sons, Ltd, Cambridge, England), 2nd ed.

Moore, B., Stone, M., Füllgrabe, C., Glasberg, B., and Puria, S. (**2008**), "Spectro-Temporal Characteristics of Speech at High Frequencies, and the Potential for Restoration of Audibility to People with Mild-to-Moderate Hearing Loss," Ear Hear **29**, 907–922.

Nabelek, A. and Pickett, J. (**August 1974**), "Recption of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids," Journal of the Acoustic Society of America **56**, 628–639.

Nielsen, J. and Dau, T. (**2011**), "The Danish hearing in noise test," International Journal of Audiology **50**, 202–208.

Noble, W. and Perrett, S. (**2002**), "Hearing speech against spatially separate competing speech versus competing noise," Perception and Psychophysics **64(8)**, 1325–1336.

Ohl, B. (**2009**), "Externalization versus Internalization of Sound in Normal-hearing and Hearing-impaired Listeners," Master's thesis, Technical University of Denmark.

Phatak, S. and Allen, J. (**April 2007**), "Consonant and vowel confusions in speech-weighted noise," Journal of the Acoustic Society of America **121**, 2312–2326.

Phatak, S., Lovitt, A., and Allen, J. (**August 2008**), "Consonant confusions in white noise," Journal of the Acoustic Society of America **124**, 1220–1233.

Schroeder, M. (**1965**), "New Method of Measuring Reverberation Time," Acoustical Society of America , 409–412.

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

Zurek, P. (**1979**), "Measurements of binaural echo suppression," Journal of the Acoustical Society of America **66**(6), 1750–1757.