Spatial release from masking in hearing-impaired listeners

Author: Baljeet Rana

Supervisors: Dr. Jorg M. Buchholz and Dr. Harvey Dillon

This thesis is presented for the degree of

Doctor of Philosophy (PhD)

National Acoustic Laboratories (NAL)

and

Audiology section (Department of Linguistics), Faculty of Human Sciences, Macquarie University, Sydney, Australia

Table of content

Declaration of Authorship

I, Baljeet Rana, declare that this thesis titled, "Spatial release from masking in hearing-impaired listeners" and the work presented in it are my own. I confirm that:

- This thesis has been submitted solely to Macquarie University for consideration for the doctoral degree.
- Ethics approvals have been obtained from the Australian Hearing Human Research Ethics Committee and Macquarie University Human Research Ethic Committee. The signed approvals with number AHHREC2014-24, AHHREC2016-23 and 5201401150, can be found in Appendix.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed:

Date: 24th April 2017.

Abstract

The improvement in speech understanding in noise due to the spatial separation of a target source from masking sources is known as spatial release from masking (SRM). SRM observed when fluctuating distractors are present on both sides of the listener, and target speech at the front, can be largely attributed to a phenomena known as better-ear glimpsing (BEG). BEG, which utilizes interaural level differences (ILDs) to help understanding speech in noise, is limited by reduced audibility at high frequencies in hearing-impaired (HI) listeners. Unfortunately, it is difficult to provide adequate amplification at high frequencies to fully restore BEG. Therefore, the idea of extending ILDs to low frequencies was proposed here. The results of a ***first study*** showed that ILDs can be effectively utilized at low frequencies by both normal-hearing (NH) and HI listeners. However, the performance noted in HI listeners was still poorer than in NH listeners, which might have been due to differences in audibility between the two groups. To test this hypothesis, in a ***second study***, both groups were tested at different equal audibility levels. Results showed that if audibility is carefully controlled then HI listeners can utilize ILDs as effectively in BEG as NH listeners. However, it was also observed that not all HI listeners could accommodate the required signal levels due to loudness discomfort. Therefore, in a ***third study***, the effect of wide dynamic range compression (WDRC) on the utilization of ILDs was investigated. The results revealed that low-frequency extended ILDs combined with WDRC can provide BEG cues that are significantly higher than the one provided by natural ILDs. Therefore, this approach may be considered as a viable option for HI listeners with hearing aids to improve speech intelligibility in noise. Further, BEG has been well studied in hearing aid users but very little is known about it in cochlear implants (CIs) recipients. This group relies solely on ILDs to localize sounds as well as to understand speech in noise. As a consequence, CI recipients may benefit the most from any improvement in ILD processing. However, before any ILD enhancement method should be considered, it needed to be clarified if CI recipients are at all able to take advantage of BEG cues. The results of the ***fourth study*** revealed that at least some CI recipients can utilize BEG to some extent.

Acknowledgments

As they say, it's not only about the destination, it's also about the journey to get there!

This journey of Ph.D. has been one of the most memorable parts of my life and I think without thanking them this journey would be completely incomplete.

No work can be done efficiently without the support of others. Well! The list of "others" is long but to start with, first and foremost I would like to thank the strongest pillar of support of this work my principal supervisor "Jorg Buchholz". I had never thought even in my wildest dream that I would ever get to work with a supervisor like him. The most humble, the most approachable and the most down to earth personality I have ever come across. I learned a lot from him throughout my candidature and wish to learn even more throughout my professional career. Thank you is not the word for his unstinting support and endless efforts. Seriously, I couldn't have asked for more.

I would like to thank my adjunct supervisor "Harvey Dillon" for his timely help and opinion when needed. It was nice to have him around and his supervision when this work was conducted.

A Big thanks to all the participants (volunteers) of this research. Without their support this research wouldn't have been possible at all.

I would like to thank Mark Seeto for his suggestions on statistical analysis, confirming my statistics and his instant help during my stay in India.

I would like to thank Scott for his 24X7 helpline service ☺ John Seymour for all the IT help, James G for going through all the hassles while setting up the experiments and for being patient while answering all of my stupid questions. Thanks to Greg for fixing my glasses ☺ and helping me in all of my trivial issues.

My sincere and heartfelt thanks to Katrina for making my stay so easy and convenient when I arrived in Sydney, and for introducing me to the culture of Australia; Pam for my first birthday present at NAL (you touched my heart and I missed my mom); Lyndal and Ingrid for considering me as your pet and feeding me all the time during our stay in Vancouver; Warwick for all the funny conversations that we had (stop me if you can, I am still laughing); Michael F, thanks for voting "yes" three times!; Bram for your jokes and for being so kind; Liz Convery, Jess and Margot for feedback on various things; Liz B, Jermy (for free cakes), Viv, Miriam, Nicky, Vicky for all the coffee-chats in the kitchen.

I would like to thank my teachers (especially Rajalakshmi ma'am, Animesh sir and Ajith sir) and friends (especially Nisha, Nandu, Sreelakshmi, Jyotsna, Akbar, Sachchidanand, Astha, Neha, Ujjwal, Abhishek Sehta, Avinash, Nandan, Deepthi, Vinni, and many more) at AIISH for their motivation and trust on me.

Finally ! here comes the list of crazy people without their presence my Ph.D. life wouldn't have been so memorable and filled with so much of fun. It was nice knowing such amazing lovely people. Everyone had some or the other lessons to teach me, some good for the good and some bad, yet again, for the good!

Starting with the bosses of madness the "Fabrice Bardy" –trust me! I don't think I could have ever asked for a better office mate cum friend than him, I am so happy that I shared my office space with him, and his presence just made me laugh and laugh and laugh all the time. A big thanks Fabrice for being such a Mr. Bean! I think I can say that you were one of the reasons that my Ph.D. journey was so much of fun. Thanks again, Brother Bardy (read it in my thick accent).

Javier, Kiri, and Remi- I must say that your entry at NAL changed my life to a very great extent and you know what I mean! You people made me so comfortable so comfortable that I could just be myself. Thanks to all the conversations during lunch time, Friday drinks, Blue Mountains trip and much more. I am glad I had friends like you who were so understanding and so patient that they could not only tolerate my nonsense talks but also could add to that. Cheers to you three! I loved it to the core! Kiri-thanks for understanding me so well, I just say half of the sentence and you would already know what I am going to speak next or what I didn't speak! You are smart! So just call me may be, simply because it's always a good time! ☺

Joaquin- I am not sure if I ever I said you this but you have always been my inspiration in many ways. Since I come from an army family background, I always find you as a young army officer who has passion, creativity, friendliness, the skill to tackle problems, and much more. Thanks for motivating me, your random "well done" worked well for me all the time, it kept me going. Thanks mate!

Els- aww! You are cute! Thank you so much for your very polite gesture of care, and help (and for chocolates, of course! You are amazing and you can do it too!). Thanks for keeping contact with me especially during the last few days of my Ph.D., trust me, it made me feel so good! A big thanks ☺

Tim B- I think our friendship is none less than Tom and Jerry, I enjoy fighting with you when you are around but I also miss you badly when you are not around ☺ oops! Thanks for all the good time, and funny moments! Many more to come! You are awesome!

Thanks, Kirsty and Paola for your professional guidance and helping me in choosing my career paths. It helped me a lot! Kirsty-thanks for keep motivating me! Your random words "you are getting there", "You can do it" made me feel good.

Mat Lavandier- you came to NAL for few months but you found a place for yourself in the list of my friends. Thanks for everything! Nemat – thanks, bro! I enjoyed chatting with you and thanks for all the food that we shared (the one that your wife cooked for us).

Adam W (thanks for everything and the falafels, I loved it), Pragati and Vishnu (thanks for every little gesture of care especially before my arrival in Sydney). Shivali (its fun working with you), Mariana, Rakshita, Viv F, Titia, Mrdiula-Thanks for being around during these years, thanks for every little help and kind gesture. It was definitely nice to know you all.

Wei Ren and Juan- Can't ask for better housemates. Thanks for introducing me to the different culture and showing me the world just within the four walls of our house! Finally Kamalika and Esther- there is no way that life would have been so entertaining without your company. Can't thank you enough for everything that you have done for me. You both were none less than a blessing in my life. Kamalika and Esther, I think I met you both accidentally and now we are friends forever. Last but not least I would like to thank my parents, grandparents, brother, sister-in-law and relatives for being my support and motivation throughout this journey.

Oh sorry, I forgot to thank myself. Thank you, *Baljeet Rana.* Thanks for being who you are! Thanks for believing in your dreams, thanks for working towards your destination, thanks for staying stronger yet calm and composed! I just love the way you are-stupid yet smart, funny yet serious, innocent yet clever but sweetheart *your journey has just begun*…. **The woods are lovely, dark and deep, But I have promises to keep, And miles to go before I sleep, And miles to go before I sleep…………………… (by Robert Frost)**

Thank you, one and all!

List of publications

Journal publications

- Baljeet Rana and Jorg M Buchholz **(2016)**. "Better-ear glimpsing at low frequencies in normal-hearing and hearing-impaired listeners," The Journal of the Acoustical Society of America**, 140** (2), 1192-1205.
- Baljeet Rana and Jorg M Buchholz **(Under review)**. "Effect of audibility on better-ear glimpsing as a function of frequency in normal-hearing and hearing-impaired listeners," The Journal of the Acoustical Society of America.
- Baljeet Rana and Jorg M Buchholz **(Under review)**. "Effect of audibility on better-ear glimpsing using non-linear amplification," The Journal of the Acoustical Society of America.
- Baljeet Rana, Jorg M Buchholz, Catherine Morgan, Mridula Sharma, Tobias Weller, Shivali A. Konganda, Kyoko Shirai and Atsushi Kawano **(2017)**. "Bilateral versus unilateral cochlear implantation in adult listeners: speech-on-speech masking and multi-talker localization," Trends in hearing, **21**, 2331216517722106.

Conference proceedings publications

- Baljeet Rana, Jorg M Buchholz, Catherine Morgan, Mridula Sharma, Tobias Weller, Shivali A. Konganda, Kyoko Shirai and Atsushi Kawano (2016). Better-ear glimpsing in Japanese adult bilateral cochlear implant users. The Journal of the Acoustical Society of America, 140(4), 3436-3436.
- Baljeet Rana and Jorg M Buchholz (2016). Effect of audibility, frequency region, and hearing loss on better-ear glimpsing. The Journal of the Acoustical Society of America, 140(4), 3437-3437.

Conference presentations

- Baljeet Rana, Jorg M Buchholz, Catherine Morgan, Mridula Sharma, Tobias Weller, Shivali A. Konganda, Kyoko Shirai and Atsushi Kawano (2016). Better-ear glimpsing in Japanese adult bilateral cochlear implant users. Oral presentation at the 5th joint meeting of the Acoustical Society of America and the Acoustical Society of Japan, 28 November-2 December, Honolulu, USA.
- Baljeet Rana and Jorg M Buchholz (2016). Effect of audibility, frequency region, and hearing loss on better-ear glimpsing. Oral presentation at the 5th joint meeting of the

Acoustical Society of America and the Acoustical Society of Japan, 28 November-2 December, Honolulu, USA.

- Baljeet Rana and Jorg M Buchholz (2016). Artificially enhancing better-ear glimpsing cues to improve understanding of speech in noise for listeners with hearing loss. Poster presented at the HEARing CRC Research Symposium, 17-19th October 2016, Australian Hearing Hub, Macquarie University, Sydney.
- Baljeet Rana and Jorg M Buchholz (2016). Artificially enhancing better-ear glimpsing cues: effect of frequency region and sensation level. Poster presented at the 33th World Congress of Audiology, Vancouver, Canada.
- Baljeet Rana and Jorg M Buchholz (2016). Understanding the role of audibility on artificially enhanced cues for better-ear glimpsing. Presented a talk at the seminar organized by Hearing Research center, Boston University and also in the meeting at Department of Speech, Language and hearing sciences, Boston University, USA.
- Baljeet Rana and Jorg M Buchholz (2016). Maximizing cues for better-ear glimpsing: improving speech intelligibility in noise for hearing-impaired individuals. Oral presentation at the Audiology Australia National Conference, Melbourne, Australia.
- Baljeet Rana and Jorg M Buchholz (2015). Improving speech understanding in noise for hearing-impaired listeners. Presented a talk at the higher degree research (HDR) lounge, November 2015, Macquarie University.
- Baljeet Rana and Jorg M Buchholz (2015). Maximizing spatial cues: Leading to better understanding of speech in noise for hearing-impaired listeners. Presented a talk at the annual Macquarie Linguistics Higher Degree Research student conference (the “showcase”), 2015.
- Baljeet Rana and Jorg M Buchholz (2015). Artificially magnified spatial cues: an attempt to improve speech understanding in noise for hearing-impaired individuals. Oral presentation at the Audiology Australia general meeting, November 2015, Sydney, NSW Chapter.

Abbreviations

4FAHL: Four-Frequency Average Hearing Loss

BEG: Better-Ear Glimpsing

BTE: Behind-The-Ear

CIs: Cochlear Implants

CIC: completely-In-The-Canal

CO: CO-located

CRM: Coordinate Response Measure

EM: Energetic masking

FIR: Finite Impulse Response

HI: Hearing Impaired

HRTFs: Head-Related Transfer Function

ILDs: Interaural level differences

IM: Informational Masking

ITDs: Interaural time differences

JASA: The Journal of the Acoustical Society of America

LiSN-S: Listening in Spatialized Noise-Sentence test

MOCA: Montreal cognitive assessment

NAL-RP: National Acoustic Laboratories – Revised Profound

NAL-NL2: National Acoustic Laboratories – Non Linear

NH: Normal Hearing

NHMRC: National Health and Medical Research Council

RMS: Root Mean Square

SCTs: Speech Comprehension Thresholds

SD: Speech Discourse

SDTs: Speech Detection Thresholds

SNR: Signal-to-Noise Ratio

SRM: Spatial Release from Masking

SRTs: Speech Reception Thresholds

SS: Spatially Separated.

SSN: Speech Shaped Noise

SII: Speech Intelligibility Index

STD: Standard Deviation

SPL Sound Pressure Level

STFT: Short-Term Fourier transform

ULC: Upper Limit of Comfort

VS: Vocoded Speech

WDRC: Wide Dynamic Range Compression

Chapter 1: General Introduction

In this contemporary world, most people are exposed to noisy acoustic environments in their everyday lives, for example, in cafeterias, supermarkets, schools, cinemas, parties and so forth. Verbal communication in such noisy environments can be challenging even for listeners with normal hearing (NH). However, due to an intact structural and functional auditory system, NH listeners are much more capable of overcoming these challenges than listeners with hearing impairment (HI). Some HI listeners are fitted with hearing aids or cochlear implants (CI), but often they still face severe difficulties in understanding speech in noise. Cherry in 1953 reported various important cues that help understanding speech in noise and one of them was spatial cues. Researchers have shown that speech intelligibly improves when a distractor is spatially separated from target speech relative to a condition when they are co-located. This improvement is due to a phenomenon referred to as spatial release from masking (SRM, e.g. Bronkhorst and Plomp, 1988; Festen and Plomp, 1990; Zurek, 1993; Peissig and Kollmeier, 1997; Arbogast *et al.,* 2002; Hawley *et al.,* 2004). SRM has generally been reported to be larger in NH listeners than HI listeners. Thereby, the reduction in the spatial advantage received by HI individuals, and the resulting difficulties in understanding speech in noise, can be attributed to many factors (see chapter 2 for details). However, the contribution of each factor to the reduced performance in HI listeners is still unclear and varies across studies.

The cues that facilitate SRM can be divided into monaural cues and binaural cues (e.g. Blauert, 1997; Kidd *et al.,*1998; Hawley *et al.,* 1999, 2004; Culling *et al.,* 2004). Monaural cues depend on the filtering effects of head, torso, shoulders, and pinna and have been reported to be contributing for vertical (e.g. Butler, 1969a, 1969b) and front-back localization (e.g. Batteau, 1967), and can help in SRM when the signal and noise are spatially separated along the vertical plane, in particular from the front versus back. Since these cues are mainly provided by the pinna, they are predominantly at high frequencies (e.g. Blauert, 1997), and thus may not be available to HI listeners with a high frequency hearing loss. Moreover, hearing aids, such as behind-the-ear (BTE) hearing aids, often apply microphones that are placed above the pinna and thereby remove most of the pinna cues. Hence, these devices do not provide the cues that are relevant for vertical plane localization, provide much reduced differential filtering, and specifically deteriorate front-back localization accuracy (e.g., Best *et al.,* 2010). However, this detrimental effect can be reduced by applying directional hearing aids (Keidser *et al.,* 2006) or by placing microphones inside the ear canal, which is the case in in-the-ear (ITEs) or completely-in-the-canal (CIC) hearing aids (e.g., Dillon, 2012). Intriguingly, it is important to note that even though cues for localization and SRM are similar, the process underlying these two mechanisms may be entirely different. Hence, an improvement noted in localization may not always be accompanied by an improvement in SRM.

Binaural cues are particularly important for localization in the horizontal plane and for SRM. An acoustic signal arriving from any side direction (left or right) reaches earlier the ear that is closer to the source of the signal (termed the ipsilateral ear) relative to the ear which is further from the source of the signal (termed the contralateral ear). This difference in the arrival time of the signal between the ears leads to the occurrence of interaural time difference (ITD) cues (e.g. Blauert, 1997). Likewise, the intensity of the signal has been reported to be significantly higher for the ipsilateral ear than for the contralateral ear. This difference in the intensity between the ears is due to the acoustic shadow introduced by the listener's head, torso, and pinna and leads to the occurrence of interaural level difference (ILD) cues (e.g. Blauert, 1997). Due to the longer wavelength of low frequency signals relative to the size of the human head, ITD cues are mainly available at low frequencies. ITD cues are either carried in the fine structure of the signal's waveform at the left and right ear or in the transients of the signals' envelope. Since the phase of the signal is ambiguous above about 750 Hz, ITD fine structure cues are only evaluated by the auditory system up to about 1.5 kHz. Above that frequency mainly ITD envelope cues are evaluated (Blauert, 1997). However, ITD fine-structure cues are much more salient than ITD envelope cues (Plack, 2005). Moreover, even though any binaural cue can contribute to SRM, in fluctuating noise, which is observed in most common realistic environments, ILD cues have been reported to be contributing more than ITD cues to SRM (Bronkhorst and Plomp, 1988; Glyde *et al.*, 2013c), as they facilitate better-ear glimpsing i.e., the ability of the auditory system to utilize differences in short-term signal-to-noise ratio (SNR) that exist between the two ears to improve speech intelligibility in an almost spatially symmetrically placed fluctuating noise. As mentioned above, ILDs are provided by the head shadow effect, which due to the size of the human head are limited to frequencies above about 1.5 kHz (Blauert, 1997). Unfortunately, this is exactly the frequency region where most HI listeners have the strongest hearing loss, which limits their ability to take advantage of better-ear glimpsing (BEG) cues and often results in poor understanding of speech in noise. There are solutions revolving around high frequencies that have been applied to resolve the issue of smaller SRM in HI listeners (e.g. Glyde *et al.*, 2015), but most of the solutions have been largely limited by the smaller dynamic range available at these frequencies (see chapter 2). Therefore, in this research, as an alternative solution, the idea was explored of artificially extending high frequency cues towards low and mid frequencies where the hearing loss is typically less pronounced or even absent. As observed by Glyde et al. 2015 and other researchers (see chapter 3), restoring audibility for HI listeners does help in improving SRM. However, the obtained SRM is still not equivalent to NH listeners. The idea of further controlling the audibility and its impact on the utilization of natural and artificial BEG cues was therefore also investigated in this research.

The phenomenon of BEG has been investigated in listeners with aidable hearing but it still remains untapped in listeners with unaidable hearing who rely on devices like CIs. Since listeners with CIs have much better access to ILDs than ITDs, it may be predicted that they are able to also take advantage of short-term ILDs and hence show substantial SRM in a symmetrical, spatially-separated masker paradigm. On the contrary, since the CIs worn on both the ears work independently unlike ears of NH listeners, it may be similarly predicted that they are not able to receive any substantial SRM from BEG. These predictions have also been investigated in this research. Further, if listeners with CIs are able to take advantage of BEG cues, then the above proposed method of artificially extending BEG cues to low frequencies should also be tested in this population. Such an approach may then further improve the spatial advantage provided by BEG, and moreover, may also improve localization ability in CI users, which is solely based on ILD cues.

1.1 Aims of the thesis

The ultimate goal of this study was to improve SRM in HI listeners and thereby improve speech intelligibility in noisy environments. The specific goals were:

- To measure the effect of artificially extending ILD cues on BEG in NH and HI listeners.
- To measure the effect of a linear increase in sensation level (or audibility) on BEG as a function of frequency.
- To measure the effect of non-linear amplification on BEG with natural and artificially extended ILD cues
- To measure SRM in a spatially symmetric masker paradigm for listeners with CIs to investigate if they can utilize ILDs for BEG.

1.2 Overview of chapters

This thesis consists of four different experiments. Chapter 2 to 4 describe the first three experiments that aim to understand BEG in depth for listeners with aidable hearing. To start with **chapter 2,** the problem of limited audibility at high frequencies for listeners with hearing loss was dealt with. The smaller dynamic range and the associated loudness discomfort limit the hearing aid gain that can be provided at high frequencies for restoring BEG cues. Therefore, in this experiment ILDs were artificially extended to low and mid frequencies and the ability of NH and HI listeners to utilize them was investigated. Further, since the main phenomenon (i.e. BEG) under investigation is energetic in nature, it was also investigated how far the applied methods gave rise to informational masking.

Generally, the availability of ILD cues for BEG is limited in HI listeners when compared to NH listeners. The limited availability of ILD cues can be largely attributed to the limited audibility at high frequencies due to hearing loss. This was confirmed by Glyde et al. (2015) as they showed that reduced BEG in HI listeners can be improved (or partly compensated) by providing extra linear amplification. In this research, since ILDs were artificially made available across all frequencies, the performance of HI listeners can be affected by the difference in audibility across frequencies due to the difference in hearing thresholds across frequencies. Therefore, in the experiment described in **chapter 3**, the audibility was carefully controlled for HI listeners and then the effect of a linear increase in audibility on the utilization of artificially extended ILDs was investigated across frequency.

A linear increase in amplification may not distort ILDs but may not allow the gain that needs to be provided to obtain the full advantage of ILDs across frequencies. Utilization of non-linear amplification may help to provide sufficient amplification to soft sounds while at the same time ensuring loudness comfort, and may thus be a better and more practical approach than linear amplification for restoring short-term ILD cues for BEG. Therefore, in the experiment presented in **chapter 4**, the effect of non-linear amplification on the utilization of natural and artificially extended ILD cues in BEG was investigated. The audibility was varied by presenting distractors at three different levels, which also allowed the assessment of the interaction between the provided increase in audibility and compression. Further, the extra benefit provided by extended ILDs on top of natural spatial cues was also measured.

Since listeners with unaidable hearing that are fitted with CIs have very limited access to ITDs provided by the signal's fine structure, they mainly rely on ILDs for speech understanding in noise as well as localization. Hence, one might predict that CI recipients are able to take advantage of ILD based phenomena such as BEG. However, BEG relies on taking advantage of short-term SNR differences between the two ears, which may require good coordination between the ears. Since CIs that are worn on both the ears work independently, such coordination may not be provided and therefore CI users may receive a very limited benefit from BEG. To investigate these predictions, the experiment described in **chapter 5** assesses SRM using a spatially symmetrical distractor paradigm with CIs worn in one ear only and in both the ears. Further, the ability of CI users to localize single talkers versus two talkers was also assessed. The utilization of a two-talker paradigm has not been investigated before and since it inherently requires the segregation of two spatially-separated talkers it may assess similar processes as involved in BEG. The experiment helped us to better understand the role of binaural hearing in BEG and localization.

Chapter 6 summarizes the overall findings of this research, its implications, and limitations.

Chapter 2: Better-ear glimpsing at low frequencies in normal-hearing and hearing-impaired listeners

Better-ear glimpsing is an auditory process that takes advantage of short-term interaural level differences (ILDs) to improve the understanding of speech in spatial fluctuating noise. Since ILDs are mainly present at high frequencies, where most hearing-impaired (HI) listeners have the strongest hearing loss, HI individuals cannot fully utilize ILDs for better-ear glimpsing, which may lead to a poorer understanding of speech in noise. This problem may be alleviated by hearing aids that artificially generate ILDs at low frequencies where hearing is typically less impaired. The present study, therefore, investigated the spatial benefit in speech intelligibility that is provided by better-ear glimpsing with low-frequency extended ILDs in a symmetric two-distractor speech background. Speech reception thresholds were measured in a spatially co-located and separated condition as a function of a frequency region in ten normal-hearing (NH) and ten mild-to-moderate sensorineural HI subjects. In both groups, the extended ILDs provided a substantial spatial advantage on top of the advantage already provided by natural ILDs. Moreover, the spatial advantage was largely independent of frequency region, suggesting that both NH and HI subjects can utilize low-frequency ILDs for improving speech understanding in noise. Overall performance, as well as spatial advantage, was reduced in the HI group.

Note: Aspects of this work were presented at the Audiology Australia National Conference, Melbourne, 2016. This chapter has been published in the The Journal of the Acoustical Society of America.

2.1 Introduction

Speech intelligibility improves when the target and distracting speech arrives from different directions relative to when they arrive from the same direction; an auditory phenomenon called spatial release from masking (SRM). SRM can provide an advantage of up to 20 dB in normal-hearing (NH) listeners (e.g., Bronkhorst and Plomp, 1988) but is strongly reduced in hearing-impaired (HI) listeners (e.g., Peissig and Kollmeier, 1997; Marrone *et al*., 2008; Best *et al*., 2012; Glyde *et a.l*, 2013a). A number of different factors contribute to SRM, including acoustic benefits due to head shadow (e.g., Blauert, 1997, pp.50-93), neural processing of interaural time and level differences (e.g., Durlach, 1963), and perceived spatial separation of target and distractor signals (e.g., Freyman *et al*., 1999). Even though all of these factors are important for understanding speech in spatialized noise (e.g., Bronkhorst, 2000), the current study solely focuses on the benefits provided by head shadow.

Head shadow typically results in signal levels at the two ears that are different for spatially separated distractor and target signals, which in turn can result in an advantage in the signal-to-noise ratio (SNR) at one ear of the listener, commonly termed the "better-ear". In the case that only one spatially separated distractor is involved, the better-ear is typically the ear opposite (or contralateral) to the distractor. In conditions where fluctuating distractors (e.g., speech) are on the left and right side of the listener, the better-ear constantly changes over time and frequency. In such conditions the auditory system can take advantage of the frequency-dependent, short-term (or local) SNR differences, a process termed better-ear glimpsing (e.g., Brungart and Iyer, 2012). Better-ear glimpsing (BEG) is very apparent in conditions in which the target speech is presented from the front and two fluctuating (speech) distractors are presented symmetrically from the left and right (e.g., at azimuth angles of $\pm 60^o$ or $\pm 90^o$) of the listener (e.g., Brungart and Iyer, 2012; Glyde *et al*, 2013b). In this condition, which is frequently used in the laboratory and also in the present study, head-shadow creates only short-term SNR differences between the two ears but no long-term differences.

Although the auditory processes underlying BEG are still unclear, it is typically considered to be a purely signal-energy driven process that utilizes short-term interaural level differences (ILDs) at the two ears to improve the effective overall SNR (Brungart and Iyer, 2012; Glyde *et al*, 2013b). Existing auditory glimpsing models apply a short-term frequency analysis separately to the signals at the two ears from which they then generate an enhanced (mono) signal by always picking the time-frequency bin from the ear with the better local SNR (e.g., Brungart and Iyer, 2012; Glyde *et al*, 2013b, Best *et al*, 2015). Using a symmetric, two-talker distractor paradigm, Brungart and Iyer (2012) demonstrated that the improvement in intelligibility gained with such modern approach is very similar to the spatial benefit achieved by their test subjects. However, Culling and Mansell (2013) showed that the required ILD processing of the auditory system is by far too sluggish to realize the rather short time constants of about 20 ms used in these BEG models. Even though this argument may be alleviated by the observation of Glyde *et al* (2013b) who found that increasing the time constant in their glimpsing model from 20 ms to 100 ms had no noticeable effect on the spatial benefit achieved with a symmetric two-talker distractor, it is still unclear how BEG is realized within the auditory system. An alternative approach, for instance, may assume that the auditory system is taking advantage of glimpses from either of the two ears separately, whereby in each ear one distractor is highly attenuated by head shadow. Hence, the auditory task may effectively be simplified from understanding speech in a two-talker background to understanding speech in a single talker background (per ear), and thereby providing a SRM. In any case, the auditory system is able to take advantage of physical glimpses provided by fluctuating distractors, which in the case of a symmetric two-talker distractor have a duration of around 50-250 ms and a highly varying bandwidth of 1 to more than 20 consecutive auditory channels (Brungart and Iyer, 2012). The number or duration and bandwidth of the physical glimpses available in the ear signals, and thus the benefit expected from BEG, will depend on factors such as the number and type of distracting talkers and their talking style, the amount of room reverberation that is present, the applied sentence material, and the synchronization between target and distracting talkers, as for instance provided by the coordinate response measure (CRM: Bolia *et al.*, 2000).

Since BEG relies on intensity differences between the two ears, it is typically observed at high frequencies (above about 1.5 kHz) where head shadow effects are significant. Unfortunately, most HI individuals have a greater degree of hearing loss at high frequencies (Dillon, 2012, pp. 286-335) and therefore, have limited access to the cues that are relevant for BEG. As a consequence, they often show substantially reduced SRM (Glyde *et al*., 2013a) which makes it harder for them to understand speech in noisy conditions. Even though amplification in hearing aids may help to at least partially restore the audibility of the required high frequency cues (Glyde *et al*, 2013a), the amplification prescribed by standard (non-linear) fitting rules such as NAL-NL2 from the National Acoustic Laboratories (Dillon, 2012, pp.290-297) or CAM2 from

Cambridge (Moore *et al*., 2010) are designed to restore intelligibility and/or loudness perception in quiet and are not designed to restore the cues required for BEG. But even if the latter was the case, then the very limited dynamic range available in most HI listeners at high frequencies would make it difficult to apply appropriate amplification that on the one hand provides NH audibility of the important cues and on the other hand provides acceptable loudness levels. Moreover, the required non-linear (compressive) amplification in hearing aids may further reduce important level fluctuations at the two ears.

An alternative to modifying amplification in hearing aids at high frequencies could be to transform the BEG cues to lower frequencies where the hearing loss is typically less severe or hearing is even normal. This may be achieved by applying similar signal processing techniques that are used in current hearing aids to transpose high-frequency speech cues to lower frequencies (e.g., Robinson *et al.*, 2007) or by digitally controlling the directivity of hearing aids using multiple microphones (e.g., Kates, 2008, pp. 93-98). Either way, bilateral hearing aids may be able to generate substantial ILDs at frequencies well below which natural ILDs occur. However, it is unclear to what extent the normal and impaired auditory system would be able to utilize these artificially generated ILD cues in BEG to improve speech intelligibility in spatialized noise. Thereby, even though reduced audibility will be less of a concern at low frequencies, the gained spatial benefit may still be limited by other factors of hearing loss or age such as reduced spectral and temporal resolution, reduced ability to utilize (supra-threshold) spatial information, or reduced cognitive performance or spatial attention (e.g., Glyde *et al.*, 2015).

To the best knowledge of the authors, no study exists that systematically investigates the strength of BEG at low frequencies (as compared to high frequencies) and its effect on speech intelligibility in noise, and only a few studies have considered the auditory processing of low-frequency ILDs in more general terms. According to ANSI S3.5 (1997) the importance of speech is frequency dependent, and is typically reduced at low frequencies. Hence, it may be expected that the benefit in speech intelligibility provided by low-frequency ILDs follows a similar frequency dependency. Brungart and Rabinowitz (1999) reported that ILDs generally increase as a lateral source approaches the listener, and even at low frequencies can exceed 20 dB when the source is at a distance of less than 0.3m from a listener. Brungart *et al*. (1999) have shown that these low-frequency ILD cues are essential for distance localization of nearby, lateral sources. Shinn-Cunningham *et al*. (2001) showed that the amount of SRM changes substantially when target and distractor sources are moved closer (i.e. 0.15 m instead of 1 m) to a listener, but did not explicitly address the effect of low-frequency ILDs. Blauert (1997, pp-155-164) reported that ILDs down to very low frequencies are successfully used by the auditory system to lateralize sounds. Finally, utilization of ILD cues at low frequencies has also been supported by an electrophysiological study by Krishnan and McDaniel (1998) who reported a reduction in the frequency following response interaction component (FFR-BIC) for 500 Hz tone bursts when ILDs were gradually increased to 30 dB. Hence, it is shown in the literature that the auditory system is in general sensitive to low-frequency ILDs, but the effect of low-frequency ILDs on BEG is still unknown. Therefore, the main goal of this study is to first investigate (1) the extent the normal auditory system can utilize low-frequency ILDs for BEG and thus, for enhancing speech intelligibility in noise, and (2) to what extent this mechanism (or advantage) is preserved in the impaired auditory system and can, therefore, be potentially utilized in bilateral hearing aids.

As mentioned above, BEG is assumed to be a signal-energy based auditory mechanism and is therefore directly linked to a release from energetic masking (EM). However, in most speech-on-speech masking tasks informational masking (IM) is involved in addition to EM (e.g., Freyman *et al.,* 1999; Brungart *et al.,* 2001; Watson, 2005; Kidd *et al.,* 2007). Whereas EM occurs at the level of the peripheral auditory pathway due to spectral and temporal overlap between target and distractor, IM occurs at a more central level due to similarities or uncertainties between the target and distractor signals (e.g. Durlach *et al*., 2002; Watson, 2005; Kidd *et al.,* 2007) and typically involves cognitive mechanisms such as selective (spatial) attention (Shinn-Cunningham, 2008). IM is particularly strong in stimulus conditions where target and distractor signals are co-located, but is highly reduced or even absent when target and distractor signals are presented from different locations. In the latter case, it is typically assumed that the provided localization cues remove (or at least highly reduce) IM and thereby leaving EM the remaining limiting factor (Best *et al*., 2013). Since SRM is often measured as the difference between the speech intelligibility performance in the spatially co-located and separated condition, SRM is typically influenced by both EM and IM. Hence, care must be taken when drawing conclusions on BEG from SRM measures. In the BEG studies by Glyde *et al.* (2013b) the influence by IM was therefore manipulated by varying the similarity between the target and distracting talkers. In Brungart and Iyer (2012) the applied modified rhyme test together with the chosen speech distractors resulted in minimal IM. In Westermann and Buchholz (2015) a brief overview is provided on methods and limitations that have been used throughout the literature to minimize

IM or to segregate IM and EM, respectively. In the present study, the influence of IM was minimized by applying a largely unintelligible, noise-vocoded speech distractor.

2.2 Methods

The present study consists of two experiments. The first experiment utilized a speech-on-speech masking task to measure the overall effect of artificially-generated broadband ILDs on SRM in NH and HI listeners and then used a noise-vocoded version of the applied speech distractor to estimate the involvement of both IM and BEG. The second experiment investigated the effect of BEG on SRM in NH and HI subjects as a function of frequency region. Thereby the sensation level was kept constant across frequency region to allow conclusions on the frequency dependency of the provided SRM. To estimate the achievable SRM at comfortable sound levels, the NH listeners were tested at substantially higher sensation levels than the HI subjects. In this way the NH data also provided a rough estimate of the maximal advantages that may be achievable in HI subjects by providing adequate ILD enhancement techniques in hearing aids. Ethical clearance was received from the Australian Hearing Human Research Ethics Committee and the Macquarie University Human Research Ethics Committees.

2.2.1 Subjects

Ten NH (hearing thresholds < 15 dBHL at least up to 8 kHz in each ear) listeners aged between 23 and 42 years (mean age of 33.1 years) and ten sensorineural HI listeners aged between 49 and 77 years (mean age of 66.9 years) participated. All subjects received a hearing test in the beginning of their first appointment to either confirm NH or to determine their degree and type of hearing loss. All HI subjects had a symmetric (threshold difference between ears < 10 dB for audiometric frequencies up to 4 kHz), mild to moderate, and sloping hearing loss. The individual and mean hearing thresholds averaged over the left and right ear are shown in figure 2.1 (left panel). Their corresponding four-frequency (0.5, 1, 2, 4 kHz) average hearing loss (4FAHL) was 37.8 ± 7.1 dB and was not correlated with age ($p = 0.674$; right panel of Fig. 2.1). All participants had English as their first language and had no reported attention deficit disorder or intellectual disability.

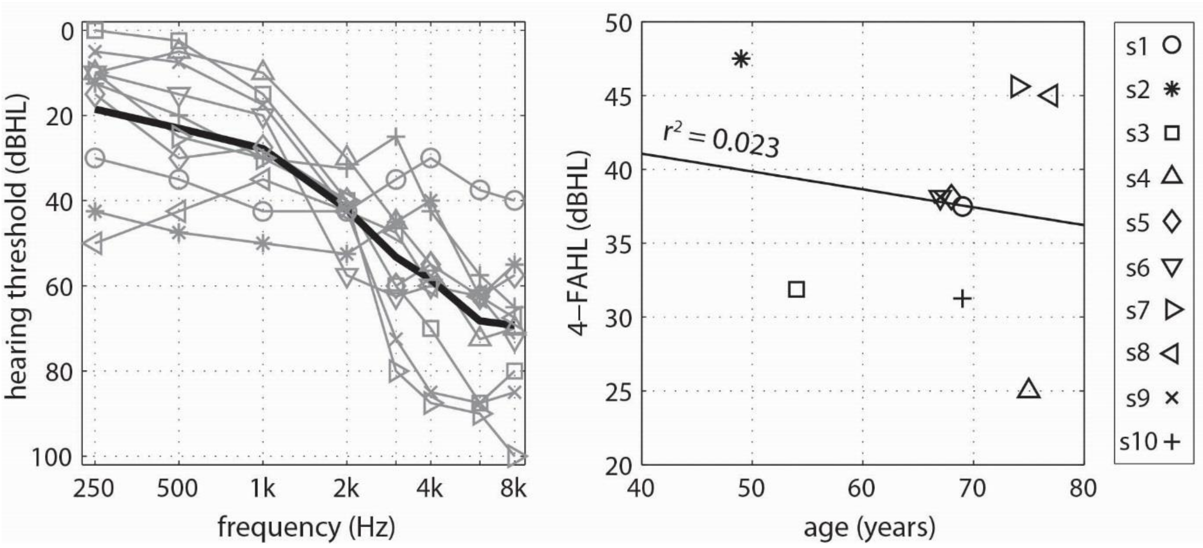

FIG. 2.1: The left panel shows the mean (black line) and individual (grey lines) audiograms for the ten HI test subjects averaged over the left and right ear. The right panel shows the dependency of their 4FAHL with age, which was not correlated ($r^2 = 0.023$, $p = 0.674$).

2.2.2 Procedure

Speech reception thresholds (SRTs) were measured using a Matlab program installed on a personal computer. Similar to the LiSN-S test (Cameron and Dillon, 2007), the participant heard short meaningful target sentences in the presence of two ongoing distractor signals and the task for them was to repeat as many words as they heard in each target sentence. An experimenter then entered the number of correctly identified words into a provided user interface. An adaptive one-up one-down procedure was used to measure the signal-to-noise ratio (SNR) at which 50% correct word identification was achieved by keeping the distractor level constant and varying the target level. The starting SNR was 7 dB and the initial step-size was 4 dB. Once at least five sentences were presented and an upwards reversal occurred the step-size was reduced to 2 dB and the measurement phase started. The adaptive procedure was stopped either when the maximum of 30 sentences was reached or at least 17 sentences were measured and the standard error was below 1 dB. The SRT was then calculated as the average SNR over all measurement trials. Further details can be found in Cameron and Dillon (2007).

Two spatial conditions were considered, a spatially co-located and a spatially separated condition. The difference in SRT between these two conditions provided a measure of the spatial advantage or SRM. To maximize the effect of BEG on the SRM across all frequencies, "infinite" broadband ILDs were applied and ITDs were excluded. In the spatially separated condition, this was simply achieved by presenting one distractor only to the left ear and the other only to the right ear. In the co-located condition both distractors were presented to both ears realizing a diotic stimulus presentation. The target sentences were always presented diotically and were taken from 81 lists of 16 BKB-Like sentences (Bench *et al*., 1979) spoken by a native Australian female talker.

Stimuli were presented through equalized Sennheiser HD215 circumaural headphones connected to a RME fireface UC USB sound card. The spectrum and RMS level of all target and distractor signals was equalized separately in each ear and their RMS-level was calibrated using a Bruel & Kjaer artificial ear. In any tested condition the SRT was averaged over two measurements, which also allowed the calculation of test-retest variability. All testing was conducted in a sound-treated booth at the National Acoustic Laboratories.

2.2.3 Stimuli

In the first experiment, two types of distractors were considered, which were realized by (1) two continuous female speech discourses taken from the different-voice condition of the LiSN-S speech corpus and recorded at a sampling frequency of 44.1 kHz (speech discourse: SD) and (2) largely unintelligible, noise-vocoded versions of the two speech discourses (vocoded speech: VS). Distractor 1 included both EM and IM and provided a reference condition that could be compared to normative LiSN-S data (Cameron *et al*., 2011). Distractor 2 was considered a purely energetic distractor that maintained most of the BEG cues of the SD distractor 1 but at the same time minimized the influence of IM (see Westermann and Buchholz, 2015). The difference in SRT measured with distractor 1 and distractor 2 was used to estimate the amount of IM provided by the SD distractor.

Similar to Westermann and Buchholz (2015), the VS distractor was realized by applying a short-term Fourier transform (STFT) with 20 ms long Hanning windows and 75% overlap separately to the two speech discourse signals. The magnitude of the resulting short-term spectra was then smoothed in the power-domain using the power spectrum of a Gammatone filter with a bandwidth of four Equivalent Rectangular Bandwidths (Patterson *et al.*, 1988). The smoothed (short-term) magnitude spectra were combined with a random phase, transformed into the time domain using an inverse Fourier Transform, and added over time to provide the final VS distractor. This process realized a noise vocoder with about eight effective frequency channels. The resulting (combined) VS distractor was relatively unintelligible and elicited highly reduced IM. However, some words were still intelligible (in particular in the spatially separated conditions), some minor IM may have still remained, and the temporal and spectral smoothing applied within the vocoding process may have reduced some of the dip-listening or BEG cues that are provided by the SD distractor (see figure 2.7 and section 2.4.2).

The long-term spectrum (and RMS level) of all 81 target lists and all distractor signals in either ear were finally equalized to the average spectrum of the entire BKB-like sentence material. This process removed any frequency-dependent differences in long-term SNR across ears and conditions. SRTs were measured with the distractor level fixed at 60 dB SPL and for the HI subjects individual, frequency-specific amplification was provided according to the National Acoustic Laboratories – Revised Profound (NAL-RP) prescription formula (Dillon, 2012, pp. 290-297). The prescription was extended here to 22 kHz by simply setting the required parameter k to -2 dB. At 8 kHz a 16th-order Butterworth lowpass filter was applied. The mean amplification applied to the HI subjects is shown in the left panel of figure 2.2 together with ±1 standard deviation (grey area).

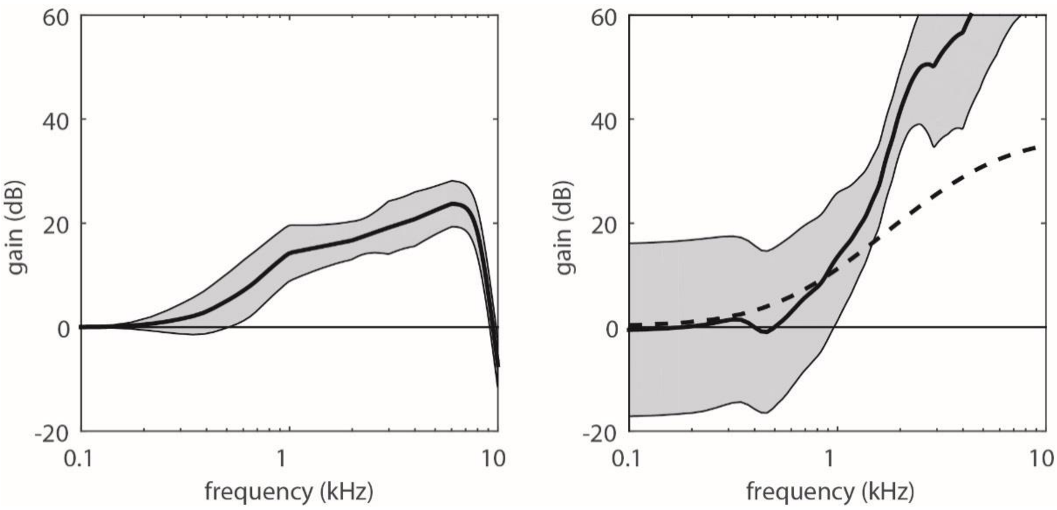

FIG. 2.2: Illustration of the individual, linear gains applied to the test subjects in experiment 1 (left panel) and experiment 2 (right panel). The mean gain applied across all HI subjects is shown by the solid black lines and +/-1 standard deviations are indicated by the grey-shaded areas. The linear gain applied to all NH subjects in experiment 2 is indicated by the dashed line.

The second experiment applied the same methods as described for the first experiment, but differed from the first experiment in three ways. First, to minimize any influence from IM, only the energetic distractor 2 was used i.e. the VS distractor. This was done to focus on BEG, which is considered a purely energy-based auditory mechanism (see section 2.1). Second, instead of applying a gain according to NAL-RP, the target and distractor signals were amplified such that they provided equal audibility across frequency within the NH and HI group. The required amplification was derived by first applying a NH auditory bandpass filterbank (Patterson *et al.*, 1987) with the centre frequencies given in ANSI S3.5 (1997, table 1) to the target and distractor signals and then adjusting the gain of a filter such that the resulting output levels in each frequency channel equalled the individual threshold in quiet for pure tones. Thereby the threshold in quiet, as a function of frequency, was determined by the thresholds given in the SII standard (ANSI S3.5-1997, table 1) to which the individual pure-tone audiogram levels (figure 2.1) were added. The audiogram levels were interpolated on a double-logarithmic frequency scale to match the center frequencies of the auditory filterbank. The overall filter gain was then adjusted such that for a target/distractor level of 60 dBSPL the speech level in each auditory channel for the NH subjects was 35 dB above threshold and 10 dB above threshold for the HI subjects. The overall gain provided across frequencies is shown in figure 2.2 (right panel) for both NH and HI individuals. Third, the distractor and target signals were filtered into different frequency regions using 8th order Butterworth band pass filters. For the NH subjects the target and distractor signals were divided into the following four frequency regions: *Low*: 100-770 Hz; *Mid*: 770-2.000 Hz; *High*: 2.000–5.300 Hz and *Broadband*: 100-5.300 Hz, and for the HI subjects into the following three frequency regions: *Low*: 100-770Hz; *Mid*: 770-2000 Hz; and *Broadband*: 100-2000 Hz. Each of the narrow band frequency regions contained six critical bands as defined by the SII standard (ANSI S3.5-1997, table 1). Originally, the idea was to test the HI group in the same four frequency regions as the NH group. Unfortunately, in the *High* and *Broadband* condition it was not possible to provide adequate (i.e., 10 dB) audibility for many of our subjects without exceeding comfortable loudness levels. Hence, for the HI subjects, we removed the *High* condition and reduced the bandwidth of the *Broadband* condition accordingly.

The entire test took about two and a half hours. Within each experiment all conditions were tested twice forming two successive testing blocks. Within each block the conditions were randomized. Subjects could take breaks as required, but were asked to take short breaks at least every 15 minutes.

2.3 Results

Statistical analysis was done using IBM SPSS statistics version 22.

2.3.1 Experiment 1

Mean SRTs with 95 % confidence intervals are shown in figure 2.3 for the two different distractor conditions. Performance was consistently worse for the HI group than for the NH group for both distractor conditions.

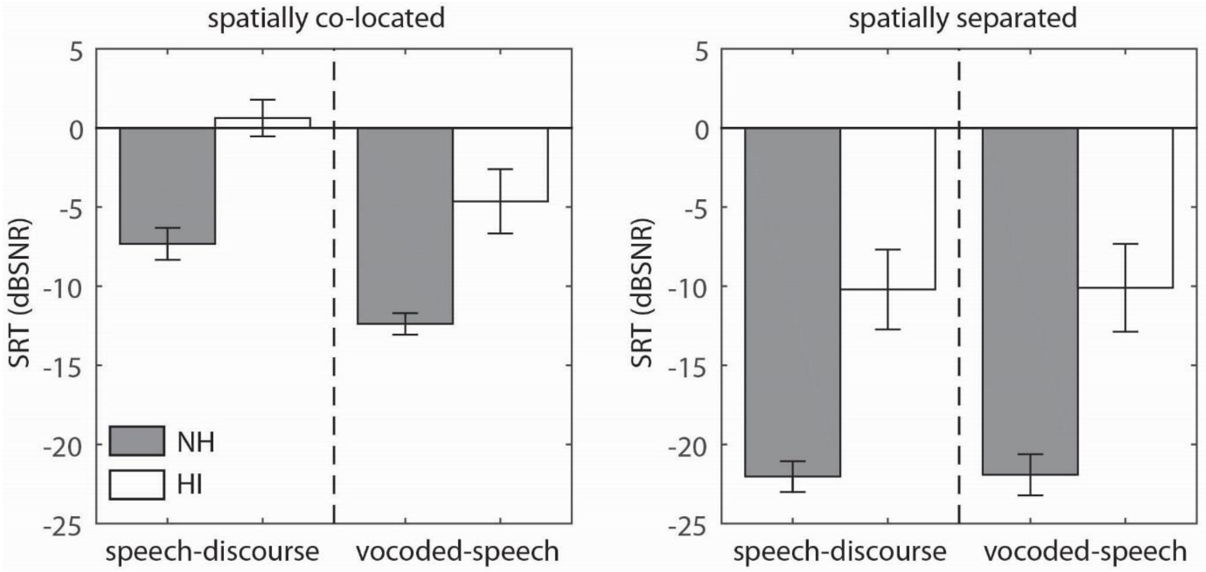

FIG. 2.3: Mean SRTs and 95 % confidence intervals for NH (grey bars) and HI (white bars) subjects obtained with different distractors in a spatially co-located (left panel) and separated (right panel) condition.

In figure 2.4, the individual data from figure 2.3 for the spatially separated SRTs are plotted against the corresponding co-located SRTs. For the HI subjects, these SRTs are correlated for both the SD distractor (left panel, $r^2 = 0.90$, $p < 0.001$) and the VS distractor (right panel, $r^2 = 0.47$, $p = 0.028$). The slopes of the corresponding linear regression lines (solid lines) are $\beta = 1.29$ and $\beta = 1.49$, respectively. However, the correlation for the VS distractor is mainly driven by the very high SRTs of subject s5 (diamonds). Excluding s5 from the analysis results in an insignificant correlation for the VS distractor ($r^2 = 0.08$, $p = 0.46$) and has only a minor effect for the SD distractor ($r^2 = 0.79$, $p = 0.001$, $\beta = 1.23$). Including the NH data in the regression analysis (dashed lines) results in a highly increased correlation while exhibiting very similar slopes (figure 2.4, dashed lines). The NH data alone showed no significant correlations ($p > 0.1$).

Although not shown here, for the HI subjects individual SRTs were also significantly correlated between the SD and VS distractor in both the spatially separated ($r^2 = 0.90$, $p < 0.001$, $\beta = 0.86$) and co-located condition ($r^2 = 0.54$, $p = 0.016$, $\beta = 0.42$). However, in the co-located condition the correlation was solely due to the very high SRTs of subject s5. The NH data showed no significant correlations ($p > 0.1$). None of the SRTs were correlated with age ($p > 0.46$) or 4-FAHL ($p > 0.25$). In the following the SRT data is further analyzed by “extracting” the involved: (1) amount of IM, (2) SRM, and (3) test-retest variability.

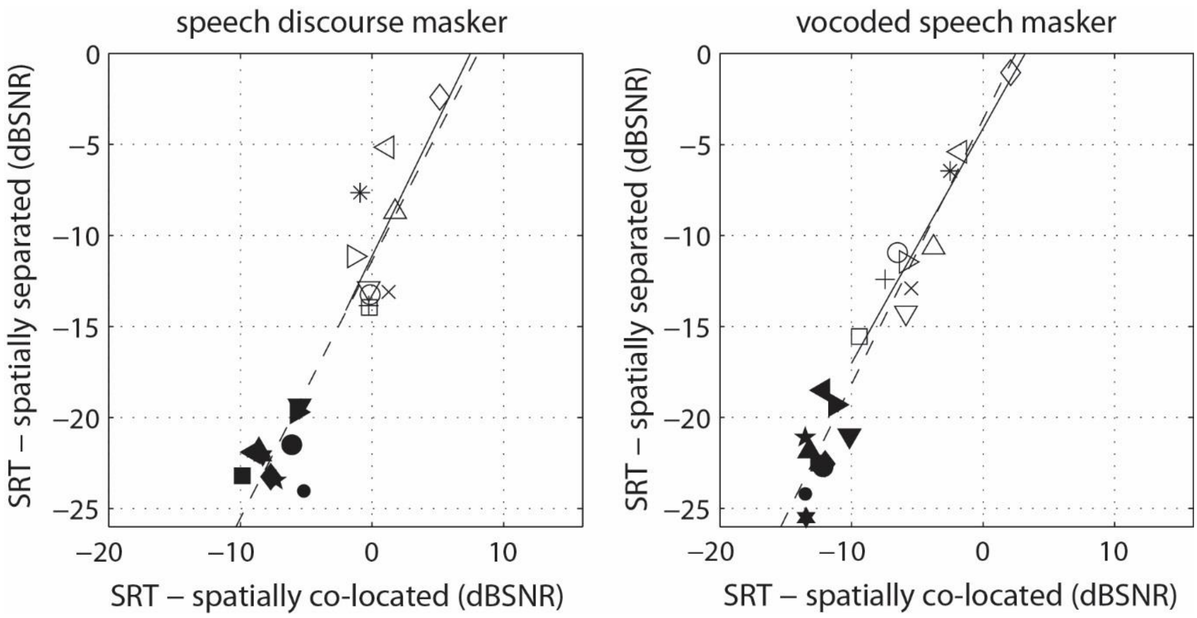

FIG. 2.4: Comparison of the individual SRTs (from figure 2.3) for the spatially separated and co-located condition. The left panel shows the results for the SD distractor and the right panel for the VS distractor. The solid lines present linear regression lines that were fitted only to the HI SRTs (open symbols) and the dashed lines present the case when HI, as well as NH (filled symbols), SRTs were considered.

2.3.1.1 Informational masking

Similar to Westermann and Buchholz (2015), the involved amount of IM was calculated as the difference between the SRTs obtained with the VS distractor and the SD distractor and the results are shown in figure 2.5 (left panel). A linear mixed-effects model with IM as the dependent variable, spatial separation, hearing status, and their interaction as fixed effects and a subject-specific intercept as the random effect revealed a significant effect of spatial separation [F (1, 36) = 80.92, p<0.01] but not of hearing status [F (1, 36) = 0.32, p>0.01]. No significant interaction was noted [F (1, 36) = 0.02, p>0.01]. For both the NH and HI group a paired t-test revealed that the amount of IM in the co-located condition was significant with 5.5 dB (NH: t(9) = 7.47, p <0.01; HI: t(9) = 7.21, p <0.01) but not in the spatially separated condition with -0.1 dB. Hence, a significant part of the IM contained in the co-located SD interferer was removed by applying noise-vocoded speech, irrespective of an individual's hearing status. Moreover, the spatial cues provided in the spatially separated condition resolved any notable IM. The amount of IM was neither correlated with age nor 4FAHL ($p > 0.1$).

2.3.1.2 Spatial release from masking (SRM)

The SRM calculated from the SRT data given in figure 2.3 is shown in figure 2.5 (right panel) for the NH and HI group for both distractors. The SRM obtained in the HI group was significantly smaller than for the NH group with both VS and SD distractors. Also, the SRM obtained with the VS distractor was significantly smaller than with the SD distractor in both the NH and HI group. This was confirmed by a linear mixed-effects model with SRM as the dependent variable, distractor type, hearing status, and their interaction as fixed effects and a subject-specific intercept as the random effect, which revealed a significant effect of hearing status [$F (1, 36) = 33.15$, $p<0.01$] as well as distractor type [$F (1, 36) =58.07$, $p<0.01$] but no significant interaction [$F (1, 36) = 0.02$, $p>0.01$].

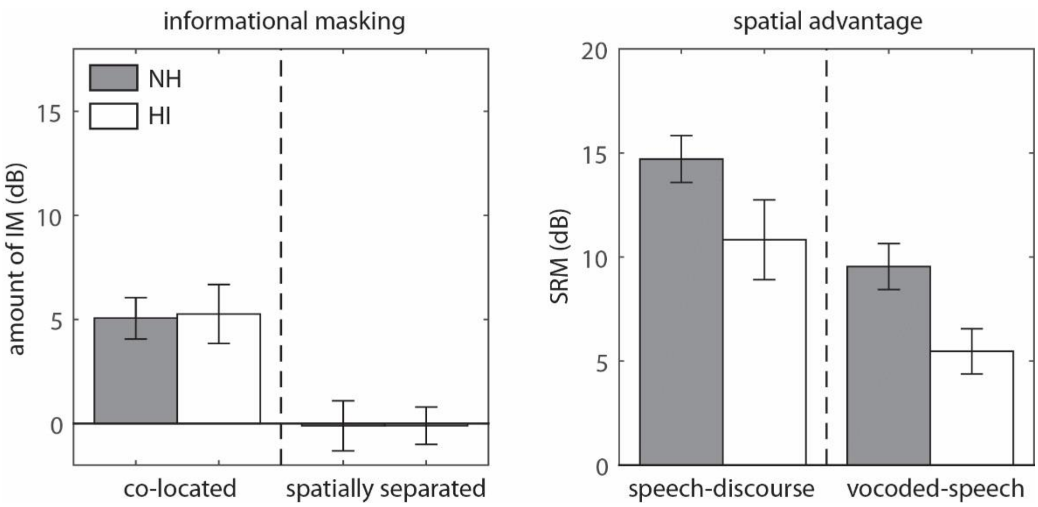

FIG. 2.5: Mean IM (left panel) and SRM (right panel) calculated from the individual SRT data shown in figure 2.3 for NH (grey bars) and HI (white bars) subjects. Error bars indicate 95% confidence intervals.

The mean SRM for the SD distractor was 14.7 dB in the NH group and 10.8 dB in the HI group. Since the amount of IM as shown in figure 2.5 (left panel) provided by the SD distractor was about 5.5 dB larger in the co-located than in the spatially separated condition (section 2.3.1.1), the SRM for the mainly energetic VS distractor was reduced to 9.2 dB in the NH group and to

5.3 dB in the HI group. The amount of SRM with SD and VS distractors for HI group was neither correlated with age nor 4FAHL ($p > 0.05$). From the slopes (i.e., $\beta > 1$) of the linear regression analysis of the individual SRT data shown in figure 2.4 it can be deduced that the higher an individual's SRT in the spatially co-located condition the even higher that individual's SRT in the spatially separated condition. Hence, the higher the SRT in the spatially separated (or co-located) condition the smaller the observed SRM.

2.3.1.3 Test-retest variability

Test-retest reliability, as well as the mean difference between the first and second SRT measurements (i.e., first SRT minus second SRT), was calculated separately for each condition, and the results are summarized in table 2.1. For NH subjects a paired t-test between the first and second SRT measurements revealed no significant differences ($p > 0.01$) for all conditions. For HI subjects a significant ($p < 0.01$) difference (i.e., increase) was found for the two SD distractor conditions as well as for the VS distractor in the spatially separated condition, indicating a small but significant training effect for these conditions of about 1-2 dB. The test-retest reliability given by the intra-subject standard deviation was between 0.92 dB and 1.64 dB for the NH subjects, except for the SD distractor in the co-located condition where a substantially higher value of 3.18 dB was observed. This lower reliability is most likely due to the involvement of IM (see section 2.3.1.1). For the HI subjects, the intra-subject standard deviation showed a similar variation across conditions as observed for the NH subjects and ranged from 1.05 dB to 2.27 dB.

Most of the observed intra-subject standard deviations are very similar to the ones reported by Keidser *et al*. (2013). They measured SRTs with BKB sentences in diffuse babble noise and found a standard deviation of 1.3 dB for NH and 1.4 dB for HI subjects. Similarly, Cameron *et al*. (2011) reported an intra-subject standard deviation of 2.2 dB in co-located SRTs and 1.6 dB in spatially separated SRTs obtained using the same voice speech distractors from the LiSN-S test in NH participants. Cameron *et al.* (2011) also reported a small but significant learning effect of about 1 dB.

TABLE 2.1: Mean difference with intra-subject standard deviation (STD) between the first and second SRT measurement. SD: speech-discourse distractor; VS: vocoded-speech distractor.

Conditions	**Co-located**		**Spatially separated**	
Distractors	SD	VS	SD	VS
NH subjects				
Mean difference (dB)	2.99	0.28	-0.46	1.52
Intra-subject STD (dB)	3.18	0.92	1.60	1.64
HI subjects				
Mean difference (dB)	1.32*	-0.18	2.10*	1.71*
Intra-subject STD (dB)	1.05	2.27	1.56	1.51

* indicates a significant effect ($p < 0.01$)

2.3.2 Experiment 2

Mean SRTs with 95% confidence intervals are shown in figure 2.6 for the NH and HI group as a function of frequency region for both the spatially co-located (left panel) and separated condition (right panel). Individual SRTs are given in the Appendix and were compared across all frequency regions, both between and within the two groups. Due to loudness discomfort, only eight out of ten HI subjects completed the low-frequency condition (subjects s2 and s8 were excluded) and only nine HI subjects completed the broadband condition (subject s2 was excluded). Again, subject s5 showed substantially higher SRTs than all other HI subjects, with SRTs almost two standard deviations above the mean. However, subject s5 was not excluded from the subsequent data analysis since its presence (or absence) had no major effect on the main conclusions.

A linear mixed-effects model with frequency region, spatial separation, hearing status, and their two- and three-way interactions as fixed effects and a subject-specific intercept as the random effect confirmed significant effects of frequency region [$F(2, 84.56) = 323.41$, $p<0.01$], spatial separation [$F(1, 84.11) = 524.35$, $p<0.01$], and hearing status [$F(1, 18.10) = 27.41$, $p<0.01$]. A significant interaction was only observed for spatial separation and hearing status [$F(1, 84.11) = 21.52$, $p<0.01$] as well as for frequency region and hearing status [$F(2, 84.56) = 11.21$, $p<0.01$].

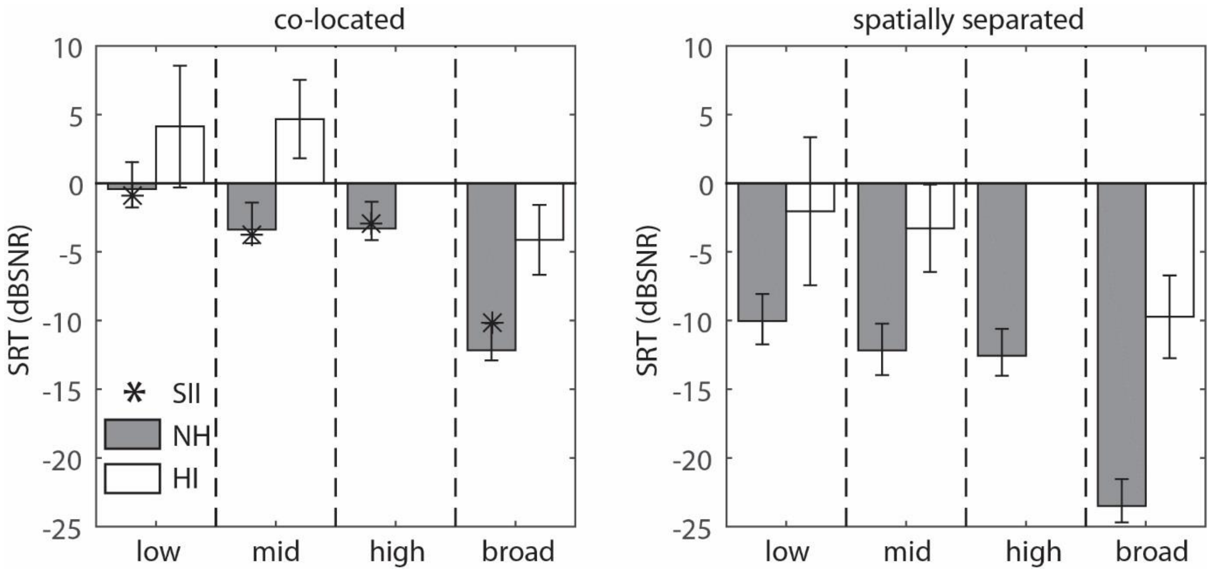

FIG. 2.6: Mean SRTs and 95 % confidence intervals for the NH (grey bars) and HI (white bars) group as a function of frequency region. The results for the co-located condition are shown in the left panel and for the spatially separated condition in the right panel. SII predictions for the NH subjects in the co-located conditions are indicated by stars.

An independent t-test with adjusted p-values using Holm-Bonferroni method (Holm, 1979) revealed that the SRTs were significantly poorer (familywise type I error rate was kept fixed at 0.01) in the HI group than in the NH group for all conditions except for the *Low* frequency region in both the co-located and spatially separated condition. The statistical similarity in SRTs at low frequencies between the NH and HI group was mainly due to the large standard deviation observed in the HI group. When only the mean values shown in figure 2.6 are considered, the SRTs in the *Low* frequency region for the HI group are clearly above the ones for the NH group. Please note that the *High* frequency condition was neither considered in the mixed analysis nor in the independent t-test, since it was measured only for the NH group. The broadband condition was considered in the analysis even though the actual bandwidth was different for the two groups. It can be noted that the 95% confidence intervals (error bars) shown in figure 2.6 are in general larger than in experiment 1 (figure 2.3), highlighting an increased SRT variation across subjects. The SRTs for the HI subjects in the spatially separated condition were all significantly correlated with the corresponding SRTs in the co-located condition ($p < 0.01$) and a linear regression analysis revealed slopes of around one (Low: $\beta = 1.14$; Mid: $\beta = 1.07$; Broad: $\beta = 1.14$). None of the SRTs were correlated with age or 4FAHL ($p > 0.1$). No significant correlations were found for the NH subjects ($p > 0.1$).

2.3.2.1 Spatial release from masking (SRM)

The mean SRM and 95% confidence intervals are shown in figure 2.7 as a function of frequency for the NH and HI group. A linear mixed-effects model with hearing status, frequency, and their interaction as fixed effects and a subject-specific intercept as the random effect revealed a significant effect of hearing status [$F(1, 18.09) = 28.55$, $p<0.01$], a non-significant effect of frequency [$F(2, 35.01) =0.34$, $p> 0.01$] and a significant interaction [$F(2, 35.01) = 5.46$, $p< 0.01$] between the two. No significant correlations were found with age or 4FAHL ($p > 0.1$).

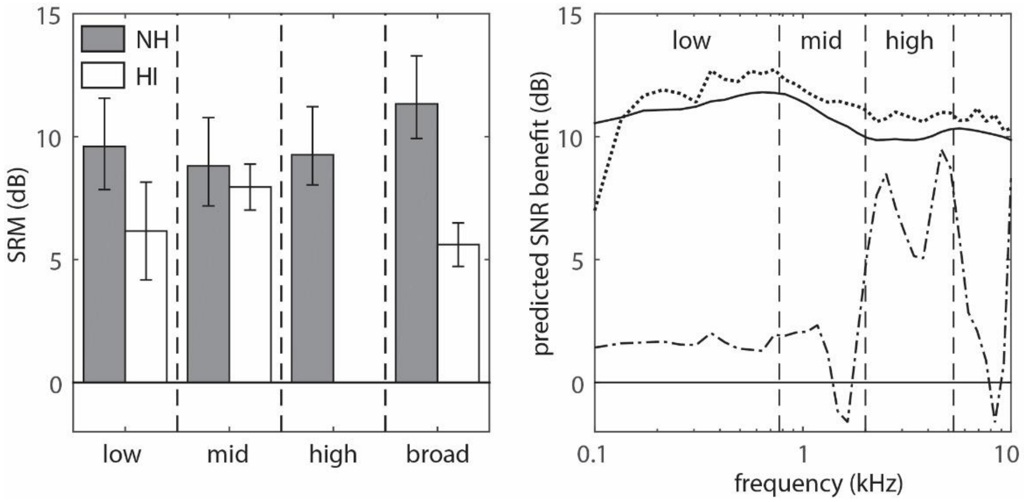

FIG. 2.7: Mean SRM and 95 % confidence intervals for NH and HI subjects as a function of frequency region (left panel) and corresponding SNR benefit predicted by a BEG model (right panel). The solid line refers to the case that infinite ILDs are applied to the VS masker (see experiment 1 and 2) and is directly related to the SRM data shown in the left panel, The dotted line refers to the case that infinite ILDs are applied to the SD masker (see experiment 1) and, for reference purposes, the dashed-dotted line refers to the case that "standard" HRTFs are applied to the VS masker. See section 2.4.2 for details.

The SRM for the HI group was consistently smaller than for the NH group, but an independent t-test with corrected p-values (Holm-Bonferroni method) showed significance only for the broadband condition. Alpha values were larger than 0.01 for the *Low* ($p = 0.02$) and *Mid* frequency condition ($p = 0.38$). A paired t-test with corrected p-values (Holm-Bonferroni method) revealed no significant differences between the SRM in the broad band and all narrow band conditions within both groups.

This indicates that NH individuals can utilize the provided frequency-independent ILD cues equally well at low, mid, and high frequencies, and the same accounts for the HI group who showed a similar sensitivity to the provided ILD cues at low and mid frequencies. As shown in figure 2.7, HI individuals cannot only utilize low and mid frequency ILDs for BEG, and thus for improving speech intelligibility in spatialized noise, they can utilize them almost as well as NH individuals. This is true even though the audibility was significantly lower for the HI than NH subjects (see section 2.2.3).

2.3.2.2 Test-retest variability

Similar to experiment 1 (section 2.3.1.3), the test-retest reliability, as well as the mean difference between the first and second SRT measurements (i.e., first SRT minus second SRT), were derived and the results are summarized in table 2.2. Paired t-test between the first and second SRTs revealed no significant differences ($p > 0.01$) for all conditions for both groups.

TABLE 2.2: Mean difference and intra-subject standard deviation (STD) between first and second SRT measurements.

Conditions	**Co-located**				**Spatially separated**			
Frequency bands	Low	Mid	High	Broadband	Low	Mid	High	Broadband
NH subjects								
Mean difference (dB)	0.64	0.92	0.94	-0.32	0.15	-0.19	0.66	1.46
Intra-subject STD (dB)	1.92	2.14	2.36	2.11	2.17	2.51	2.64	2.03
HI subjects								
Mean difference (dB)	0.34	2.04	NA	-1.41	-2.35	1.45	NA	1.18
Intra-subject STD (dB)	3.20	3.80	NA	1.86	2.69	3.25	NA	2.58

The intra-subject standard deviation for the NH subjects was between 1.92 and 2.64 dB and thus, substantially higher than in experiment 1 as well as for other data reported in literature (section

2.3.1.3). For the HI subjects, the intra-subject standard deviation was between 1.86 and 3.80 dB and thus, substantially larger than in experiment 1 as well as for the NH subjects.

2.4 Discussion

The purpose of experiment 1 was mainly to investigate the spatial benefit (i.e., SRM) that can be achieved in NH and HI subjects by maximizing ILDs across all frequencies, in particular at low frequencies. Furthermore, the extent of the observed spatial benefit due to a release from EM (as opposed to IM) and thus, due to BEG was investigated. In experiment 2 the effect of BEG was then further investigated as a function of frequency and hearing status. The results and implications of the two experiments are discussed in the following sections.

2.4.1 Maximized ILDs

To investigate the potential increase in SRM that can be achieved by (artificially) maximizing ILDs across the entire frequency range a speech intelligibility test was conducted that was very similar to the LiSN-S test (Cameron *et al.*, 2011). The main differences were: (i) the head-related transfer function (HRTFs) were replaced by artificially generated transfer functions that provided maximal possible ILDs across all frequencies and no ITDs, (ii) the different-voice speech-distractors from the LiSN-S test were used but equalized to provide the same long-term spectrum as the target speech material, and (iii) the distractor level was increased from 55 to 60 dBSPL to provide a more realistic speech level.

Comparing the present results with the corresponding LiSN-S results (Cameron *et al.*, 2011, figure 2.2, age 30-39 years), the SRTs for the NH subjects are higher here by about 3 dB in the co-located condition and lower by about 4 dB in the spatially separated condition. Comparing the results for the HI subjects to the results provided by Glyde *et al.* (2013a) for subjects with a similar 4FAHL of on average 37.8 dB, similar observations can be made. The SRTs are higher here by about 4 dB in the co-located condition and lower by about 1.5 dB in the spatially separated condition. Hence, by providing enhanced (low-frequency) ILDs, NH and HI subjects received a significant spatial benefit of 7 or 5.5 dB, respectively, on top of the one already provided by realistic ILDs (as used in the LiSN-S).

The higher SRTs in the co-located condition (when compared to the LiSN-S) are most likely due to an increase in both EM and IM that was introduced by the equalization of the distractor, which made the long-term spectrum of the distractor equal to the target speech. The lower SRTs in the spatially separated condition are most likely explained by the enhanced ILDs improving the spatial release from EM achieved by BEG, in particular at low frequencies. Hence, the observed increase in SRM is due to an increase in both EM and IM in the co-located condition as well as an enhanced spatial release from EM in the spatially separated condition. Whereas the increase in the SRTs in the co-located condition will have exaggerated the increase in SRM provided by the low-frequency ILDs, limited target audibility (or floor effects) may have limited the SRT in the spatially separated condition and therefore resulted in an underestimation of the increase in SRM.

The observation that the measured SRM was not only significantly smaller in HI subjects than in NH subjects, but also that HI listeners did not benefit as much as NH listeners from the enhanced ILDs (when compared to the LiSN-S), may be explained by a number of factors, such as reduced audibility due to the different sensation levels that were applied to the NH and HI subjects (i.e., the applied NAL-RP amplification does not restore normal audibility), reduced temporal and spectral resolution, and reduced cognitive performance or spatial attention due to age differences; although no significant correlation was found in section 2.3.1.2 between SRM and age. For a further discussion on these factors see Glyde *et al.* (2015).

2.4.2 BEG with maximized ILDs

To estimate the extent the spatial benefit (or SRM) that was observed with the SD distractor in experiment 1 was provided by a spatial release from EM (and not by a spatial release from IM), a second, largely unintelligible and mostly energetic VS distractor was considered, which was generated by noise-vocoding the SD distractor. The SRM measured with the VS distractor was used as an estimate of the contribution of BEG to the SRM measured with the SD distractor. The difference in SRTs measured between the SD and VS distractor was considered as an estimate of the involved IM. In this way, a significant amount of IM of about 5 dB was revealed for both the NH and HI subjects in the co-located condition, but no IM was observed in the spatially separated condition. Hence, out of the 14.7 and 10.8 dB of SRM measured with the SD distractor for the NH and HI subjects BEG contributed 9.5 and 5.5 dB, respectively. The observation that IM is basically resolved once sufficient spatial (or other) cues are available to reliably segregate the target from the distractor is in agreement with the previous literature (e.g., Westermann and Buchholz, 2015; Best *et al.,* 2013; Kidd *et al.,* 2007). Glyde *et al.* (2013b) predicted that in the different-voice condition of the LiSN-S about 5.4 dB of the measured SRM can be attributed to BEG. Similarly, Brungart and Iyer (2012) observed an SRM due to BEG of about 6 dB. In both studies, the effect of BEG for NH listeners was smaller than the 9.5 dB observed here, which most likely can be explained by the ILDs that were (artificially) increased in experiment 1, in particular at low frequencies. To further evaluate the effect of the increased ILDs on the measured SRM, a BEG model was utilized that was very similar to the one described in Glyde et al. (2013b). Within this model, a NH auditory spectrogram is derived separately for the left and right ear signals, by first applying a Gammatone bandpass filterbank with 1-ERB wide filters and then, in each channel, calculating the short-term power within 20-ms long time segments using a Hanning window with 50% overlap. To simulate the effect of BEG, the short-term power within each time-frequency bin is then compared between the two ears, and only the bin with the smaller power is further considered. To estimate the SRM provided by BEG, the described processing was applied to the co-located as well as spatially separated distractors and the difference (in dB) of the two resulting spectrograms was calculated. Within each frequency channel, this difference was then clipped to the range [-5 dB, 20 dB] and averaged over time to derive a metric similar to the segmental SNR (benefit) described by Hansen and Pellom (1998). The resulting predictions of the benefit provided by BEG is shown in Figure 2.7 (right panel) as a function of frequency for the VS distractor applied in experiment 1 and 2 (solid line), as well as the SD distractor, applied in experiment 1 (dotted line). The different gains that were applied in the two experiments had no effect here because audibility was not considered in the model. For reference purposes, the case when the HRTFs used in the LiSN-S test are applied to spatialize the noise-vocoded speech discourses is shown by the dashed-dotted line. For the VS distractor with maximized ILDs the predicted BEG benefit is about 10-11 dB independent of frequency (solid line), which is in rather good agreement with the corresponding measured benefit of 9.5 dB. Comparing the solid and dotted line illustrates the predicted reduction in BEG due to the temporal and spectral smoothing applied by the noise-vocoding process described in section 2.2.3. Comparing the solid and dashed-dotted line illustrates the increase in better ear-glimpsing that is potentially achieved by maximizing ILDs, in particular at low frequencies (section 2.2.3).

It should be noted that the noise-vocoded distractors that were applied here in the spatially separated condition did not only differ from the co-located condition by their (fluctuating) ILDs and thus, in their potential for BEG. The fact that two independent noise carriers were used to generate the two distractors resulted in a spatially separated distractor signal that was uncorrelated between the two ears (i.e., the interaural coherence was equal to zero). In contrast, for the (diotic) target as well as the co-located distractor the signals at the two ears were perfectly correlated (i.e., the interaural coherence was equal to one). The difference in interaural correlation between the target and the spatially separated distractor may have contributed to the SRM observed with the noise-vocoded distractors (e.g., Blauert, 1997, pp. 238-271) and may be falsely attributed here to BEG. To quantify this potential SRM component, an additional test condition was run with the same group of listeners and procedures described in section 2.2, but in this condition the two distractors were independently realized by speech-shaped noise (SSN). The SSN was created by applying a filter with the same long-term spectrum as the target speech to white noise, which resulted in a stationary distractor with minimal level (or ILD) fluctuations. As a consequence, the spatially separated distractors produced a spatially diffuse percept. This was very different for the target as well as the co-located distractors, which were both sharply localized in the center of the head. The resulting mean SRT and ±1 standard deviation for the spatially co-located and separated condition were -9.0 ± 1.0 dB and -11.1 ± 0.8 dB for the NH listeners and -4.9 ± 1.9 dB and -5.9 ± 2.9 dB for the HI listeners. The corresponding SRM for the NH and HI group was thus 2.1 ± 0.9 dB and 1.0 ± 1.3 dB, respectively. Hence, the contribution of BEG to SRM as derived in section 2.3.1.2 with the noise-vocoded distractors (as discussed above) may have been overestimated by 1-2 dB. However, this effect is larger than the one reported by Licklider (1948), who found that the intelligibility of diotic target speech is only improved by 0.5-1 dB when a diotic noise is replaced by interaurally uncorrelated noise. Applying the BEG model described above to the SSN noise stimuli, a SRM of about 1-1.5 dB is predicted. This in turn suggests that the considered SSN noise stimuli may still involve some BEG, either in addition or alternatively to a purely coherence-based process. Either way, it cannot be ruled out that a small part of the SRM measured with the vocoded-noise distractors is provided by other auditory processes than BEG.

Not many studies exist that have measured speech intelligibility in noise and applied artificially enhanced (or infinite) ILDs as done in the present study. Best *et al.* (2013) measured both SRTs and SRM in NH and HI listeners using a very similar, two-distractor ILD-only condition (amongst other conditions) as applied here. The SRM noted by them was approximately 2-2.5 dB larger in NH individuals than measured here but around 3 dB smaller in HI individuals when speech distractors were used, and around 1-2 dB smaller in both groups when noise-vocoded speech distractors were used. The higher SRM for NH subjects in the speech-distractor condition was most likely due to the very high IM that was provided by the applied coordinate response measure (CRM) speech corpus in the co-located condition, which was then resolved in the spatially separated condition. The reduced SRM in the noise-vocoded speech distractor condition was most likely due to reduced ILD fluctuations or BEG, respectively, because Best *et al* applied a single channel noise-vocoder with a 50-ms long smoothing window to generate the noise-vocoded speech distractor whereas in the present study a multi-channel noise-vocoder was applied with a 20-ms long smoothing window. The reduced SRM observed by Best *et al* in HI subjects for both distractor conditions was mainly due to higher SRTs in the spatially separated condition, which may have been limited by reduced audibility of the target speech (or floor effects) as well as of the spatial cues (i.e., ILDs) required for BEG. In the present study, 50 % of HI subjects had pure-tone thresholds much better than 20 dB at low and mid frequencies, whereas in Best *et al* most of the HI subjects had thresholds around 20 dB or higher. The potential effect of reduced audibility on the SRM is extensively discussed by Glyde *et al.* (2015).

2.4.3 Frequency-dependency of BEG

In experiment 2, SRTs were measured in NH and HI listeners as a function of frequency region using only the noise-vocoded (mostly energetic) distractor. Due to loudness discomfort, SRTs in HI subjects could only be measured in the *Low*, *Mid*, and (reduced) *Broadband* frequency region, but not in the *High* frequency region. NH listeners were tested in all four frequency regions. Both groups showed a substantial SRM in all tested frequency regions (Fig. 2.7), which was (slightly) larger in the NH group than in the HI group. Moreover, overall speech intelligibility was consistently poorer for the HI subjects than for the NH listeners.

In both groups the SRTs did not vary significantly between the different narrowband conditions, except for the NH listeners in the *low* frequency region where SRTs were elevated by about 2-3 dB in the co-located condition. The SRTs in the broadband condition were substantially lower (performance improved) than in the narrowband conditions for both subject groups and both spatial conditions. In the NH subjects, this difference due to increased bandwidth was about 9-12 dB in the co-located condition and 11-14 dB in the spatially separated condition. In the HI subjects, this difference was about 8-9 dB in the co-located condition and around 7 dB in the spatially separated condition.

To better understand the observed speech intelligibility variation across frequency region and bandwidth, the speech intelligibility index (SII) (ANSI S3.5-1997, table 1) was applied to the spatially co-located (diotic) conditions and compared to the corresponding NH results. Thereby, to remove any processing artifacts, the SII band-importance function (I_i) was set to zero for all critical bands outside the considered frequency ranges. Hence, in the SII calculation (equation 14) only six critical bands were considered in each of the three narrow band conditions and 18 bands in the broadband condition. The SRTs were then predicted by first deriving psychometric functions (i.e., the SII as a function of SNR) for each of the four conditions and then finding the SII value at which the corresponding SNRs fit best the measured SRTs (in a least squared error sense). The resulting SII value was 0.101 and the predicted SRTs for the *Low*, *Mid*, *High*, and *Broadband* conditions were -0.9, -3.7, -3, and -10.2 dB, respectively. The SII predictions corresponded very well to the measured SRTs (Fig. 2.6, left panel, grey bars versus stars), with an overall RMS error of 1.1 dB. Hence, the observed increase in SRT in the *Low* frequency region can be explained by the reduced importance of frequency bands below 450 Hz, and the observed decrease in SRT in the *Broadband* condition can be explained by the increased number of bands contributing to overall intelligibility.

In a similar way, the SII was applied to predict the measured SRTs for the HI subjects in the co-located conditions, taking into account the average audiogram shown in figure 2.1. However, the resulting SRT predictions did strongly over-predict (by up to 8 dB) the effect of hearing loss on speech intelligibility performance and are therefore not considered any further. Nevertheless, it is expected that the reduced benefit noted between the broadband and narrowband conditions (when compared to NH subjects) is due to the reduced sensation level of the distractor (i.e., 10 dB versus 35 dB in the NH group), which limited the adaptive SRT towards lower SNRs, and maybe also due to the reduced bandwidth of the broadband condition (i.e., 100-2.000 Hz versus 100-5.300 Hz in the NH group). The overall increase in SRTs between HI and NH subjects in the co-located condition of 5-8 dB may be explained by factors such as reduced audibility, reduced temporal and spectral resolution, and reduced cognitive function due to an increased age of the HI subjects (section 2.2.1).

The above observation that the NH, as well as HI subjects, showed a substantial SRM across all tested frequency regions, which was not very different between the different narrowband conditions (i.e., of about 9 dB for NH and about 7 dB for HI subjects), indicates that the auditory system can utilize BEG cues equally well across frequency to improve speech intelligibility in spatial noise. This is in agreement with the benefit predicted by the BEG model for NH listeners described in section 2.4.2 and shown in figure 2.7 (right panel, solid line), which predicts a frequency independent benefit of about 10-11 dB. Since the SRM for NH listeners increased from about 9.2 dB in the narrowband conditions to about 11.3 dB in the broadband condition, the (normal) auditory system seems to be able to combine ILD information across frequency to further improve the spatial advantage provided by BEG. This latter finding is in principle agreement with Kidd *et al*. (2010) who also found that the auditory system integrates spatial information across frequency; although in their case this spectral integration process also involved different spatial cues (i.e., ITDs at low frequencies and ILDs at high frequencies).

In contrast to the NH results, the SRM observed in the HI subjects decreased from about 7.1 dB in the narrowband conditions to about 5.6 dB in the broadband condition. However, this decrease in performance does not necessarily suggest that the impaired auditory system cannot combine ILD information across frequency to improve BEG. Due to the rather low sensation level of the distractor (i.e., 10 dB), and given that the co-located SRT in the broadband condition (with -4.2 dB) was already about 8.5 dB lower than in the narrowband conditions, the SRT in the spatially separated conditions may have been limited by insufficient audibility of the target speech. This assumption is further supported by the experimental data (not shown here) that was measured with six of the HI subjects who participated in experiment 2, who were initially presented with an increased distractor level of 15 dBSL. At this increased sensation level the observed SRM in the broadband condition was about 2 dB higher than at 10 dBSL. Unfortunately, due to loudness discomfort, only six subjects were tested at this increased sensation level. For the NH subjects in experiment 2, limited audibility would have played a less significant role due to the rather high distractor levels of 35 dBSL. However, at least for the spatially separated broadband condition with an average SRT of -23.5 dB, floor effects cannot be fully ruled out. Finally, it should be noted that the results for the broadband condition of experiment 2 were very similar to the corresponding results for the VS distractor in experiment 1 for both the HI and NH group, with mean differences being within about 1 dB of each other. The similarity was confirmed by a paired t-test ($p > 0.1$). Given that in both experiments the usable bandwidth was rather large and differences mainly occurred at high frequencies (i.e., above 2.000 Hz for HI and above 5.300 Hz in NH subjects) where at least for the HI subjects audibility will have limited access to the provided speech information, the similarity may not be surprising and rather confirm the reliability of the applied methods.

2.4.4 Concluding discussions

In section 2.3.1.3 (table 2.1), it was shown that the intra-subject test-retest standard deviation for most of the conditions applied in experiment 1 was in the same range (around 1-1.5 dB) as the one reported by Keidser *et al*. (2013) or Cameron *et al.* (2011). Since these studies applied the same sentence material and test procedures as well as distractor signals with very similar temporal, spectral and spatial energy fluctuations, this observation may not be surprising. However, it confirms that the psychoacoustic properties of the applied procedures and stimuli were not significantly affected by the modifications applied in experiment 1, such as maximizing (and extending) ILDs, removing ITDs, frequency equalization, and noise-vocoding of speech distractors. In experiment 2 (section 2.3.2.2), the test-retest standard deviation already increased in the broadband conditions to about 2 dB in NH and 2-2.5 dB in HI subjects, and further increased in the narrowband conditions to up to 2.6 dB in NH subjects and up to 4 dB in HI subjects (see table 2.2). Besides potential fatigue effects, the increased intra-subject standard deviation may be mainly explained by the increased RMS level variation across sentences, which was introduced by the applied sensation level equalization (see applied gain shown in figure 2.2) as well as the bandpass filtering into the *Low*, *Mid*, and *High* frequency regions. In the original BKB speech corpus all sentences were normalized to the same (broadband) RMS level, which was maintained in experiment 1. In experiment 2, the RMS level for NH subjects in the broadband condition varied over a range of about ±5 dB, which was increased to about ±8 dB in the *Mid* and *High* frequency region (in the *Low* frequency region it was ±2 dB). For the HI subjects this RMS level variation was even larger and increased with increasing hearing loss. The variance in speech intelligibility in the narrowband conditions may have been further increased by the fact that the frequencies that mainly contribute to intelligibility varies from word to word (or even phoneme to phoneme), which may not matter when the entire speech spectrum is available but increases the variance in word (and thus sentence) recognition when only a narrow frequency channel is considered.

The large level variations had also the side-effect at high frequencies that due to the substantial hearing loss of the HI subjects (see audiograms shown in figure 2.1) it was very difficult (or even impossible) to provide sufficient target audibility while guaranteeing comfortable loudness levels. This was also the reason for why the *High* frequency condition was excluded from this study for HI subjects. Hence, future studies should look into better methods to individually control the sensation level and thereby allow conclusions on the (maximal) spatial benefit that can be achieved by BEG in HI subjects at high frequencies. In this regard, amplitude compression (as provided by hearing aids) would simplify the control of the applied sensation levels, but the distortions that are potentially introduced to the ILD cues (e.g. Byrne and Noble, 1998; Moore, 2008) could interfere with the BEG process and thereby result in an underestimation of the real (achievable) spatial benefit. The general effect of amplitude compression on BEG, however, is an interesting research topic for fitting bilateral hearing aids and should be addressed in future studies.

It should be highlighted that in experiment 2 the sensation level of the distractor was much higher for the NH subjects (i.e., 35 dBSL) than for the HI subjects (i.e., 10 dBSL). As already discussed, the higher sensation level may at least partly explain the better SRTs in the NH than HI subjects, in particular, in the spatially separated conditions, and thus, may have also contributed to the increased SRM. Hence, this difference in sensation level does not allow a direct comparison of the spatial advantage achieved by BEG between groups. However, keeping the sensation level constant across frequency allowed a direct comparison of the effectiveness of BEG across frequency, which was the main purpose of experiment 2. Even though at low frequencies the sensation level could have been increased in the HI subjects, this was not the case at higher frequencies due to loudness discomfort. Similarly, the sensation level in the NH group could have been reduced to the same level as for the HI subjects (as for instance done in Glyde *et al.*, 2015), but here the main idea was to measure the maximal possible spatial benefit that can be achieved by BEG as a function of frequency at comfortable loudness levels. This goal would have been jeopardized by such low sensation level due to audibility problems. Future studies should therefore aim at comparing BEG performance (as a function of frequency) between NH and HI subjects at equal sensation levels. However, this will require careful level control to avoid loudness discomfort, in particular at high frequencies.

It is interesting to note that the mean difference in SRT between the test and retest measurements (table 2.1 and 2.2) suggests at least in some conditions a noticeable training effect, which was stronger in the HI group (up to about 1-2 dB) than in the NH group (up to about 1 dB). In this regard, the co-located condition in experiment 1 with the SD distractor was particularly interesting, which in the NH group showed a training effect, though non-significant, of about 3 dB. This noticeable behavior is also reflected in a highly increased intra-subject test-retest standard deviation of 3.2 dB (instead of 1-1.5 dB as observed in all other broadband conditions).

Since this condition is highly influenced by IM, it may suggest that subjects are not very experienced with listening to stimuli that contain a high amount of IM and thus, need to learn to process such stimuli. It is unclear why this specific training effect is not observed in HI subjects, but maybe they are not sensitive enough to the subtle cues that are utilized by trained NH subjects to (partially) resolve IM.

The potential effect of age on SRM has been widely discussed in the literature, with some studies showing a significant age effect (e.g., Gallun *et al.*, 2013; Murphy *et al.*, 2006) and others not (e.g., Glyde *et al.*, 2013a). In the present study a correlation analysis was applied between the subjects' individual SRM results and their age which showed neither a significant effect in experiment 1 (section 2.3.1.2) nor in experiment 2 (section 2.3.2.2). Even though this may be partly explained by the rather small number of subjects and limited test-retest reliability, this analysis still suggests that SRM, at least as measured here, is not substantially affected by age. This is a promising result since it suggests that the older age of most HI subjects will not limit the potential spatial benefit provided by ILD enhancement methods in hearing aids.

Finally, it should be highlighted that both NH and HI listeners were able to successfully utilize ILD cues for BEG at low and mid frequencies, even though in real life these ILD cues are rarely available (see section 2.1). Hence, it is expected that if adequate low and mid frequency ILDs can be (artificially) provided, then HI listeners can utilize them to improve speech intelligibility in spatial noise. As already mentioned in section 2.1, one way to provide such ILDs could be by directional hearing aid microphones, which can be created by combining the output of multiple microphones placed around the ears (or head) of the listener (e.g., Kates, 2008, pp.75-109). However, neither the optimal directivity for such directional microphones is known (as a function of frequency) nor the number and placement of microphones that is required to create it. Independent of that, directional microphones will typically increase the internal noise level in the hearing aid and amplify wind noise (Kates, 2008, pp.75-109). Both can cause significant problems in hearing aids, in particular at low frequencies. Moreover, a benefit provided by any signal enhancement method can only be utilized by a HI listener if the enhanced signal is audible and dominates the signal to the listeners' ears. Especially for open fittings, hearing aids cannot provide sufficient amplification at low frequencies (i.e., below about 1000 Hz) and as a consequence the acoustic signal that is circumventing the hearing aid is dominant (e.g., Dillon, 2012, pp.127-169). Applying a closed fitting would improve the output level that can be provided by the hearing aid at low frequencies (and attenuate the acoustic path), but at the same time cause other problems such as occlusion (Dillon, 2012, pp.127-169). Even though occlusion can be reduced (Mejia *et al*., 2008), all the mentioned constraints need to be considered when developing a method that provides enhanced ILDs in hearing aids, and most of these aspects are already taken into account in the design of modern hearing aids.

2.5 Conclusions

It was found that both NH and HI subjects can successfully utilize BEG at low frequencies to enhance speech intelligibility in spatial noise. In experiment 1 it was shown that in a "common" symmetric two-speech-distractor scenario the SRM in NH listeners is increased by about 6.7 dB when maximized (low frequency extended) ILDs are applied instead of natural ILDs. For the considered HI group with a moderate degree of hearing loss and linear amplification according to NAL-NL2 this additional increase in SRM was about 3.8 dB. In experiment 2, this spatial advantage was further investigated as a function of frequency region and bandwidth. In NH listeners the achieved spatial benefit measured within six critical band wide frequency channels was around 9 dB and independent of frequency. For broadband stimuli this advantage increased to about 11 dB, suggesting that the spatial advantage provided by BEG is integrated across frequency. For HI listeners the spatial benefit could only be measured at low (100-770 Hz) and mid (770-2000 Hz) frequencies and, compared to NH listeners, was slightly reduced to about 7 dB. At high frequencies (2000-5300 Hz) the available dynamic range provided by the considered hearing losses did not allow reliable measurements of SRTs (and SRM) without exceeding uncomfortable loudness levels during the adaptive testing. In particular, in the broadband condition, audibility (due to the low distractor sensation levels) limited the SRT in the spatially separated condition and thus the observed SRM. Additional aspects of hearing loss (e.g., reduced temporal and spectral resolution) as well as reduced cognitive performance due to age differences between groups may have had also an impact on the results, but could not be further evaluated.

Future research should systematically study the effect of sensation level (or audibility) on BEG in HI listeners and compare results to corresponding NH data. In particular, at high frequencies, this will require improved methods for controlling the target speech level individually to avoid loudness discomfort. Moreover, the effect of amplitude compression in (bilateral) hearing aids on BEG needs to be studied, methods need to be developed that can generate ILDs that are optimized across the entire frequency range, and the benefit on speech intelligibility needs to be investigated in more realistic conditions. Thereby, besides speech intelligibility, other aspects need to be considered, such as spatial perception or the acceptance by the listener.

2.6 Acknowledgements

The authors acknowledge the financial support of Macquarie University through an international research excellence (iMQRES) scholarship and the Australian National Health and Medical Research Council (NHMRC) via grant APP1056332.

Chapter 3: Effect of audibility on better-ear glimpsing as a function of frequency in normal-hearing and hearing-impaired listeners

Better-ear glimpsing (BEG) is an auditory phenomenon that helps understanding of speech in noise by utilizing interaural level differences (ILDs). The benefit provided by BEG is limited in hearing-impaired (HI) listeners by reduced audibility at high frequencies. Rana and Buchholz [JASA (2016), 140(2), 1192-1205] have shown that artificially enhancing ILDs at low and mid frequencies can help HI listeners understanding speech in noise, but the achieved benefit is smaller than in normal-hearing (NH) listeners. To understand how far this difference is explained by differences in audibility, audibility was carefully controlled here in ten NH and ten HI listeners and Speech Reception Thresholds (SRTs) in noise were measured in a spatially separated and co-located conditions as a function of frequency and sensation level. Maskers were realized by noise-vocoded speech and signals were spatialized using artificially generated broadband ILDs. The spatial benefit provided by BEG and SRTs improved consistently with increasing sensation level. Moreover, no significant differences were found between groups, indicating that HI listeners can achieve similar performance as NH listeners when differences in audibility are compensated. The results help to understand the hearing aid gain that is required to maximize the spatial benefit provided by ILDs as a function of frequency.

Note: Aspects of this work were presented at the World congress of Audiology (Vancouver, Canada, 2016) and at the 5th Joint Meeting, Acoustical Society of America and Acoustical Society of Japan (Honolulu, Hawaii, 2016)

3.1 Introduction

The ease with which a listener can pick up information from target speech and simultaneously suppress a background noise when they arrive from different locations versus when they arrive from the same location is attributed to a phenomenon known as spatial release from masking (SRM) or spatial advantage. In general, SRM depends on various factors, including the number, type and location of the maskers (e.g., Peissig and Kollmeier, 1997; Brungart *et al.*, 2001; Arbogast *et al*., 2005; Marrone *et al*., 2008; Best *et al*., 2013), and is pronounced in conditions where informational masking is involved (e.g., Kidd *et al*., 2007; Glyde *et al*., 2013b). Whereas energetic masking is associated with the spectral and temporal overlap of the target and masker signals within the auditory periphery, informational masking refers to more central auditory mechanisms that are associated with auditory scene analysis and attention (e.g., Shinn-Cunningham, 2008). The present study focuses on energetic aspects of SRM and applies methods similar to Rana and Buchholz (2016) to minimize the influence of informational masking.

There are a number of "energetic" auditory mechanisms that contribute to SRM. In the case of a single masker, spatially separating the masker from the target leads to a consistent improvement in the signal-to-noise ratio (SNR) in one of the ears due to acoustic head shadow. This is commonly referred to as the "better-ear effect" and, for single maskers, is often the dominant contributor to SRM. However, the better-ear effect diminishes with increasing number of spatially separated maskers, and is absent in spatially symmetric masker conditions (e.g., Hawley *et al*., 2004; Misurelli and Litovsky, 2012). In the case that the (multiple) maskers fluctuate over time, such as speech, the SNR will fluctuate between ears and thereby provide a better-ear that rapidly switches between ears. The auditory system can take advantage of the SNR fluctuations at the two ears by utilizing a process termed better-ear glimpsing (e.g., Brungart *et al*., 2012; Glyde *et al*., 2015). Besides these head-shadow or interaural level difference (ILD) based mechanisms, the auditory system can also take advantage of interaural time differences (ITDs) by applying a mechanism similar to the equalization-cancellation process (e.g., Durlach, 1963; Breebaart *et al*., 2001).

Even though all the above auditory mechanisms will help understanding speech in the real-world, their individual contribution and relevance is not known. Glyde *et al*. (2013c) provided evidence that ILD cues, due to better-ear glimpsing (BEG), contribute more to SRM than ITD-based processing. However, Culling *et al*. (2004), as well as Kidd et al. (2010), observed the opposite. Most likely these differences can be explained by methodical differences, in particular by the chosen target and masker signals and their spatial configurations. However, since people experience a great variety of acoustic environments in their daily lives, any of these stimulus conditions may be equally relevant or irrelevant, respectively.

Most studies agree that the SRM observed in HI listeners is significantly smaller than in NH subjects and, as a consequence, makes it harder for HI subjects to understand speech in noisy conditions (e.g., Gelfand *et al*., 1988; Dubno *et al*., 2002; Arbogast *et al*., 2002; Marrone *et al.*, 2008). Glyde *et al.* (2013a) tested a large cohort of subjects with varying degrees of hearing loss and found that, as the degree of hearing loss increases, the SRM decreases. This was the case even though linear amplification according to the National Acoustic Laboratories – Revised Profound (NAL-RP: Dillon, 2012, pp 290-297) was applied to (partially) compensate the hearing loss. In a follow up study, Glyde *et al*. (2015) showed that the SRM in HI listeners improves when increased (linear) amplification is provided on top of NAL-RP. They also compared their HI data with NH data measured at equal sensation levels and found that the SRM difference between groups is substantially reduced when compared to the SRM measured at original (i.e., different) stimulus levels. Overall they concluded that the reduction in SRM seen in many HI subjects is largely due to reduced audibility. However, they also highlighted that additional factors may have been involved, including increased masking due to wider auditory filters, reduced temporal resolution, age, and reduced cognitive performance (see also: Dubno *et al*., 2002; Gelfand *et al*., 1988; Gallun *et al*., 2013).

Hearing aids mainly address the loss of audibility, but they cannot address any of the other factors that are likely to contribute to the reduced spatial advantage seen in HI listeners. Since better-ear processing, including BEG, relies on the head shadow effect, it is mainly observed at frequencies above about 1.5 kHz (e.g., Rayleigh, 1876; Blauert, 1997, pp. 36-200; Macpherson and Middlebrooks, 2002). Unfortunately, this is the frequency range in which most HI listeners show the strongest hearing loss (e.g., Dillon, 2012, pp. 286-335). Given also that speech has a low-pass characteristic (i.e., the long-term average speech spectrum rolls off above about 1 kHz), better-ear processing is particularly vulnerable to reduced audibility. The detrimental effect of audibility on BEG has already been highlighted by a number of studies (e.g., Best *et al*., 2015; Glyde *et al*., 2015), and Glyde *et al*. (2015) have shown that it can be (partly) restored in HI listeners by increasing standard (linear) amplification. In contrast, the spatial advantage provided by ITDs is less likely to be limited by audibility. This is because the required temporal fine-structure cues mainly occur at frequencies below about 1.5 kHz (e.g., Sandel *et al*., 1955; Wightman *et al*., 1992), i.e., at frequencies where most hearing losses are less severe and speech contains its main energy. Hence, hearing aids may be able to improve the benefit provided by BEG, but they are less likely to address any limitation that is caused by impaired ITD processing. Even though hearing aids can help restore audibility, they also have to ensure that incoming sounds do not become uncomfortably loud.

To deal with dynamic range limitation, the overall gain provided by hearing aids is limited and wide dynamic range compression (WDRC: Kates, 2008, 221-259) is applied. However, applying WDRC can also alter the temporal and spectral behaviour of the incoming sound signals, which may counteract the spatial advantage that may be provided by the provided audibility increase. To avoid this potential impact of WDRC when investigating the effect of hearing loss on SRM, many studies have applied linear, often frequency-dependent, amplification (e.g., Glyde *et al*., 2015; Jakien *et al*., 2017). To avoid loudness discomfort they then chose a "suitable" stimulus sensation level (e.g., 20 or 30 dB-SL) and, if required, decreased the level for individual subjects (e.g., Glyde *et al*., 2015; Jakien *et al*., 2017). Following such approach, Glyde *et al*. (2015) have shown that standard linear amplification according to NAL-RP does not provide sufficient audibility at mid and high frequencies to fully restore SRM in HI listeners, which in their case was mainly provided by BEG. Although they demonstrated that adding additional gain on top of NAL-RP improved SRM, they also highlighted that, in the real world, the loudness provided by such increased gain would not be accepted by HI listeners and furthermore, may be limited by technical issues such as acoustic feedback, occlusion, and saturation effects. Even though the loudness discomfort may be alleviated by appropriate WDRC, it is unclear how far the resulting changes to the temporal and spectral behaviour of the incoming sounds will affect BEG.

As an alternative solution, Rana and Buchholz (2016) proposed to shift the ILD cues that are underlying BEG to lower frequencies, i.e. to frequencies at which audibility is far less of an issue. They showed that both NH and HI subjects can take advantage of such (artificially generated) low-frequency ILDs, exhibiting a SRM that was very similar to the one observed at higher frequencies (when similar audibility was provided across frequency). However, the average SRM observed in the HI group was still smaller than in the NH group, which may have been due to the different sensation levels that were applied to avoid loudness discomfort in the

HI group at high frequencies: whereas the HI listeners were tested at 10 dBSL (relative to their individual pure-tone thresholds) the NH listeners were tested at 35 dBSL. The present study builds upon the methods and findings described by Rana and Buchholz (2016) to address the following goals: (i) to investigate whether SRM due to BEG can be restored in HI listeners if audibility equivalent to NH listeners is provided, (ii) to better understand the sensation level (or hearing aid gain) that is required to maximize the spatial benefit provided by ILDs across frequency, and (iii) to provide a conclusive data-set that guides future developments in hearing aid technology as well as gain prescription methods that improve speech intelligibility in noise by maximizing the SRM provided by BEG.

3.2 Methods

The present study investigated the effect of audibility on BEG in NH and HI subjects as a function of frequency and sensation level. Speech reception thresholds (SRTs) were measured in both a co-located and spatially separated condition as well as in different frequency regions using a noise-vocoded two-talker masker. The difference in SRTs measured in the co-located and spatially separated conditions was considered an estimate of the spatial benefit provided by BEG. Audibility of the target and masker stimuli was equalized across both frequency and ears for each subject using individual speech detection thresholds (SDTs) measured in nine narrow frequency bands. Four different sensation levels were tested as defined relative to the individual's SRT in quiet. Loudness comfort was ensured for each subject by taking into account their individual upper limit of comfortable loudness for the applied stimuli. Ethical clearance was received from the Australian Hearing Human Research Ethics Committee and the Macquarie University Human Research Ethics Committees.

3.2.1 Subjects

Ten NH (hearing thresholds < 15 dBHL) listeners with a mean age of 23.2 ± 3.2 years and ten sensorineural HI listeners with a mean age of 70.3 ± 7.8 years participated in this study. All subjects received a hearing test in the beginning of their first appointment to either confirm normal hearing or to determine their degree and type of hearing loss. All HI listeners had symmetric (threshold difference between ears < 10 dB), mild to moderate (Clark, 1981), sloping hearing loss. Their four-frequency (0.5, 1, 2, 4 kHz) average hearing loss (4FAHL) was 29.1 ±

8.0 dBHL, and the mean hearing thresholds converted into dBSPL and averaged over the left and right ear are shown in figure 3.1. All participants had English as their first language and had no reported attention deficit disorder or intellectual disability.

3.2.2 Stimuli

Similar to Rana and Buchholz (2016), speech reception thresholds were measured using BKB-like sentences (Bench *et al*., 1979) spoken by a native Australian English female speaker. The entire corpus consisted of 80 lists of 16 sentences each. The RMS level of all sentences was normalized and the average spectrum of each sentence list was equalized to match the average spectrum of the entire corpus. The distractor signals were created by noise-vocoding the two continuous female speech discourses taken from the different-voice condition of the LiSN-S speech corpus (Cameron and Dillon, 2007). Each speech discourse was noise-vocoded individually using the same methods as described in chapter 2 and equalized to match the average spectrum of the target sentences. The derived distractors minimized the potential influence of informational masking on the measured SRTs while preserving most of the temporal-spectral properties of speech (e.g., Westermann and Buchholz, 2015), and when presented together they were largely unintelligible.

The target sentences were always presented diotically and the distractors were either presented diotically (i.e., both distractors co-located with the target) or spatially separated. The spatial separation was realized by presenting one distractor only to the left ear and the other distractor only to the right ear, i.e., by applying infinite ILDs. The increase in the combined distractor level relative to the target level in the co-located condition was compensated by attenuating each distractor by 3 dB.

To evaluate BEG as a function of frequency region, all target as well as masker stimuli were bandpass (BP) filtered into a low, mid, high, and broad frequency region with details described in table 3.1. The BP-filters were realized by 4-th order Butterworth filters and derived in Matlab. All stimuli were presented via equalized Sennheiser HD215 circumaural headphones connected to a RME fireface UC USB sound card. The headphone equalization was realized by a 3000-taps long minimum-phase FIR filter at a sampling frequency of f_s = 44.100 Hz, which was derived in Matlab and using a Bruel & Kjaer artificial ear.

3.2.3 Audibility equalization

In order to equate audibility across frequency of both the target and distractor stimuli used throughout the SRT measurements, equalization filters were derived for each subject individually using the following procedure: speech shaped noise (SSN) was first generated with a spectrum that matched the average spectrum of the BKB-like sentences and then BP-filtered into nine different frequency regions. The cut-off frequencies of these BP filters were taken from table 1 of the SII standard (ANSI S3.5-1997) and realized by 4^{th}-order Butterworth filters. Each filter encompassed two critical bands with the lower and upper cut-off frequencies (f_1 and f_2) as well as the centre frequencies (f_0) summarized in table 3.1.

TABLE 3.1: Frequency channels used in the audibility equalization process as well as the frequency regions used in the SRT measurements. f_0: centre frequency; f_1 *and* f_2: lower and upper cut-off frequency.

channel	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
f_1 (Hz)	100	300	510	770	1080	1480	2000	2700	3700
f_2 (Hz)	300	510	770	1080	1480	2000	2700	3700	5300
f_0 (Hz)	200	400	630	920	1270	1720	2320	3150	4400
		low			mid			high	
					broad				

Afterwards, speech detection thresholds (SDTs) were measured in quiet for each of the nine BP-filtered SSNs using an adaptive three-alternative forced-choice method in which one randomly chosen interval contained a BP-filtered SSN and the other two intervals contained silence. In order to establish the starting level for this adaptive procedure, individual audiometric pure tone thresholds measured (in dBHL) for each ear (see section 3.2.1) were converted into ear drum levels (in dBSPL) by adding Reference Equivalent Sound Pressure Levels for THD 39 headphones (ISO 389-2 1994) and nominal values for the transformation from 6 cc coupler to ear drum levels (Bentler and Pavlovic, 1989). Since the resulting thresholds are given at standard audiometric frequencies (i.e., 250 Hz, 500 Hz, 1 kHz, 2 kHz, 3 kHz, 4 kHz, 6 kHz and 8 kHz), they were mapped onto the centre frequencies (f_0) given in table 3.1 using a linear interpolation on a double-logarithmic scale. Mean values and 95 % confidence intervals of the pure tone thresholds averaged across the left and right ear are shown in figure 3.1 for the NH and HI group.

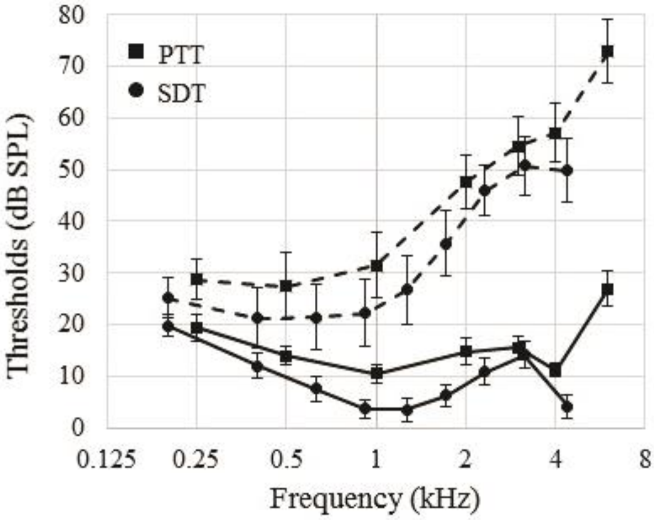

FIG. 3.1: Mean pure tone thresholds (PTT) and mean speech detection thresholds (SDT) with 95 % confidence intervals for NH and HI listeners. Dashed lines indicate thresholds of HI listeners and solid lines indicate thresholds of NH listeners.

The starting level in the adaptive procedure was calculated such that the output level of the BP-filtered SSN in the considered frequency channel was 6 dB above the corresponding interpolated individual pure-tone threshold averaged across ears. Any difference in the pure-tone thresholds between ears was compensated by reducing the level of the BP-filtered SSN at the one ear and increasing the level at the other ear by half the difference. The listener's task was to select the interval that contained the BP-filtered SSN, whereby the current interval was visually highlighted by coloring the corresponding response button. The threshold was measured adaptively using a 1-up 2-down method with a starting step-size of 6 dB which after 2 reversals was decreased to the final step-size of 3 dB at which six reversals were measured.

The SDT was then determined as the average level calculated over the last six turn-points. Since the applied SSN had the same long-term spectrum as well as RMS level as the target and distractor stimuli used in the SRT measurements (see above), the gain that was applied to the BP-filtered SSN at SDT provided a direct estimate of the gain that should be applied at the corresponding centre frequency to equate audibility of the stimuli used in the SRT measurements. Based on these gains, minimum phase FIR equalization filters were derived in

Matlab with a length of 1024 taps at a sampling frequency of f_s = 44.100 Hz. Within the filter derivation process the gains were extrapolated to 0 Hz as well as $f_s/2$ by setting the corresponding gains to the values of the gains at the lowest and highest BP center frequency (i.e., 200 Hz and 4400 Hz; table 3.1). It should be noted that the equalization filters were derived separately for the left and right ear to compensate for any audiometric pure-tone threshold differences between ears as described above.

The mean values (in dB-SPL) and 95 % confidence intervals of the measured SDTs are shown in figure 3.1 for both the NH and HI group. Paired t-test revealed no significant difference (p>0.01) between right and left ear SDTs for any frequency for both the groups. Hence, an average of right and left ears are plotted in figures (3.1 and 3.2) but the gain considered for target and distractor stimuli was still ear specific. From figure 3.1 it can also be deduced that, on average, the SDTs for the NH and HI listeners are about 7 dB and 12 dB below the corresponding (frequency-mapped) audiometric pure-tone thresholds. However, individual differences varied substantially across subjects and frequency as shown by the error bars which indicate 95 % confidence intervals.

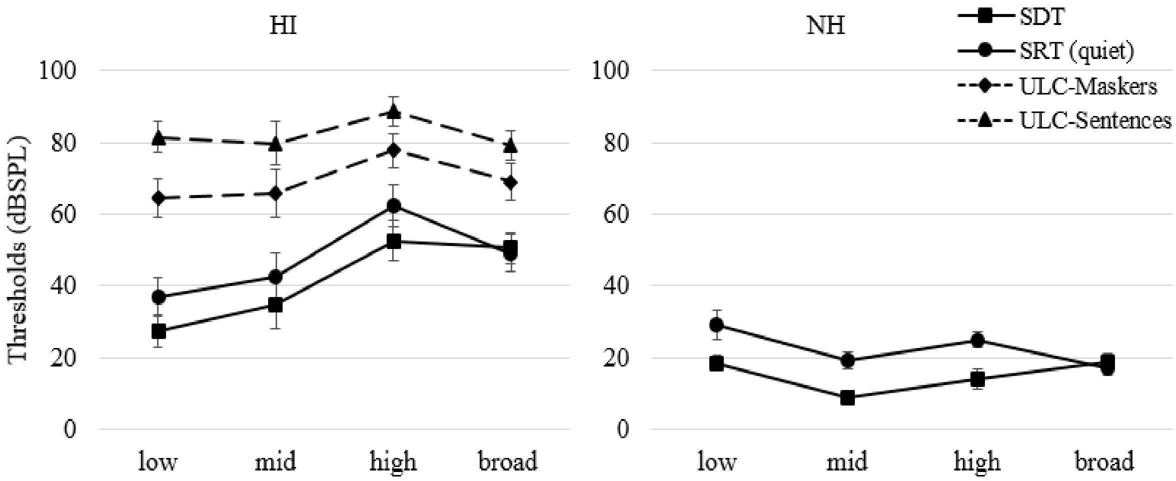

FIG. 3.2: Mean speech detection thresholds (SDTs) and mean speech recognition thresholds (SRTs) in quiet with 95 % confidence intervals for NH and HI listeners. Mean upper limit of comfort (ULC) with 95 % confidence intervals of maskers and sentences for HI listeners.

3.2.4 Upper limit of comfortable level

To avoid presenting stimuli at uncomfortably loud levels during the SRT measurements, the level of the upper limit of comfort (ULC) was measured for the audibility equalized (see section 3.2.3) target and distractor stimuli separately using an 8-point rating scale (1-very soft; 2-soft; 3-comfortable, but slightly soft; 4-comfortable; 5-comfortable, but slightly loud; 6-loud, but ok; 7-slightly uncomfortable; 8-uncomfortably loud). For target sentences, the ULC was defined as the minimum level above which the subjects' response shifted from "loud, but ok" to "slightly uncomfortable". For distractors, the ULC was defined as the minimum level at which the subjects first responded "comfortable, but slightly loud" and then they were asked if they could happily tolerate that level for at least 5 minutes. The ULC for the distractors was measured in both the co-located and the spatially separated condition. The ULCs were defined differently for the target speech and the distractors because within the SRT measurements the distractors were presented continuously for around 5 minutes per condition whereas each target sentence was presented only for a few seconds and the level typically decreased quickly throughout an adaptive track. The ULCs were obtained separately for all four frequency regions (i.e., low, mid, high and broad) in which the SRTs were measured (see table 3.1 and section 3.2.5).

To determine a starting level for the ULC measurements, SDTs were first measured separately for the low, mid, high, and broad frequency region using the same methods as described in section 3.2.3 except that here the applied SSN was audibility equalized. The resulting mean SDTs (in dB-SPL) with 95 % confidence intervals are shown in figure 3.2 for the NH as well as HI group. ULC measurements were done using an ascending method wherein a test run began at 5 dB above the corresponding SDT. The level of each successive stimulus was raised step-by-step, with the participant furnishing a loudness category judgment for each stimulus presentation from the above mentioned 8-point rating scale. In the beginning, a larger stimulus increment size of 8 and 4 dB was used which was reduced to 2 dB and 1 dB as the participant progressed towards the louder levels. Once a ULC level was obtained, the stimulus level was reduced to a comfortable level to start another run. The final ULC levels were calculated as the median value over three runs. The procedure was similar to the one described by Cox *et al.* (1997).

Mean ULCs measured for the audibility equated stimuli are shown in figure 3.2. Since ULCs measured for distractors in co-located and spatially separated condition were statistically similar ($p>0.01$), averaged results are shown in figure 3.2. Please note that results obtained were same even when alpha of 0.05 was considered. ULCs measured for target sentences were higher than the ULCs measured for the distractors due to the difference in the set criteria to define ULCs. Further, ULCs were measured for HI subjects only, because loudness tolerance was not an issue for the NH subjects within the SRT measurements.

3.2.5 Speech reception thresholds

As the goal of this study was to measure the SRM attributed to BEG as a function of both frequency and sensation level, first, the audibility-equalized target and distractor signals (see section 3.2.3) were BP-filtered into four different frequency regions: low, mid, high, and broad (see section 3.2.2 and table 3.1 for details). Afterwards, SRTs for the audibility-equalized target sentences were measured in quiet for the four different frequency regions. The SRTs were measured with a Matlab program installed on a personal computer and the subjects' task was to repeat as many words as they heard in each target sentence. In each condition, up to 32 sentences were presented and the target level was adjusted adaptively to achieve 50% correct word identification. The number of correctly identified morphemes were entered into a provided user interface. Details can be found in Keidser *et al*. (2013). All SRTs were measured twice and the results were averaged. Mean SRTs measured in quiet with 95 % confidence intervals for the audibility-equalized BKB-like sentences are shown in figure 3.2 for NH and HI subjects as a function of frequency region. As expected, the difference between the SRTs in quiet and ULCs, i.e., the available dynamic range, was reduced at high frequencies as well as in the broad condition when compared to the low and mid frequency regions. The smaller dynamic range limited the number of HI listeners that could be tested in all conditions (figure 3.4). Further, the SDTs shown in figure 3.2 were consistently better than the SRTs in quiet across all narrow band conditions, but this difference disappeared in the broad condition. This may be attributed to the different stimuli and tasks used to measure these thresholds: SDTs were measured with SSN and SRTs with speech. The difference between the SRT in quiet and the corresponding SDT is very similar between groups for all four frequency regions, with a mean difference (NH minus HI value) of: Low: 1.4 dB; Mid: 2.5 dB; High: 1.2 dB; Broad: -0.3 dB.
Once the individual SRTs in quiet were derived, SRTs were measured at different distractor levels using the same methods as described above. The SNR was adjusted adaptively by varying the target level and keeping the distractor level constant. SRTs were measured with the audibility-equalized distractors presented at 0, 10, 20, and 30 dB above the SRTs measured in

quiet. At each distractor level, SRTs were measured in all four frequency regions as well as in the spatially co-located and separated condition (see section 3.2.2). The difference in SRT between the two spatial conditions, at each sensation level and for each frequency region, provided an estimate of the spatial benefit provided by BEG. The resulting SRTs as well as the derived spatial benefit are presented in section 3.3. Throughout the SRT measurements the investigators made sure that the overall intensity of the presented stimuli did not exceed uncomfortable loudness levels. In the case that a distractor level was above the subject's individual ULC level (see section 3.2.4) the condition was dropped. Similarly, if the target level within an adaptive track exceeded the corresponding ULC level, then the target level was set to that ULC level. If this occurred multiple times then the track was stopped and the condition dropped.

Finally, it should be noted that the long-term spectrum (evaluated in 2 CB-wide bands) of the audibility-equalized distractors at a sensation level of 0 dB-SL was equal to the narrow-band SDTs shown in figure 3.1 in dBSPL for NH and HI subjects.

3.3 Results

Statistical analysis was done using IBM SPSS statistics version 22.

3.3.1 Speech reception thresholds

Individual and mean SRTs for the HI subjects as well as the mean SRTs for the NH subjects are shown in figure 3.3 as a function of sensation level separately for the four different frequency regions as well as the two spatial configurations. A linear mixed-effects model with frequency region, spatial separation, hearing status, sensation level and their two-and three-way interactions as fixed effects and a subject-specific intercept as the random effect showed significant effects of frequency region [$F(3,470) = 129.90$, $p<0.01$)], spatial separation [$F(1,470) = 221.88$, $p<0.01$)], and sensation level [$F(3,470) = 352.75$, $p<0.01$)], but no significant effect of hearing status [$F(1,470) = 1.31$, $p>0.01$)]. A significant interaction was observed only for hearing status and frequency region [$F(3,470) = 7.35$, $p<0.01$)], frequency region and sensation level [$F(9,470) = 11.69$, $p<0.01$)], and sensation level and spatial separation [$F(3,470) = 19.37$, $p<0.01$)].

As one main goal of this study was to examine the effect of audibility on SRTs, multiple paired t-tests with adjusted p-values (Holm, 1979) were conducted at each frequency region and spatial configuration to compare the SRTs at the different sensation levels. Results revealed that both co-located and spatially separated SRTs improved with increasing sensation level within all frequency regions for both groups, but the effect was generally more evident for spatially separated SRTs. Moreover, the significant improvement in the co-located SRTs with increasing sensation level was generally more evident for NH than HI listeners (figure 3.3). An independent t-test with adjusted p-values (Holm, 1979) revealed no significant difference in SRTs between NH and HI listeners in any condition, indicating that, for sensation levels up to 30 dB, once audibility is taken into account HI subjects perform very similar to NH subjects.

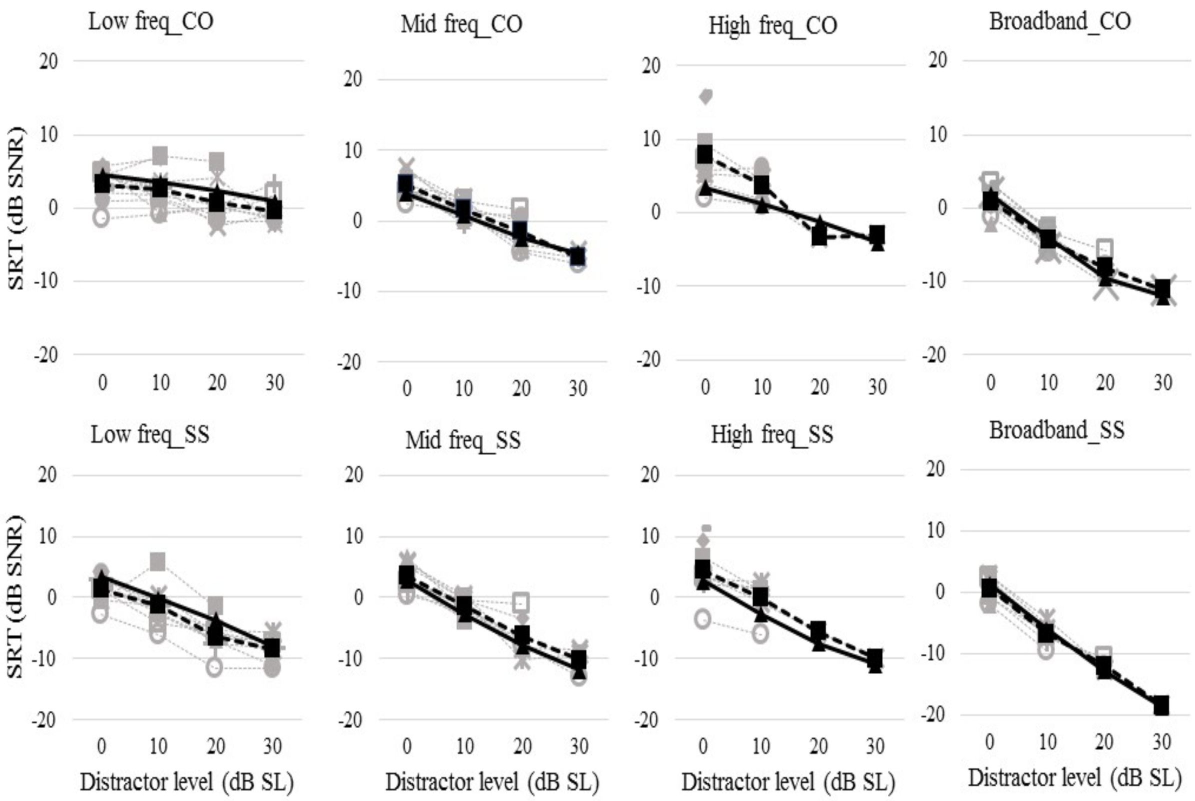

FIG. 3.3: Individual and mean SRTs as a function of sensation level. Black triangles connected with solid lines indicate mean SRTs for listeners with NH. Black squares connected with dashed lines indicate mean SRTs for HI listeners. Symbols in grey represent individual SRTs for HI listeners. CO: Co-located, SS: spatially separated.

3.3.2 Spatial release from masking

Spatial release from masking was calculated by subtracting the individual SRTs measured in the spatially separated condition from the corresponding SRTs measured in the co-located condition. Figure 3.4 shows the mean SRM as a function of sensation level for the NH and HI group separately for the different frequency regions. For the HI group the individual data is additionally shown. A linear mixed-effects model with frequency region, hearing status, sensation level and their two- and three-way interactions as fixed effects, and a subject-specific intercept as the random effect showed significant effects of frequency region [$F(3,234) = 7.15$, $p<0.01$)] and sensation level [$F(3, 234) = 61.40$, $p<0.01$)], but no significant effect of hearing status [$F(1,234) = 0.12$, $p>0.01$)]. Further, no significant interaction was observed for any combination.

To further examine the significance of the observed SRM as a function of frequency and sensation level, a paired t-test with adjusted p-values (Holm, 1979) was conducted at each sensation level and frequency region. Results revealed significant differences between co-located and spatially separated SRTs at 10, 20 and 30 dBSL for both groups at most frequency regions. Even though a clear and consistent improvement in average SRM with increasing sensation level can be seen in figure 3.4 for both groups within most frequency regions, an independent t-test revealed that for the NH group, this increase was only significant when the sensation level was increased by at least 20 dB. No significant effect was observed for the HI group. The missing significance may be due to the reduced number of HI participants with increasing sensation level, which was due to loudness discomfort and is indicated in figure 3.4 by the numbers in brackets. An independent t-test revealed no significant difference in SRM between NH and HI listeners at any sensation level for any frequency region. At equal sensation level, the mean difference and +/- 1 STD in SRTs between groups and across all frequency regions and sensation levels was equal to 0.11 ± 1.45 dB.

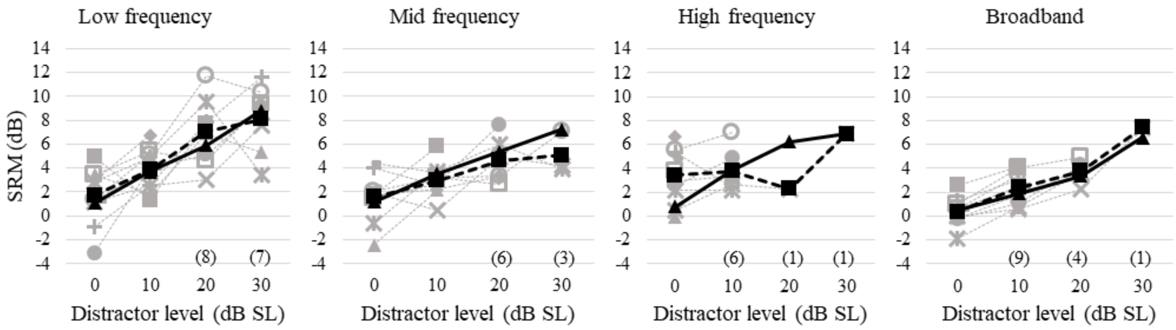

FIG. 3.4: SRM as a function of distractor level for all four frequency regions. Grey lines represent the individual SRM for the HI listeners with the mean SRM indicated by black dashed lines and squares. Black triangles connected with solid lines indicate the mean SRM for NH listeners. When the number of participants for the HI group was less than ten, then the number is shown in parentheses. All ten NH listeners were tested in all conditions.

3.3.3 Test-retest variability

Test-retest variability was calculated by subtracting the individual SRTs in the second trial from the corresponding SRTs in the first trial. A paired t-test revealed no statistically significant difference ($p > 0.01$) between the first and second trials in all conditions for both groups. For NH listeners, the mean difference ranged from -1.74 to 1.28 dB in the co-located and from -0.73 to 1.30 dB in the spatially separated conditions. The intra-subject standard deviation was between 0.71 to 3.15 dB in the co-located and between 1.05 to 3.39 dB in the spatially separated conditions. For HI listeners, the mean difference ranged from -1.93 to 2.36 dB in the co-located and from -2.32 to 1.78 dB in the spatially separated conditions. The intra-subject standard deviation was between 0.23 to 3.14 dB in the co-located and between 0.84 to 3.08 dB in the spatially separated conditions. These results are similar to the ones reported by Rana and Buchholz (2016).

3.4 Discussion

The current study applied an energetic, spatially-symmetric noise-vocoded speech masker to investigate the role of audibility on SRM due to BEG in NH and HI listeners as a function of frequency region. The results revealed a monotonic increase in both SRTs and SRM with increasing sensation level for all considered frequency regions. Moreover, no significant differences were observed between the HI and NH group for both SRTs and SRM. The results suggest that at masker sensation levels of up to 30 dB speech intelligibility as well as SRM due to BEG can be restored in HI listeners if adequate audibility is provided.

3.4.1 Effect of audibility normalization

Speech intelligibility as well as SRM was measured here in a group of ten older HI subjects (mean age of 70.3 years) and compared to the results measured in a group of ten younger NH subjects (mean age of 23.2 years). Audibility of all stimuli was normalized for each subject individually, first as a function of frequency and then relative to their individual SRT in quiet (see section 3.2.3). After this normalization, the results revealed no significant differences in performance between the HI and NH group, neither in SRTs nor SRM (see section III). This suggests that, when audibility is carefully controlled, older HI listeners are able to achieve similar speech intelligibility performance as well as SRM (due to BEG) as young NH adults.
These findings contradict the observations reported by other studies which typically find that even though audibility plays an important role, it cannot fully explain the reduction in performances seen in most HI listeners. Glyde *et al*. (2015), for instance, filtered (i.e., attenuated) the speech stimuli presented to their NH subjects in such a way that the sensation level within critical bands and relative to the audiogram, was equal across frequency for both groups. They then compared speech intelligibility performance in noise as well as SRM between the two groups and found that, even though differences between groups decreased substantially after audibility normalization, significant differences still remained. In a similar fashion, Jakien *et al*. (2017) investigated speech intelligibility performance in noise in a group of subjects with a wide range of hearing losses (and age) using speech stimuli that were normalized in sensation level across frequency using seven (overlapping) two-octave wide frequency bands. Their data did not only reveal a remaining (significant) decrease in performance with increasing four-frequency average hearing loss (4-FAHL), but also showed a substantial variance in performance within subjects with equal 4-FAHL. However, it is unclear how far the latter can be explained by the test-retest variability of their speech test. Moreover, the 4-FAHL does not take into account the shape of the audiogram, which may significantly influence performance. More appropriate measures may be the speech intelligibility index (SII: ANSI S3.5-1997) or related speech intelligibility models (e.g., Beutelmann and Brand, 2006; Rhebergen *et al*., 2010). Best *et al*. (2017) applied a very different approach as the above studies to investigate the effect of spatial auditory processing on speech

intelligibility in noise. They applied a bilateral glimpsing model to isolate target speech from noise and then provided the processed speech to their NH and HI subjects. In this way the subjects were not required to perform any stream segregation or spatial processing. The results were in agreement with the assumption that the performance of the HI (as well as NH) subjects was limited by their performance to identify the target speech within the available glimpses, which was most likely limited by audibility, and not by a deficit in spatial processing. Finally, Best *et al*. (2015) investigated BEG in a cohort of rather young adult HI listeners and found that even though speech intelligibility was significantly reduced in their HI subjects when compared to NH performance, the spatial benefit provided by BEG was not affected by hearing loss. The latter observation is in agreement with the present findings.

A number of studies have demonstrated that other factors than audibility play an important role in speech intelligibility in noise, including spectro-temporal resolution, the ability to process temporal fine-structure, cognitive abilities (e.g., selective attention), and age (e.g., Singh *et al*., 2008; Ahlstrom *et al*., 2009; Neher, *et al*., 2012; Gallun *et al*., 2013; Glyde *et al*., 2013a; Besser *et al*., 2015). Glyde *et al*. (2013a), for instance, used a correlation analysis to investigate the factors affecting speech intelligibility in noise for a cohort of 65 HI adults with a wide range of hearing loss and age. They applied linear amplification according to the NAL-RP prescription formula (Dillon, 2012, pp. 290-297) to partially restore audibility and showed that the 4-FAHL was still the main predictor for speech intelligibility as well as SRM. However, they also showed that age and cognitive abilities played a significant (though minor) role. Marrone *et al*. (2008) as well as Gallun *et al*. (2013) measured speech intelligibility in noise as well as SRM in HI subjects and reported that performance was mainly related to hearing loss but also found age an important factor. However, even though they presented their speech stimuli at a rather high sensation level, relative to SRTs in quiet, such an approach did not ensure audibility at mid to high frequencies where most hearing losses are strongest and speech has the least energy. Hence, audibility may have played a stronger role than anticipated and the observed effect of age may have been partly due to its correlation with hearing loss.

It is in no way suggested here that reduced audibility is the only factor that influences speech intelligibility in HI listeners in general. The stimuli used in this study were optimized in a rather specific way to focus on the effect of BEG in HI listeners, and the conclusions may well be different if other (maybe more realistic) stimuli were applied. The applied noise-vocoded speech maskers, for instance, minimized the influence of informational masking, which will have reduced the overall difference in SRM between HI and NH listeners (e.g., Best *et al*., 2012; Best *et al*., 2015). Moreover, stimuli were spatialized using artificial (infinite) ILDs and no ITDs, whereas most other studies used either loudspeaker for playback or Head-Related Transfer-Functions (HRTFs) that contained natural ILDs as well as ITDs. Finally, the fact that no significant differences were found between groups may be partly due to the limited statistical power provided by the data, which was limited by the number of subjects as well as the variance (or test-retest variability) of the SRTs, in particular within the HI group. Even though an increased number of subjects may have revealed a significant difference between groups, it should be highlighted that, when audibility is carefully controlled (i.e., more carefully than in most studies), then reduced audibility may play a more important role in speech intelligibility in noise than generally anticipated.

3.4.2 Effect of sensation level

The results presented in section 3.3 revealed that the SRTs increased monotonically in all conditions when the sensation level was increased, i.e., when the overall audibility was improved. For both groups and all for frequency regions, the SRT decreased (performance improved) in the co-located condition by about 4 dB with each 10 dB increase in sensation level and by about 6 dB in the spatially separated condition. As a result of these different growth-rates, the SRM increased on average by about 2 dB per 10 dB increase in sensation level. Considering the results as a function of frequency, it can be observed that the increase in both co-located and spatially separated SRTs was slightly steeper in the broadband condition than in the three narrowband conditions. Since the SRTs in all frequency regions were separately normalized to their corresponding SRT in quiet, the behaviour of the broadband SRTs suggests that the integration of speech information across frequency is improved with increasing audibility. This effect seems to be independent of the spatial configuration of the target and masker signal. Finally, the increase in SRT with increasing sensation level was slightly shallower in the low-frequency region than in the other frequency regions, which was more pronounced in the co-located condition. As a consequence, the SRM in the low-frequency region increased slightly faster with increasing sensation level than in the other frequency regions. This observation, together with the fact that audibility plays only a minor role at low frequencies (see section 3.1), may suggest that HI listeners may benefit more from BEG at low frequencies than at mid or high frequencies where the underlying ILD cues are naturally available.

The improvement in speech intelligibility observed in the HI (as well as NH) subjects when the overall (sensation) level is increased, is in general agreement with the relevant literature on speech intelligibility in noise (e.g., Moore *et al*., 1985; Ahlstrom *et al*., 2009; Kuk *et al*., 2015; Woods, *et al*., 2015) and is one of the main reasons for why amplification (with hearing aids) is prescribed as the main remediation for hearing loss. However, only a few studies have systematically investigated the effect of increasing sensation level in HI subjects with maskers that also allowed conclusions on the effect of sensation level on SRM. Glyde *et al*. (2015), for instance, provided different levels of linear, frequency-dependent amplification to their HI subjects, with the lowest amplification level equal to the gain prescribed by NAL-RP and the highest amplification level providing an additional gain of about 10 to 20 dB (depending on frequency). Whereas the increase in amplification had no effect on the co-located SRTs, the spatially separated SRTs (and thus also the SRM) improved by about 5 dB over the range amplification was varied. Hence, the general behaviour of both the spatially separated SRTs and SRM was similar to the present study, which was not the case for the co-located SRTs. The difference may be explained by the significant involvement of informational masking in Glyde *et al* due to the application of the same female talker as target and masker signals. This may have elevated the co-located SRTs and thereby reduced the influence of audibility. Moreover, it shifted the listener's task more from a speech identification task to a speech discrimination task, since informational masking is typically due to the subject confusing target speech with masking speech and not due to problems understanding the target speech. Similar to the present study, Jakien *et al*. (2017) reported on an experiment (their experiment II) in which they first normalized audibility of their speech stimuli as a function of frequency and then varied the overall sensation level (either 19.5 dB-SL or 39.5 dB-SL). They observed no significant change in average SRTs in the co-located condition and only a small improvement of about 1 dB in the spatially separated condition (and thus, in SRM). It is unclear why they found much smaller (or even negligible) effects of sensation level, but this is most likely explained by the difference in methods that they applied. In contrast to the present study, informational masking will have played a significant role, their audibility normalization was not as frequency specific as realized in the present study, and they spatialized their stimuli using natural ILDs and ITDs instead of using infinite ILDs (and no ITDs).

Finally, it should be mentioned that Rana and Buchholz (2016, experiment 2) followed a very similar approach to the present study, except that they only considered a single sensation level, which was 10 dB-SL for HI subjects and 35 dB-SL for NH subjects. Even though they used the individual audiograms (averaged over left and right ear) to normalize audibility across frequency as well as to derive sensation levels, which is different to the ear-specific, measurement-based approach followed here (see section 3.2), their stimuli and sensation levels are still more or less comparable to the ones applied here. Hence, their HI data can be compared to the present HI data at 10 dB-SL and their NH data to the present NH data at 30 dB-SL. Following such comparison, the SRTs as well as SRM measured in the two studies are in a very similar range, except for the SRM in the HI listeners which was slightly lower in the present study. This was due to slightly lower co-located SRTs (about 1 dB) and slightly higher spatially separated SRTs (about 1-2 dB). Nevertheless, this comparison provides at least some confirmation of the reproducibility of the present data and given that HI and NH subjects in Rana and Buchholz were tested at very different sensation levels, this comparison provides also some confirmation of the sensation level dependency observed in the present study.

3.4.3 Implications for hearing aids

The results obtained from this study can be used to better understand the amplification needed for enhancing speech intelligibility using ILDs. As described in section 3.2, the audibility of the present stimuli was normalized across both subjects and frequency and adjusted relative to individual SRTs in quiet. This makes it difficult to relate the stimuli to natural speech as well as to interpret the applied level modifications in terms of hearing aid gain. Hence, in order to better relate the results to hearing aid applications (or gain), the average spectrum in dB-SPL of the stimuli presented to the HI group, as measured at the ear drum of a HATS and analysed within two critical-band wide frequency bands (see table 3.1), is shown in the left panel of figure 3.5 for all four sensation levels (grey lines and filled circles). Here, only the spectra for the narrowband conditions are shown, but the spectra for the broadband condition are very similar, except that they are lower by about 11-14 dB. It should be noted that the shown spectra represent the target speech as well as the masker signals, because the long-term spectrum of the maskers was adjusted to the average spectrum of the target speech (see section 3.2.2). For reference purposes the average (ear drum) spectrum of speech for “normal” vocal effort (see ANSI S3.5-1997, table 1), with an unweighted free-field level of 62.4 dBSPL, is also shown in figure 3.5 (left panel) by the solid line with stars. The same speech spectrum after amplification according to NAL-RP is shown by the solid line with diamonds, whereby the average pure tone thresholds shown in figure 3.1 (dashed line and squares) were considered in the gain calculations. The upper limit of comfort (ULC) for the (ongoing) masker is shown by the light-grey area, and for the target sentences by the dark-grey area. The right panel of figure 3.5 shows the same information, except that all the spectra shown in the left panel (in dBSPL) were mapped into linear (hearing aid) gains by subtracting the (unaided) speech spectrum for normal vocal effort from all spectra (i.e., turning the solid line with stars shown in the left panel to a straight line at 0 dB gain).

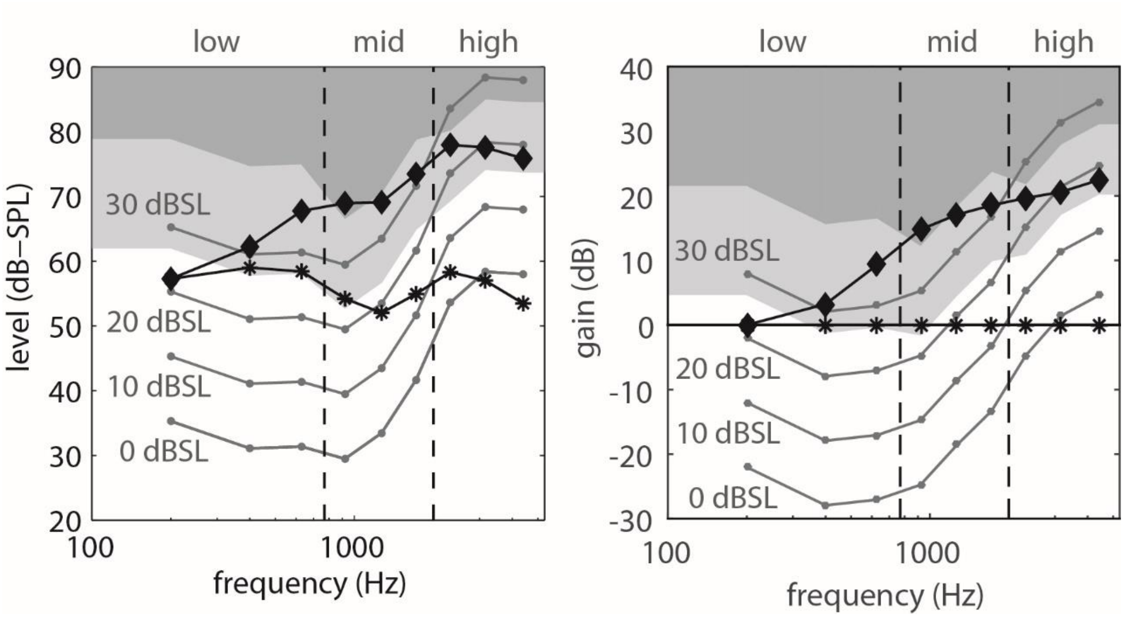

FIG. 3.5: Illustration of the stimulus spectra that were applied throughout the present study as measured at the ears of a dummy head (left panel). The corresponding gains are shown in the right panel. The grey lines refer to the stimuli at the four considered sensation levels. The solid lines with stars refer to natural speech for normal vocal effort according to ANSI S3.5-1997 and the solid lines with diamonds to the same speech but after linear amplification according to NAL-RP. The dark grey area indicates the ULC for the target sentences and the light grey area the ULC for the ongoing masker.

Comparing the spectra of the stimuli that were applied throughout this study (grey lines) with the spectrum of natural speech (solid line and stars), it can be deduced that the applied spectra were strongly amplified towards higher frequencies. This increase in gain with increasing frequency was even stronger as provided by NAL-RP, and directly reflects the sloping hearing loss of the HI subjects (see figure 3.1), which is very representative for HI listeners and is only partially compensated for by NAL-RP. Moreover, it can be seen in figure 3.5 that most stimuli at low and mid frequencies were softer than normal speech, even at the highest considered sensation levels. However, this was not the case at high frequencies, where most of the stimuli reached (or even exceeded) the ULC. In this regard it should be highlighted that all the curves shown in figure 3.5 reflect data averaged across all HI subjects, and individual spectra (or gains) varied substantially across subjects. Moreover, subjects were only tested in a given condition if the corresponding stimulus level was within their individual ULC (see section 3.2.4). As a consequence, only very few subjects were tested at the highest sensation levels, in particular in the high frequency region. Hence, any sensation level that in figure 3.5 is within the shaded areas indicates that a substantial part of the subjects could not be tested at that level due to loudness discomfort.

With respect to the right panel of figure 3.5, it can be deduced that the (average) gain provided here in the low and mid frequency region was always below (or just below) the gain provided by NAL-RP. This suggests that with standard amplification a SRM due to BEG of more than 8 dB (see figure 3.4) can be achieved (or even exceeded) in many HI listeners. However, natural ILDs, in particular at low frequencies, are rather small (or even absent), and thus will not provide the observed spatial advantage. This highlights the potential advantage of extending ILDs towards low frequencies in hearing aids as done in this study. Within the high frequency region (figure 3.5, left panel), any gain above about 15-20 dB already reached the ULC for the ongoing masker in a significant number of subjects, which is also about the gain provided by NAL-RP. Since such gain corresponds to a sensation level of about 10-15 dBSL (see grey lines in figure 3.5), it can be deduced from figure 3.4 that better-ear glimpsing at high frequencies can only provide a spatial benefit of up to about 2 dB, at least in subjects with similar hearing loss as considered here. Considering that natural ILDs are smaller than the applied infinite ILDs, it is expected that BEG with natural ILDs may even be smaller than 2 dB. Hence, it is rather unlikely that hearing aids can provide a gain at high frequencies that is sufficient to restore a significant spatial advantage from better-ear glimpsing without incurring loudness discomfort.

It should be highlighted that the above considerations are valid for any reference speech signal, since the gain shown in the right panel is always relative to that speech level. If the reference level is increased the gains are automatically decreased and vice versa. This general applicability of the results was one of the main reasons for why (individual) sensation levels were considered here rather that absolute speech levels. This also highlights that for soft speech a stronger amplification should be provided than for louder speech to maximise (or to keep constant) the spatial benefit provided by BEG; which is also true for speech intelligibility in general (see SRTs in figure 3.3). This is exactly the behaviour of the gain provided by WDRC in hearing aids. However, here only linear level manipulations (or gains) were considered, and it is unclear if the behaviour seen in figures 3.3-3.5 would still apply if non-linear amplification were provided using WDRC. Depending on the implementation of the WRDC, temporal modulations as well as spectral contrast of the incoming (speech) signals may be reduced, which may result in reduced speech intelligibility (e.g., Plomp, 1988; Boike *et al*., 2000; Bor *et al*., 2008; Dillon, 2012., pp-170-197). Moreover, binaurally unlinked hearing aids (i.e., hearing aids that operate independently at the left and right ear) may distort ILDs and thereby reduce the potential benefit provided by BEG (e.g., Wiggins and Seebeer, 2013). Binaurally linked hearing aids do not distort ILDs, but they may provide insufficient gain in at least one of the two ears, in particular when asymmetric hearing losses are considered. However, since at low frequencies most hearing losses are rather mild and speech also contains most of its energy, very little gain, and thus, very little (or even no) compression, is required to provide substantial benefits from BEG at low frequencies. However, this assumed that the acoustic path circumventing the hearing aid is sufficiently attenuated (e.g., by applying a tight ear mold), such that the hearing aid signal dominates the acoustic signal that arrives at the listener's ears.

3.4.4 Further considerations

Even though this study demonstrated that HI listeners can take advantage of (artificially generated) low- and mid-frequency ILDs in BEG to achieve a significant improvement in speech intelligibility in spatially separated and fluctuating noise, there are a number of factors that may reduce the benefit that can be achieved in the real world:

1. The present stimuli did not include any ITDs, but the auditory system utilizes ITDs, in particular at low frequencies (i.e., below about 1.5 kHz), to improve speech intelligibility in noise (e.g., Bronkhorst and Plomp, 1988). The spatial benefit provided by ITDs may partially offset the benefit provided by low-frequency ILDs.
2. The considered spatially-symmetric condition with two (noise-vocoded) speech maskers is known to provide rather large spatial benefits from BEG. In the real world, often more than two (dominant) sound sources may be present that are very differently located, and room reverberation may further deteriorate the benefit achieved by BEG (e.g., Marrone *et al*., 2008).

3. Speech at frequencies below about 500 Hz contributes less to overall speech intelligibility than higher frequencies, which is reflected in the reduced weighting in the speech intelligibility index (SII: ANSI S3.5-1997). This effect is qualitatively confirmed in figure 3.3, in which the SRTs are generally higher in the low frequency region than in the mid or high frequency region. Hence, it may be speculated that even if a substantial SRM is provided at (very) low frequencies, this benefit may contribute little to overall speech intelligibility.
4. The SRM measured in this study used linear signal processing to normalize the audibility of the stimuli as a function of frequency as well as to adjust sensation level. Since hearing aids need to ensure loudness comfort, they typically apply WDRC, which can modify the temporal and spectral behaviour of the signals arriving at the listener's ears, and, at least for binaurally unlinked hearing aids, can also distort ILDs (e.g., Wiggins and Seeber, 2012). However, since most hearing losses are rather mild at low frequencies and speech contains a lot of energy, very little amplification may already provide adequate audibility (see also section 3.4.3 for discussion).
5. Even though there are a number of approaches that may be used in hearing aids to enhance ILDs at low frequencies, they are not expected to provide ILDs as large as the infinite ILDs applied throughout this study. As a consequence, the provided SRM may be smaller than observed here. Potential candidates for realizing low-frequency ILDs are directional processing with multiple microphones (e.g. Kates, 2008, pp. 93-98) or frequency transposition of natural ILDs at mid and high frequencies to low frequencies (e.g. Robinson *et al*., 2007). However, details are out of the scope of the present study.

Future studies will need to address the impact of these potentially detrimental effects on the spatial advantage provided by better-ear glimpsing before adequate ILD enhancement methods as well as improved amplification schemes can be successfully implemented in hearing aids.

3.5 Conclusions

Results indicate that if audibility is compensated in HI listeners then they can utilize ILDs as well as NH listeners (at low sensation levels) to improve speech intelligibility in noise. However, due to the limitations in both user acceptance and available hearing aid technology, it is not possible to provide the gain that is required to provide significant SRM at high frequencies. Nevertheless, it is possible to introduce substantial SRM at low and mid frequencies by artificially enhancing ILDs, which may be achieved by using either multiple microphones or signal processing strategies similar to frequency transposition. The present results will help developing future hearing aid technologies as well as amplification strategies that can improve speech intelligibility in noise by maximizing SRM. Further, overall results also highlight the important role of audibility in utilizing ILD cues as well as to understand speech in noise in general.

3.6 Acknowledgements

The authors acknowledge the financial support of Macquarie University through an international research excellence (iMQRES) scholarship and the Australian National Health and Medical Research Council via grant APP1056332.

Chapter 4: Effect of audibility on better-ear glimpsing using non-linear amplification

Better-ear glimpsing (BEG) utilizes interaural level differences (ILDs) to improve speech intelligibility in noise. This spatial benefit is reduced in most hearing-impaired (HI) listeners due to their increased hearing loss at high frequencies. Even though this benefit can be improved by providing increased amplification, the improvement is limited by loudness discomfort. An alternative solution, therefore, extends ILDs to low frequencies, which has been shown to provide a substantial benefit from BEG. In contrast to previous studies, which only applied linear stimulus manipulations, wide dynamic range compression was applied here to improve the audibility of soft sounds while ensuring loudness comfort for loud sounds. Performance in both speech intelligibility and spatial release from masking was measured in 13 HI listeners at three different masker levels and compared to their performance with linear amplification as well as to normal-hearing listeners. The results revealed that at low signal levels, performance substantially improved with increasing masker level, but this improvement was reduced by the compressive behaviour at higher levels. Moreover, extending ILDs to low frequencies provided an extra spatial benefit of up to 5 dB on top of the one provided by natural interaural time differences and ILDs, which increased with increasing signal level.

4.1 Introduction

Spatial release from masking (SRM) refers to the improved perception of a target signal when it is spatially separated from a masker signal versus when it is co-located. SRM helps in understanding of speech in noise for listeners with normal hearing (NH) as well as listeners with hearing impairment (HI). However, the effect of SRM is typically reduced in HI listeners for a number of reasons, including loss of audibility, poor spectral and temporal resolution, impaired processing of the signal's temporal fine structure, and aging-related effects such as reduced selective attention, short-term memory, and executive function (e.g., Dubno *et al.*, 2002, Gallun *et al.*, 2013 Glyde *et al.*, 2013). In general, SRM can be due to a release from energetic masking as well as informational masking, depending on the nature of the involved stimuli as well as the listener's task (e.g., Freyman *et al.,* 1999; Brungart *et al.,* 2001; Kidd *et al.,* 2007). Energetic masking occurs when similar set of neurons are excited by the masker and the target stimuli due to their temporal and spectral overlap within the auditory periphery, whereas informational masking occurs either due to similarities in the target and masker signals or due to maskers interfering with auditory stream segregation processes or selective attention (e.g. Pollack, 1975; Shinn-Cunningham, 2008; Kidd *et al.,* 2010; Culling *et al.,* 2017).

The SRM obtained when the distractors are placed only on one side of the head and the target signal arrives from the front (or the other side of the head) can be largely attributed to the head shadow effect resulting in a single ear with a consistent better signal to noise ratio (SNR). This situation becomes more complex when listeners are surrounded by more than one fluctuating distractor, and when distractors are placed not only on one side of the head but rather on both sides of the head. In such conditions, the ear with the better SNR is not consistent but rather keeps fluctuating between the ears. The SRM obtained in such scenario can be largely attributed to a phenomenon known as better-ear glimpsing (BEG; Brungart and Iyer, 2012). BEG is best studied in symmetric masker conditions (i.e., with the target speech from the front and a single fluctuating masker from either side of the head (Brungart and Iyer, 2012; Glyde *et al.,* 2013b) and provides a release from energetic masking by relying on the listener's ability to take advantage from either ear that provides the better short-term SNR. This process relies on the occurrence of short-term interaural level differences (ILDs, Glyde *et al.,* 2013b) which, in the real world, are always accompanied by interaural time differences (ITDs). However, it has been reported by Glyde *et al.* (2013c), that, at least in spatially-symmetric masker conditions, ILDs (due to BEG) contribute more to SRM than ITDs. Although some studies have reported contrary results (Culling *et al*., 2004; Kidd *et al*., 2010). Independent of the relative contribution of ILDs to SRM, it is known that ILDs mainly exist at high frequencies but, unfortunately, that is where most HI listeners show the strongest hearing loss (e.g., Dillon, 2012, pp. 286-335). Thereby, the more severe the hearing loss the lesser will be the audibility of the signal. As a consequence, HI listeners cannot take full advantage of BEG, leading (or contributing) to their difficulties in understanding speech in noise.

The most commonly applied solution to compensate for the loss of audibility is to provide hearing aid amplification, which for most hearing losses requires increased amplification at high frequencies. In this regard, Glyde *et al*. (2015) have shown that by providing extra (linear) amplification on top of what is recommended by common prescription rules such as National Acoustic Laboratories – Revised Profound (NAL-RP; Dillon, 2012, pp. 290-297), the SRM provided by BEG can be significantly improved. However, providing amplification that provides audibility that is similar to NH listeners, in particular at high frequencies, is challenging. On the one hand, this is due to the reduced auditory dynamic range leading to loudness discomfort issues and, on the other hand, due to technical limitations such as acoustic feedback. In this regard as an alternative to increasing amplification (and thereby improving audibility) at high frequencies, Rana and Buchholz (2016) proposed to artificially provide BEG cues (i.e. ILDs) at low and mid frequencies, where hearing loss is usually less pronounced. They found that both NH and HI listeners can utilize these artificially extended BEG cues to provide a substantial amount of SRM. However, the performance obtained in HI listeners was poorer than for NH listeners, which was most likely due to differences in audibility. Whereas the distractors for the HI listeners were presented at 10 dBSL (relative to pure tone thresholds), the distractors for NH listeners were presented at 35 dBSL. Therefore, in a follow-up study, Rana and Buchholz (submitted) carefully controlled the audibility of their speech signals in a cohort of HI listeners across frequency by providing gain equivalent to their individual speech detection thresholds (SDTs).They then tested their speech intelligibility performance in noise with artificially maximized (broadband) ILDs at four different sensation levels (0, 10, 20 and 30) relative to their speech recognition thresholds (SRTs) in quiet. The results revealed that HI listeners can utilize these artificial ILD cues in BEG to the same extent as NH listeners, as long as both groups are tested at equal audibility (or sensation) level. Interestingly, the improvement in both SRTs and SRM increased linearly with the increase in sensation level up to 30 dB, highlighting the important role played by audibility on BEG and the possibility to provide substantial SRM to HI listeners, at least at low and mid frequencies. The observed increase of

SRM with increasing sensation level is in general agreement with other studies that utilized natural spatial cues (e.g., Best *et al.,* 2017; Jakien *et al.,* 2017). One common problem that was raised by all the referenced studies was the inability of some HI listeners to perform at higher audibility levels due to loudness tolerance problems.

Wideband Dynamic Range Compression (WDRC: Kates, 2008, pp. 221-259) is one of the commonly used methods in hearing aids to avoid loudness tolerance issues and, at the same time, provide increased audibility to soft sounds. Even though WDRC provides increased audibility to soft sounds, it is unclear how far the increase in audibility can improve the effectiveness of BEG. Depending on how WDRC is implemented in a hearing aid as well as fitted to a listeners hearing loss, it can reduce temporal fluctuations as well as the spectral contrast of the incoming signals or provide insufficient amplification (Dillon, 2012, pp. 170-197). It may well be that the detrimental effect of these signal distortions counteract the benefit provided by the increase in audibility of soft sounds, resulting in a negligible or even negative effect on BEG. This potential problem may be further aggravated when two independently operating hearing aids are fitted to the left and the right ear of a listener, which can result in distorted ILD cues (e.g., Wiggins and Seebeer, 2012; Buchholz, 2013). To the best knowledge of the authors, there are no studies that have systematically investigated the effect of increasing audibility on BEG using non-linear amplification. However, there are a few studies that have investigated the effect of audibility using non-linear amplification on speech recognition tasks. For instance, Davies-Venn et al. (2009) tested listeners with different degrees of hearing loss on a nonsense syllable recognition task in noise who were fitted with hearing aids with fast acting WDRC. They reported a significant improvement in performance when the stimulus level was increased from 50 to 65 dBSPL, but the performance reduced when the sensation level was further increased from 65 to 80 dBSPL. Hence, it may be similarly expected that SRM provided by BEG is improved by non-linear amplification at soft signal levels, but reduced at high signal levels.

Overall, this study progresses the work done by Rana and Buchholz (2016) as well as Rana and Buchholz (submitted) in terms of understanding the role of audibility on BEG and the ways that BEG can be maximized in HI listeners to improve speech intelligibility in noise. The specific aim of the present study was to investigate the effect of non-linear amplification using WDRC on BEG and thus, to better understand the interaction between the provided increase in audibility and the inherent signal distortions. The results obtained from this study will help to better understand and to design non-linear amplification schemes that optimize the benefit provided by BEG in real-world environments.

4.2 Method

Speech intelligibility, as well as SRM, was measured for NH and HI listeners using stimuli with different combinations of spatial cues as well as using different amplification schemes.

4.2.1 Participants

Ten NH listeners (hearing thresholds < 15 dB hearing level) aged between 25 and 41 years (mean age of 33.5 years) and 13 HI listeners aged between 68 years and 79 years (mean age of 74 years) with a symmetric (<10 dB difference between ears from 250 Hz to 4 kHz), sensorineural, mild to moderate-severe hearing loss participated in this study. Hearing thresholds were tested prior to the experiment for all participants to either confirm normal hearing or to establish their degree of hearing loss. The mean and ±1 standard deviation of the four-frequency (500, 1000, 2000, 4000 Hz) average hearing loss (4-FAHL) of the HI subjects was 31 ± 8 dB. Individual and mean pure tone thresholds are shown in figure 4.1 averaged across the left and right ear. All participants had Australian English as their first language and had no reported attention deficit disorder or intellectual disability. The complete testing was conducted in a sound-treated audiological test booth at the National Acoustic Laboratories and took about 4 hours per subject, which was divided into two appointments of two hours each. Participants received a small gratuity for their participation. Ethical clearance was taken from the Australian Hearing Human Research Ethics Committee and the Macquarie University Human Research Ethics Committees.

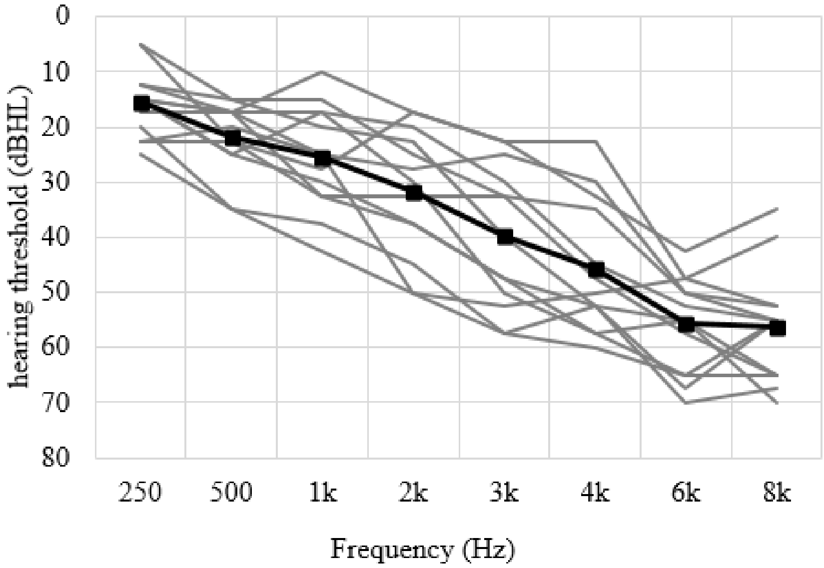

FIG. 4.1: Mean (black line) and individual (grey line) pure tone audiograms of the 13 HI listeners averaged over the left and right ear.

4.2.2 Stimuli

Speech intelligibility was assessed with a corpus of 80 lists of 16 BKB-like target sentences (Bench *et al*., 1979) spoken by a native Australian English female speaker. The RMS level of all sentences was normalized and the average spectrum of each sentence list was equalized to match the average spectrum of the entire corpus, which is shown in the left panel of figure 4.2. The masker was realized by two separate noise-vocoded single talker speech distractors, which were identical to the ones described by Rana and Buchholz (2016). In brief, the two different-voice speech discourses from Cameron and Dillon (2007) were noise vocoded using a short-term Fourier Transform with 20 ms long time windows and four critical bands wide spectral smoothing. This process realized a noise vocoder with about 5 effective frequency channels (within a bandwidth of 8 kHz) and made the two distractors, when combined, largely unintelligible. Each distractor was equalized to match the long-term spectrum of the target speech. The noise vocoding was applied to minimize the influence of informational masking and thereby to focus on energetic auditory processes such as BEG (Brungart and Iyer, 2012) or equalization-cancellation (Durlach, 1963).

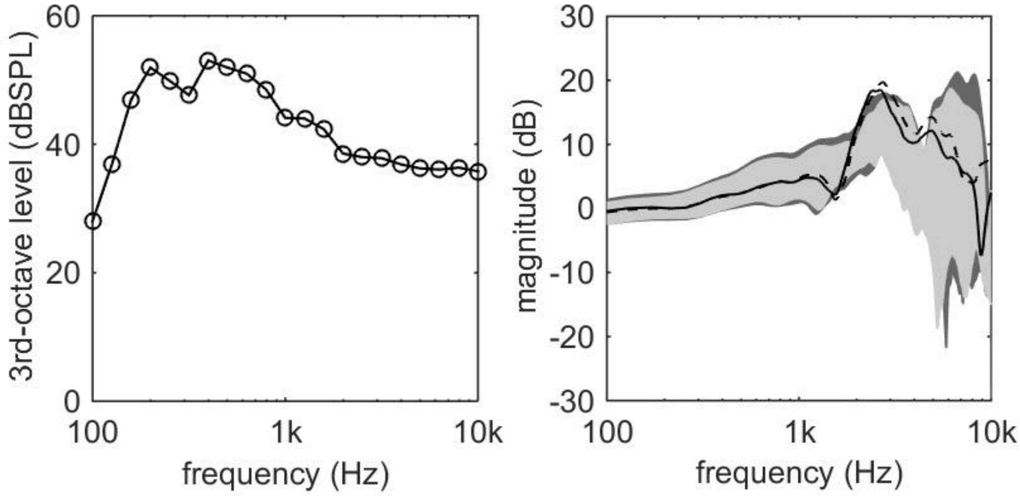

FIG. 4.2: The left panel shows the third-octave spectrum of the target speech in free-field (i.e., before spatialization) at 60 dBSPL, which was identical for the distractor signals. The right panel shows the magnitude spectrum of the applied natural HRTFs for the front (0°) direction (left ear: solid line; right ear: dashed line) as well as for the +90° (light-grey shaded area) and -90° (dark grey-shaded area) direction, whereby the corresponding spectra at the left and right ear are indicated by the edges of the shaded areas. The shaded areas provide a direct representation of the involved ILDs.

All stimuli were spatialized using non-individualized Head-Related Transfer Functions (HRTFs) and stored as binaural wave-files with a sampling frequency of 44.1 kHz. The target signal was always presented from the front of the listener (0° direction) and the two distractor signals were either co-located with the target or spatially separated at +90 and -90 degrees. The stimuli were spatialized using the HRTFs described by Cameron and Dillon (2007), which were measured on a Bruel & Kjaer Head and Torso Simulator (HATS). The magnitude spectra of the HRTFs are shown in the right panel of figure 4.2 separately for the left and right ear. For the spatially separated HRTFs the ILDs are indicated by the shaded areas. The corresponding ITDs are not shown here, but for the 0° direction the ITD was less than 23 μs (i.e., less than 1 sample) and around -680 μs and +680 μs for the -90° and +90° direction, respectively.

To investigate the individual contributions of ILDs and ITDs on speech intelligibility as well as SRM, in one spatially-separated condition the ITDs were removed and only the ILDs were preserved. This ILD-only condition was realized by deriving minimum-phase versions of the original HRTFs for the +90° and -90° directions. Finally, a condition with infinite ILDs and no ITDs was created by using the HRTFs for the 0° direction, but removing (i.e., multiplying with zero) either the left ear or the right ear of the HRTF. In the co-located (i.e., 0°) condition the same HRTF as for the target speech was utilized, but a gain of -3 dB was applied to each ear to compensate for the level increase that resulted from adding two distractors. In summary, speech intelligibility was assessed either with the distractor co-located with the target at 0° or with the distractors spatially separated at ±90° using (a) natural HRTFs (i.e., natural ITDs and ILDs), (b) natural HRTFs without ITDs (i.e., using ILDs only), and (c) extended (or infinite) ILDs.

The binaural target and distractor stimuli were presented to a real-time master hearing aid (Yeend et al., 2014) using a standard Windows computer connected to a RME™ Fireface UC USB sound card and running purpose-built speech test software in Matlab (see below). The hardware of the master hearing aid consisted of a standard Windows computer connected to another RME™ Fireface UC USB sound card and received the binaural input from the test computer via a balanced audio cable. The amplified output signal of the master hearing aid was presented to the test subjects through Sennheiser HD215 circumaural headphones, which were calibrated using a Bruel & Kjaer artificial ear.

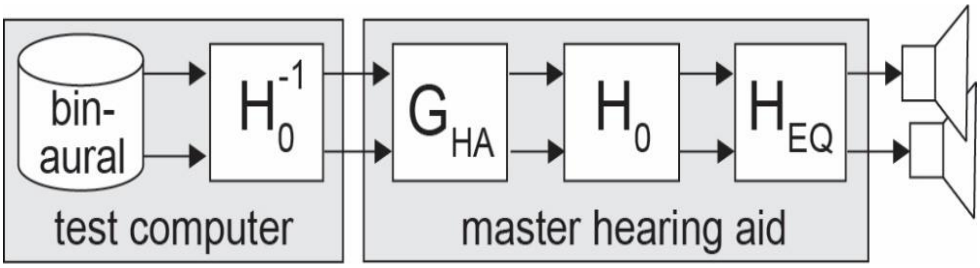

FIG. 4.3: Illustration of the signal processing relevant to the stimulus playback and hearing aid signal processing. Binaural stimuli are generated for different spatial test conditions, transformed to free-field equivalent levels using a filter with transfer function H_0^{-1}, and presented to a real-time master hearing aid using a test computer. The master hearing aid applies either linear or non-linear amplification as indicated by the gain G_{HP}, maps the resulting signals back to ear-drum levels, and plays them to the subjects using equalized (via a filter with transfer function H_{EQ}) headphones. Further details are described in the text.

The main signal processing associated with the stimulus playback and master hearing aid processing is illustrated in figure 4.3, with H_0 the free-field-to-ear-drum transfer function, H_0^{-1}

the ear-drum-to-free-field transfer function (or the inverse of H_0), G_{HA} the main hearing aid processing (or gain), and H_{EQ} the headphone equalization filter. The free-field-to-ear-drum transfer function H_0 approximated the absolute spectrum of the HRTF for the 0° direction shown in figure 4.2 (right panel) averaged across the left and right ear. This function was realized in the master hearing aid by applying appropriate gains to the individual frequency channels of the hearing aid filterbank. The ear-drum-to-free-field transfer function H_0^{-1} was realized by a 1024-samples long minimum-phase Finite Impulse Response (FIR) filter that was applied offline to the binaural signals on the test computer. The headphone equalization filter was realized by a 512-samples long minimum-phase FIR filter that was derived with a Bruel & Kjaer artificial ear and averaged across the left and right ear. Hearing aid amplification was set in 16 1/3-octave wide frequency channels either by applying linear amplification according to NAL-RP prescription or by applying non-linear amplification according to the NAL-NL2 prescription (e.g., Dillon, 2012, pp. 313-314). In the latter case, fast-acting wideband dynamic range compression (WDRC) was applied to each frequency channel separately with attack and release times of 10 ms and 100 ms, respectively. Amplification was derived individually for the subject's hearing loss averaged across the left and right ear.

4.2.3 Procedure

Similar to Rana and Buchholz (2016), adaptive speech reception thresholds (SRTs) were measured for target sentences in the presence of different distractors using a Matlab program installed on a personal computer. The participant's task was to repeat as many words as they heard in each target sentence while ignoring the distracting signals. At least 17 and up to 30 sentences were presented and the signal-to-noise ratio (SNR) was adjusted adaptively to achieve 50% correct word identification (SNR50) by keeping the distractor level constant and varying the target level. Further details can be found in Keidser *et al.* (2013).

In order to investigate the effect of non-linear amplification on the utilization of spatial cues in understanding speech as well as in SRM, all HI listeners were tested at an overall masker level of 50, 60, and 70 dB-SPL, as measured in free-field before amplification was applied. As a reference condition, all HI listeners were also tested with linear amplification at a single masker level of 60 dBSPL. The conditions tested with the HI subjects are summarized in table 4.1. The NH subjects were tested at all four spatial distractor configurations with linear amplification of 0 dB and a combined distractor level of 60 dBSPL. For each condition, the SRTs were averaged over two measurements, and the SRM was calculated by subtracting SRTs in the spatially separated conditions from the SRTs in the corresponding co-located condition.

TABLE 4.1: Overview of conditions tested with HI subjects.

Spatial presentation	Amplification	Masker level (dBSPL)
	Linear	60
	Non-linear	50, 60, 70
	Linear	60
	Non-linear	50, 60, 70
	Linear	60
	Non-linear	50, 60, 70
	Linear	60
	Non-linear	50, 60, 70

Before any speech testing started, loudness assessments were done in each test subject for the target and masker signals separately to ensure loudness comfort for all tested stimuli using a 8-point rating scale (1-very soft; 2-soft; 3-comfortable, but slightly soft; 4-comfortable; 5-comfortable, but slightly loud; 6-loud, but ok; 7-slightly uncomfortable; 8-uncomfortably loud). None of the subjects rated any of the stimuli as uncomfortable.

4.3 Results

The data was analyzed using IBM SPSS version 22.

4.3.1 Speech reception thresholds (SRTs)

The mean SRTs for the 13 HI subjects are shown in figure 4.4 (left panel, filled symbols) together with 95% confidence intervals for the four different spatial cue conditions with amplification as parameter (i.e., non-linear amplification at 50, 60, and 70 dBSPL masker level and linear amplification at 60 dBSPL masker level). For reference purposes the SRTs for the NH subjects are also shown (open symbols). Since the main interest of this study was to examine the effect of restoring audibility using non-linear amplification on the utilization of artificially extended ILDs and other combinations of spatial cues, a two-way repeated measure ANOVA was conducted with masker level and type of spatial cues as main effects. The results revealed a significant main effect of masker level [$F(2, 24) = 46.03$, $p < 0.01$) as well as type of spatial cues [$F(3, 36) = 282.36$, $p < 0.01$], but no significant interaction. As it is shown in figure 4.4 (left panel), the SRTs improved with an increase in masker level across all conditions. However, the rate of improvement was much faster when the level was increased from 50 to 60 dBSPL than when increased from 60 to 70 dBSPL. The results of a paired t-test with Holm-Bonferroni correction (Holm, 1979) revealed a significant difference in SRTs between 50 and 60 dBSPL across all conditions ($p < 0.01$) but no significant difference between 60 and 70 dBSPL, except when extended ILDs were used. Further, based on an independent t-test, it was also found that even at the highest masker level of 70 dBSPL with non-linear amplification, the SRTs of the listeners with HI were on average 4 dB higher (i.e., worse) than for the listeners with NH at 60 dBSPL masker level ($p < 0.01$).

Another main interest of this study was to measure the extra benefit that artificially extended ILDs can provide on top of natural spatial cues. Results of a paired t-test with Holm-Bonferroni correction showed that SRTs with extended ILDs were significantly ($p<0.01$) lower than the corresponding SRTs with natural ILDs alone as well as when natural ILDs were combined with natural ITDs for HI listeners as well as NH listeners. The contribution of ITDs to the SRT on top of natural ILDs was also analyzed by subtracting SRTs with natural ILDs and ITDs from SRTs with natural ILDs alone. Based on a paired t-test with Holm-Bonferroni correction, a significant contribution of ITDs was found only for HI listeners for a masker level of 60 dBSPL, for both linear and non-linear amplification, and for NH listeners.

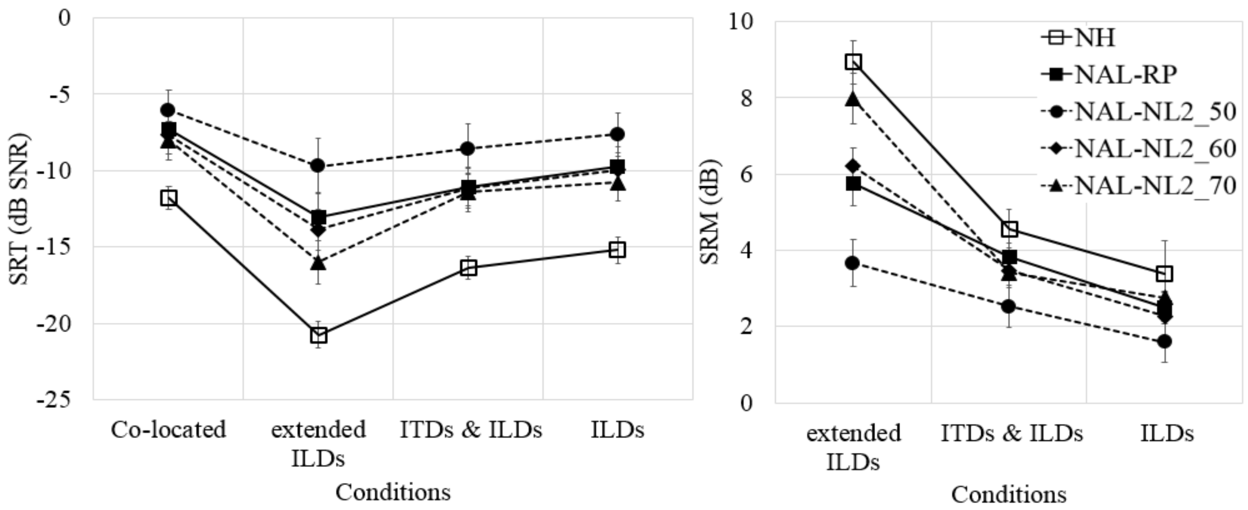

FIG. 4.4. Mean SRTs, mean SRM and 95 % confidence intervals across conditions.

The extent to which WDRC in non-linear amplification can negatively or positively affect the utilization of spatial cues when compared to linear amplification was also investigated. The results of a paired t-test revealed no significant differences (p>0.01) between SRTs with linear and non-linear amplification across all conditions when the same masker level (i.e., 60 dBSPL) is considered. Finally, the test-retest variability was assessed by subtracting the SRTs of the second trial from the SRTs of the first trial. According to a paired t-test, the mean intra-subject difference across all conditions, for both NH and HI listeners, ranged from -0.7 to 1.32 dB, and was not significant. These results are in line with Rana and Buchholz (2016).

4.3.2 Spatial release from masking (SRM)

SRM was calculated by subtracting the spatially separated SRTs from the co-located SRTs. The mean SRM with 95 % confidence intervals is shown in figure 4.4 (right panel) for the HI (filled symbols) and NH (open symbols) listeners. Similar to the SRT analysis, a two-way repeated measure ANOVA was applied to the SRM for non-linear amplification with masker level and type of spatial cues as main effects. Results revealed a significant main effect of level [F (2, 24) = 22.72, p <0.01) as well as type of cues [F (2, 24) =148.9, p < 0.01], but no significant interaction. A paired t-test with Holm-Bonferroni correction revealed a significant effect of masker level on SRM only for masker with extended ILDs. To investigate if the best performance of listeners with HI, i.e. the SRM at a masker level of 70 dBSPL, is similar to the performance of the listeners with NH, an independent t-test was conducted. Results showed no significant difference in SRM between the two groups. This observation is different to the SRTs, which, in HI listeners, were significantly higher (i.e., worse) than for NH listeners in all conditions. Hence, even though SRM may be restored in HI listeners by providing sufficient non-linear amplification, this is not the case for overall performance in speech intelligibility; at least for the amplification levels considered in the present study.

To measure the extra advantage of artificially extending ILDs to low and mid frequencies in SRM, a paired t-test with Holm-Bonferroni correction was conducted. The SRM obtained with extended ILDs was significantly ($p<0.01$) better than the SRM obtained with natural spatial cues, both with and without ITDs. The SRM data shown in figure 4.4 (right panel) suggests that the SRM obtained with natural ILDs and ITDs is on average about 1 dB larger than for the case that only ILDs are applied. However, the results of a paired t-test revealed that this contribution of ITDs was only significant for listeners with HI at 60 dBSPL, for both linear and non-linear amplification, and for NH listeners. Finally, using a paired t-test to compare the SRM between linear and non-linear amplification at a masker level of 60 dBSPL found no significant difference for all conditions. This is in agreement with the SRT data and indicates a negligible impact of WDRC on SRM across all spatial cues.

4.4 Discussion

The effect of sensation level on the ability of HI listeners to utilize artificially extended ILDs in BEG with controlled audibility across frequency using linear amplification (or attenuation) has been investigated by Rana and Buchholz (submitted). The present study builds upon their methods and results with the new aim to better understand (a) how far the utilization of natural ILD cues can be restored for BEG in HI listeners when audibility is controlled by applying non-linear amplification, as commonly used in hearing aids, and (b) how far this spatial benefit can be further improved by applying artificially extended ILDs. The findings of this study are discussed in the following sections.

4.4.1 Effect of audibility

The role of audibility on speech intelligibility has been investigated in a number of studies. For example, Rana and Buchholz (submitted) controlled audibility in HI (and NH) listeners carefully across frequency and then measured SRTs in a noise-vocoded two-talker masker at different masker sensation levels. For each increase in sensation level of 10 dB they found an average improvement of 2 dB in co-located SRTs, 4 dB in spatially-separated SRTs, and a corresponding 2 dB in SRM. Glyde et al. (2015) provided 50 % extra amplification on top of the frequency-dependent amplification prescribed by the NAL-RP formula, and HI listeners showed an improvement in co-located SRTs by about 1 dB, in spatially separated SRTs by about 4 dB, and in SRM by about 3 dB. Similarly, Jakien et al. (2017) investigated speech intelligibility at low (19.5 dBSL) and high (39.5 dBSL) sensation levels relative to their individual SRTs in quiet, and depending on their stimulus conditions, they found different degrees of improvements in SRTs and SRM with increasing sensation level.

In all of these studies the benefit provided by increased sensation level was presumably due to improved access (due to increased audibility) to the softer components of the target speech as well as to the available spatial cues (e.g., Glyde *et al.*, 2015). Since all of these studies, except Rana and Buchholz (submitted), did not (or not fully) equalize the audibility of their stimuli across frequency, the observed improvement in performance with increased overall sensation level may also have been caused by an increase in the overall effective frequency bandwidth of their stimuli. This is because the considered hearing losses were more severe towards higher frequencies and at the same time the energy of the applied speech stimuli decreased with increasing frequency. As a consequence, audibility plays a larger role at higher frequencies and therefore, provides also an increased potential for receiving a benefit from amplification. This observation can be also linked to the findings by Glyde *et al.* (2015) that HI listeners may largely benefit from amplification in spatially-separated masker conditions due to the increased access to (natural) ILD cues that are mainly available at high frequencies (i.e., above about 1.5 kHz). It is likely, that in all the above studies, the utilization of linear amplification played a crucial role in the observed benefits in speech intelligibility as well as SRM, since it not only preserved the temporal and spectral behaviour of the signals at the listeners' ears, but also preserved the available ILDs (Dillon 2012, pp.170-197). Unfortunately, the limited auditory dynamic range that is available in HI listeners, in particular at high frequencies, does not provide enough scope for improving the audibility of the incoming signals with linear amplification without incurring loudness discomfort (Rana and Buchholz, submitted). Accordingly, all the studies referenced above reported loudness discomfort problems in at least some of their HI listeners.

Therefore, in the present study, two possible solutions were investigated which aimed at improving access to ILD cues for BEG and at the same time controlled loudness comfort: (a) non-linear amplification using WDRC and (b) extension of (artificial) ILD cues towards lower frequencies where audibility is typically less of an issue. Non-linear amplification is commonly applied in hearing aids (e.g., Kates, 2008, pp.221-259; Dillon, 2012, pp.170-197) and frequency-extension of ILDs has already been shown to be effective in combination with linear amplification (Rana and Buchholz, 2016; submitted). To provide different levels of audibility using non-linear amplification, HI listeners were tested at three different masker levels (i.e., at 50, 60 and 70 dBSPL) using a master hearing aid with multi-channel, fast-acting WDRC that was fitted to the individual test subjects according to the NAL-NL2 prescription formula. It was found that, when the masker level was increased from 50 to 60 dBSPL, the co-located SRTs improved by about 1.6 dB, whereas the spatially separated SRTs improved by about 3.5-5.5 dB, with the larger increase observed with the artificially extended ILDs. Hence, the SRM was increased by about 2-4 dB. When the masker level was further increased from 60 to 70 dBSPL, the co-located SRTs improved by only 0.4 dB, whereas the spatially separated SRTs improved by only 1-3 dB, again with the larger increase observed with the artificially extended ILDs. The improvement in SRTs was found to be significant for all conditions when the masker level was increased from 50 to 60 dBSPL, but when the masker level was increased from 60 to 70 dBSPL the effect was only significant for the spatially-separated condition with artificially extended ILDs. The improvement in co-located and spatially-separated SRTs (and thus SRM) with an increase in masker level at soft levels is very similar to the improvement found with linear amplification as described by Rana and Buchholz (submitted) as well as in Glyde et al. (2015) and Jakien et al. (2017) The observed reduction of this improvement at higher signal levels is in agreement with other studies that measured the effect of non-linear amplification on speech intelligibility in noise (e.g., Davies-Venn *et al.*, 2009) and reflects the increase in compression at higher signal levels as further described below.

4.4.2 Effect of compression

The effect of different compression parameters (such as compression threshold, compression ratio, the number of frequency channels, attack time, release time) on speech intelligibility has been discussed extensively throughout the relevant literature (Dillon, 2012, pp. 170-197). In the present study, a real-time master hearing aid was used to realize fast-acting WDRC in 16 independent 1/3-octave wide frequency channels. In each frequency channel, the compression characteristics were set according to the NAL-NL2 prescription formula, with the compression thresholds set to the 1/3-octave levels of LTASS speech (Byrne *et al.*, 1994) with a broadband level of 52 dBSPL. The average compression ratio for the HI group was between 1:1 and 1.5:1 at frequencies below 1 kHz and around 2:1 at frequencies above 1 kHz. In figure 4.5, the provided gain according to NAL-NL2 and averaged across the HI subjects is shown in the left panel for the applied speech signals at 50, 60, and 70 dBSPL. The resulting free-field levels are shown in the right panel. The applied (linear) gain according to NAL-RP is also shown as well as the corresponding speech spectrum before and after linear amplification with an input level of 60 dBSPL.

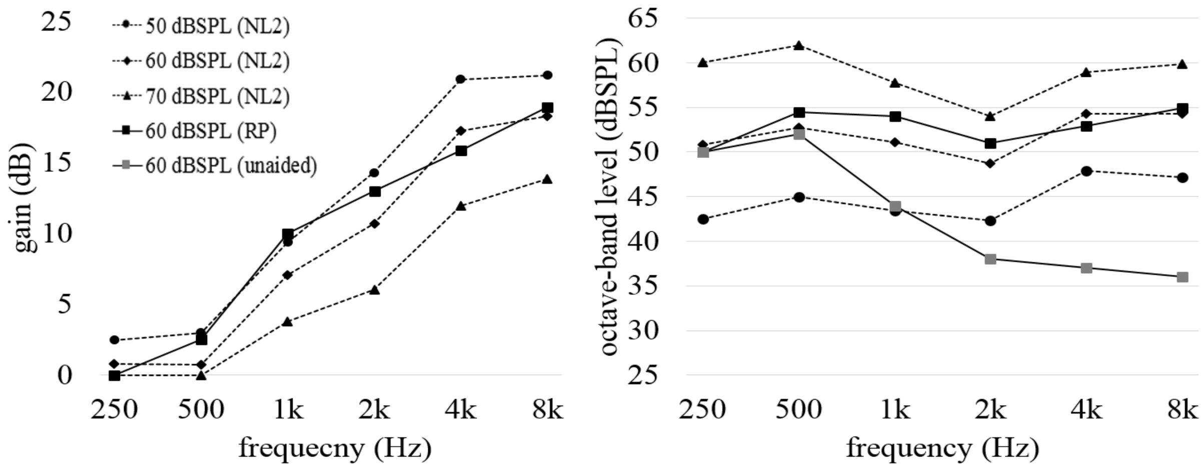

FIG. 4.5: The average gains prescribed by the (non-linear) NAL-NL2 formula for the tested speech levels of 50, 60, and 70 dBSPL are shown in the left panel and the corresponding speech spectra after amplification are shown in the right panel. The corresponding gains and resulting speech spectrum are also shown together with the unaided speech spectrum at a free-field level of 60 dBSPL.

In the previous section it was mentioned, that when the masker level is increased from 50 to 60 dBSPL, a similar large increase in both SRTs and SRM is achieved as previously observed with linear amplification. When the masker level was further increased from 60 to 70 dBSPL, this increase was substantially reduced in the co-located as well as all the spatially separated conditions except for the case of artificially extended ILDs. This behavior can be explained by considering the details of the applied WDRC implementation described above and figure 4.5. Since the compression thresholds were set to the hearing aid frequency band levels of speech with a broadband level of 52 dBSPL, the 50 dBSPL masker was basically processed linearly, which was also the case for a substantial part of the masker at 60 dBSPL. Given that the measured SRTs were at highly negative SNRs (i.e., around -6 to -16 dB), the corresponding target speech was also in the linear range of the WDRC. Hence, it is not surprising that, due to the rather linear behavior of the WDRC, a similar increase in performance can be observed with increasing masker level as previously reported with linear amplification. However, this is not the case for the highest masker level of 70 dBSPL, for which most of the masker components and a substantial part of the target speech would have been in the compressive region of the WDRC. As a consequence, the benefit achieved by increasing the masker level from 60 to 70 dBSPL was reduced. In the same vein, it can be explained why the condition with artificially extended ILDs was less affected by the WDRC. In this condition, a substantial part of the SRM is provided by BEG at low and mid frequencies. At these frequencies the compression ratios were less than 1.5:1 (see above) and thus, the WDRC acted rather linearly over the entire input level range. Finally, the equal performance observed for the non-linear and linear amplification at a masker level of 60 dBSPL can be explained by the similarity in hearing aid gains that is provided by NAL-NL2 and NAL-RP at this level (figure 4.5, left panel) together with the above observation that the target signal, as well as a substantial part of the masker signal, was mainly in the linear range of the WDRC.

The observation that WDRC has a negligible effect at low (and mid) signal frequencies, at least for the rather typical hearing losses considered in this study (i.e., with a rather mild low-frequency and sloping hearing loss), supports the idea that extending ILDs towards low-frequencies may provide an interesting solution for improving BEG with hearing aids and thereby improving speech intelligibility in noise. However, more recent hearing aids tend to apply more linearly-acting WDRC due to the application of rather slow (or dual) time constants. Even though those devices may show less effect of compression they also will not provide the same effective gain (and thus audibility) that was applied here. Hence, the advantage of reduced distortions by the more recent WDRC schemes will be paid for a less effective compensation of audibility and thus, by a reduced spatial benefit.

4.4.3 Extra benefit provided by extended ILDs

Besides understanding the spatial benefit that can be provided by BEG when audibility is increased using non-linear amplification, the interest of the present study was also to investigate the spatial benefit that artificially extended ILDs to low frequencies can provide on top of natural ILDs. Therefore, speech intelligibility was measured with two spatially-separated, noise-vocoded speech-distractors, in which the distractors were either spatialized using extended ILDs or natural ILDs, and no ITDs. The results showed that extending ILDs to low frequencies provided an extra benefit in SRTs of about 3 dB when linear amplification was used at a masker level of 60 dBSPL. When non-linear amplification was used, SRTs with extended ILDs were better by about 2 dB at 50 dBSPL, 4 dB at 60 dBSPL, and by about 5 dB at 70 dBSPL. Similarly, listeners with NH also showed an extra benefit in SRTs by about 6 dB when ILDs were extended to low frequencies.

Interestingly, the additional benefit provided by extended ILDs with non-linear amplification increased with increasing masker level. This behavior can be explained by the non-linear characteristics of the WDRC that is described in the previous section. Since natural ILDs only exist at high frequencies, they are strongly affected by the applied compression. As a consequence, the rate of improvement in SRTs with an increase in masker level was rather small (see figure 4.4). This was different for the case of extended ILDs, which most likely provided most of their spatial benefit at low and mid frequencies, at which the applied WDRC acted rather linearly. Hence, it can be concluded that extending ILDs to low frequencies would be helpful in improving speech intelligibility in noise and can provide an extra benefit on top of what natural ILDs provide.

However, the noted extra benefit might be questionable when artificially extended ILDs are combined with natural ITDs, which may provide a substantial spatial advantage via noise cancellation at frequencies below about 1.5 kHz (e.g., Durlach, 1963). Hence, to investigate the interaction between the spatial benefit provided by natural ITDs and the one provided by extended ILDs to low frequencies, speech intelligibility was also measured in a condition with natural ILDs and ITDs. Providing ITDs in addition to ILDs improved SRTs by about 1 dB across all the different amplification methods for both NH and HI listeners. Consequently, the extra benefit provided by extended ILDs on top of natural ITDs in combination with natural ILDs, was about 2 dB when linear amplification was used. With non-linear amplification the extra benefit was about 1 dB at 50 dBSPL, 3 dB at 60 dBSPL and about 4 dB at 70 dBSPL. Similarly, NH listeners also showed an extra benefit in SRTs by about 4 dB. However, this extra benefit may increase or decrease when extended ILDs are combined with natural ITDs, an aspect that should be investigated in future. In the present setup, ITDs cannot be added to the extended ILDs because they were realized by "infinite" ILDs (i.e., each distractor is only applied to one ear). Since this is an artificial best case scenario, future studies should also consider more realistic methods for extending ILDs to low frequencies that can also be implemented in hearing devices. This may be done by applying directional processing with multiple microphones (e.g., Kates, 2008, pp.75-109), transposing high frequency speech cues to low frequencies (e.g., Robinson *et al.*, 2007), or using ITD information to generate ILDs (Moore *et al.*, 2016). Further, it would also be interesting to investigate the obtained spatial benefit using non-vocoded speech masker, although the existence of informational masking and its influence on SRTs in a spatially separated condition may be negligible (e.g., Rana and Buchholz, 2016).

4.4.4 NH vs HI listeners

The improvement of SRM and SRT with an increase in audibility can help reduce the performance gap between NH and HI listeners, though this depends on the way audibility is controlled. Glyde et al. (2015) reported an average difference in SRM of about 2.5 dB between NH and HI listeners when performance was compared at equal audibility levels (re. pure tone thresholds), which may be explained by age-related differences or other aspects of hearing impairment that are not related to a loss of audibility. In contrast, Rana and Buchholz (submitted) equalized audibility carefully in a group of NH and HI listeners across frequency by measuring speech detection thresholds (SDTs), and found that the difference in both SRTs and SRM between NH and HI listeners reduced to less than 0.5 dB. This suggests that, at least for their applied stimuli and methods, carefully controlling audibility across frequency can remove the performance gap between NH and HI listeners. In the present study, SRTs for the HI listeners at the highest amplification or audibility level (i.e., 70 dBSPL masker level and non-linear amplification) were about 4 dB higher than for NH listeners, which were tested at a more or less realistic masker levels of 60 dBSPL. This difference between groups decreased to a non-significant difference of less than 1 dB in SRM. This suggests that providing sufficient non-linear amplification can improve speech intelligibility performance in particular in spatially-separated conditions, and can almost restore SRM. However, the spatially-separated conditions also exhibit the lowest SRTs and are therefore more affected by reduced audibility. The gap of 4 dB in SRTs between the HI and NH groups may have been reduced if audibility would have been carefully controlled across frequency using a procedure similar to Rana and Buchholz (submitted). However, since prescriptive formulas such as NAL-NL2 have not been designed with the sole aim of restoring audibility, it is difficult to use a standard approach for investigating the effect of non-linear amplification for controlling audibility in the same systematic way as described by Rana and Buchholz (submitted). Future studies should further investigate how WDRC could be improved in hearing aids to optimize the benefit provided by BEG in noisy conditions.

4.5 Conclusion

The present study confirmed that non-linear amplification using WDRC improves speech intelligibility in noise at low signal levels by increasing the audibility of the incoming signals, but this improvement is counteracted by the compressive behavior at high signal levels. Similarly, the SRM provided by natural ILD and ITD cues is significantly improved at low but not at high signal levels. Moreover, artificially extending ILDs to low frequencies provided a substantial increase in SRM, due to BEG, on top of the SRM already provided by natural ILDs and ITDs. This extra benefit increased with increasing signal level, which was mainly due to the fact that most of the SRM was achieved at low and mid frequencies where the applied WDRC scheme behaved rather linearly (due to the rather mild hearing losses of the HI subjects at those frequencies). These results confirm that extending ILDs to low frequencies may be an interesting solution for hearing aids to improve speech intelligibility in noise. Future research should further investigate methods that optimize ILDs as well as non-linear amplification as a function of frequency such that hearing aids can provide the best benefit to HI listeners when communicating in noisy conditions.

4.6 Acknowledgements

The authors acknowledge the financial support of Macquarie University through an international research excellence (iMQRES) scholarship and the Australian National Health and Medical Research Council (NHMRC) via grant APP1056332.

Chapter 5: Bilateral versus unilateral cochlear implantation in adult listeners: speech-on-speech masking and multi-talker localization

Binaural hearing helps normal-hearing listeners localize sound sources and understand speech in noise. However, it is not fully understood how far this is the case for bilateral cochlear implant (CI) users. To determine the potential benefits of bilateral over unilateral CIs, speech comprehension thresholds (SCTs) were measured in 7 Japanese bilateral CI recipients using Helen test sentences (translated into Japanese) in a two-talker speech interferer presented from the front (co-located with the target speech), ipsilateral to the first-implanted ear (at +90° or -90°), and spatially symmetric at ±90°. Spatial release from masking (SRM) was calculated as the difference between co-located and spatially separated SCTs. Localization was assessed in the horizontal plane by presenting either male or female speech or both simultaneously. All measurements were performed bilaterally and unilaterally (with the first implanted ear) inside a loudspeaker array. Both SCTs and SRM were improved with bilateral CIs, demonstrating mean bilateral benefits of 7.5 dB in spatially asymmetric and 3 dB in spatially symmetric speech-mixture. Performance of localizing a single talker varied strongly between subjects but was clearly improved with bilateral over unilateral CIs with the mean localization error reduced by 27°. Surprisingly, adding a second, simultaneous talker had only a negligible effect on localization.

Note: Aspects of this work were presented at the 5th Joint Meeting, Acoustical Society of America and Acoustical Society of Japan (Honolulu, Hawaii, 2016).

5.1 Introduction

Listening to speech in the presence of noise is an integral part of our daily lives. In noisy situations, individuals with normal hearing take advantage of listening with two ears (rather than with only one ear), a phenomenon known as binaural hearing. Binaural hearing plays an important role in localizing sounds as well as in segregating target speech from distracting speech or noise (e.g., Cherry, 1953; Bronkhorst, 2000). When a distractor is located on one side of the head and is spatially separated from the target signal then the head shadow will typically improve the signal-to-noise ratio (SNR) in one ear (termed the "better-ear") and thereby improve speech intelligibility. However, the situation becomes more challenging when there is more than one distracting source. In such complex "cocktail-party" scenarios, the auditory system takes advantage of interaural time differences (ITDs) as well as interaural level differences (ILDs) to improve the "effective" SNR as well as to spatially attend to the signal of interest and to suppress interfering signals (Glyde *et al*., 2013c). For spectro-temporally fluctuating interferers such as speech the head shadow typically results in an SNR that continuously changes over time, frequency, and between ears. The auditory system can take advantage of these SNR variations either by glimpsing (Cooke, 2006) within each ear separately or by a process termed better-ear glimpsing (Brungart and Iyer, 2012). Additionally, the auditory system can utilize ITDs to improve the effective SNR by a process similar to the equalization-cancelation theory (Durlach, 1963). Independent of the involved mechanism, the improvement in effective SNR is commonly attributed to a spatial release from energetic masking, whereas the benefit related to spatial attention and stream segregation is commonly attributed to a spatial release from informational masking or a "perceived" (spatial) segregation of target and interferer signals, respectively (e.g. Kidd *et al*., 2007; Shinn-Cunningham, 2008). Spatial release from masking (SRM) in general is decreased in individuals with hearing loss as well as increased age, leading to difficulty in understanding speech in noise (e.g. Glyde *et al*., 2013a; Best *et al*., 2015).

The most common treatment for hearing loss is either a hearing aid or a cochlear implant (CI) depending on the severity of the hearing loss. SRM has been studied in hearing aid users (e.g., Glyde *et al*., 2013a; Glyde *et al*., 2013b) as well as in CI users (e.g., van Hoesel and Tyler, 2003; van Hoesel, 2012), utilizing single as well as multiple interferers. For the case that for a frontal target a single interferer is moved from the front to the side of the listener, a substantial

SRM can be observed in bilateral CI users (Müller *et al*., 2002; Tyler *et al*., 2002; van Hoesel and Tyler, 2003; Laszig *et al*., 2004; Buss *et al*., 2008). However, as mentioned above, this spatial benefit is mainly due to the better-ear effect and does not involve any sophisticated binaural processes. Nevertheless, two CIs are generally required to take full advantage of this effect in the real-world.

In contrast to this spatially asymmetric masker condition, which has been extensively studied in CI users, very little is known about the SRM in spatially symmetric conditions with fluctuating (speech) interferers (e.g., Schön *et al*., 2002). As mentioned above, in these conditions neither ear provides a consistent SNR advantage, but ITD as well as ILD cues can provide a SRM. Since ITD cues are basically not available to CI users (e.g., van Hoesel, 2012), they are also not expected to provide any contribution to SRM in these spatially symmetric speech-mixtures. However, ILDs are reasonably well preserved in the implanted ear and provide the main cue for localization in bilateral CI users (Grantham *et al*., 2008; Seeber and Fastl, 2008; Aronoff *et al*., 2012). Hence, ILDs may also provide SRM in these conditions, either by utilizing within-ear or across ear (i.e., better-ear) glimpsing or by perceptually segregating the target from the interfering talkers and thereby providing a spatial release from IM. Either way, in comparison to the healthy auditory system it is expected that the achieved SRM will be limited by the reduced spectral and temporal resolution of the implanted ear, mismatch in tonotopicity and loudness between ears, the limited dynamic range that is available in the implanted ear, and the distortion of the ILDs created by the wide-dynamic range compressors (and other adaptive processes) that operate independently at the left and right ear (e.g., Dillon, 2012, pp. 170-193). Moreover, it should be noted that any SRM that is achieved with bilateral CIs may be offset by an overall reduction in performance due to adding a second CI to a poor-performing ear; a phenomenon known as "binaural interference" that at least in hearing aid users can result in a negligible bilateral benefit or even in a detrimental effect (e.g. Walden and Walden, 2005; Mussoi and Bentler; 2017). This study investigated the SRM and the bilateral benefit achieved by CI users in a spatially symmetric as well as in a spatially asymmetric condition using an ongoing two-talker interferer. Whereas the spatially symmetric condition was of main interest here, the spatially asymmetric condition was included as a reference "best-case" condition that also allowed direct comparison to results reported in the literature. To derive the involved SRM, speech comprehension was measured in a co-located condition and compared to the two spatially separated conditions. To estimate the bilateral benefit, the performance achieved with a single CI fitted to the first implanted ear was compared to the performance achieved with CIs on both the ears.

ITDs and ILDs do not only provide a spatial advantage for understanding speech in noise, they are also the basic cues for localizing sounds. Localization of sounds is not only important for identifying the direction of a sound source, but also for participating in (multi-talker) conversations, being aware of the surroundings, and for protection from dangerous situations such as road accidents. Generally, adults with CIs on both the ears perform better in localization tasks than adults with a single CI (Tyler *et al*., 2007; Dunn *et al*., 2008; Mosnier *et al*., 2008). However, most studies have only investigated the localization of single sound sources, even though in real-life a listener is often surrounded by multiple sound sources or wants to participate in (or attend to) a conversation with more than one partner. In such cases, listeners need to segregate as well as to localize the different sound sources. Therefore, in the present study, the ability to localize a single talker as well as two spatially-separated simultaneous talkers of different gender was evaluated in bilateral CI users and compared to the performance achieved in a unilateral condition. It was assumed that the localization performance in the two-talker condition was not only affected by the listening mode (i.e., bilateral versus unilateral listening) but also by the participant's performance to segregate the two talkers. The first aspect is mainly affected by the availability and utilization of binaural (mainly ILD) cues. The latter aspect will also depend on the ability to analyze and utilize pitch cues as well as other talker difference cues.

To study how far localization and speech understanding in noise are limited by the same (spatial) auditory mechanisms (i.e., the auditory sensitivity to ITDs and ILDs and their supra-threshold utilization in spatial hearing), Rychtáriková *et al*. (2011) compared the performance of normal-hearing and hearing-impaired subjects in both tasks, but they did not find any significant correlation. This may not be surprising because speech intelligibility requires a continuous evaluation of the speech signal whereas localization may only require a few signal glimpses that provide sufficient spatial information, as illustrated by the Franssen effect (e.g., Hartmann and Rakerd, 1989). Moreover, localization in normal-hearing listeners may primarily depend on temporal fine-structure-based ITD cues at frequencies below about 1.5 kHz (e.g., Wightman and Kistler, 1992; Blauert, 1997; Macpherson and Middlebrooks, 2002) and speech intelligibility on ILD (or head shadow) cues that are mainly available at frequencies above about 1.5 kHz (e.g., Glyde *et al*., 2013b; Glyde *et al*., 2013c). This may be different in CI users who only have access to ILD cues and therefore rely on ILD cues for both localization and understanding speech in noise. However, except for Litovsky *et al*. (2009) most existing studies either did not find any significant correlations between localization and speech intelligibility performance in CI users (e.g., Litovsky and Misurelli, 2016) or did not report on any correlation results (e.g., Litovsky *et al*., 2004; van Hoesel and Tyler, 2003; Laszig *et al*., 2004). This may be explained by the limited statistical power of the studies due to the small number of subjects as well as the large variability of the data, or by the applied performance measures. In particular, the applied localization tasks only considered a single sound source, whereas the speech intelligibility tasks generally required segregating the target speech from interfering noise or talkers. In this regard, we hypothesize that the results from the two-talker localization task that was applied in the present study, which inherently required the spatial segregation of two simultaneous talkers, may provide stronger correlations with the results measured in the applied speech comprehension task. This is in particular expected in the spatially symmetric two-talker condition, because it is the only noise condition in which "true" binaural processing is expected to be involved.

Hence, the overall aims of this study were to (1) investigate the extent to which SRM can be observed in bilateral CI users in spatially symmetric speech mixtures, (2) investigate the extent to which localization performance is affected in CI users when a more realistic source-segregation task is included, (3) investigate if performance in localization and speech intelligibility in noise are correlated if both tasks involve segregating multiple spatially separated talkers, and (4) measure the benefit of providing two CIs over a single CI on the considered sound localization as well as speech comprehension tasks.

5.2 Methods

5.2.1 Participants

Seven adults (mean age of 62.9 years) with post-lingual deafness were recruited. All participants except of one had at least 6-months experience with their bilateral cochlear implants and had less than or equal to eight years of severe to profound hearing loss prior to bilateral cochlear implantation (Table 1). They all used both CIs regularly, scored more than 60% at +10 dB SNR in the CI-2004: Adult Everyday Sentence Test (Megumi *et al*., 2011) and spoke Japanese as their first language. None of the participants reported any cognitive impairment that would prevent or restrict participation in the audiological evaluations. This was confirmed for all subjects by administering a Japanese version of the Montreal cognitive assessment (MOCA) screening test (Nasreddine *et al*., 2005)**.** The mean and standard deviation of the scores obtained in the MOCA test was 25 ± 2.4. All participants were implanted with devices from Cochlear Limited and traveled from Japan to Sydney for testing. Biographical details of all participants are given in table 1. Written consent was obtained from all participants and ethical clearance was received from the Macquarie University Human Research Ethics Committees (Reference No**:** 5201401150).

TABLE 5.1: Biographical details of all seven participants. Note: numbers in bold represent the preferred ear; yr =year; mo=months

Subject code	Age (yr)	Sex	Age of first CI surgery (yr.mo)	Implant type	Speech processor	Age of second CI surgery (yr.mo)	Implant type	Speech processor	Cause of hearing loss
S1	34	F	31.10	CI422	CP 900	**32.1**	CI422	CP 900	Unknown
S2	68	M	**61.10**	CI24RE (CA)	CP 900	65.11	CI24RE (CA)	CP 900	Unknown
S3	62	M	**47.11**	CI24M	SPrint	60.3	CI422	CP 900	Unknown
S4	71	M	**63**	CI24R (CS)	CP 900	69.4	CI422	CP 900	Genetic
S5	78	M	**74.11**	CI422	CP 900	77.2	CI422	CP 900	Unknown
S6	60	F	**58.5**	CI422	CP 900	58.8	CI422	CP 900	Unknown
S7	67	F	**63.10**	CI24RE (CA)	N5 CP 800	64.8	CI422	N5 CP 800	Meniere's disease

5.2.2 Speech comprehension in noise

5.2.2.1 Stimuli

Speech comprehension thresholds (SCTs) were measured by asking the participants to answer brief questions in the presence of various background noises. The questions were taken from the English Helen sentence test (Ludvigsen, 1974) and were extended for this study. The resulting test contained 8 categories (colors, numbers, opposites, days of the week, addition and subtraction, multiplication and division, size comparison, and how many) of 20-51 questions each, providing 227 questions in total. The questions were all brief and easy to answer and included non-bibliographical questions from the above mentioned categories, such as: “what colour is a polar bear?”, “what day comes after Monday?”, or “what is two plus five?”. A speech comprehension task was applied here instead of a more common sentence (or word) recognition task because it was assumed to provide a more realistic performance measure (see Best *et al*., 2016a). This is because the task involves extraction of meaning as well as the formulation of a reply, which is very different to a simple speech recognition task, and may address higher level auditory functions that are relevant for communication in daily life (Kiesling *et al*., 2003). Additionally, it was particularly important for the spatially symmetric noise condition that the performance measure involved a substantial amount of glimpses, which is not really the case in a word test.

The Helen questions were translated into Japanese, spoken by a native Japanese female speaker with a mean fundamental frequency (±1 standard deviations) of 226 (± 49) Hz. The questions contained between 5 and 7 words and had a mean duration (±1 standard deviation) of 1.9 (±0.3) seconds. The recording took place in a double-walled audiological test booth using a Rhode NT-1A microphone connected to a desktop computer via a RME QuadMic microphone preamplifier and a RME Fireface USB soundcard. The questions were recorded and edited using Adobe Audition 5.5 software. All sentences were RMS level normalized in Matlab.
The speech comprehension test was administered inside a double-walled, acoustically treated audiological test booth containing an array of 16 Genelec 8020C loudspeakers that were used to present the different stimuli via a purpose-built Matlab interface. All loudspeakers were placed equidistantly on a circle with a radius of 1 m and connected to a desktop computer inside a control room via two RME ADI-8 DS analog-to-digital converters and a RME fireface USB sound card. The participants were wearing a lapel microphone connected to a high quality intercom to communicate with the experimenter inside the control room. The participants were seated in the center of the loudspeaker array facing the frontal (0°) loudspeaker with their ears at the height of the loudspeakers. The target questions were always presented from the frontal loudspeaker.

Three different noise conditions were created: (a) two speech discourses presented from the loudspeaker in front of the listener (the co-located condition), (b) two speech discourses presented from the side of the first-implanted ear (loudspeaker either at -90° or +90°), i.e., the ear with the unilateral CI (the spatially asymmetric condition), and (c) one speech discourse presented from the left side (loudspeaker at -90°) and one from the right side (loudspeaker at +90°) of the listener (the spatially symmetric condition). The two speech discourses were realized by a male and a female native Japanese talker reading different 5 minutes long popular children stories. The mean fundamental frequency (±1 standard deviation) of the male talker was 108 (±21) Hz and 239 (±41) Hz for the female talker and both talkers spoke with a rather slow speech rate of about 3.5 Hz, as calculated by the main maximum of their speech modulation spectrum. The discourses were recorded, processed, and RMS level normalized in the same way as the Helen questions described above. In every noise condition two SCTs were measured and averaged. Within the spatially symmetric noise condition, the first SCT was measured with the female distractor on the left and the male distractor on the right side of the listener and the second SCT with interchanged distractor locations. The different noise conditions are illustrated in figure 5.1.

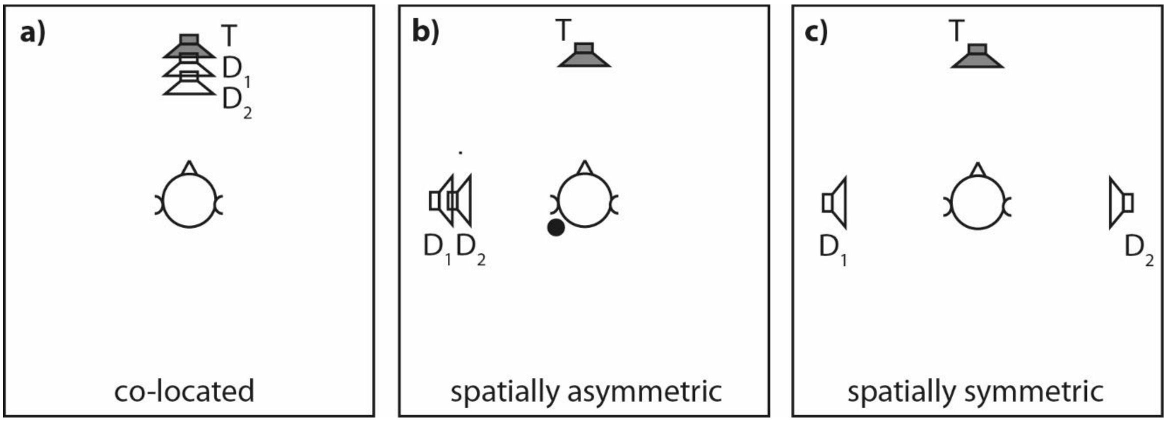

FIG. 5.1: The three different noise conditions applied in the speech comprehension test. The target source is indicated by the grey filled loudspeakers and noise sources are indicated by the open loudspeakers. Note that the spatially asymmetric condition shown in panel b) represents the case when the left ear is tested in the unilateral condition (as indicated by the dot) and needs to be mirrored for the right ear. T: Target speech; D: speech distractor.

5.2.2.2 Procedure

SCTs were adaptively measured with a 1-up 1-down procedure using up to 32 questions in the presence of the three different background noises described above. The questions were organized in 4 successive blocks of 8 questions each, with the order of the questions in each block randomized. Each block was generated by randomly selecting one question from each of the 8 categories. The different background noises were presented continuously at a constant intensity of 60 dBSPL for all participants except one. Participant S4 could not tolerate 60 dBSPL and hence the intensity was reduced to 50 dBSPL in 5 dB steps until the participant reported it to be comfortable.

Participants were instructed to first repeat the question before providing the answer verbally, but only the answers were considered in the adaptive procedure. The repetition of the question was mainly to monitor if subjects had problems with answering even though they repeated the question correctly. This could have indicated a cognitive problem and would have disqualified the listener from this study. For all listeners these errors were extremely rare. The SCT was defined as the signal-to-noise-ratio (SNR) at which the listener could answer the questions correctly at least 50 % of the time. To obtain SCTs, the adaptive track started with an SNR of 10 dB and the target level was varied with a step-size of 4 dB. When at least 5 questions were answered and an upward reversal occurred, the measurement phase started wherein the step-size was reduced to 2 dB. The SCT was then calculated as the mean value of all SNRs that were tested during the measurement phase. An adaptive track finished when all 32 questions were presented or when the standard error was below 1 dB and at least 17 questions were presented during the measurement phase. Each SCT was repeated once and the average was calculated as well as the test-retest accuracy. If the two SCTs differed by more than 4 dB, a third SCT was measured and the closest two SCTs were averaged. All the responses given by the subjects were translated by a Japanese translator and were scored by the experimenter.

All measurements were done with CIs in both ears (bilateral condition) as well as with only one CI in the first implanted ear (unilateral condition). For all subjects except S1 the first implanted ear was also the preferred ear. The speech comprehension testing was conducted over two consecutive days and each subject took about 1-1.5 hours per day. All conditions were measured once on the first day and repeated on the second day. On each day measurements were done in two blocks with bilateral conditions first and unilateral conditions after an extensive break. Within each block the conditions were randomized. Before any testing started the procedures were explained to the participants and practiced until they felt comfortable with it. Also, prior to any testing, loudness balancing was done for each participant to ensure that the perceived loudness was equal across the two ears. This was achieved by changing the volume and sensitivity control of the CIs. Since the study focused on measuring the spatial benefit provided by the auditory system, any adaptive noise suppression or directional features such as SCAN were turned off.

5.2.3 Sound localization

5.2.3.1 Stimuli

The subject's ability to localize speech was tested using 15 seconds long snippets of speech, which were extracted from the original 5 minutes of discourse used in the speech comprehension test and recorded with a male and a female native Japanese talker. The male and female talkers were either presented individually (one-talker condition) or simultaneously (two-talker condition) in randomized order at a constant intensity of 60 dBSPL. Two different genders were applied here to provide a unique identification of the two sources within the localization task. However, this inherently assessed also the accuracy of the subjects to identify the individual talkers (or voice genders).

The participants were seated in the center of a three-dimensional (spherical) loudspeaker array with a radius of 1.85 m located inside the anechoic chamber of the Australian Hearing Hub. The array consisted of 41 Tannoy V8 loudspeakers that were controlled by a desktop computer with an RME MADI PCI sound card located outside the anechoic chamber. The loudspeakers were connected to the sound card via two RME M-32 Digital-to-Analog converters and 11 Yamaha XM4180 amplifiers. Only 13 loudspeakers were used in this study (see highlighted buttons in figure 2): nine loudspeakers in the frontal horizontal plane (at azimuth angles from -90° to +90° with an angular spacing of 22.5°), three loudspeakers in the horizontal plane behind the subject (at -135°, +135°, and 180°), and one loudspeaker directly above the subject at an elevation angle of 90°.

5.2.3.2 Procedure

The participants were seated such that the head was in the centre of the loudspeaker array and facing the frontal loudspeaker (0°). They were asked to wear a small lapel microphone in order to be heard clearly by the experimenter who was seated outside the chamber with headphones on. The experimenter monitored participants via a webcam to ensure they maintained a fixed head position, and could talk to them via an intercom when required. All subjects were aware that they may hear either one or two talkers, but within each trial no information on the number of presented talkers was provided. The participant's task was to indicate both the direction of the talker(s) and the gender of the voice(s) on a handheld touch screen (iPad) showing the user interface given in figure 5.2.

Touching any of the 16 loudspeaker buttons or the subject button (to indicate the elevated loudspeaker) once turned the button red to indicate a female talker from the corresponding loudspeaker direction. Touching the button twice turned it blue to indicate a male talker. Touching the button three times turned the button half blue and half red to indicate a male and female talker from the same direction. Touching the button four times reset the button. The participants were able to respond as soon as they heard the stimuli and had additional 5 seconds before the next stimulus was presented. At any time they could move to the next condition by touching the start button.

The number of trials completed was shown by a counter in the left part of the user interface. The user interface was programmed in Matlab and controlled from the computer outside the anechoic chamber, which was connected to the iPad via WIFI and the Splashtop software. The interface was either controlled by fingers or a stylus pen as preferred by the participant. The experimental procedure was described to the participants at the beginning of the experiment by a written information sheet as well as by verbal communication with the experimenter and interpreter.

Prior to commencing the experiment each participant performed a number of familiarization trials until they and the experimenter were both confident that they understood the task. Each of the 13 possible directions (i.e., the loudspeaker locations described above and highlighted in figure 5.2) was tested five times in the single-talker condition and ten times in the two-talker condition, resulting in 130 two-talker items (65 male and 65 female) and 65 single-talker items (33 male and 32 female). Therefore, for each direction this resulted in 2-3 trials per gender in the single-talker condition and 5 trials per gender in the two-talker condition. Within the two-talker condition the talker directions were randomly combined, which resulted in 5 trials in which the two talkers were presented from the same (randomly chosen) loudspeaker. All 130 trials were randomized and measured in both a bilateral condition as well as in a unilateral condition.

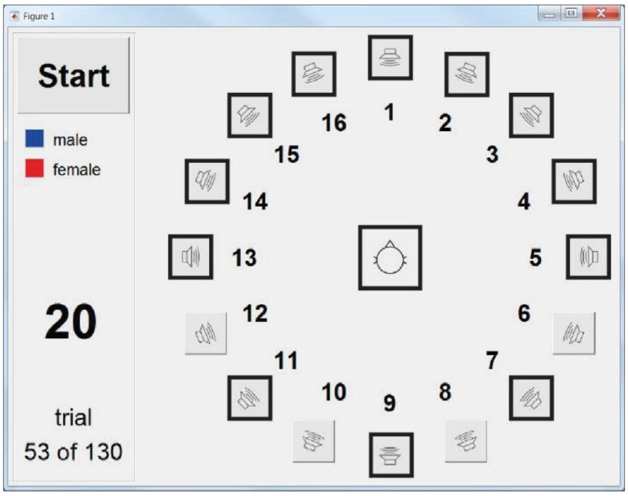

FIG. 5.2: Graphical user interface for the localization and voice gender identification experiment provided to the subjects on a handheld touch screen (iPad). The highlighted loudspeaker and listener buttons indicate the 13 source directions that were tested.

5.3 Results

5.3.1 Speech comprehension in noise

Figure 5.3 shows the individual SCTs obtained for all seven participants across the three different noise configurations in the unilateral (left panel) and bilateral (right panel) condition. Since the SCTs of subject 4 (left-pointing triangles), in particular in the bilateral conditions, were largely affected by fatigue effects, the corresponding data was considered unreliable and therefore excluded (as indicated by the round brackets) from the mean values shown in figure 5.3 (filled circles) as well as the subsequent speech comprehension data analysis. Many of the adaptive tracks of this subject showed variations of more than 20 dB which did not necessarily improve after taking extensive breaks.

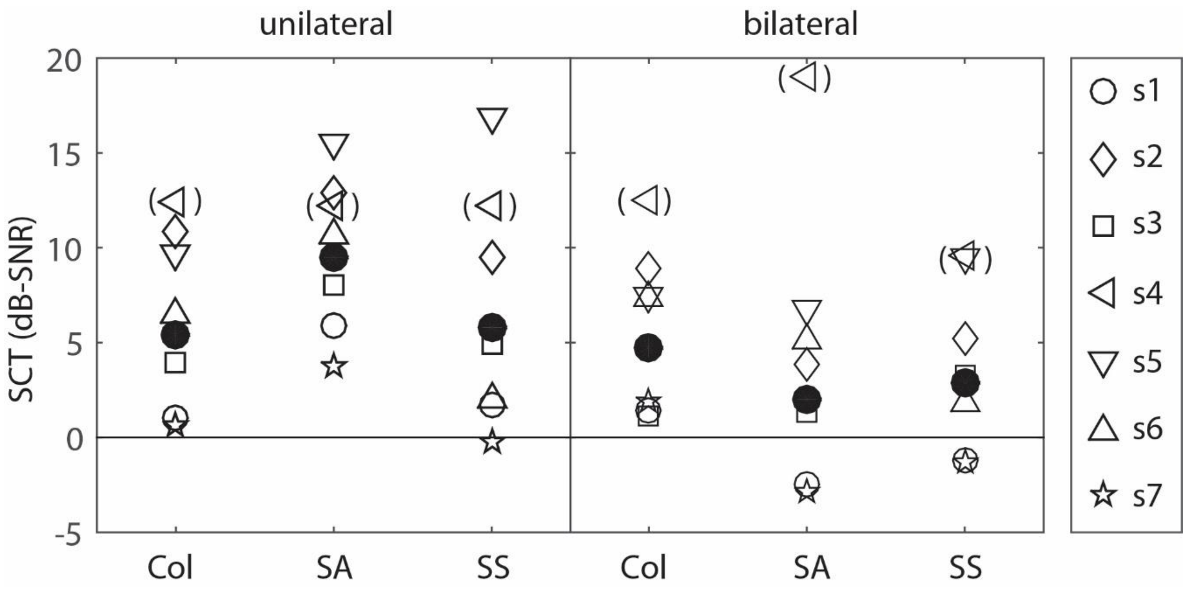

FIG. 5.3: Mean and individual SCTs obtained in the three different background noise configurations in the unilateral (left panel) and bilateral (right panel) condition. The data of subject 4 (left-pointing triangles in round brackets) is neither considered in the mean value (solid circles) nor in the subsequent statistical analysis. Col: co-located; SA: spatially asymmetric; SS: spatially symmetric.

As described in the methods section, the individual SCTs shown in figure 5.3 were averaged over two measurements. Additionally, test-retest variability was calculated by subtracting the second from the first SCT measurement and the resulting mean and intra-subject standard deviations are summarized in table 2. Mean values were within ±1 dB and a paired t-test revealed no significant differences ($p > 0.05$) for all conditions. The intra-subject standard deviation was less than 1.5 dB for the bilateral conditions and increased to up to 3 dB for the unilateral conditions. The intra-subject standard deviation in the bilateral conditions was very similar to one of the more common sentence tests, as for instance reported by Keidser *et al*. (2013) for the BKB sentence test, indicating a sufficient test-retest reliability of the applied speech comprehension test.

TABLE 5.2: Mean differences between the first and second SCT measurement and corresponding intra-subject standard deviation (STD).

	CIs	Co-located	Spatially asymmetric	Spatially symmetric
	Unilateral	-0.13 dB	0.52 dB	0.52 dB
	Bilateral	0.56 dB	0.97 dB	0.8 dB
	Unilateral	3.03 dB	2.73 dB	1.22 dB
	Bilateral	0.90 dB	1.48 dB	1.36 dB

Two-way, repeated measures ANOVA with noise configuration and listening mode as independent variables revealed a significant main effect of listening mode [F(1, 5) = 44.82, p = 0.001] but not of noise configuration [F(2,10) = 0.72, p = 0.51]. A significant interaction between noise configuration and listening mode was also observed [F (2, 10) = 21.18, p = <0.01]. Within the unilateral mode, a paired t-test with adjusted p-values (Holm, 1979) revealed a significant difference between the co-located and spatially asymmetric SCTs (p < 0.001), but neither between the co-located and spatially symmetric SCTs (p = 0.82) nor the spatially symmetric and spatially asymmetric SCTs (p = 0.06). Within the bilateral mode, no significant differences were found for any of the three SCT comparisons (p > 0.05).

To investigate if the relative performance between subjects was consistent across all noise conditions, an interclass correlation analysis in terms of consistency was applied using a two-way model (McGraw and Wong, 1996). The resulting intraclass correlation coefficient (ICC) indicated a consistent subject effect, with ICC = 0.87 for the unilateral conditions, ICC = 0.91 for the bilateral conditions, and ICC = 0.94 for the unilateral and bilateral conditions combined. Hence, subjects performed either consistently well or consistently poorly across all conditions.

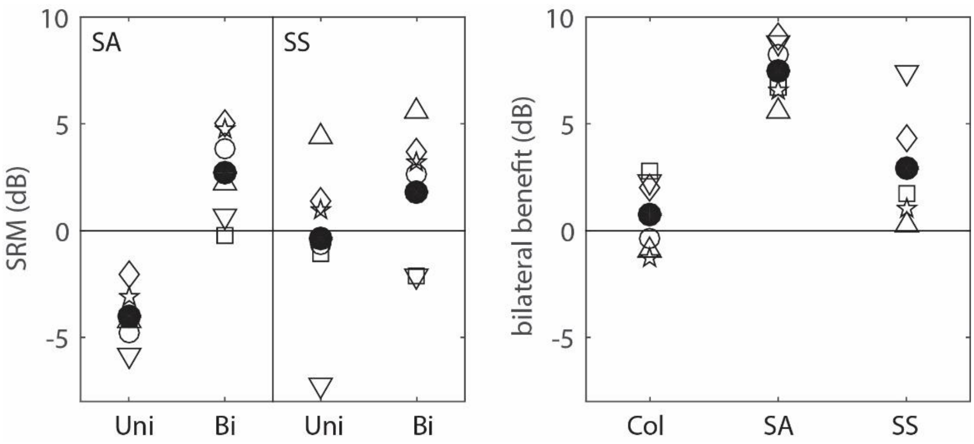

FIG. 5.4: Mean and individual SRM (left panel) as well as the bilateral benefit obtained in the three different background noise configurations (right panel). Data of subject 4 is not considered here. Col: co-located; SA: spatially asymmetric; SS: spatially symmetric; Uni: unilateral CIs; Bi: bilateral CIs.

In figure 5.4, a number of performance measures are shown that were derived from the individual SCT data shown in figure 5.3. The left panel shows the spatial benefit (or SRM) that is achieved when the two distracting talkers are spatially separated from the target speech and derived by subtracting the individual SCTs measured in the spatially separated condition(s) (figure 1b and 1c) from the individual SCTs measured in the co-located condition (figure 5.1a). To analyze the significance of the SRM observed in the different conditions, a paired t-test with adjusted p-values (Holm, 1979) was applied to compare the SCTs in the co-located and spatially-separated conditions. For the spatially asymmetric noise, a significant negative SRM (i.e., a disadvantage) of, on average, -4 dB was found in the unilateral condition ($p < 0.001$). In the bilateral condition the observed SRM of, on average, 2.7 dB was not significant ($p = 0.03$). For the spatially symmetric noise, the SRM was neither significant in the unilateral nor in the bilateral condition ($p > 0.05$).

The right panel of figure 5.4 shows the bilateral benefit for all three noise conditions, which was calculated by subtracting the individual SCTs measured in the bilateral conditions from the corresponding SCTs measured in the unilateral conditions. The bilateral benefit directly quantifies the advantage in speech comprehension that is provided by two CIs over one CI. To analyze the significance of the bilateral benefit, a paired t-test with adjusted p-values (Holm, 1979) was applied to compare bilateral with unilateral SCTs. A significant bilateral benefit of, on average, 7.5 dB was found for the spatially asymmetric noise condition ($p < 0.01$). Neither the bilateral benefit in the spatially symmetric noise condition of, on average, 3 dB, nor the bilateral benefit in the co-located condition of, on average, 0.8 dB was significant ($p > 0.01$).

5.3.2 Sound localization

Most subjects were very accurate in identifying the number of talkers (i.e., one or two) that were presented within each trial (error rate < 1%). Only subjects s2, s3, and s5 wrongly estimated the number of talkers in the beginning of their first experiment (bilateral condition), which was most likely due to initial problems with handling the user interface. These trials were disregarded in the following analysis. The gender of the talker(s) was correctly identified by all subjects in at least 98.5 % of the trials, suggesting that all subjects were able to utilize the voice-gender or voice identity within the two-talker localization task. In the unilateral condition, all subjects except subject s1 (with 10.9 seconds) spent the entire available time of 20 seconds to provide their responses. In the bilateral condition, subject s1, s3, and s6 had significantly shorter average response times than in their unilateral condition with 6.5, 10.6, and 13.3 seconds. This indicates that at least for these three subjects, providing a second CI made the localization task easier. Only subject s1 and s7 managed to correctly identify when a talker was presented from the elevated loudspeaker with a sufficient reliability, i.e., achieving sensitivity values between $d' = 1.3$ and $d' = 2.5$. However, they were only able to do that in the bilateral condition.

Horizontal localization performance varied strongly between subjects but always improved when a second CI was applied, indicating a clear bilateral benefit in localization. Example response patterns for two subjects (s1 and s7) are shown in figure 5 for the unilateral (left panels) as well as the bilateral (right panels) condition. Response directions are plotted against the presentation directions for each of the 130 trials indicated by a circle. To avoid circles masking each other due to the discretization of the presented and responded directions (the angular resolution was 22.5°), the presented directions were shifted horizontally in figure 5 within the white-and-grey shaded areas. In the unilateral condition subject s1 showed a rather large variation in the responses but somehow managed to utilize the entire horizontal plane. In contrast, subject s7 localized all the presented sources in the direction of the left ear where the unilateral CI was fitted. Localization performance clearly improved in both subjects when a second CI was fitted. However, the response pattern across the horizontal plane was very different between subjects. Whereas subject s1 still showed rather poor localization performance for directions to the side and back, subject s7 was rather accurate across the entire horizontal plane. The data of the other subjects followed either type of pattern and are not shown here due to space limitations.

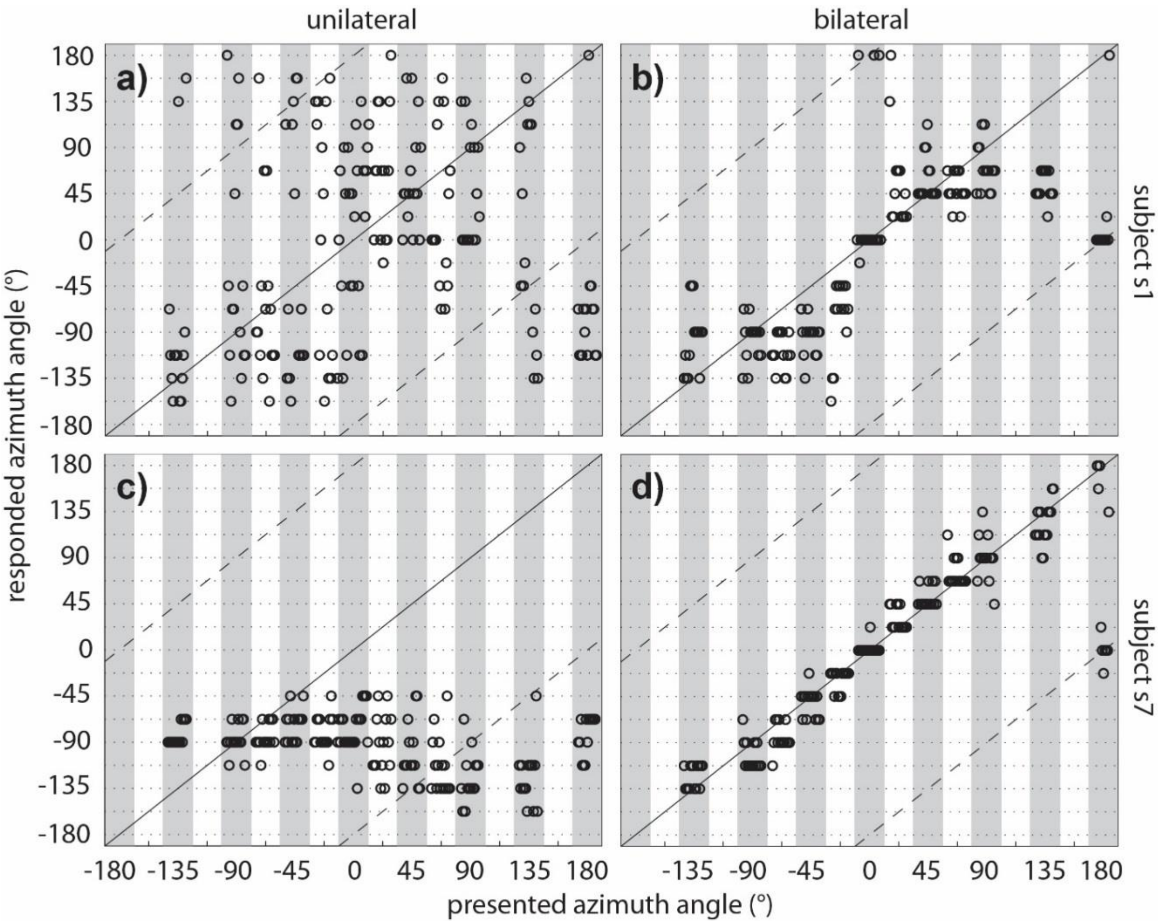

FIG. 5.5: Horizontal localization performance for two example subjects (s1 and s7) for the unilateral (left panels) as well as the bilateral condition (right panels). The circles indicate individual trials and are shifted horizontally within the grey-and-white shaded area for clarity.

Even though front-back confusions were apparent in all subjects' responses (see responses around the dashed diagonal lines in figure 5.5), the large overall RMS error observed in particular in the unilateral conditions made it impossible to reliably segregate front-back confusions from actual localization errors. Hence, front-back confusions were not further considered here. To quantify the localization performance in the horizontal plane the root-mean-squared (RMS) localization error was therefore applied as given by:

$$E_{RMS} = \sqrt{\frac{1}{N}\sum_{i-1}^{N}\left(\arcsin\left(\sin\left(\eta_i\right)\right) - \arcsin\left(\sin\left(\phi_i\right)\right)\right)^2} \qquad (1)$$

with arcsin referring to the arcus sine operation, N the total number of considered items (i.e., individual talker directions), η_i the presented azimuth angle in radians of item i, and ϕ_i the corresponding responded azimuth angle. The number of the considered items (or scoring units) N depended on weather one or two simultaneous talkers were presented and whether male and female talkers were evaluated separately or together. The RMS error disregards front-back confusions by “folding” the directions behind the participants to the front. The individual RMS errors are shown in the left panel of figure 5.6 averaged across the male and female talker but separately for the one-talker and two-talker condition. The unilateral results are indicated by open symbols and the bilateral results by filled symbols.

The average RMS error (±1 standard deviation) was 4.2° (±8.5°) lower for the one-talker than for the two- talker condition and, even though not shown here, 4.2° (±8.5°) lower for the male than for the female talker. However, a paired t-test neither found the effect of the number of talkers significant nor the effect of their gender ($p > 0.05$). Hence, in the following analysis, the RMS error was combined across the male and female talker as well as across the one-talker and two-talker condition.

The RMS error is improved by, on average, 27° when a second implant is fitted, but this improvement (or bilateral benefit) varied strongly between subjects, i.e., between 9° for subject s3 and 68° for subject s7. The variation of the localization performance between subjects was significantly smaller in the bilateral condition than in the unilateral condition, with an inter-subject standard deviation of 9.4° in the unilateral and 19° in the bilateral condition. This was largely due to subject s1, s6, and s7, who received a much larger bilateral benefit of, on average, 46° than the other 4 subjects with an average benefit of only 12.9°.

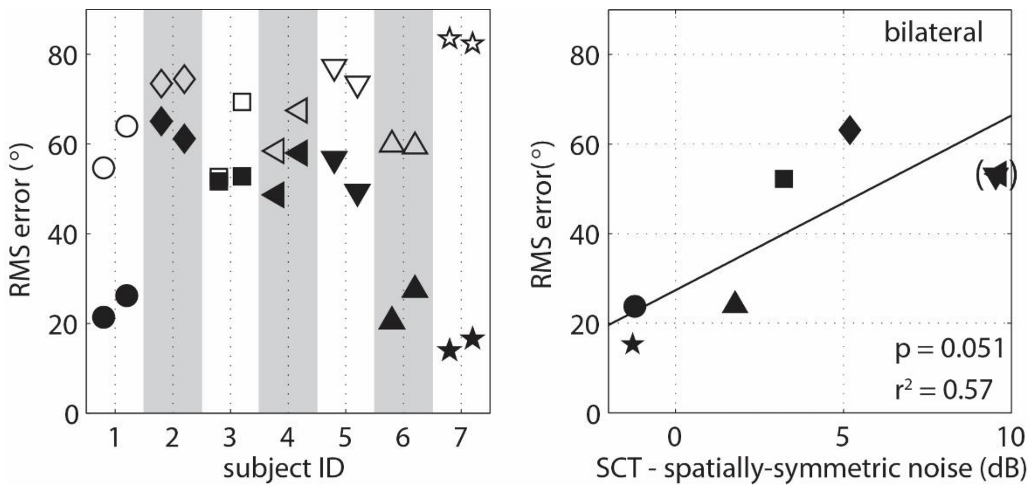

FIG. 5.6: In the left panel, individual RMS localization errors are shown for the unilateral (open symbols) and the bilateral (filled symbols) condition, with the data for the single-talker condition plotted on the left and for the two-talker condition on the right of the dashed lines. In each condition, the RMS errors were averaged across the male and female voices. In the right panel, the individual RMS localization errors in the bilateral condition are plotted against the corresponding SCTs measured in the bilateral condition for the spatially symmetric two-talker noise. The linear regression line is shown by the solid line and excluded subject s4 as indicated by the round brackets.

To get an indication whether the individual performances in the speech comprehension and localization tasks are limited by similar processes within the implanted ear, the individual RMS localization errors were correlated with the individual SCTs measured in the three different noise conditions. An almost significant correlation ($r^2 = 0.57$, $p = 0.051$) was only found in the bilateral condition for the spatially symmetric noise (see figure 5.6, right panel). However, the correlation became significant ($r^2 = 0.65$, $p = 0.029$) when subject s4 was included, who in this specific condition showed reliable SCTs (i.e., the adaptive tracks showed very small variations across test trials with a standard error of about 1 dB). Multiple comparisons were not considered in the correlation analysis.

5.4 Discussion

5.4.1 Speech comprehension in noise

The performance in speech comprehension varied strongly between subjects with an average inter-subject standard deviation of about 5.3 dB in the bilateral and 4.7 dB in the unilateral condition (see figure 5.3). In contrast, the intra-subject standard deviation between repeated measures was on average only 1.3 dB in the bilateral conditions and 2.3 dB in the unilateral conditions. Hence, the differences between subjects were strongly influenced by the differences in individual performance. These differences could be attributed to many factors such as the age of implantation, chronological age, the amount of bilateral exposure, talker variability etc. (e.g. Tamati *et al.,* 2017). As a consequence, the inter-subject standard deviation of the performance measures that were based on within-subject differences were strongly reduced (see figure 5.4), with an average inter-subject standard deviation of 2.6 dB in SRM and 3.3 dB in bilateral benefit. This was further confirmed by the intraclass correlation analysis described in the results section, which revealed that the large inter-subject variability was due to a consistent subject effect, i.e., subjects performed either consistently well or consistently poorly across all noise conditions as well as listening modes (i.e., unilateral versus bilateral). The differences in SCTs between the co-located and spatially separated conditions (i.e., the SRM) as well as the differences between the bilateral and unilateral conditions (i.e., the bilateral benefit) are discussed in the following sections.

5.4.1.1 Spatial release from masking

Spatial release from masking (SRM) is here defined as the individual SCTs measured in the co-located two-talker noise conditions minus the SCTs measured in either of the two spatially separated two-talker noise conditions. In the case that the two-talker noise was presented only from the side of the listener with the first-implanted ear, i.e. the spatially asymmetric condition (figure 5.1b), a non-significant SRM of, on average, 2.7 dB was observed in the bilateral condition and a significant SRM, on average, of -4 dB in the unilateral condition. The SRM observed in the bilateral condition was due to head shadow improving the SNR at the ear contralateral to the interferer, i.e., the “better-ear”. In the unilateral condition, the CI was fitted only to the ear ipsilateral to the interferer, i.e., the ear with the poorer SNR. The (non-significant) SRM of 2.7 dB that was observed in the bilateral condition was much smaller than the SRM of about 7-9 dB that is typically observed in such spatially asymmetric conditions in normal-hearing listeners, which to some extent, depends on the type of masker that is applied (e.g., Misurelli and Litovsky, 2012; Hawley *et al*., 2004; Arbogast *et al*., 2002). The observed spatial benefit is in agreement with studies that investigated SRM in similar spatial configurations in bilateral adult CI users. Loizou *et al*. (2009) found an SRM of about 4 dB when a three-talker interferer is used (from 30, 60 and 90 degree), van Hoesel and Tyler (2003) found an SRM of about 4 dB when a speech-shaped noise masker is used (from ±90 degree.), and Kokkinakis and Pak (2014) found an SRM of about 2-4 dB when a four-talker babble is used (±90 degree).

In the case that one distracting talker was presented from the left and one from the right side of the listener, i.e., the spatially symmetric condition (figure 5.1c), an average SRM of 1.8 dB was observed in the bilateral condition and -0.4 dB in the unilateral condition. However, both effects were not significant ($p > 0.05$). Considering the small number of subjects that were tested and the rather significant spread of the individual SRM data both in the unilateral (i.e., from -7.2 Db to 4.4 dB) and bilateral condition (i.e., from -2.1 dB to 5.6 dB), it may still be that at least some bilateral CI users were either able to take advantage of the fluctuating SNR, either by within ear glimpsing or BEG (i.e., providing a release from energetic masking), or by utilizing ILDs to perceptually segregate the target talker from the interfering talkers (i.e., proving a release from IM). Either way, the resulting SRM is much smaller than the SRM of more than 7 dB that is observed in normal-hearing listeners in similar conditions (e.g., Brungart *et al*., 2012; Glyde *et al*., 2013b). To the best knowledge of the authors, SRM in a spatially symmetric two-talker noise has not been investigated before in bilateral adult CI users. However, there are a few relevant studies in children, which also did not find any significant SRM but showed smaller average values than observed here of between -2 and 1 dB (Misurelli and Litovsky, 2012; Misurelli and Litovsky, 2015). However, these results cannot be directly compared with the present study, because the detailed procedures and stimuli were very different and moreover, the SRM that is observed in, at least, normal-hearing children is generally smaller than in adults (Cameron *et al*., 2011). Since bilateral CI recipients in the spatially-symmetric condition need to integrate information across ears, this condition may be particularly sensitive to differences in both the loudness matching and the frequency mapping between ears (due to differences in electrode placement) as well as the distortion of the relevant ILD cues by the non-linear and independently operating devices. These differences may at least partly explain the large inter-subject variability of the SRM observed in the spatially-symmetric noise condition. Future studies should, therefore, investigate how these (and other) factors influence SRM in speech-mixtures and find solutions to maximize it.

5.4.1.2 Bilateral benefit

The bilateral benefit was calculated here as the individual SCTs measured in the unilateral condition minus the SCTs measured in the corresponding bilateral condition, which was separately derived for all three background noises (see figure 5.4, right panel). A significant bilateral benefit of, on average, 7.5 dB was observed in the spatially asymmetric noise condition and a non-significant bilateral benefit of, on average, 3 dB in the spatially symmetric noise condition. The small bilateral benefit seen in the co-located condition of about 0.8 dB was also not significant. The large bilateral benefit of around 7.5 dB seen in the spatially asymmetric condition is in general agreement with other studies with bilateral CI users, who, dependent on the type of masker, reported bilateral benefits of 5-7 dB (e.g. Schleich *et al*., 2004; Litovsky *et al*., 2006; Kokkinakis and Pak, 2014). The observed benefit mainly reflects the SNR difference between the ear ipsilateral to the masker (the “poor ear”) and the ear contralateral to the masker (the “better-ear”). In the unilateral condition, the CI users had only access to the “poor ear”, whereas in the bilateral condition the CI users had also access to the “better-ear”.

The mean bilateral benefit of 3 dB that was observed in the spatially symmetric two-talker noise condition was either due to the increased availability of glimpses (within ear or between ear) or due to improved perceptual segregation of the target talker and the interfering talkers. Even though this effect was non-significant ($p=0.04$), the individual data showed a substantial range of benefits from 0 to 7 dB. To the best knowledge of the authors, no other study exists that explicitly measured the bilateral benefit in adult bilateral CI users using such spatially-symmetric noise condition. Only Misurelli and Litovsky (2012 and 2015) measured speech intelligibility in children with bilateral CIs in a very similar noise condition, but they did not include a unilateral condition. Even though the underlying auditory mechanisms are not fully understood, the present results suggest that at least some CI users can receive a substantial bilateral benefit in spatially symmetric speech mixtures where neither ear provides a consistent SNR advantage. Future studies should, therefore, further investigate the underlying auditory mechanisms and find solutions that optimize the resulting benefit in bilateral CI users.

The small but not significant bilateral benefit of about 0.8 dB that was observed in the co-located two-talker noise condition is consistent with other CI studies that compared bilateral to unilateral performance in the preferred ear (e.g., Laske *et al*., 2009; van Hoesel and Tyler, 2003; Loizou *et al*., 2009) as well as with studies with normal-hearing listeners (Mac Keith and Coles, 1971; Bronkhorst and Plomp, 1989; Hawley *et al*., 2004).

5.4.2 Sound localization

The main novelty of the applied localization test was the inclusion of a two-talker localization task, which was expected to improve the ecological validity of the results. In this regard, it was surprising that the subjects' localization performance in the two-talker condition was only slightly (on average 4.2°) worse than in the one-talker condition. Since the subjects in the two-talker condition had not only to localize but also to segregate the two simultaneous talkers, it was expected that the performance would be substantially worse than in the single-talker condition. However, the observed similarity in the RMS errors highlights that none of the subjects had difficulties in segregating the two simultaneous talkers, which may be an interesting observation on its own. However, the similarity in performance may have been due to the rather long speech segments of 15 seconds that were provided in this task, which may have allowed the subjects to localize one talker at a time. The results may have been different if shorter speech stimuli were applied, such as the sentences used in the speech comprehension task. Moreover, the rather large angular separation of 22.5° of the applied loudspeakers may have further contributed to the negligible differences in performance. Given the not significant effect of number of talkers (as well as the gender of the talker) only the RMS error averaged over number and gender of talkers was further analyzed.

Localization performance in the horizontal plane was found to be consistently poor for all subjects in the unilateral condition with a mean RMS error of 68° and an inter-subject standard deviation of 9.4°. Considering other studies that utilized the entire horizontal plane in the localization experiment, i.e. that applied a loudspeaker array span of 360°, the observed RMS error is much lower than commonly reported with mean values of around 90° (e.g., Neuman *et al*., 2007; Laszig, *et al*., 2004). Similar low RMS errors of around 50°-70° have only been reported for loudspeaker array spans that are 180° or less (e.g., van Hoesel, 2012). The difference may be due to the rather long speech signals of up to 15 seconds duration that were applied here, which gave the listeners much more time to make their decisions than given in previous studies, or due to the rather wide loudspeaker spacing of 22.5°. Considering the example localization pattern shown in figure 5.6 (left panels), which were representative for most subjects, it can be deduced that the large RMS errors seen in the unilateral condition were either from subjects providing rather random localization responses (panel a) or from subjects mainly localizing sources at their unilateral ear (panel c).

The overall performance for all subjects improved substantially in the bilateral condition with a resulting mean RMS error of 40.7° and a mean bilateral benefit of 27.7°. However, at the same time, the inter-subject standard deviation increased from 9.4° to 19°, indicating that some subjects received larger bilateral benefits than others. Considering the individual RMS errors shown in figure 5.6 (left panel) it can be deduced that the increased inter-subject variation in the bilateral condition was mainly due to three subjects (s1, s6, and s7) that received a much larger bilateral benefit of, on average, 46.3° than the other four subjects with 12.9°. The resulting mean RMS errors for these two groups were 21° and 55.5°. The RMS error and the individual differences are in good agreement with Laszig *et al*. (2004), who found for a loudspeaker array span of 360° an average RMS error of about 50°. However, the first group showed RMS errors that have only been reported for loudspeaker array spans of less than 90° (e.g., van Hoesel, 2012). Nevertheless, the present study is in general agreement with the existing literature in that horizontal localization performance is substantially improved when a second CI is provided. Considering the example localization pattern shown in figure 5.5 (right panels), some subjects were able to equally well localize sound sources over the entire horizontal plane (panel d) and others showed rather good performance in the front of the listener that then deteriorated towards the side of the listener (panel b). The behavior seen in figure 5.5b can most likely be explained by the broadband ILD function, which exhibits larger changes with direction for frontal sources than for lateral sources (van Hoesel, 2004). The behavior seen in figure 5.5d may suggest that subjects are able to utilize frequency-specific ILD cues for localization (e.g., van Hoesel, 2012).

Even though it was not the main goal of this study, the applied localization task also showed that subjects were able to reliably determine the number of talkers, with an error rate of less than 1 %, and also to identify their gender, with an error rate of less than 1.5 %. Existing studies on gender identification typically found by far higher error rates in CI users of between 5 % and 56 % (e.g., Fu *et al*., 2004; Kovacic and Balaban, 2009; Massida *et al*., 2013). However, the applied listening tasks were very different, stimuli durations were much shorter, the talker differences (e.g., fundamental frequency, formant frequencies, spectrum) were smaller and often manipulated using speech transformation software, and the number of options (i.e., talkers) was larger. In particular, the fact that only two different talkers were applied here may have allowed the CI users to utilize cues that are not directly related to talker gender identification.

Finally, it should be mentioned that in the bilateral condition two out of the seven subjects were able to rather reliably detect when speech was presented from the (elevated) loudspeaker above the subject. Since the spectral cues that are relevant for elevation perception are very subtle (e.g., Blauert, 1997) it is rather unlikely that these subjects were actually able to localize the elevated source. It is more likely that the subjects interpreted the ambiguous ILD cues together with some (learned) broad spectral or level cues as an indicator that the source could not come from the horizontal plane, leaving the loudspeaker above the only remaining option. The latter is in general agreement with Majdak *et al*. (2011) who found that, within the vertical plane, CI users were only able to identify the correct hemifield (and not localize within a hemifield) which was mainly achieved by using level rather than spectral cues. Even though participants were not allowed to move their head, small head movements may have also helped localizing the elevated source.

5.4.3 Sound localization versus speech comprehension

As already mentioned in the introduction, existing studies with normal-hearing and hearing-impaired subjects could not find any significant correlation between the subjects' individual performance in localization and speech intelligibility in noise (Rychtáriková *et al*., 2011). For bilateral CI users, Litovsky *et al*. (2009) is the only study that found significant correlations between both in adults. In children with bilateral CIs, Litovsky and Misurelli (2016) could not find any significant correlations. Most other studies that measured both in bilateral CI users did not report any correlation results (van Hoesel and Tyler, 2003; Litovsky *et al*., 2004; Laszig *et al*., 2004; Smulders *et al*., 2016; Mosnier *et al*., 2008). In the present study, a significant correlation was observed only between localization and speech comprehension performance in the bilateral condition for the spatially symmetric two-talker interferer (and only if all 7 subjects were considered). Even though this correlation was slightly higher when only the localization performance in the two-talker condition was considered, the correlation was very similar in the single-talker condition. This similarity is rather surprising, because it was expected that the auditory stream segregation process involved in the two-talker condition, which is not required for localizing a single talker in quiet, would address auditory processes that are more similar to the processes involved in the speech comprehension task and thus, performance should also be more similar. However, this was just a direct consequence of the very small differences that were observed in the localization results for the single-talker and two-talker conditions. It is unclear why the localization performance was so similar, but it may have been due to the rather long speech stimuli of 15 seconds that were used, which may have made the stream segregation task too easy for the test subjects. Future studies should therefore investigate if making the stream segregation task more difficult results in a stronger correlation between the individual performance in speech intelligibility in noise and localization, for instance by reducing the stimulus duration to a duration similar to the sentences used in the speech task (i.e. around 1.9 seconds) or by adding background noise. Such shorter speech stimuli have been utilized in a number of studies although only for single talker localization (Neuman *et al*., 2007; Laszig *et al*., 2004).

5.5 Conclusions

The present study confirmed that a second CI provides a clear (bilateral) benefit over only one device for the understanding of speech in noise as well as the localization of sounds. The largest bilateral advantage in speech intelligibility of about 7.5 dB was observed when a two-talker distractor was presented from the side of the listener that was ipsilateral to the unilateral CI and target speech was presented from the front. In this case, the bilateral benefit was mainly due to the better (long-term) SNR provided by head shadow at the contralateral ear, and is in good agreement with existing literature. On the one hand, this agreement confirms the validity of the applied comprehension task, which is different from the more commonly applied sentence/word recall task, but on the other hand, it suggests that the realism added by the comprehension task did not affect the outcomes. However, this may be different for other noise conditions or for other types of comprehension measures (e.g., Best *et al*., 2016b), which should be further evaluated in the future.

A novel finding was the bilateral advantage (of, on average, 3 dB) that was seen with a spatially symmetric two-talker interferer, where one talker was presented from the left and another talker from the right side of the listener. Even though the effect was not significant, some subjects showed a bilateral benefit of up to 7 dB. Since the long-term SNR at the two ears was identical in this condition, the benefit provided by the second CI was either due to BEG or ILD cues providing a perceived (spatial) separation of the different talkers. Moreover, a clear bilateral benefit was also observed in the localization of a single talker as well as of two simultaneous talkers. Thereby the localization performance in the two-talker condition was only slightly (and not significantly) poorer than in the single-talker condition, suggesting that the stream segregation processes that were inherently involved in the two-talker condition did not play a major role. This rather surprising result may have been due to some methodological details that should be improved in future studies. Finally, an almost significant correlation was found between the performance in sound localization and speech intelligibility in a spatially symmetric two-talker noise condition with bilateral CIs.

Even though the performance in localization as well as speech intelligibility in noise was improved by a second CI, the overall performance, as well as the benefit provided by the second device, was still significantly poorer than generally observed in normal-hearing listeners. There are a number of well-known reasons for this discrepancy, including differences in the spectral maps between the left and right ear (due to differences in electrode placement), loudness differences between ears, limited temporal and spectral resolution, insufficient temporal fine-structure coding of ITDs as well as pitch, and independently operating devices. Even though these issues may be largely resolved by adequate technologies, implantation techniques, and fitting procedures, the subsequent neural auditory pathway may also be different between ears and further disrupt binaural processing, which may be improved by adequate training procedures.

Future studies should further investigate the SRM observed in spatially symmetric speech-on-speech masking and develop methods that maximize its benefit in CI users. Moreover, it is unclear how far the current findings, as well as the findings reported in the existing literature, reflect the performance experienced in real-life. Therefore, future measures of localization and speech intelligibility need to consider more realistic environments, including room reverberation, background noise, and multiple talkers at different distances and head orientations. Moreover, the applied tasks need to be more realistic, address more cognitive processes, and provide visual cues as well as context information. For instance, speech comprehension may be measured by asking subjects to answer questions while listening to monologues or dialogues in a noisy environment (e.g., Best *et al*., 2016 a); and localization in quiet may be replaced by measures of auditory spatial awareness (e.g., Brungart *et al*., 2014; Weller *et al*., 2016). Finally, to draw stronger conclusions about the investigated processes within the implanted ears, the statistical power needs to be improved by applying more subjects.

5.6 Acknowledgements

The authors acknowledge the financial support of Macquarie University through an international research excellence (iMQRES) scholarship and the Australian National Health and Medical Research Council (NHMRC) via grant APP1056332. The National Acoustics Laboratories and Cochlear Ltd. are part of the Australian Hearing Hub, an initiative of Macquarie University that brings together Australia’s leading hearing and healthcare organizations to collaborate on research projects. The Authors declare that there is no conflict of interest.

Chapter 6: General Summary and discussion

Understanding speech in noise has always been a challenging issue for HI listeners even with their hearing aids or CIs. This thesis presents a series of experiments that were conducted to understand and enhance BEG in HI listeners with the ultimate goal to improve their speech intelligibility performance in noise. Thereby, one of the main problems that are encountered is their limited dynamic range at high frequencies where natural ILD cues are available for BEG. To overcome this problem, an alternative approach was also investigated which extended the high-frequency ILD cues to low frequencies (chapter 2, 3 & 4).

In the first study (chapter 2), the ability of NH listeners and HI listeners was investigated to utilize artificially extended ILDs in a speech recognition task. The results indicated that both NH and HI listeners can utilize artificially extended ILDs, but the performance observed in HI listeners was poorer than in NH listeners. Since the distractors were presented at higher sensation level for NH listeners compared to HI listeners, the difference in performance was most likely attributed to the difference in audibility of the provided signals. Based on this experiment it was predicted that audibility might be an important factor to be considered if one aims to investigate or improve BEG in HI listeners. In order to test this prediction, in the second study (chapter 3), both HI listeners and NH listeners were tested at the same audibility levels. Importantly, the audibility was carefully controlled across frequencies for NH and HI listeners by providing gain equivalent to the measured SDTs across nine bands of two critical bandwidth each. In a speech recognition task using audibility equated stimuli, the ability of listeners in taking advantage of ILDs across frequency was assessed at different audibility levels. The distractors were presented at four different sensation levels relative to the SRTs in quiet. SNR required to achieve 50 % intelligibly of target sentences was measured at each sensation level. Results indicated that if the loss of audibility is carefully compensated for, HI listeners can take advantage of artificially extended ILDs as well as NH listeners. No significant difference in performance was noted between HI and NH listeners at all sensation levels across all frequencies confirming the important role played by audibility. Further, the trend of improvement in SRTs and SRM with an increase in audibility was also noticed across all frequencies. However, due to the smaller dynamic range especially at high frequencies, not all HI listeners could perform at all sensation levels due to loudness tolerance issues highlighting the limited benefit that can be achieved with BEG at high frequencies.

It is known that a linear increase in amplification helps to preserve natural ILDs but may not accommodate a large range of sounds within the reduced hearing dynamic range of HI listeners. The utilization of non-linear amplification using WDRC does help in amplifying soft sounds and attenuating loud sounds, hence, resolving the issue of loudness tolerance at high levels and the inaudibility at softer levels. Consequently, this helps to accommodate the large range of sounds within the dynamic range of an HI listener. However, WDRC can distort the temporal and spectral behavior of the incoming signals and may have some deleterious effect on ILDs. As a result any increase in audibility may or may not lead to an improvement in speech perception. Therefore, the third study (chapter 4) aimed at investigating the effect of non-linear amplification, and of its inherent interaction between an increase in audibility and compression, on BEG. To systematically vary the influence of both audibility and compression, SRTs were measured for target sentences with a constant masker level of 50, 60 and 70 dBSPL. Thereby, non-linear amplification with fast-acting WDRC was provided to two independently operating hearing aids fitted to each listener individually using the NAL-NL2 prescription. For reference purposes, the HI listeners were also tested with linear amplification and a NH group was tested without amplification, both at a realistic masker level of 60 dBSPL. Results revealed that, SRTs for stimuli with artificially enhanced ILDs improved by larger amounts when the level of distractors was increased from 50 to 60 dBSPL than when it was increased from 60 to 70 dBSPL. This observation may be attributed to the higher amplification provided by WDRC at softer levels and compressive amplification at high levels. The similar trend of improvement in SRTs with an increase in audibility was also noticed in co-located condition, and other reference conditions (i.e. stimuli with natural ILDs only and, stimuli with both natural ILDs and ITDs). Interestingly, the SRM for HI listeners at the highest level almost reaches NH values across all conditions. However, most of the SRM was provided at low frequencies where the considered hearing losses were rather mild. Further, the significant extra benefit provided by artificial ILDs on top of natural ILDs was also noted.

Overall results indicated that if current or future hearing aid technology can provide ILDs at low and mid frequencies, then they can be utilized by HI listeners to achieve larger SRM and hence improve understanding of speech in noise. Therefore, future research should look into ways of enhancing low-frequency ILDs in hearing aids, for instance by applying appropriate directional multi-microphone processing, transposing natural ILDs to lower frequencies, or utilizing information provided by (low-frequency) interaural time differences (ITDs).

Most studies on BEG, including the studies described in chapter 2, 3 and 4 considered either NH listeners or HI listeners with mild to moderate or moderately severe hearing loss that can be treated by hearing aids. In contrast, very little is known about BEG in listeners with more profound hearing loss who are recipients of CIs. CI users cannot access ITDs reliably and purely rely on ILDs to localize sounds as well as to improve speech intelligibility in noise. Hence, one might expect that CI recipients are able to also use ILD cues for BEG. Hence, in chapter 5, the effect of BEG on SRM was investigated in a group of bilateral CI recipients. Further, the ability of CI users to utilize ILDs for a basic task, simpler than understanding speech in noise, was also assessed using a two-talker localization task. The ability to localize two spatially separated talkers may have a close resemblance to the ability to understand speech in a multi-talker background, since it requires not only detection of sounds but also segregating multiple talkers arriving from different directions. The bilateral benefit obtained in speech understanding in noise and in localization task was also assessed by comparing performances with a CI only on one ear versus CIs on both the ears. Results revealed that average SRM was smaller than observed in NH listeners, but the performance varied strongly between subjects, with some subjects showing a substantial SRM. Further, SRM noted with two CIs was better than noted with one CI highlighting the importance of bilateral CIs in understanding speech in noise. Similarly, localization performance was better with CIs on both the ears than just on one ear, but surprisingly there was no significant difference in localizing a single talker versus two simultaneous talkers. However, the latter findings may have been largely due to the application of rather long speech stimuli of 15 seconds, which were substantially longer than the short target sentences used in the speech task. This might have given enough time to the CI users to segregate and localize one talker after the other rather than at the same time. Hence, future research should consider using short sentences within the two-talker localization task. Overall, even though the data suggests that at least some CI recipients can utilize BEG to improve speech intelligibility in noise, this needs to be further investigated by testing more subjects as well as by providing other, more realistic, acoustic scenarios. Moreover, the applied localization paradigm needs to be further developed to better capture the complex auditory scene analysis that is applied in noisy environments.

6.1 Main conclusions of this research

- Audibility plays an important role in the utilization of ILD cues for BEG.
- HI listeners can utilize artificially extended ILDs as well as NH listeners when the loss of audibility is carefully compensated for.
- A linear increase in sensation level (and thus audibility) does not only improve speech intelligibility in noise but also the spatial benefit provided by BEG.
- WDRC does not have any impact on the utilization of ILDs (and ITDs) at low signal levels, but limits the spatial benefit provided by natural (high-frequency) ILDs within BEG.
- Implementations of artificially extended ILDs provides an extra benefit on top of natural ILDs and ITDs, in particular at high signal levels.
- Some CI users showed a substantial amount of BEG, though overall performance is not as good as in NH listeners.

6.2 Limitations of this research and future recommendations

Even though hearing aids have come a long way in helping HI listeners in understanding speech, their performance is still not adequate in noisy conditions. The main objective of this thesis was to understand and improve the effect of BEG in HI listeners to improving speech intelligibility in noise. In particular, it was investigated how far artificially extending ILD cues, which are naturally occurring only at high frequencies, towards low frequencies can be utilized by HI listeners in BEG and thereby improve speech intelligibility in noise above the benefit already provided by natural cues. Results showed that HI listeners can receive a substantial extra benefit from BEG using artificially extended ILDs, in particular with non-linear amplification. Since BEG is a signal energy based phenomenon, maskers utilized in this research were mainly energetic in nature (especially chapter 3 & 4). However, in a real-life scenario, listeners come across maskers that exhibit both energetic masking and informational masking (e.g. speech). Therefore, it might be interesting to investigate the experiments in chapter 3 and 4 by using speech maskers. However, it is unclear how far informational masking is involved in conditions where maskers are spatially separated from target speech, in particular under realistic conditions.

The number of talkers used as speech maskers in this research were two and were placed symmetrically at the left and right side of the listener (i.e., at ±90 degree azimuth). In a natural scenario, the likelihood of two talkers speaking at the same time in such scenario is very minimal. However, in a larger group of people, there are more chances of many people speaking at the same time such as in cocktail parties. Therefore, in future research, it might be more helpful to include more than two speech maskers, each arriving from different directions and distances inside reverberant environments. Ultimately, more realistic noise samples such as cafeteria noise should be used.

The amplification provided and the resulting benefit noted in this study (chapter 2, 3 & 4) was under headphones. However, since one of the main objectives was to investigate the ability of HI listeners to use ILDs at low frequencies, this was only investigated in a best case scenario. Hence, a number of factors that may be relevant when actual hearing aids are used for implementation are not taken into account. This may include factors such as hearing aid microphone placements, vent size, feedback cancellation methods as well as other advanced signal processing features. Hence, studies need to implement the applied ideas in real hearing aids and study their benefit in more realistic environments. This includes the development of a technique that can realize the low frequency extension of ILDs in hearing aids.

SRM investigated in bilateral CI users can be attributed to both release from EM and release from IM depending on the nature of the stimuli and methods used (chapter 5). The influence of IM on understanding speech in a speech-mixture is not very well understood in CI users. The (partially) intelligible speech of the interfering talkers may make it hard for the CI users to attend to (or concentrate on) the target speech. But even if the interferers are not intelligible, they still contain very similar spectro-temporal fluctuations that may introduce confusions (or uncertainties) in the CI users about what information belongs to the target signal and what information belongs to the individual interferers. Either aspect may interfere with the bottom-up as well as top-down processes involved in auditory scene analysis (e.g., Westermann and Buchholz, 2015) and potentially result in a decreased performance in the CI users. In normal-hearing subjects, IM is highly reduced (or even absent) when sufficient cues are available to perceptually segregate the target from the interferers. For example, IM is highly reduced when talker-difference cues are available (e.g., Glyde *et al.*, 2013b) and basically absent when the target signal is spatially separated from the distractor signals (e.g., Best *et al.*, 2013). However, since talker-difference cues, as well as spatial cues, are highly distorted in the implanted ear (e.g., van Hoesel and Tyler, 2003) it is unclear how far they contribute to auditory stream segregation in CI users and thus, help to reduce IM. Unfortunately, very few studies exist that have explicitly investigated IM in CI users. Bernstein *et al.* (2016) provided evidence that bilateral CI users can benefit from the perceived (spatial) separation of a target talker from one or two interfering talkers, indicating a spatial release from IM. They presented the target together with the interferer to one ear and observed a small but significant benefit when the interferer alone (not the target) was presented to the other ear. Misurelli and Litovsky (2015) showed that performance in speech intelligibility decreases when distracting talkers are of a different gender as the target talkers are replaced by talkers of the same gender, which may be explained by an increase in IM. However, the resulting similarity in the speech spectra and fundamental frequency may have also increased energetic masking. Currently, in this and other studies, it is hard to segregate the amount of release from energetic masking and informational masking due to the type of distractors used. Therefore, future studies should further investigate the effect of informational masking on CI recipients, utilizing controlled stimuli exhibiting only energetic or informational masking.

Further, in this study (chapter 5), speech comprehension thresholds (SCTs) were measured in order to get a better reflection of real-life performance since speech comprehension involves the extraction of meaning (i.e., the questions need to be comprehended) as well as the formulation of a reply. This task is different from simple recall of isolated sentences that it is used in measuring sentence recall, since it involves higher level (e.g. cognitive) processes. During the speech comprehension task, subjects were additionally asked to repeat the heard question to the experimenter. But this was only introduced to check if the subjects could actually perform the task. If the subjects were able to repeat the question reasonably well but not to respond correctly then this might have highlighted a potential cognitive problem and disqualified them from the study. But such an error happened extremely rarely and was mainly due to fatigue effects. Interestingly, there are no studies in the literature that have compared the difference in performance between a standard sentence recall task and this specific speech comprehension task. However, few studies have used comprehension performance scored by subjects answering question while (or after) listening to a 3-5 minutes long speech passage. In this regard, Best *et al*. (under revision) reported that speech comprehension assesses cognitive factors that are not assessed by simple sentence recall. Unfortunately, in this study, the scoring for repetition of the sentences/questions was not done and should be investigated in the future.

The other limitations of this research, which are rather common in this field of research, were the difficulty in matching groups for age, hearing and cognitive status as well as only considering the average results of two trials due to time limitations, which produced a not negligible intra-subject variability. Unfortunately, it was not possible for this research to find enough listeners that allowed matching for age, cognition and hearing status. Applying matched subjects in future studies would be important to understand the contribution of these factors on the findings of this study (especially chapter 5).

Appendix

A. Individual data

Table 1: Individual SRTs for the mean HI data shown in figure 2.3.

Experiment 1				
	Co-located		Spatially separated	
	Speech-discourse	Vocoded-speech	Speech-discourse	Vocoded-speech
HI01	-5.20	-13.50	-24.05	-24.20
HI02	-7.30	-13.50	-23.45	-21.10
HI03	-9.85	-12.50	-23.20	-22.50
HI04	-8.60	-13.20	-21.80	-21.85
HI05	-7.70	-12.00	-23.25	-22.55
HI06	-5.50	-10.15	-19.35	-21.00
HI07	-5.70	-11.20	-19.70	-19.30
HI08	-9.00	-12.15	-21.90	-18.50
HI09	-8.30	-13.45	-22.10	-25.50
HI10	-6.10	-12.15	-21.50	-22.70
Mean	**-7.33**	**-12.38**	**-22.03**	**-21.92**

Table 2: Individual SRTs for the mean HI data shown in figure 2.6.

Experiment 2						
	Co-located			Spatially separated		
	Low	Mid	Broad	Low	Mid	Broad
HI01	-0.6	7.1	-5.6	-4.5	-0.1	-8.8
HI02	NA	1.8	NA	NA	-4.85	NA
HI03	-3,65	-1.95	-9	-9.1	-10.45	-16.45
HI04	7.75	3.35	-4.5	-2.05	-3	-11.65
HI05	17.3	14.85	4.25	16.3	8.5	-0.25
HI06	0	5.05	-4.5	-6.35	-3.7	-9
HI07	4.1	2.4	-6.9	-5.45	-8.05	-13.15
HI08	NA	8.65	-0.5	NA	-1.15	-6.1
HI09	4.05	2.9	-6.3	-2.85	-5.9	-12
HI10	4.05	2.55	-4	-2.3	-4.05	-10.1
Mean	**4.13**	**4.67**	**-4.12**	**-2.04**	**-3.28**	**-9.72**

Table 3: Individual SRTs for the mean HI data shown in figure 4.4.

	NAL-RP			
Subject code	Co-located	Extended ILDs	ITDs & ILDs	ILDs
HI01	-7.83	-13.91	-11.40	-11.30
HI02	-10.17	-16.84	-14.56	-13.07
HI03	-5.57	-10.78	-9.00	-7.19
HI04	-4.89	-10.33	-8.23	-6.88
HI05	-8.27	-12.71	-12.30	-10.35
HI06	-6.41	-12.53	-10.13	-9.45
HI07	-9.88	-16.52	-13.78	-11.80
HI08	-3.09	-7.61	-8.30	-4.91
HI09	-8.32	-16.51	-13.91	-10.80
HI10	-10.54	-15.94	-12.04	-11.63
HI11	-4.21	-9.25	-7.95	-7.74
HI12	-7.10	-13.38	-11.30	-10.37
HI13	-8.34	-12.93	-11.24	-11.49
Mean	**-7.28**	**-13.02**	**-11.09**	**-9.77**
	NAL-NL2_50			
Subject code	Co-located	Extended ILDs	ITDs & ILDs	ILDs
HI01	-4.79	-8.32	-7.99	-8.97
HI02	-9.41	-13.49	-12.41	-11.42
HI03	-4.24	-7.41	-5.11	-5.45
HI04	-4.31	-8.63	-7.71	-5.69
HI05	-7.42	-12.06	-11.09	-9.65
HI06	-6.18	-9.74	-8.26	-6.97
HI07	-8.15	-12.70	-12.38	-10.56
HI08	-1.06	-2.47	-2.94	-3.25
HI09	-8.16	-12.38	-11.53	-8.82
HI10	-7.47	-13.33	-9.65	-8.71
HI11	-2.86	-5.41	-4.24	-3.64
HI12	-6.22	-9.27	-7.80	-6.69
HI13	-8.33	-10.95	-10.32	-9.48
Mean	**-6.04**	**-9.70**	**-8.57**	**-7.64**

	NAL-NL2_60			
Subject code	Co-located	Extended ILDs	ITDs & ILDs	ILDs
HI01	-7.27	-13.18	-11.74	-11.53
HI02	-9.98	-14.59	-13.59	-11.90
HI03	-6.75	-12.82	-9.94	-8.04
HI04	-6.68	-12.68	-9.54	-7.93

HI05	-11.25	-16.54	-13.29	-11.18
HI06	-6.87	-13.76	-9.97	-9.35
HI07	-10.04	-17.12	-13.84	-12.40
HI08	-3.25	-8.38	-5.48	-6.04
HI09	-9.15	-16.08	-13.43	-11.11
HI10	-8.82	-14.85	-12.94	-12.00
HI11	-3.86	-10.81	-8.12	-7.53
HI12	-7.32	-13.48	-11.57	-9.88
HI13	-8.42	-16.00	-11.22	-10.34
Mean	**-7.66**	**-13.87**	**-11.13**	**-9.94**

	NAL-NL2_70			
Subject code	Co-located	Extended ILDs	ITDs & ILDs	ILDs
HI01	-9.09	-16.85	-12.03	-11.88
HI02	-13.03	-19.89	-14.12	-14.52
HI03	-6.80	-14.21	-12.24	-9.60
HI04	-6.44	-13.25	-7.59	-7.41
HI05	-6.54	-15.40	-8.58	-8.73
HI06	-7.99	-14.96	-11.97	-11.54
HI07	-9.21	-18.55	-14.64	-12.42
HI08	-3.84	-11.09	-8.09	-7.50
HI09	-8.46	-19.36	-11.91	-11.71
HI10	-10.42	-17.68	-12.53	-13.02
HI11	-5.44	-13.70	-9.29	-9.27
HI12	-8.79	-15.86	-13.75	-12.12
HI13	-8.28	-17.18	-11.77	-10.39
Mean	**-8.02**	**-16.00**	**-11.42**	**-10.78**

Note: There were only two participants who participated in both the second study and the third study. Subject A labeled as subject 7 in the second study and as subject 8 in the third study. Similarly, subject B labeled as subject 6 in the second study and as subject 12 in the third study.

Appendix (Ethics Approvals) of this thesis has been removed as it may contain sensitive/confidential content

Bibliography

ANSI S3.5 **(1997)**. Revision of ANSI S3.5-1969 (**R** 1986): Methods for Calculation of the Speech Intelligibility Index (American National Standards Institute, New York).

Ahlstrom, J. B., Horwitz, A. R., and Dubno, J. R. **(2009)**. "Spatial benefit of bilateral hearing Aids," Ear Hear. **30**, 203-218.

Arbogast, T. L., Mason, C. R., and Kidd, G. Jr. **(2002)**. ''The effect of spatial separation on informational and energetic masking of speech,'' J. Acoust. Soc. Am. **112**, 2086-2098.

Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. **(2005)**. "The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. **117**(4), 2169-2180.

Aronoff, J. M., Freed, D. J., Fisher, L. M., Pal, I., and Soli, S. D. **(2012)**. "Cochlear implant patients' localization using interaural level differences exceeds that of untrained normal hearing listeners," J. Acoust. Soc. Am**. 131**(5), EL382–EL387.

Batteau, D. W. **(1967)**. "The role of the pinna in human localization," Proc. Roy. Soc. (London B) 158: 158–180.

Bench, J., Kowal, A., and Bamford, J. **(1979)**. "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," Brit. J. Audiol. **13**, 108-112.

Bentler, R. A., and Pavlovic, C. V. **(1989)**. "Transfer functions and correction factors used in hearing aid evaluation and research," Ear Hear. **10**(1), 58-63.

Bernstein, J. G., Goupell, M. J., Schuchman, G. I., Rivera, A. L., and Brungart, D. S. **(2016).** "Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees," Ear Hear. **37**(3), 289-302.

Best, V., Kalluri, S., McLachlan, S., Valentine, S., Edwards, B., and Carlile, S. **(2010)**. "A comparison of CIC and BTE hearing aids for three dimensional localization of speech," Int J Audiol, **49**, 723–732.

Best, V., Marrone, N., Mason, C.R., and Kidd, G. Jr. **(2012)**. "The influence of non-spatial factors on measures of spatial release from masking," J. Acoust. Soc. Am. **131**, 3103-3110.

Best, V., Thompson, E.R., Mason, C.R., and Kidd, G. Jr. **(2013)**. "An Energetic Limit on Spatial Release from Masking," J. Assoc. Res. Otolaryngol. **14**(4), 603–610.

Best, V., Mason, C.R., Kidd, G. Jr., Iyer, N., and Brungart, D. S. **(2015)**. "Better-ear glimpsing in hearing-impaired listeners," J. Acoust. Soc. Am. **137**(2), EL213-EL219.

Best, V., Streeter, T., Roverud, E., Mason, C. R., and Kidd Jr, G. **(2016 a)**. "A Flexible Question-and-Answer Task for Measuring Speech Understanding," Trends Hear. **20,** 2331216516678706

Best, V., Keidser, G., Buchholz, J. M., and Freestonn, K. **(2016b)**. "Development and preliminary evaluation of a new test of ongoing speech comprehension," Int J Audiol. **55**(1), 45-52.

Best, V., Mason, C. R., Swaminathan, J., Roverud, E., and Kidd Jr, G. **(2017)**. "Use of a glimpsing model to understand the performance of listeners with and without hearing loss in spatialized speech mixtures," J. Acoust. Soc. Am. **141**(1), 81-91.

Best, V., Keidser, G., Freeston, K., Buchholz, J. M. (under revision). Evaluation of the NAL Dynamic Conversations Test in older listeners with hearing loss. Int. J. of Audiol.

Besser, J., Festen, J. M., Goverts, S. T., Kramer, S. E., and Pichora-Fuller, M. K. **(2015)**. "Speech-in-speech listening on the LiSN-S test by older adults with good audiograms depends on cognition and hearing acuity at high frequencies," Ear hear. **36**(1), 24-41.

Beutelmann, R., and Brand, T. **(2006)**. "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **120**(1), 331-342.

Blauert, J. **(1997)**. Spatial Hearing: The Psychoacoustics of human sound localization, revised edition (The MIT Press, Cambridge, MA).

Boike, K. T., and Souza, P. E. **(2000)**. "Effect of compression ratio on speech recognition and speech-quality ratings with wide dynamic range compression amplification," J Speech Lang Hear Res. **43**(2), 456-468.

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. **(2000)**. "A speech corpus for multi-talker communications research," J. Acoust. Soc. Am. **107**, 1065–1066.

Bor, S., Souza, P., and Wright, R**. (2008)**. "Multichannel compression: Effects of reduced spectral contrast on vowel identification," J *Speech* Lang Hear Res. **51**(5), 1315-1327.

Breebaart, J., van de Par, S., and Kohlrausch, A. **(2001)**. "Binaural processing model based on contralateral inhibition. I. Model structure," J. Acoust. Soc. Am. **110**, 1074–1089.

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

Bronkhorst, A. **(2000)**. "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," Acta. Acust. United. Ac. **86**, 117–128.

Brungart, D.S., and Rabinowitz, W. M. **(1999)**. “Auditory localization of nearby sources. Head related transfer functions,” J. Acoust. Soc. Am. **106** (3), 1465-79.

Brungart, D.S., Durlach, N. I., and Rabinowitz, W. M. **(1999)**. “Auditory localization of nearby sources. II. Localization of a broadband source,” J. Acoust. Soc. Am. **106**, 1956–1968.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. **(2001)**. “Informational and energetic masking effects in the perception of multiple simultaneous talkers,” J. Acoust. Soc. Am. **110**, 2527–2538.

Brungart D. S., and Iyer, N. **(2012)**. “Better-ear glimpsing efficiency with symmetrically-placed interfering talkers,” J. Acoust. Soc. Am. **132**(4), 2545-2556.

Brungart, D.S., Cohen, J., Cord, M., Zion, D., and Kalluri, S. **(2014)**. “Assessment of auditory spatial awareness in complex listening environments,” J. Acoust. Soc. Am. **136**, 1808-1820.

Buchholz, J. M. **(2013)**. “A real-time hearing-aid research platform (HARP): realization, calibration, and evaluation,” ACUSTICA/Acta Acustica **99**, 477-492.

Buss, E., Pillsbury, H. C., Buchman, C. A., Pillsbury, C. H., Clark, M. S., Haynes, D. S., and Novak, M. A. **(2008)**. “Multicenter US bilateral MED-EL cochlear implantation study: Speech perception over the first year of use,” Ear Hear, **29**(1), 20-32.

Butler, R. A. **(1969a). “**Monaural and binaural localization of noise bursts vertically in the median sagittal plane,” Journal of Audiology Research, **3**, 230-235.

Butler, R. A. **(1969b).** “On the relative usefulness of monaural and binaural cues in locating sound in space,” Psychon Sci, **17**(4), 245-246.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C. and Kiessling, J. **(1994)**. “An international comparison of long-term average speech spectra,” J. Acoust. Soc. Am. **96**(4), 2108-2120.

Byrne, D., and Noble, W. **(1998)**. “Optimizing sound localization with hearing aids,” Trends. Amplif. **2**, 51-73.

Cameron, S., and Dillon, H. **(2007)**. “Development of the Listening in Spatialized Noise-Sentences Test (LiSN-S),” Ear Hear. **28**, 196-211.

Cameron, S., Glyde, H., and Dillon, H. **(2011)**. “Listening in Spatialized Noise-Sentences Test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age,” J. Am. Acad. Audiol. **22**, 697-709.

Cherry, E. C. **(1953).** “Some experiments on the recognition of speech, with one and with two Ears,” J. Acoust. Soc. Am. **25**(5), 975-979.

Clark, J. G. **(1981).** “Uses and abuses of hearing loss classification,” Asha, **23**(7), 493-500.

Cox, R. M., Alexander, G. C., Taylor, I. M., and Gray, G. A. **(1997)**. “The contour test of loudness perception,” Ear Hear, **18**(5), 388-400.

Culling, J. F., Hawley, M. L., and Litovsky, R. Y. **(2004)**. “The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources,” J. Acoust. Soc. Am. **116**(2), 1057-1065.

Culling, J. F., and Mansell, E. R. **(2013)**. “Speech intelligibility among modulated and spatially distributed noise sources,” J. Acoust. Soc. Am. **133**(4), 2254–2261.

Davies-Venn, E., Souza, P., Brennan, M., and Stecker, G. C. **(2009).** “Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing losS,” Ear Hear, **30**(5), 494-504.

Culling, J. F., & Stone, M. A. (2017). Energetic Masking and Masking Release. In *The Auditory System at the Cocktail Party* (pp. 41-73). Springer International Publishing.

Dillon, H. **(2012).** “Hearing aids,” 2nd Edition. (Boomerang Press, Sydney).

Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. **(2002).** “Spectral contributions to the benefit from spatial separation of speech and noise,” J. Speech Lang. Hear. Res. **45**, 1297–1310.

Dunn, C. C., Tyler, R. S., Oakley, S., Gantz, B. J., and Noble, W. **(2008)**. “Comparison of speech recognition and localization performance in bilateral and unilateral cochlear implant users matched on duration of deafness and age at implantation,” Ear Hear, **29**(3), 352.

Durlach, N. I. **(1963)**. “Equalization and cancellation theory of binaural masking-level differences,” J. Acoust. Soc. Am. **35**, 1206-1218.

Durlach, N. I., Kidd, Jr., G., Mason, C. R., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. (**2002**). ‘‘Informational masking: Toward improved understanding,’’ J. Acoust. Soc. Am. **111**, 2337-2337.

Festen, J. M., and Plomp, R. **(1990)**. “Effect of fluctuating Noise and interfering speech on the speech receptionthreshold for impaired and normal hearing,” J. Acoust. Soc. **Am. 88,** 1725-1736.

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. **(1999)**. “The role of perceived spatial separation in the unmasking of speech,” J. Acoust. Soc. Am. **106**(4), 3578–3588.

Fu, Q. J., Chinchilla, S., & Galvin, J. J. **(2004).** “The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users," J Assoc Res Otolaryngol. **5**(3), 253-260.

Gallun, F. J., Diedesch, A. C., Kampel, S. D., and Jakien, K. M. **(2013)**. "Independent impacts of age and hearing loss on spatial release in a complex auditory environment," Front. Neurosci. **7**, 252.

Gelfand, S. A., Ross, L., and Miller, S. **(1988)**. "Sentence reception in noise from one versus two sources: Effects of aging and hearing loss," J. Acoust. Soc. Am. **83**, 248–256.

Glyde, H., Cameron, S., Dillon, H., Hickson, L., and Seeto, M. **(2013a)**. "The effects of hearing impairment and aging on spatial processing," Ear Hear. **34** (1), 15-28.

Glyde, H., Buchholz, J. M., Dillon H., Best, V., Hickson, L., and Cameron, S. **(2013b)**. "The effect of better- ear glimpsing on spatial release from masking," J. Acoust. Soc. Am. **134** (4), 2937-45.

Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., and Hickson, L. **(2013c)**. "The importance of interaural time differences and level differences in spatial release from masking," J. Acoust. Soc. Am. **134,** EL147-EL152.

Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., Hickson, L., and Best, V. **(2015)**. "Effect of audibility in spatial release from speech-on-speech masking," J. Acoust. Soc. Am. **138**(5), 3311-3319.

Grantham D. W., Ashmead D. H., Ricketts T. A., Haynes D. S., and Labadie R. F. **(2008).** "Interaural time and level difference thresholds for acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+processing," Ear Hear. **29,** 33–44

Hansen, J. H. L., and Pellom, B. L. **(1998)**. "An Effective Quality Evaluation Protocol for Speech Enhancement Algorithms," in Proc. of the International Conference on Spoken Language Processing, ICSLP (Sydney).

Hartmann, W. M., and Rakerd, B. **(1989).** "Localization of sound in rooms IV: The Franssen Effect," J. Acoust. Soc. Am. **86**(4), 1366-1373.

Hawley, M., Litovsky, R., and Colburn, H. **(1999)**. "Speech intelligibility and localization in a multi-source environment," J. Acoust. Soc. Am. **105,** 3436-3448.

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**(2), 833-843.

Holm, S. **(1979)**. "A simple sequentially rejective multiple test procedure," Scand J Stat. **6** (2), 65–70.

ISO 389-2. Acoustics—Reference zero for the calibration of audiometric equipment—Part 2: Reference equivalent threshold sound pressure levels for pure tones and insert earphones. International Organization for Standardization, Geneva. 1994.

Jakien, K. M., Kampel, S. D., Gordon, S. Y., and Gallun, F. J. **(2017)**. “The Benefits of Increased Sensation Level and Bandwidth for Spatial Release from Masking,” Ear Hear. **38**(1), e13-e21.

Kates, J. M. **(2008)**. Digital hearing aids. Plural Publishing, San Diego, USA.

Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., and Convery, E. **(2006).** The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers,” Int J Audiol, **45**(10), 563-579.

Keidser, G., Dillon, H., Mejia, J., and Nguyen, C.V. **(2013)**. “An algorithms that administers adaptive speech-in-noise testing to a specified reliability at selectable points on the psychometric function,” Int. J. Audiol. **52**(11), 795-800.

Kidd, G. Jr., Mason, C. R., Rohtla, T. L., and Deliwala, P. S. **(1998)**. “Release from masking due to spatial separation of sources in the identification of non-speech auditory patterns,” J. Acoust. Soc. Am. **104**, 422-431.

Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. **(2007)**. “Informational masking”, in Auditory Perception of Sound Sources, Springer Handbook of Auditory Research, edited by W. A. Yost, A. N. Popper, and R.R. Fay (Springer, New York) , pp.143–189.

Kidd, G., Mason, C. R., Best V., and Marrone N. **(2010)**. “Stimulus factors influencing spatial release from speech-on-speech masking,” J. Acoust. Soc. Am. 128, 1965–1978.

Kiessling J., Pichora-Fuller M. K., Gatehouse S., Stephens D., Arlinger S., Chisolm T., et al. **(2003).** “Candidature for and delivery of audiological services: special needs of older people,” Int J Audiol. **42**(Suppl. 2), S92–S101.

Kokkinakis, K., and Pak, N. **(2014)**. “Binaural advantages in users of bimodal and bilateral cochlear implant devices,” J. Acoust. Soc. Am. **135**(1), EL47-EL53.

Kovacic, D., and Balaban, E. **(2009)**. “Voice gender perception by cochlear implantees,” J. Acoust. Soc. Am. **126**(2), 762-775.

Krishnan, A., and McDaniel S. S. **(1998)**. “Binaural interaction in the human frequency-following response: effects of interaural intensity difference,” Audiol. Neuro-otol. 3, 291–299.

Kuk, F., Lau, C. C., Korhonen, P., and Crose, B. **(2015)**. “Speech intelligibility benefits of hearing aids at various input levels," J Am Acad Audiol. **26**(3), 275-288.

Laske, R. D., Veraguth, D., Dillier, N., Binkert, A., Holzmann, D., and Huber, A. M. **(2009)**. "Subjective and objective results after bilateral cochlear implantation in adults," Otol. Neurotol. **30**(3), 313-318.

Laszig, R., Aschendorff, A., Stecker, M., Müller-Deile, J., Maune, S., Dillier, N., and Battmer, R. D. **(2004).** "Benefits of bilateral electrical stimulation with the nucleus cochlear implant in adults: 6-month postoperative results," Otol Neurotol, **25**(6), 958-968.

Litovsky, R. Y., Parkinson, A., Arcaroli, J., Peters, R., Lake, J., Johnstone, P., and Yu, G. **(2004)**. "Bilateral cochlear implants in adults and children," Arch Otolaryngol Head Neck Surg. **130**(5), 648-655.

Litovsky, R., Parkinson, A., Arcaroli, J., and Sammeth, C. **(2006)**. "Simultaneous bilateral cochlear implantation in adults: a multicenter clinical study," Ear Hear. **27**(6), 714.

Litovsky, R. Y., Parkinson, A., and Arcaroli, J. **(2009)**. "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear, **30**(4), 419.

Litovsky, R. Y., and Misurelli, S. M. **(2016)**. "Does bilateral experience lead to improved spatial unmasking of speech in children who use bilateral cochlear implants?," Otol. Neurotol, **37**(2), e35-e42.

Licklider, J. C. R. **(1948).** "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. **20**,150 –159.

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., and Roland, P. **(2009)**. "Speech recognition by bilateral cochlear implant users in a cocktail-party setting," J. Acoust. Soc. Am. **125**(1), 372-383.

Ludvigsen, C. **(1974)**. "Construction and evaluation of an audio-visual test, the Helen test,". Scand Audiol Supplementum, **3,** 67-75.

MacKeith, N. W., and Coles, R. R. A. **(1971)**. "Binaural advantages in hearing of speech," J Laryngol Otol. **85**(03), 213-232.

Macpherson, E.A. and Middlebrooks, J.C. **(2002)**. "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," J. Acoust. Soc. Am. **111**(5), 2219–2236.

Majdak, P., Goupell, M. J., and Laback, B. **(2011)**. "Two-dimensional localization of virtual sound sources in cochlear-implant recipients," Ear Hear. **32**, 198-208.

Marrone, N., Mason, C. R., Kidd, G., Jr. **(2008)**. "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," J. Acoust.

Soc. Am. **124,** 3064-3075.

Massida, Z., Marx, M., Belin, P., James, C., Fraysse, B., Barone, P., and Deguine, O. **(2013)**. “Gender categorization in cochlear implant users,” J Speech Lang Hear Res, **56**(5), 1389-1401.

McGraw, K. O., and Wong, S. P. **(1996)**. “Forming inference about some intraclass correlation Coefficients,” Psychol Methods I, 30-46.

Megumi I., Fumiai K., Kozo K., Kumiko S., and Hidehiko T. **(2011).** “Evaluation of speech perception for cochlear implant users using the Japanese speech recognition test battery,”CI-2004,” Audiology Japan. **54**(4), 277-284.

Mejia, J., Dillon, H., and Fisher, M. **(2008)**. “Active cancellation of occlusion: An electronic vent for hearing aids and hearing protectors,” J. Acoust. Soc. Am. **124**(1), 235-240.

Misurelli, S. M., and Litovsky, R. Y. **(2012)**. “Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry,” J. Acoust. Soc. Am. **132**(1), 380-391.

Misurelli, S. M., and Litovsky, R. Y. **(2015)**. “Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity,” J. Acoust. Soc. Am. **138**(1), 319-331.

Moore, B. C. J., Laurence, R. F., and Wright, D. **(1985)**. “Improvements in speech intelligibility in quiet and in noise produced by two-channel compression hearing aids,” Br J Audiol. **19**(3), 175-187.

Moore, B. C. J. (2008). “The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people,” J. Assoc. Res. Otolaryngol. 9, 399–406.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (**2010**). "Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF," Int. J. Audiol. **49**(3), 216-227.

Moore, B. C. J., Kolarik, A., Stone, M. A., and Lee, Y. W. **(2016)**. “Evaluation of a method for enhancing interaural level differences at low frequencies,” J. Acoust. Soc. Am.**140**(4), 2817-2828.

Mosnier, I., Sterkers, O., Bebear, J.-P., Godey, B., Robier, A., Deguine, O., and Fraysse, B. **(2008)**. “Speech performance and sound localization in a complex noisy environment in bilaterally implanted adult patients,” Audiol Neurotol. **14**(2), 106-114.

Müller, J., Schon, F., and Helms, J. **(2002)**. “Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system," Ear Hear. **23**(3), 198-206.

Murphy, D. R., Daneman, M., and Schneider, B. A. **(2006)**. "Why do older adults have difficulty following conversations?," Psych. Aging **21**, 49–61.

Mussoi, B. S., and Bentler, R. A. **(2017)**. "Binaural Interference and the Effects of Age and Hearing Loss," J Am Acad Audiol, 28(1), 5-13.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. **(2005)**. "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," J Am Geriatr Soc. **53**(4), 695-699.

Neher, T., Lunner, T., Hopkins, K., and Moore, B. C. **(2012)**. "Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners," J. Acoust. Soc. Am. **131**(4), 2561-2564.

Neuman, A. C., Haravon, A., Sislian, N., and Waltzman, S. B. **(2007)**. "Sound-direction identification with bilateral cochlear implants," Ear Hear. **28**(1), 73-82.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. **(1987)**. "An efficient auditory filterbank based on the gammatone function," In *a meeting of the IOC Speech Group on Auditory Modelling at RSRE* (Vol. 2, No. 7).

Patterson, R. D., Holdsworth, J., Nimmo-Smith, I., and Rice, P. **(1988)**. "Svos final report: The auditory filterbank," Tech. Rep. APU report 2341.

Peissig, J., and Kollmeier, B. **(1997)**. ''Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners,'' J. Acoust. Soc. Am. **101**, 1660–1670.

Pollack, I. **(1975)**. Auditory informational masking. J. Acoust. Soc. Am. **57**(S1), S5-S5.

Plack C. J. **(2005).** The sense of hearing, London, England: Lawrence Erlbaum Associates

Plomp, R. **(1988)**. "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function," J. Acoust. Soc. Am. **83**(6), 2322-2327.

Rana, B., and Buchholz, J. M. **(2016)**. "Better-ear glimpsing at low frequencies in normal-hearing and hearing-impaired listeners," J Acoust. Soc. Am. **140**(2), 1192-1205.

Rana, B., and Buchholz, J. M. **(Submitted)**. "Effect of audibility on better-ear glimpsing as a function of frequency in normal-hearing and hearing-impaired listeners," J Acoust. Soc. Am.

Rayleigh L **(1876)**. "Our perception of the direction of a source of sound," Nature 14: 32–33

Rhebergen, K. S., Versfeld, N. J., de Laat, J. A., and Dreschler, W. A. **(2010)**. “Modelling the speech reception threshold in non-stationary noise in hearing-impaired listeners as a function of level,” Int J Audiol. **49**(11), 856-865.

Robinson, J. D., Baer, T., and Moore, B.C. **(2007)**. “Using transposition to improve consonant discrimination and detection for listeners with severe high-frequency hearing loss,” Int. J. Audiol. **46**, 293-308.

Rychtarikova, M., Van Den Bogaert, T., Vermeir, G., and Wouters, J. **(2011)**. “Sound source localisation and speech intelligibility in virtual rooms,”Proceedings of the Jaarvergadering ABAV.

Sandel, T. T., Teas, D. C., Feddersen, W. E., and Jeffress, L. A. **(1955)**. “Localization of sound from single and paired sources,” J Acoust. Soc. Am. **27**(5), 842-852.

Schleich, P., Nopp, P., and D'Haese, P. **(2004)**. “Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant,” Ear Hear. **25**(3), 197-204.

Schön, F., Müller, J., and Helms, J. **(2002)**. “Speech reception thresholds obtained in a symmetrical four-loudspeaker arrangement from bilateral users of MED-EL cochlear implants,” Otol. Neurotol. **23**(5), 710-714.

Seeber B. U., and Fastl H. **(2008)**. “Localization cues with bilateral cochlear implants,” J. Acoust. Soc. Am. **123**, 1030–1042.

Shinn-Cunningham, B. G., Schickler, J., Kopco, N., and Litovsky, R. **(2001)**. “Spatial unmasking of nearby sources in a simulated anechoic environment,” J. Acoust. Soc. Am. **110**(2), 1118-1129.

Shinn-Cunningham, B. G. **(2008)**. “Object-based auditory and visual attention,” Trends. Cogn. Sci. **12**, 182–186.

Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. **(2008)**. “The effect of age on auditory spatial attention in conditions of real and simulated spatial separation,” J. Acoust. Soc. Am. **124**(2), 1294-1305.

Smulders, Y. E., van Zon, A., Stegeman, I., Rinia, A. B., Van Zanten, G. A., Stokroos, R. J., and Briaire, J. J. **(2016)**. “Comparison of bilateral and unilateral cochlear implantation in adults: a randomized clinical trial,” JAMA Otolaryngol Head Neck Surg. **142**(3), 249-256.

Tamati, T. N., Janse, E., Pisoni, D. B., and Baskent, D. (2017). Talker variability in real-life speech recognition by cochlear implant users. J. Acoust. Soc. Am. 141(5), 4032-4032.

Tyler, R. S., Gantz, B. J., Rubinstein, J. T., Wilson, B. S., Parkinson, A. J., Wolaver, A., and

Lowder, M. W. **(2002)**. “Three-month results with bilateral cochlear implants,” Ear Hear. **23**(1), 80S-89S.

Tyler, R. S., Dunn, C. C., Witt, S. A., and Noble, W. G. **(2007)**. “Speech perception and localization with adults with bilateral sequential cochlear implants,” Ear Hear. **28**(2), 86S-90S.

Van Hoesel, R. J. M., and Tyler, R. S. **(2003)**. “Speech perception, localization, and lateralization with bilateral cochlear implants,” J. Acoust. Soc. Am. **113**(3), 1617-1630.

Van Hoesel, R. J. M. **(2004)**. “Exploring the benefits of bilateral cochlear implants,” Audiol Neurotol. **9**, 234-246.

Van Hoesel, R. J. M**. (2012)**. “Contrasting benefits from contralateral implants and hearing aids in cochlear implant users,” Hear Res. **288**, 300-314.

Walden, T. C., and Walden, B. E. **(2005)**. “Unilateral versus bilateral amplification for adults with impaired hearing,” J Am Acad Audiol. **16**(8), 574-584.

Watson, C.S. **(2005)**. “Some comments on informational masking,” Acta. Acust. united. Ac. **91**, 502-512.

Weller, T., Buchholz, J. M., and Best, V. **(2016)**. “A method for assessing auditory spatial analysis in reverberant multi-talker environments,” J Am Acad Audiol. **27,** 601-611.

Westermann, A., and Buchholz, J. M. **(2015).** “The influence of informational masking in reverberant, multi-talker environments,” J. Acoust. Soc. Am. **138**(2), 584-593.

Wightman, F. L., and Kistler, D. J. **(1992)**. “The dominant role of low-frequency interaural time differences in sound localization,” J. Acoust. Soc. Am. **91**(3), 1648-1661.

Wiggins, I. M., and Seeber, B. U. **(2012)**. “Effects of dynamic-range compression on the spatial attributes of sounds in normal-hearing listeners,” Ear Hear, **33**(3), 399-410.

Wiggins, I.M., and Seebeer, B. **(2013)**. “Linking dynamic-range compression across the ears can improve speech intelligibility in spatially separated noise,” J. Acoust. Soc. Am. **133**, 1004-1016.

Woods, D. L., Arbogast, T., Doss, Z., Younus, M., Herron, T. J., and Yund, E. W. **(2015)**. “Aided and unaided speech perception by older hearing impaired listeners,” PloS one, **10**(3), e0114922.

Yeend I, Mejia J, Dillon H, Loi T **(2014)**. Validation of the NAL Experiential Hearing Aid Simulator. Oral poster presented at the World Congress of Audiology, Brisbane.

Zurek, P. M. **(1993**). Binaural advantages and directional effects in speech intelligibility. In: G.A. Studebaker & I. Hockberg (eds.) Acoustical Factors Affecting Hearing Aid Performance. Allyn and Bacon: Needham Heights, MA, 255–276.