The perception of speech in noise with bilateral and bimodal hearing devices

Tim Schoof

Master's Thesis

Linguistics: The Study of the Language Faculty
Utrecht University

Supervisor: Dr. Bert M.E.H. Schouten[a]
Second reader: Prof. Stuart Rosen[b]

June 2010

[a] Utrecht University
[b] University College London

Acknowledgements

First of all, I wish to thank my supervisors Bert Schouten, Stuart Rosen, Andy Faulkner, and Tim Green. Thank you for being such great supervisors. I appreciate your valuable suggestions and comments and I very much enjoyed our insightful discussions. I also want to thank you for believing in me and always pushing me just a little bit further.

There are several people without whom it would not have been possible to run my Dutch experiments. I especially want to thank Warmolt Tonckens and Theo Veenker for helping me to get MATLAB to run on my computer. I also wish to thank Rolph Houben and Joost Festen for providing me with the recordings of the Dutch sentences.

I also want to thank Hugo Quené for helping me with the analysis of my data. Thank you for remaining optimistic, even when all the computers and software in the room simply stopped working.

Special mention needs to go to everyone who participated in the experiments. Thank you for trying so hard to decipher all those sentences. I know it wasn't always easy.

This thesis could not have been written without the support of my friends and family. I want to thank Andreas for being the best brother in the world. I want to express my thanks to Mirjam, Paul, and Marco for always being supportive of my crazy plans. I want to thank my friends Anne, Coppe, Carolina, Andréa, Liquan, Marko, Floris, Emma, Thijs, Wouter, Rob, Liesbeth, Annique, Fran, Julie, and Aditi for providing me with all the good advice and the much needed distraction. Special mention needs to go to Andréa for donating that lovely green blanket.

Last but not least, I want to thank the UK Royal National Institute for Deaf People and the Utrecht institute of Linguistics for providing funding for this study.

Contents

Glossary

AGC	automatic gain control
BM	basilar membrane
CF	centre frequency
CI	cochlear implant
F0	fundamental frequency
HA	hearing aid
High frequencies	above 1500 Hz
HRTF	head related transfer function
ILD	interaural level difference
ITD	interaural time difference
JND	just noticeable difference
Low frequencies	below 1500 Hz
LSO	lateral superior olive
MLD	masking level difference
MSO	medial superior olive
N0	noise source at 0° azimuth
N90	noise source at 90° azimuth
N-90	noise source at -90° azimuth
NHL	normal-hearing listener
PTA	pure tone average
SNR	signal-to-noise ratio
SRT	speech reception threshold

Introduction

This thesis is concerned with the perception of speech in noise through simulated bilateral cochlear implants and bimodal hearing devices (i.e. the combination of a cochlear implant in one ear and a hearing aid in the other). The main objective is to examine to what extent the use of bilateral implants and bimodal hearing devices can provide a binaural advantage to listeners who have to understand speech in noise. Several simulation experiments were conducted to address this issue.

The outline of this thesis is as follows. Chapter 1 provides some general information regarding hearing aids, cochlear implants and bimodal hearing devices. Particular attention is paid to challenges with regard to speech perception when these devices are used. Chapter 2 provides an overview of general issues related to the perception of speech in noise. It focuses on binaural advantages listeners can obtain, especially when the speech and noise sources are spatially separated. The effects of summation, head shadow and binaural unmasking are discussed. Particular attention is paid to the contribution of interaural time and level differences to the binaural unmasking effect. Chapter 3 deals with the binaural advantages users of bilateral and bimodal fittings obtain when perceiving speech in noise. Two studies are discussed, which assessed the effects of summation, head shadow and binaural unmasking for listeners with simulated bilateral and bimodal fittings. In addition, the relative contribution of interaural time and level differences to the binaural unmasking effect is examined. The findings of these studies are related to the question of whether either bilateral or bimodal fittings offer more advantages for the perception of speech in noise. The major results show that performance is significantly better with simulated bimodal hearing devices than with bilateral implants. The advantage of bimodal stimulation cannot be explained, however, in terms of binaural effects. Chapter 4 focuses on improving the perception of speech in noise for bimodal listeners. It was hypothesized that performance could be improved by creating a binaural unmasking effect. A simulation study is discussed, which aims to enhance interaural time and level differences and thereby create a binaural unmasking effect. Performance does not improve, however, when interaural time and level differences are enhanced. Lastly, other strategies are discussed that may prove more successful in improving the perception of speech in noise for bimodal listeners.

1. Hearing devices

1.1 Hearing impairment

According to estimates of the World Health Organisation (WHO), 278 million people worldwide have moderate to profound hearing loss in both ears (2005). Hearing loss can be categorized into two main types. The first type, conductive hearing loss, is caused by defects in the outer and middle ear, which reduce transmission of sounds to the inner ear. The second type, sensorineural hearing loss, is the most common type of hearing loss and arises from defects in the inner ear, or the auditory nerve. Sensorineural hearing loss is usually the result of degeneration of the hair cells of the organ of Corti in the cochlea. The basal regions along the basilar membrane (BM), which are tuned to high frequencies in the auditory input, are generally most affected. Hearing loss is commonly described in terms of reduced sensitivity to low-level sounds. However, damage to the hair cells not only elevates absolute thresholds, but also reduces the frequency selectivity of the system[1]. In this thesis the term 'hearing loss' will be taken to mean sensorineural hearing impairment caused by damage to the hair cells in the cochlea.

Even though the degree of hearing loss generally varies with frequency, its severity is commonly described in terms of an overall degree of hearing loss. A common way to quantify the severity of hearing loss is in terms of the pure tone average (PTA), in which absolute thresholds for pure tones (500, 1000, 2000 Hz) are averaged over the frequencies. The following classification system is often used: slight (16 – 25 dB HL), mild (26 – 40 dB HL), moderate (41 – 55 dB HL), moderately severe (56 – 70 dB HL), severe (71 – 90 dB HL), and profound hearing loss (>91 dB HL) *(c.f. ASHA)*. Listeners who have profound hearing loss are also referred to as deaf. This thesis focuses on listeners with severe to profound hearing loss in at least one ear.

Hearing loss caused by cochlear damage is permanent. Hearing-impaired listeners can, however, use hearing devices to aid auditory perception. Two types of hearing devices that exist are hearing aids (section 1.2) and cochlear implants (section 1.3). These devices are described in more detail below.

1.2.1 Hearing aids

Hearing aids (HA) are devices that can improve hearing through amplification. They are only useful when a region of hair cells in the cochlea is still intact and some residual hearing remains. In the case of severe to profound hearing loss residual hearing is often limited to a small frequency range, for example only up to about 1 kHz *(Moore, 1998)*. Only the frequencies within this range can be usefully amplified. For individuals with dead regions towards the base of the cochlea (i.e. regions without functioning inner hair cells), amplification of high frequencies is reported to sound distorted or noise-like *(c.f. Moore*

[1] For a more detailed description of the effects of cochlear hearing loss, see *Moore (1998)*.

1998). The parameters of the HA are in this case set to attenuate rather than amplify high frequencies.

One of the biggest challenges concerning amplification of sound for people with hearing impairment is loudness recruitment. Absolute thresholds are elevated, whereas the level of uncomfortable loudness remains the same. The dynamic range of people with hearing loss is thus reduced. If all sounds were amplified to the same degree, high level sounds would be uncomfortably loud. Several strategies have been developed to compensate for loudness recruitment. The most commonly used strategy in current HAs is automatic gain control (AGC). This strategy reduces the dynamic range of the signal by means of compression. A wide range of levels at the input is compressed into a smaller range at the output and low-level sounds are amplified more than high level sounds.

An important aspect of HA designs, especially with respect to the perception of speech in noise, is the microphone. Many HAs have omnidirectional microphones that amplify sounds from all directions to the same degree. Directional microphones, on the other hand, are more sensitive to sounds that come from a particular direction. An important benefit of this type of microphone is that it can improve the signal-to-noise ratio (SNR) and thereby aid perception of speech in noisy situations.

1.2.2 Challenges of speech perception with hearing aids

Despite the fact that HAs can successfully amplify sounds to an audible level, speech perception abilities of HA users are still greatly diminished compared to normal-hearing listeners (NHL).

An important factor contributing to the reduced speech perception abilities of HA users is that the device only improves audibility of frequencies that fall within the range where residual hearing remains. In most cases this means that only the low-frequency components of speech sounds are conveyed. Recognition of speech sounds that have prominent acoustic cues in the high frequency range, such as voiceless fricatives /s/ and /f/, may be impaired.

Another reason why HA users do not perceive speech as well as NHLs is that HAs cannot compensate for the reduced frequency selectivity that is often associated with sensorineural hearing loss. As a result of loss of outer hair cells within the cochlea, the auditory filters are broader than normal, leading to reduced sharpness of tuning. The bandwidths of the filters tend to broaden with increasing hearing loss *(Bonding, 1979)*. Listeners with no frequency selectivity perform poorly with amplified speech *(Faulkner et al., 1990)*. This is probably due to the fact that they have problems resolving formants[2] in complex waveforms.

[2] A formant indicates a frequency region with a relatively high concentration of acoustic energy in a spectrum. Formants are of particular value for the classification of vowels and vowel-like sounds, and of transitional features between vowels and adjacent sounds *(Crystal, 2008)*.

1.3.1 Cochlear implants

People with severe to profound hearing loss in both ears, who do not benefit sufficiently from hearing aids, may be provided with a cochlear implant (CI). A CI is a prosthetic device which can partly restore hearing. The device stimulates the auditory nerve directly, thus bypassing malfunctioning parts of the outer, middle and (part of the) inner ear. According to the Food and Drug Administration (FDA) there are approximately 188,000 CI users worldwide (April 2009).

CIs rely on electric signals to stimulate the auditory nerve. In 1957, the first device for electric stimulation of the auditory nerve was developed *(Djourno, Eyriès and Vallancien, 1957)*. Djourno and colleagues reported a sense of environmental sounds, although speech could not yet be perceived. The first commercial implant was developed in the 1980s. The House-3 M single electrode implant was developed in 1984 and had several hundred users *(Zeng et al., 2004).*

1.3.2 Current designs

Since the 1980s several different CIs have been developed, but all have certain features in common, including: (1) a microphone that picks up the acoustic signal; (2) a signal processor that converts the input signal into a waveform appropriate for electrical stimulation; (3) an external transmission system and an internal receiver / stimulator which are connected by means of a transcutaneous link and are responsible for the transmission of the electric signal to (4) multiple electrodes inserted into the cochlea (*Loizou, 1998; Wilson and Dorman, 2008*). The electrode array is usually implanted in the scala tympani, which is located close to the auditory nerve. The stimulator is connected to the electrodes and transmits the electrical signals to the appropriate electrodes. The mapping between the channels generated by the signal processor and the electrode array is fixed. Figure 1 provides an overview of the components of the CI.

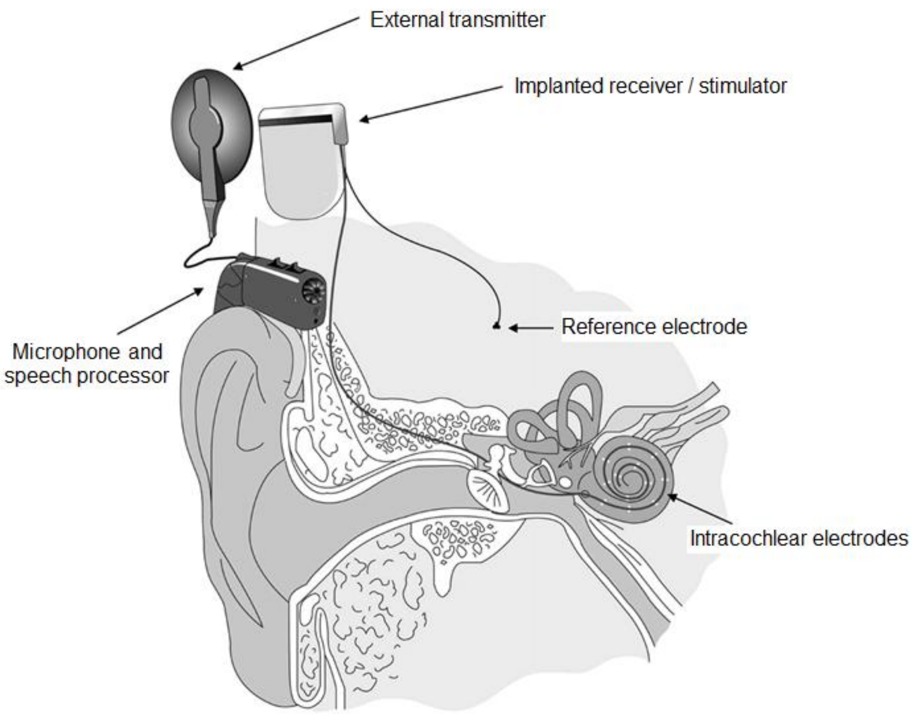

Fig. 1. Representation of the components of current CI systems (reprinted from *Wilson and Dorman (2008)*, courtesy of Med-El, Medical Electronics GmbH, of Innsbruck, Austria).

The cochlea has a tonotopic organisation, which means that specific locations along the cochlea are stimulated maximally by specific frequencies (*Greenwood, 1990*). The electrode array mimics the organisation of the healthy cochlea by stimulating nerve fibres at different places along the cochlea. The electrodes near the base of the cochlea convey the high frequencies and the electrodes nearest the apex the low frequencies. The number of implanted electrodes varies among different designs, but is usually between 12 and 22 (*Francart, 2008a*). The electrodes are generally inserted equidistantly up to 25 – 30 mm from the base of the cochlea, thus stimulating only nerve fibres with characteristic frequencies (CF) above 250 – 500 Hz (Fig. 2). The frequency selectivity of the system depends on the number of electrodes as well as the distance between these electrodes. In principle, the higher the number of electrodes, the better the frequency selectivity of the system. However, even though speech perception performance of CI users increases with the number of electrodes, it reaches an asymptote around seven to ten electrodes (*Friesen et al., 2001*). This can be explained by the fact that the frequency-to-place coding is constrained by (1) the number of surviving nerve fibres that can be stimulated and (2) the spread of activation associated with electric stimulation, which results in activation of a relatively large group of auditory neurons. In addition, the limited effect of increasing the number of electrodes may also be explained by distorted frequency-to-place mapping associated with implantation. There is generally a mismatch between the frequency outputs of the speech processor and the CF of the nerve fibres stimulated by the electrode array. The frequency alignment of the electrodes and the auditory nerves is an important factor contributing to speech understanding. A shift in frequency-to-place map leads to degeneration of speech intelligibility (*Fu and Shannon, 1999*;

Faulkner, 2006). It should be mentioned, however, that listeners can to some degree adapt to this shift in frequency-to-place mapping (*Rosen, Faulkner, and Wilkinson, 1999*).

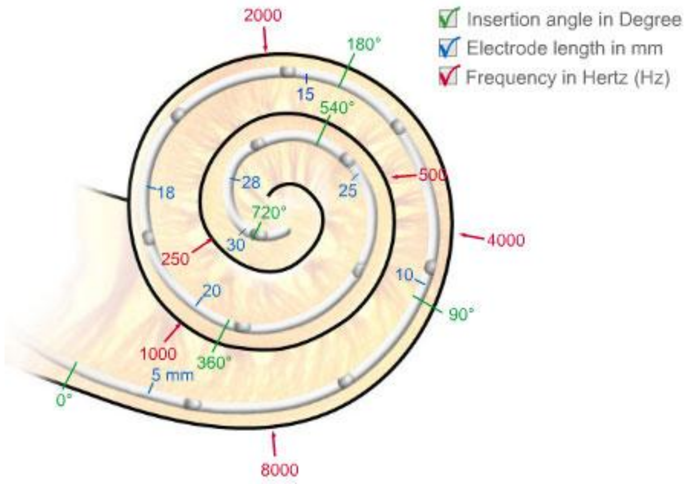

Fig. 2. Diagram of the basilar membrane showing the place of maximum displacement in response to different frequencies (Hz) and the insertion depth of electrodes (mm). In addition, the insertion angle of the electrodes is provided. This is not relevant, however, for the current discussion (reprinted from Med-El, Medical Electronics GmbH, of Innsbruck, Austria).

The signal processor forms an important aspect of implant design. The function of the processor is to decompose the signal into different frequency bands, thus attempting to simulate the frequency analysing functions of a healthy cochlea. In comparison with a healthy cochlea, however, the decomposition of the signal into its frequency components is far less precise. The implant provides a limited set of frequency components with a wider bandwidth, resulting in spectral degradation of the signal.

Various signal processing strategies have been developed over the years. One of the most common is the Continuous Interleaved Sampling (CIS) strategy. The main advantage of the CIS approach is that electrode interaction is limited because the different electrodes are stimulated nonsimultaneously (*Wilson et al., 1991*; *Favre and Pelizzone, 1993*). The signal processer first passes the incoming signal through a bank of bandpass filters, which divides the signal into several contiguous frequency channels. It is important to note that a narrowband signal at the output of a single filter can be regarded as a slowly varying envelope superimposed on a more rapidly varying temporal fine structure (TFS) (Fig. 3) (*Rosen, 1992*; *Moore, 2008a*). The envelope consists of fluctuations in overall amplitude at rates between 2 and 50 Hz, whereas amplitude fluctuations from 600 Hz to about 10 kHz are described as variations in the TFS (*Rosen, 1992*)[3]. The signal processor extracts the envelope for each channel by means of rectification and smoothing. Rectification is a process which leads to a signal with only positive amplitude values (Fig. 4). The process of smoothing refers to the application of a low-pass filter which eliminates TFS from the signal (Fig. 5). A smoothing filter of 400 Hz is typically implemented in current CIs. Compression is subsequently performed on the output of these channels. This is necessary because the range of the acoustic

[3] Note that the characterisation of the signal in terms of envelope and TFS reflects amplitude fluctuations in the time domain and does not reflect the spectral composition of the signal.

amplitudes present in the input is larger than the dynamic range of the CI users (*Loizou, 1998*). The selected envelopes are then used to modulate trains of electric pulses. In order to get a good representation of the modulation, the pulse rate should be 4 – 5 times the modulation frequency (*McKay et al., 1994*). The electric pulses introduce a new TFS with fixed-rate amplitude fluctuations which are unrelated to the rapid oscillations in the input signal. The stimulation patterns are then sent to the appropriate electrode. A schematic representation of the signal processing is provided in Figure 6.

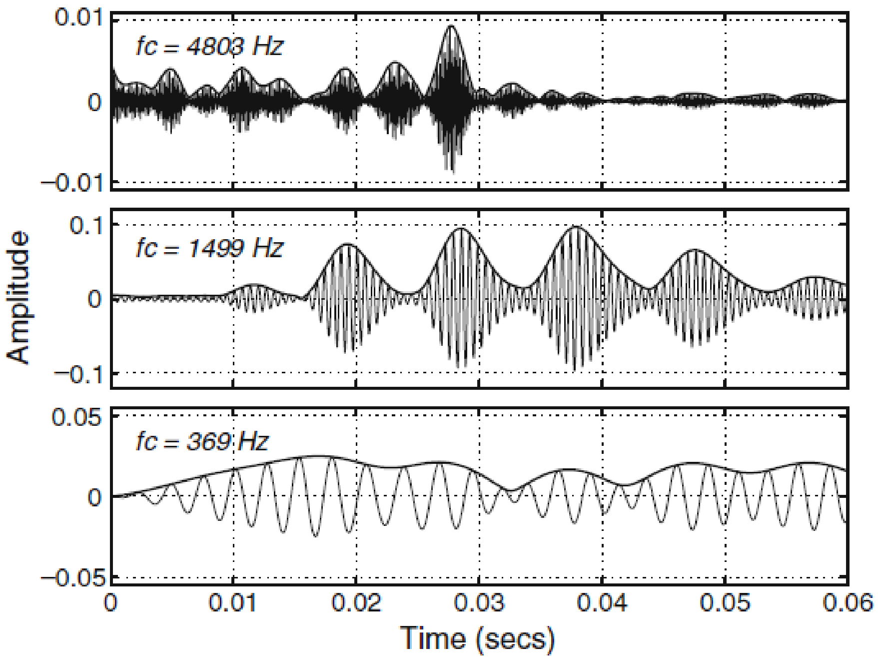

Fig. 3. Waveforms filtered at 369, 1499, and 4803 Hz in response to the sound ‘en’ in ‘sense’. The thick lines represent the envelope of the waveforms; the rapid oscillations show the fluctuations in the temporal fine structure (TFS) and are close to the centre frequency of the filter bandwidth (*Moore, 2008a*).

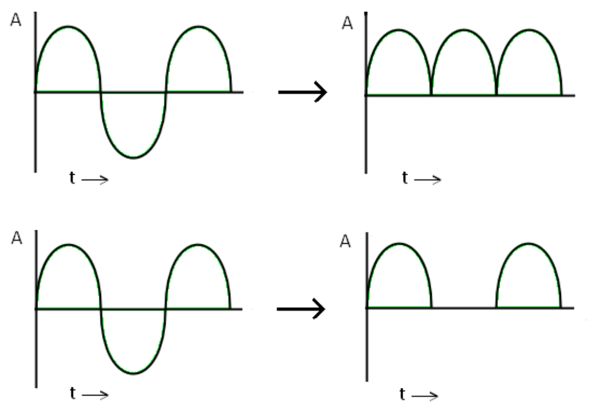

Fig. 4. Schematic representation of full- and half-wave rectification. The top graph shows the process of full-wave rectification where all negative amplitude values are converted into positive values. The bottom graph shows half-wave rectification which refers to a process in which the negative amplitude values are removed from the signal.

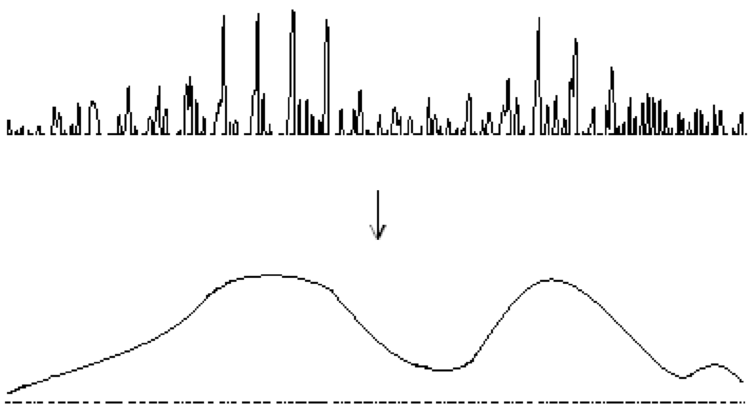

Fig. 5. Schematic representation of the process of smoothing. The illustration at the top shows a waveform after half-wave rectification. A low pass filter is applied which eliminates fast changing information from the signal. The illustration at the bottom shows the extracted envelope.

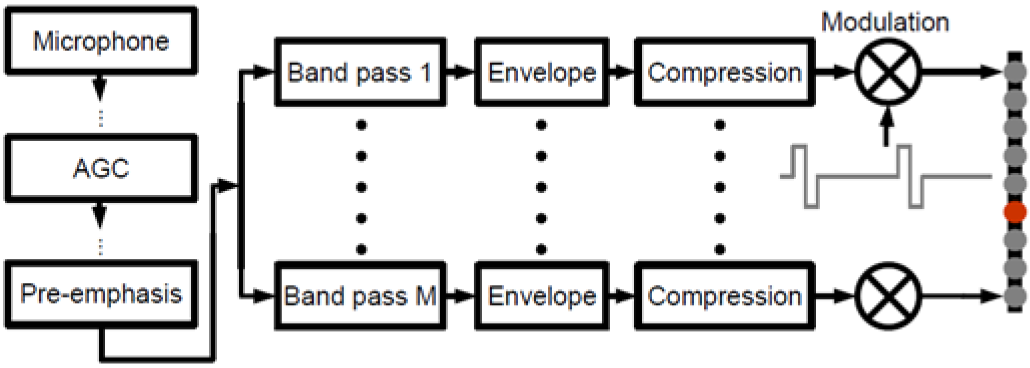

Fig. 6. A block diagram of the CIS processing strategy. The microphone picks up the acoustic signal. The amplitude of the signal is then adjusted according to the average level of the input by AGC, where a wide range of levels at the input is compressed into a smaller range at the output. Subsequently, the signal is passed through a pre-emphasis filter which is used to attenuate strong components in the signal below 1.2 kHz (Wilson & Dorman 2008). The signal is then filtered into several frequency bands; the envelope of each band is extracted, the signals are compressed and used to modulate a pulse train which has a frequency appropriate to the band. The output of each channel is then transmitted to the appropriate electrode (Block diagram has been adapted from *Francart, 2008a*).

1.3.3 Challenges of speech perception with cochlear implants

Even though current CIs have proven successful in restoring hearing to profoundly deaf individuals, the prostheses only provide listeners with a crude representation of the acoustic signal.

One of the main deficiencies of CIs is that the spectral resolution of the signal is reduced. This is the result of the limited number of frequency bands conveyed by the device. Speech recognition in quiet can be achieved with as few as four spectral bands (*Shannon et al., 1995*). However, speech recognition further improves, at least for NHLs, as the number of spectral bands increases (*Dorman et al., 1997*). Limiting the number of spectral bands causes spectral smearing, which obscures acoustic cues that would normally be available. Phonemic differences, for example, are often signalled by differences in formant frequencies that can be as small as 100 Hz. When speech is conveyed in only a limited number of spectral bands, however, this difference may fall within one spectral band and as such be undetectable. In addition, it is worth mentioning that pitch cues are difficult to perceive for CI users (*Green et al., 2005*). Since the smoothing filter (400 Hz) typically implemented in CIs is higher than the average fundamental frequency (F0) of speakers, some pitch cues are represented in the modulated waveform. Nevertheless, the availability of pitch cues is only limited as a result of spectral smearing, since the lower harmonics signalling F0 changes are unresolved. Consequently, it is thought that CI users need to rely on temporal cues instead to detect pitch.

In addition, however, the signal conveyed through a CI is also degraded in the time domain. As already mentioned, the signal processor of the CI discards the TFS of the signal when extracting the envelope. The relative importance of TFS and of envelope cues for speech perception has been studied by means of 'auditory chimeras', which are compiled of the envelope structure of one sound and the TFS of another sound. It has been shown that envelope cues are more important for the perception of English sentences in quiet than TFS (*Smith et al., 2002*). In tone languages such as Mandarin Chinese, however, TFS turns out to be more important for speech perception (*Xu and Pfingst, 2003*). Similarly, TFS tends to be the dominant cue for the perception of melody and pitch (*Qin and Oxenham, 2003; Smith et al., 2002*). Since informative TFS is discarded in the signal provided to CI users, their ability to perceive speech may be affected. It should be noted, however, that the studies described above were conducted under unnatural listening conditions (i.e. in a soundproof booth and / or with headphones). In reverberant environments, for example, TFS may be less important for the perception of speech as the reflections of the sound alter the acoustic waveform by smearing changes in the fine structure over time (*c.f. Sayles and Winter, 2008*).

Furthermore, the perception of speech in noise proves particularly problematic for CI users (*Dorman et al., 1998; Fu et al., 1998; Zeng and Galvin, 1999; Friesen et al., 2001; Stickney et al., 2004a*). This can be attributed to spectral as well as temporal degradation of the CI signal. Fu et al. (*1998*) indicate that noise susceptibility of CI users could be related to the spectral resolution of the CI signal. They measured speech recognition abilities as a function of SNR for signals with different numbers of spectral bands. They showed that performance rapidly decreased as the number of spectral bands was reduced. It has been suggested that decreased spectral resolution causes the masker to become perceptually integrated with the target signal (*Stickney et al., 2004a*). Segregating the speech and noise signals becomes increasingly difficult as the bandwidth of the frequency channels increases, since differences in the frequency domain between the speech and noise sources may fall within one spectral band and as such be undetectable. In addition, the fact that the signal processor in the CI eliminates the TFS of the input signal probably contributes to impaired perception of speech in noise. The importance of TFS with respect to the perception of speech in noise has been indicated by Drennan et al. (*2007*). They showed that even though NHLs are able to perceive speech in noise in the absence of TFS, perception improved when fine structure was introduced. This effect may in part be attributed to glimpsing. Glimpsing is sometimes also called dip-listening and refers to the ability to detect a target signal during the amplitude dips of a fluctuating masker. When the amplitude of the masker signal is relatively low compared to that of the target signal, and the SNR is at its peak, the target signal is more easily detected (*Festen and Plomp, 1990*). The inability to exploit TFS in the input signal has been linked to an inability to listen in the dips (*Lorenzi et al., 2006*). Since the signal processor of CIs discards the TFS of the input signal, it can be expected that CI users cannot benefit from amplitude dips in the masker signal.

1.3.4 Simulating cochlear implants

Research into the perception of speech with CIs sometimes makes use of simulation experiments. CIs are usually simulated with a channel vocoder, which involves processing techniques that resemble those of actual implants (e.g., *Shannon et al., 1995; Dorman et al., 1997, Friesen et al., 2001; Qin and Oxenham, 2003*). Similarly to actual CIs, channel vocoders transmit signals using a set of amplitude-modulated carriers. Two vocoding techniques which are often used to simulate the information available to CI users are noise- and tone-vocoding. The difference between the two techniques is that noise-vocoders use band-limited noise carriers and tone-vocoders rely on sine wave carriers. Consequently, the quality of the signals generated by these vocoders is very different. Even though simulations using noise and sine wave carriers have resulted in replications of speech recognition abilities commonly found for CI users, noise- and tone-vocoded speech do not provide the same information to listeners (*Whitmal et al., 2007; Souza and Rosen, 2009*). Both types of vocoding result in more-or-less realistic representations of reality, reflecting different aspects of actual CIs.

An important difference between tone- and noise-vocoding concerns the temporal representation of the signal. It can be hypothesised that temporal patterns may be more accurately preserved in tone-vocoded speech. This is explained by the fact that intrinsic temporal fluctuations of the noise carrier can interfere with the fluctuations in the speech signal. Tone-vocoded speech, on the other hand, does not have intrinsic envelope fluctuations and consequently facilitates better modulation detection (*Whitmal et al., 2007*).

Another important difference between the two techniques is related to the representation of pitch cues. Tone-vocoded speech generally gives a better representation of periodic information than noise-vocoded speech. This is due to the fact that when an envelope is modulated with a sinusoidal carrier, sidebands consisting of the sum and difference frequencies of each spectral component are created *(Souza and Rosen, 2009)*. These cues lead to a stronger perception of pitch. Sidebands also occur in noise-vocoded speech, although this does not lead to stronger pitch cues, since multiplying a white noise source by any signal still results in white noise. Since pitch cues are poorly preserved by actual CIs, noise-vocoding may provide a more accurate representation of pitch.

In addition, the cutoff frequency of the smoothing filter plays an important role when a CI signal is simulated. Carrier type in combination with envelope cutoff frequency has been shown to alter the available spectral and temporal cues in the signal (*Souza and Rosen, 2009*). The cutoff frequency of the smoothing filter determines the range of modulation frequencies available in the signal. A lower cutoff frequency results in a decrease in the rate of fluctuation in the envelope superimposed on the carrier signal. In addition, an envelope cutoff which is low in comparison to a speaker's F0 will delete F0-related modulations in the envelope. Although a smoothing filter of 400 Hz is typically implemented in current CIs, simulation techniques have also used filters with different cutoff frequencies. When modulating a sine-carrier, for example, using a smoothing filter which is high in comparison to the talker's F0 (e.g. 400 Hz) results in the presence of pitch cues that are much stronger than those available to actual CI users. As was already mentioned, pitch cues are difficult to perceive for CI users. A relatively low smoothing filter (e.g. 50 Hz) provides a more accurate representation of pitch cues. By contrast, a noise vocoder using a smoothing filter with a relatively high cutoff frequency provides pitch cues as actual CI users would perceive them. The advantage of using a low cutoff frequency, on the other hand, is that it is likely to minimise the influence of inherent fluctuations in the noise which would otherwise be superimposed on the envelope.

In this thesis, the perception of speech in noise by users of bilateral and bimodal fittings is assessed by means of simulation studies. These experiments rely on CI simulations using a noise-vocoder with a smoothing filter of 400 Hz. Even though inherent fluctuations in the noise may be superimposed on the envelope, the signal is expected to convey pitch cues as actual CI users would perceive them. In Study II, however, performance in noise is also examined for bilateral CI simulations created with a noise-vocoder with a smoothing filter of 50 Hz and a tone-vocoder with a smoothing filter of 50 Hz, since these simulations may more accurately preserve timing information.

1.4.1 Bimodal fittings

As a result of the success of cochlear implantation, CI candidacy has been extended to include individuals who have some residual hearing in at least one ear. Currently, individuals with profound hearing loss (PTA >90 dB HL at 2 and 4 kHz) who do not benefit adequately from HAs (50% correct recognition BKB sentences[4] at 70 dB SPL) are eligible for implantation in the UK (*NICE, 2009*). This has led to an increasing number of individuals with residual hearing who can benefit from a HA in the ipsilateral (implanted) or, more often, in the contralateral (non-implanted) ear. This thesis is concerned with bilateral bimodal stimulation, when a CI is implanted in one ear and a HA in the other. Bimodal stimulation is sometimes also referred to as electro-acoustic stimulation as it combines electric hearing through the CI and acoustic hearing through the HA.

There are several advantages to providing unilateral CI users with a HA in the non-implanted ear. Firstly, the addition of a contralateral HA has been shown to improve localisation abilities and the perception of speech, especially in noise (*Armstrong et al., 1997; Tyler et al., 2002; Kong et al., 2005; Dorman et al., 2008*). The advantage of providing a contralateral HA can particularly be explained by the fact that phonetic information and pitch cues are more accurately preserved in the low-frequency range of the acoustic signal. In addition, the availability of TFS in the HA signal may facilitate glimpsing and thereby improve perceptual abilities in noisy situations. The benefits of bimodal fittings, especially with regard to the perception of speech in noise, are discussed in more detail in chapter 3. Secondly, the provision of amplification in the non-implanted ear helps prevent possible

[4] A set of standard sentences often used in audiological research (*Bench, Kowal and Bamford, 1979*).

effects of auditory deprivation in that ear. A lack of auditory stimulation may in turn lead to deterioration of speech perception abilities in the unaided ear (*Silman et al., 1984; Neuman, 1996*).

1.4.2 Challenges of speech perception with bimodal fittings

Despite the advantages of bimodal stimulation, some challenges are inherent in the combination of a CI and a HA. There are some crucial differences between the two methods of auditory stimulation leading to discrepancies between the two ears.

An important difference between electric and acoustic hearing concerns the frequency range that is covered by CI and HA. CIs usually do not cover frequencies lower than 100 – 200 Hz, but can convey signals up to about 8000 Hz. It should be pointed out, though, that the range of frequencies covered by the device is a matter of choice. HAs, on the other hand, often only amplify low frequencies (e.g. below 1 kHz), since this is where residual hearing generally remains. Even when the electric and acoustic signals convey the same frequency information, however, they result in different percepts. McDermott and Sucher (*2006*) reported that acoustic pure tones are perceived as very different from electric pulse trains delivered to a single electrode position with a constant rate. This may in part be due to a mismatch in frequency-to-place mapping in CIs.

Another important issue when combining a CI and a HA concerns synchronisation of the signals. The two devices both present the signal to the listener with a small delay. This delay, however, is different for acoustic and electric stimulation. The CI, on the one hand, introduces a device-dependent delay of approximately 1 – 20 ms. The HA, on the other hand, introduces a device-dependent delay of about 1 – 12 ms. In addition, the acoustic signal has a frequency delay of 1 – 4 ms introduced by the middle and the inner ear (*Cf. Francart, 2008a*). This difference in delay may impair speech perception since the signals cannot be processed simultaneously. In addition, the synchronisation mismatch may also be detrimental to the localisation and segregation of sound sources which in part depend on time differences between the ears.

Finally, a third problem of bimodal stimulation is related to the perception of loudness. Currently, HAs and CIs both contain AGCs that reduce the dynamic range of the signal. The two systems, however, have different parameter settings and are fitted independently. Since the AGCs rely on different degrees of compression, the dynamic ranges of the HA and CI will also be different. This often leads to unbalanced loudness across the ears (*Blamey et al., 2000*). In addition, loudness recruitment associated with hearing impairment may also affect the balance of loudness. As already mentioned, the dynamic range of people with loudness recruitment is reduced. When the level of a sound is increased above the absolute threshold, the growth of loudness as a function of sound level is larger than for NHLs. If there is a mismatch in loudness growth across different frequencies in residual hearing and the different electrodes in electric hearing, the balance of loudness across the ears may be further impaired. This mismatch in loudness may impair localisation and segregation of sound sources, since this in part depends on loudness differences across the ears.

1.5 Binaural stimulation

Currently, there is increased interest in facilitating binaural hearing for people with a unilateral CI by providing them either with a second implant or with a contralateral HA. Bilateral implantation is becoming more common, particularly for children and adults who have disabilities that increase their reliance on auditory information for spatial awareness (*NICE, 2009*). In the Netherlands, however, bimodal stimulation is preferred over bilateral implantation (*CVZ, 2009*). The increased interest in bilateral and bimodal fittings is largely due to the well-known binaural advantage normal-hearing listeners (NHL) experience when perceiving speech in noise. The next chapter deals with issues concerning the perception of speech in noise.

2. Speech in noise: normal-hearing listeners

2.1 The perception of speech in noise

The issue of perceiving speech in the presence of background noise was first addressed by Cherry (*1953*), who coined the term 'cocktail party problem'. The problem concerns understanding a speaker while ignoring the competing sound sources. This requires listeners to separate simultaneously presented frequency components into different sound sources and selectively attend to the source of interest.

When speech is presented in the presence of competing speech, two types of masking can occur (*Freyman et al., 1999; Brungart, 2001*). First, energetic masking occurs when the target and masker signals simultaneously contain energy in the same auditory filter. This peripheral process results in degraded perception of the target and affects the response along the BM to this signal. Informational masking, on the other hand, occurs at higher levels of the auditory system. In this case, both target and masker signals are audible, but the listener is unable to disentangle the two sounds. This type of masking can occur even when the spectra of the two signals do not overlap. An important factor in non-energetic masking is similarity of the masker and target signals. Increased dissimilarity between the signals decreases the effect of masking (*Freyman et al., 1999; Brungart, 2001; Durlach et al., 2003*).

Several factors which indicate dissimilarity between the target and masker signals have been shown to aid the segregation of the sound sources. Two factors are of particular relevance for this thesis. First, differences in the location in space of the target and masker signals can decrease the effect of informational masking. It is easier to perceive speech in the presence of background noise when the speech and noise sources are spatially separated (*Hirsh, 1950; Bronkhorst and Plomp, 1988; van de Par, 1998*). This particularly holds under binaural listening conditions, as it is easier to localise sounds using two ears rather than one. Second, differences in pitch have been shown to improve the perception of speech in a noisy background. Informational masking is decreased when the target and masker signals differ in terms of F0 (*Brokx and Nooteboom, 1982; Brungart, 2001*).

This thesis is primarily concerned with the effect of spatial separation under binaural listening conditions in a multi-talker background. Binaural advantages for normal-hearing listeners are described in more detail below.

2.2 Binaural advantages

The binaural advantage for perceiving speech in noise can be explained in terms of three different effects: binaural summation, head shadow, and binaural unmasking. The first two effects refer to physical phenomena, whereas the binaural unmasking effect involves neurological processes dealing with binaural interactions.

The effect of binaural summation, sometimes called redundancy, refers to the benefit of receiving two inputs of the same signal. Summation on average leads to an improvement of the speech reception threshold (SRT)[5] of 1-2 dB in NHLs (*c.f. Ching et al., 2006*). The head shadow effect refers to the fact that when speech and noise sources are spatially separated, the ear that is closer to the speech and further away from the noise source has a better SNR. Listeners can selectively attend to the ear with the best SNR. The improved SNR results from the obstruction of the head, which leads to an attenuation of the noise source in the contralateral ear. This effect is restricted to high frequencies, since short wavelengths do not exceed the size of the head and diffraction occurs. The head shadow effect can lower the SRT by approximately 9 dB for NHLs (*Bronkhorst and Plomp, 1988; Arsenault and Punch, 1999*).

Of particular interest with regard to the perception of speech in noise is the effect of binaural unmasking, sometimes called binaural squelch or spatial release from masking. This effect refers to the benefit of spatially separating the speech and noise sources. In order to disentangle sounds that come from different directions, the listener needs to localise the sources in space. Listeners can determine the location of a sound source along the frontal, median, and horizontal planes. While monaural cues are used to localise sounds along the frontal and median planes, binaural cues are used to determine the location of sounds along the horizontal plane. The direction of a sound along this plane is specified by degrees of azimuth (Fig. 7). When determining the directionality of a sound source, the brain is able to exploit interaural time differences (ITD) and interaural level differences (ILD). In principle, the sound reaching the farther ear is delayed and less intense compared to the ear closer to the source. The effect size of binaural unmasking for NHLs is approximately 5 dB (*Bronkhorst and Plomp, 1988; Arsenault and Punch, 1999*).

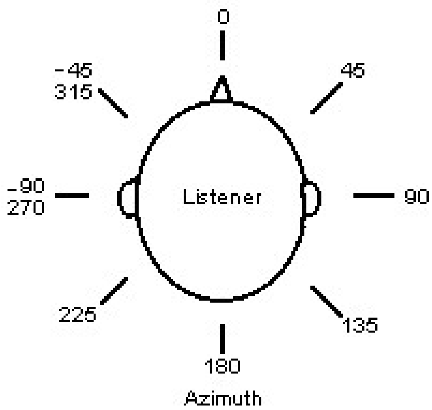

Fig. 7. Schematic representation of a head with locations along the horizontal plane specified by degrees of azimuth. A sound at 0° azimuth lies directly in front of the listener, for example, while a sound at 180° lies directly behind her. A sound at 90° lies at the right ear, whereas a sound with 270° (or -90°) lies at the left ear.

2.3.1 Interaural Time Differences

ITDs refer to the difference in arrival time of a sound between the two ears and are most informative at low frequencies, although they may also be used in high frequencies of complex waveforms (*Yost, 1974; Blauert, 1983; Henning, 1974; McFadden and Pasanen,*

[5] Speech perception can be assessed by measuring the signal-to-noise ratio required to achieve a particular level of performance. The speech reception threshold reflects the faintest level at which a speech segment can be heard and correctly repeated.

1976). In this thesis, the term ‘low’ will be taken to refer to frequencies below 1500 Hz and ‘high’ will be used for frequencies above 1500 Hz. Depending on the size of the head, ITDs range from 0 for a sound at 0° azimuth to about 600 μs for a sound at 90° azimuth (Fig. 8). NHLs can detect ITDs as small as 10 μs (*Mills, 1958*).

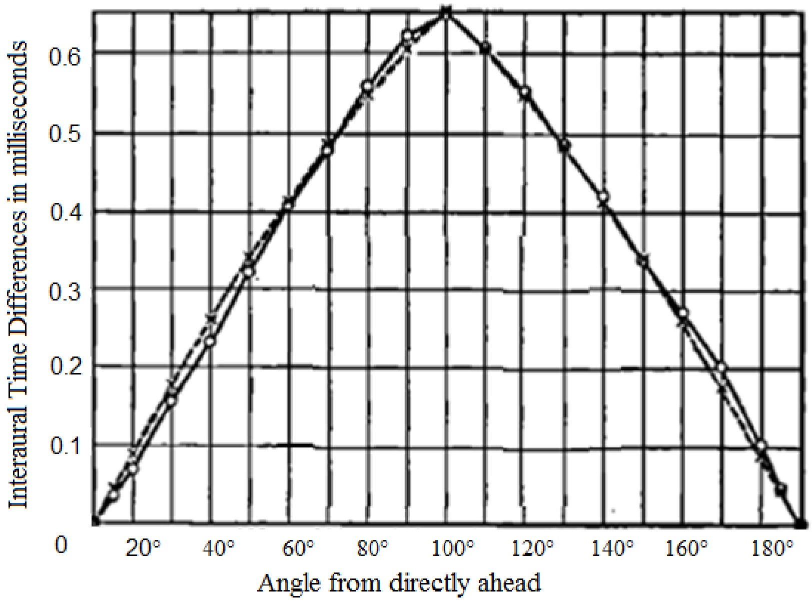

Fig. 8. Interaural time differences plotted as a function of azimuth. Measurements are for a rigid sphere with a radius of 8.75 cm (adapted from *Feddersen et al., 1957*).

The ITD can be calculated by means of the following formula:

$$\text{ITD} = \frac{r\theta + r\sin\theta}{s} \qquad -90^\circ \leq \theta \leq +90^\circ$$

where r is the radius of the head (approximately 9 cm), θ refers to degrees of azimuth and s is the speed of sound in cm/s (*c.f. Moore, 2008b*). The numerator describes the path the sound needs to travel between the two ears. This value is divided by the speed of sound to derive the ITD. Given that the speed of sound is 34300 cm/s at 20˚C, it takes 29 µs for a sound to travel 1 cm under these conditions.

ITDs for sinusoids are most informative at low frequencies. A sinusoidal wave of 1500 Hz has a wavelength about equal to the width of the head. When the wavelength of a sound is shorter than the distance between the ears (i.e., above 1500 Hz), phase ambiguities occur, since the same interaural phase difference could result from a number of different source locations. At low frequencies, listeners derive ITDs encoded in the TFS (*Henning, 1983*). TFS information depends on phase locking to frequency components in the waveform. In most mammals this is limited to frequencies up to 4 – 5 kHz (*Moore, 2008a, 2008b*). In fact, ITDs in the fine structure can only be detected up to about 1.3 kHz (*Zwislocki and Feldman, 1956*). At higher frequencies, TFS information is not encoded, since neural firing moments no longer respond to a particular phase in the stimulating waveform. The importance of ITD cues at low frequencies encoded by TFS is illustrated by Bernstein and Trahiotis (*1985a, 1985b*). They point out that, at low frequencies, delays in the envelope can be used to signal ITDs but that they interact with the dominant cues in TFS. In addition, they show that sounds of which the entire waveform (envelope and TFS) is delayed are localised more accurately than when only the high frequencies (where TFS is not available) signal an interaural delay.

Sensitivity to ITDs, however, has also been shown for frequencies above 1500 Hz in complex waveforms (*Henning, 1974; McFadden and Pasanen, 1976*). In complex waveforms, listeners are able to recruit ITDs in the envelope which are present at high frequencies. Henning (*1974*) studied the ability to detect ITDs for tones in which either the entire waveform or just the envelope of the amplitude modulated waveform was delayed. He showed that, for high frequencies, the information in the envelope rather than the fine structure within the waveform determines the apparent location of a sound source.

Models of binaural interaction (*e.g., Jeffres, 1948; Durlach, 1963; Colburn, 1973; Trahiotis and Stern, 1989*) assume that ITDs can only be derived when the outputs from nerve fibres with corresponding CFs are compared. Magezi and Krumbholz (*2008*) examined whether ITDs in the TFS of non-corresponding frequency channels could be extracted. They presented listeners with a 500 Hz tone which was partially masked by a low-pass masker in one ear and by a high-pass masker in the other. Their results show that performance in an ITD discrimination task was poor when listeners had to extract ITD information from non-corresponding frequency channels. These findings suggest that the ability to segregate concurrent sound sources is most effective when the frequency components for the signals in the two ears are identical. Magezi and Krumbholz (*2008*) suggested, however, that listeners might be able to derive ITDs from mismatched frequency channels when these channels are limited to a narrow frequency range of little more than one auditory-filter bandwidth.

The processing of ITDs involves activation of different groups of cells in the auditory brainstem. A group of cells in the medial superior olive (MSO) is especially sensitive to ITDs in the TFS in the low frequency range. A different group of cells, located in the lateral superior olive (LSO), is sensitive to envelope delays in the high frequency range as well as to onset delays (*Palmer, 1995*). Rose et al. (*1966*) showed that each cell responds to pure tones of different frequencies at a particular interaural delay. They termed the fixed interaural delay the ‘characteristic delay’ of these cells. The majority of the cells has a characteristic delay which falls within the natural range of interaural delays (i.e. up to approximately 700 µs) (*Kuwada and Yin, 1983*).

2.3.2 Interaural Level Differences

ILDs are most useful at high frequencies and refer to a difference in sound level (dB) between the two ears (*Feddersen et al., 1957; Blauert, 1983).* Level differences across the ears are the result of the acoustic shadow of the head, which attenuates high frequencies. ILDs are not as easily calculated in terms of degrees of azimuth of a particular sound source, since they are largely frequency dependent. Figure 9 shows the relationship between ILDs and the direction of the sound source in azimuth for a variety of frequencies. The figure indicates that ILDs are negligible below 500 Hz, become useful above 1800 Hz, and may be as large as 20 dB for frequencies at the upper end of the scale. NHLs have JNDs of 1 dB at 1000 Hz and 0.5 dB for higher frequencies (*Mills, 1958, 1960*).

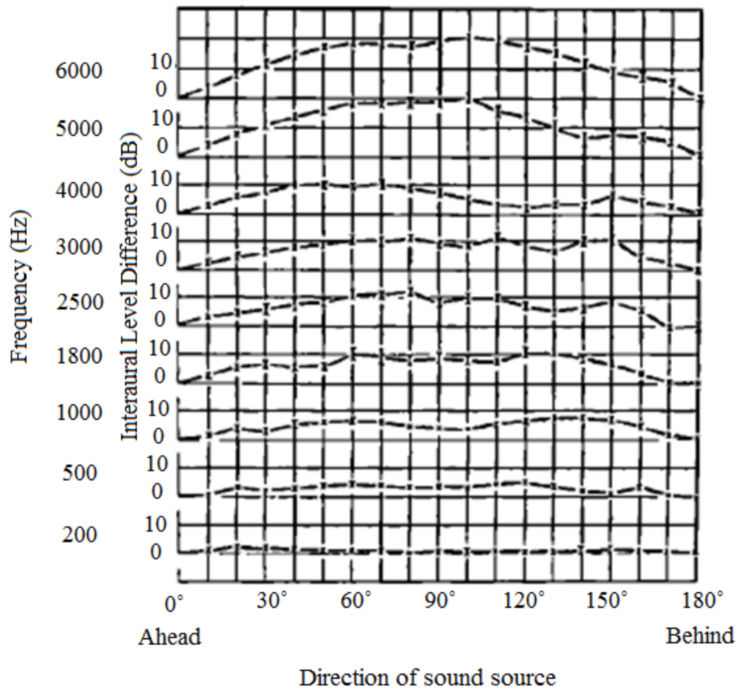

Fig. 9: Interaural intensity differences as a function of frequency and direction. Measurements are for a rigid sphere with a radius of 8.75 cm and using a distant source (*Feddersen et al., 1957*).

Even though ILDs mainly occur at high frequencies, listeners are also sensitive to ILDs in the low frequency range. When a sound source is located close to the listener's head (within 1 m), ILDs are also detectible below 1500 Hz (*Brungart and Rabinowitz, 1999*). As the distance between the source and the head decreases below 1 m, ILDs across the whole frequency range increase. Brungart and Rabinowitz (*1999*) suggest that in close proximity high as well as low frequencies are attenuated by the obstruction of the head. As the distance between the sound source and the head decreases, the area affected by a head shadow effect increases. Consequently, complex waveforms are increasingly affected by the obstruction of the head.

The processing of ILDs involves activation of a group of cells in the auditory brainstem. These cells are located in the LSO and are innervated primarily by neurons with CFs in the high frequency range. Note that these are the same cells that are sensitive to onset delays and envelope ITDs. The cells are sensitive to the balance of intensity at the ears and inhibit the response to an ipsilateral signal when the level of the contralateral input is increased (*c.f. Palmer, 1995*).

2.3.3 Relative importance of ITD and ILD cues

The relative contribution of ITD and ILD information to the localisation of a sound source was first described in terms of the 'duplex theory' (*Lord Rayleigh, 1907*). This theory states that ITD information is most useful at low frequencies and ILD information at high frequencies. This was illustrated by the fact that for pure tones ITDs are only useful below

1500 Hz, whereas ILDs only become useful above 1800 Hz. The organisation of groups of cells in the auditory brainstem seems to support the duplex theory. Cells in the MSO are innervated by neurons sensitive to low frequencies, and cells in the LSO are primarily associated with high frequencies. However, the theory is not strictly accurate since listeners are sensitive to ITD and ILD information across a range of frequencies.

It has been shown that time and level differences are not entirely equivalent. When ILD and ITD cues indicate different locations of the sound source, two sound images may be heard. Whitworth and Jeffress (*1961*) examined the trading ratio for pure tones and found that the image which primarily depended on ITD information was little affected by ILDs. They found a trading ratio of 1 µs/dB. A trading ratio of 20 µs/dB was found, however, for the ILD-dependent image, indicating that a relatively larger time difference is required to offset a 1-dB difference. The finding that time and level differences are not entirely equivalent seems to suggest that ITDs and ILDs may similarly not be of equal importance with respect to sound source localisation and the binaural unmasking effect.

It turns out that ITDs are the major cue for sound localisation on the horizontal plane. Wightman and Kistler (*1992*) examined the relative importance of ITDs and ILDs using broadband noise sources in which the ITD cues signalled one direction and the ILD and pinna cues signalled another direction. They showed that listeners almost always localised the perceptual image according to the location signalled by the ITD. ILD cues were the dominant cue only in the absence of low frequency components in the signal. These findings were consistent for stimuli presented over headphones and in the free field.

Despite the fact that ITDs have been shown to be the dominant cue for sound localisation, listeners do not depend on this cue to group simultaneous frequency components and to segregate concurrent sound sources (*Culling and Summerfield, 1995; Darwin and Hukin, 1999*). Since frequency components from the same direction share the same ITD (*Kuhn, 1977*), it was assumed that across-frequency grouping by common ITD might be useful for separating sound sources. However, Culling and Summerfield (*1995*) showed that listeners cannot group simultaneous formant-like noise bands with a common ITD. Similarly, Darwin and Hukin (*1999*) showed that listeners are unable to segregate a harmonic from a target vowel when the two objects differ in terms of ITD. Darwin and Hukin (*1999*) assume that selective attention toward auditory objects is not due to the grouping of frequency components that share a common ITD. Instead, they argue that selective attention can be attributed to the subjective locations of the auditory objects.

In fact, it may be the case that ILDs are the dominant cue with respect to the binaural unmasking effect. Bronkhorst and Plomp (*1988*) studied the relative contributions of ITDs and ILDs for the perception of speech and noise by NHLs. They found that ITD alone could lower SRTs by up to about 5 dB, ILD alone by up to about 8 dB and the combination by up to as much as 10 dB. These findings seem to suggest that ILDs are relatively more important for the binaural unmasking effect, since ILDs could lower SRTs to a larger extent than ITDs.

3. Speech in noise: users of hearing devices

3.1 Binaural advantages

Unilateral CI users generally perform relatively well in quiet listening situations, but their performance rapidly deteriorates in the presence of background noise (*Stickney et al., 2004a*). Performance is likely to improve when CI users can rely on input provided to two ears as opposed to one. Previous studies have shown that the perception of speech in noise improves when unilateral CI users are provided with a second implant (*e.g., Van Hoesel and Tyler, 2003; Garadat et al., 2009; Mosnier et al., 2009*). Similarly, the perception of speech in noise has also been shown to improve when unilateral CI users are provided with a contralateral HA (*e.g., Kong, 2005; Ching et al., 2006, 2007; Cullington and Zeng, 2010*). There is an ongoing debate about whether unilateral CI users should be provided with a second implant or a contralateral HA (*NICE, 2009; CVZ, 2009*). Bilateral implantation is becoming more common (*NICE, 2009*), although providing people with a second implant is relatively expensive. It has in fact been argued that the quality of life which is likely to be gained per unit of expenditure is smaller for bilateral implantation than for unilateral implantation (*Summerfield et al., 2002; CVZ, 2009*). Bimodal stimulation is a more cost-effective alternative to provide unilateral CI users with binaural auditory input (*c.f. Offeciers et al., 2005*).

The question that needs to be answered is to what extent people with bilateral and bimodal fittings can benefit from binaural effects when perceiving speech in noise. In order to answer this question, two simulation experiments were conducted. In addition, these experiments aimed to shed some light on the debate about whether bilateral or bimodal fittings are more advantageous for the perception of speech in noise. The first experiment looked at the effects of binaural summation, head shadow and binaural unmasking for CIs, HAs and bimodal hearing devices (section 3.2). Since there are various ways of simulating CIs, the second experiment examined the binaural benefits for the bilateral CI condition using a different simulation technique. This technique is likely to preserve timing information and pitch cues more accurately. In addition, the experiment examined more closely the relative importance of ITD and ILD information with regard to the binaural unmasking effect in the bilateral CI condition (section 3.3).

3.2 Study I

This study aims to determine the perceptual advantage of binaural summation, head shadow and binaural unmasking for speech in noise using hearing devices. Previous studies have examined the perceptual advantages of these effects for NHLs. However, the extent to which users of hearing devices can benefit from binaural stimulation remains largely unclear. In addition, the experiment addresses the question of whether bilateral or bimodal fittings are more beneficial with respect to the perception of speech in noise.

3.2.1 Listeners

Twelve normal hearing listeners (2 male, 10 female) participated in the experiment. All participants were monolingual native speakers of English. They ranged in age from 19 to 36 years (mean 23 years). All listeners had normal hearing, defined as pure-tone thresholds of 20 dB HL or better at octave frequencies between .5 and 8 kHz. None of the participants had any prior experience with the test materials. Participants signed a consent form approved by UCL/UCLH Committee on the Ethics of Human Research and were paid for their participation. The experiment consisted of a single two-hour session.

3.2.2 Stimulus materials and conditions

The target stimuli used in this experiment consisted of a selection of pre-recorded BKB sentences *(Bench, Kowal and Bamford, 1979*) containing three key words each. In addition, a selection of pre-recorded sentences from the Adaptive Sentence List (*MacLeod and Summerfield, 1990*) was used. These sentence lists are comparable in structure and more or less equally difficult. The target stimuli were all produced by the same male speaker.

Participants were tested on the perception of speech with simulated CIs and bimodal fittings with a CI in the left ear and a HA in the other ear (CI-HA). In addition, performance was also examined with simulated HAs, since this may shed some light on the results for the bimodal conditions (CI-HA). The stimuli were presented both monaurally, in the left as well as the right ear alone, and binaurally.

All sentences were presented in a background of twenty-talker babble noise. Speech was always presented at 0° azimuth (N0). The noise source was either presented at the same angle or spatially separated from the target speech, at 90° azimuth (N90). In the bimodal condition, the noise was also presented at -90° azimuth (N-90). The noise would thus be presented ipsilaterally to the CI, at -90° azimuth, as well as to the HA, at 90° azimuth.

Participants were tested in 13 different conditions, each with a unique combination of hearing condition, listening mode and noise azimuth (Table 1).

	Hearing condition	Listening mode	Noise azimuth
1	CI	left ear	N0
2	CI	left ear	N90
3	CI	right ear	N90
4	CI	both ears	N0
5	CI	both ears	N90
6	HA	left ear	N0
7	HA	left ear	N90
8	HA	right ear	N90
9	HA	both ears	N0
10	HA	both ears	N90
11	CI-HA	both ears	N0
12	CI-HA	both ears	N90
13	CI-HA	both ears	N-90

Table 1. Test conditions, each with a unique combination of hearing condition, listening mode and noise azimuth.

The stimuli were presented in blocks organised according to hearing condition (HA, CI, CI-HA). Using a Latin square design, the order of the blocks was counterbalanced across participants. In addition, the order of conditions within each block was also counterbalanced. Participants listened to 20 sentences in each condition. In order to rule out possible learning effects they were tested on all conditions twice. A short break was introduced after all conditions had been presented for the first time. In the second half of the experiment all conditions were presented again, this time in reversed order. The sentence lists (BKB and ASL) were also presented in blocks. Had the participant been presented with BKB sentences in the first half, he or she would subsequently listen to ASL sentences and vice versa.

3.2.3 Signal processing

The stimuli were created within the MATLAB environment, as follows. First, the spatial configurations were applied, using a spherical head model (*Brown and Duda, 1998*) which simulates head-related transfer functions (HRTFs) to synthesize binaural sounds from monaural sources. HRTF refers to the way a sound source is filtered by the diffraction and reflection of the head, pinnae, and torso, before it reaches the eardrum. The model applies a head shadow model simulating ILDs and a time delay model simulating ITDs. As the name of the model suggests, it provides a simplification of reality, since the head is represented as a sphere. The simplicity of the model, however, allows efficient implementation and facilitates online processing (*Brown and Duda, 1998*).

The spherical head model has several parameters which can be adjusted, such as the location of the ears and the size of the head. It would thus be possible to generate individualized HRTFs. In this thesis it is assumed that the ears of the participants are located at 100˚ azimuth and that the radius of the head is 8.75 cm, which is commonly cited as the average radius for an adult human head (*Brown and Duda, 1998; Kuhn, 1977*).

The spherical head model converted mono input files into stereo files with speech and noise waveforms presented at a specified azimuth. The speech and noise sources were then added together. A section of the noise was randomly selected from the 270 s available. An additional 100 ms was presented before the speech signal started.

Subsequently the signals were manipulated to sound as if they were heard through a CI or a HA. First, the stimuli simulating a CI covered the frequency range between 100 and 7500 Hz. The signals were bandpass filtered into 6 contiguous channels. This falls within the number of effective channels expected in actual CI users (*Wilson and Dorman, 2008*). The channels were created using a second-order Butterworth filter. This type of filter is designed to have a frequency response which is maximally flat in the passband and rolls off towards zero in the stopband (*Butterworth, 1930*). The order of the filter refers to the decrease in the response in dB per octave. Each order means a decrease of 6 dB per octave. The filter used in this study thus resulted in a decrease of the frequency response of 12 dB per octave. The different channels were spaced according to the frequency-to-place mapping along the BM proposed by Greenwood (*1990*). Band centre and cutoff frequencies are shown in Table 2. The envelope was extracted from each frequency band by half-wave rectification and smoothing at 400 Hz. Subsequently, the envelopes were used to amplitude-modulate a white noise source filtered to the channel bandwidth. Independent noise sources were used to simulate CIs in the two ears. Secondly, the stimuli simulating the HA condition were low pass filtered at 1 kHz, using a 20th order Butterworth filter, resulting in a decrease of the frequency response of 120 dB per octave above 1000 Hz. A steep filter was used to limit the influence of high frequencies in this condition. The choice to simulate the HA by filtering out high frequencies is based on the assumption that hearing generally only remains at low frequencies.

Channel	Lower bound	Upper bound	CF
1	100.0	299.5	199.7
2	299.5	648.9	474.2
3	648.9	1260.9	954.9
4	1260.9	2333.0	1797.0
5	2333.0	4210.8	3271.9
6	4210.8	7500.0	5855.4

Table 2. Filter spacing of the CI simulation (6 channels, 100 – 7500 Hz). The lower and upper bounds, as well as the centre frequencies (CF), are given for each channel.

3.2.4 Procedure

The participants were seated in a soundproof booth and listened to the stimuli over headphones (Sennheiser HD 414). The participants' task was to repeat verbatim what they heard. The experimenter was seated in the same room and scored responses using a graphical user interface (GUI) which showed the three key words. The scoring screen was not visible to the participants and no feedback was provided.

The design of the experiment followed the adaptive-up procedure as described in Plomp and Mimpen (*1979*). The first sentence was presented at a SNR of -10 dB. Until at least 2 out of 3 key words were repeated correctly, SNR was increased by 10 dB for the next presentation. The sentence was repeated until at least 2 key words were correct or SNR reached 30 dB. For each subsequent stimulus the change vector was set to increase SNR by 3 dB when at least 2 out of 3 key words were repeated correctly and decrease SNR by 3 dB when this was not the case. Since the number of trials was fixed, the number of reversals varied across participants and conditions. SRTs were tracked at 50%.

Participants received training on the binaural HA, CI, and CI-HA conditions to familiarize themselves with the task and to get used to the listening conditions. The target stimuli consisted of a set of sentences in the form of a short story. In addition, participants received training on one BKB list, using unprocessed speech, to familiarize themselves with the type of sentences used in the experiment proper. The target sentences were presented in twenty-talker babble noise and the speech and noise sources were spatially separated. Speech was always presented at 0° azimuth and noise at 90° azimuth. In addition, noise was presented at -90° azimuth in the CI-HA condition. Each participant received 5 minutes of training on each hearing device condition and listened to one BKB list in the unprocessed condition, resulting in 20-25 minutes of training in total.

3.2.5 Results

Figure 10 summarises SRT scores (in dB) for each test condition. SRTs were obtained by calculating average SNRs of all reversals (i.e. those leading to an increase or decrease in SNR in the subsequent stimulus).

The effect of the different conditions was tested by means of mixed-effects modelling. The advantage of this technique is that participants and sentence lists can be treated as random effects simultaneously. In addition, an advantage of this method compared to repeated measures analysis of variance (RM-ANOVA) is that it does not require sphericity. Moreover, mixed-effects models are robust against missing data (*for details, see e.g., Quené and Van den Bergh, 2004; Baayen et al., 2008*).

The mixed-effects model used for this analysis contains ‘condition’ as a fixed factor and includes ‘participants’ and ‘sentence lists’ as two random factors. It turns out that it is indeed useful to include these random effects since approximately 13% of the variance within the data set can be attributed to variance between participants and 17% of the total variance is due to variation between the sentence lists. The remaining 70% can be explained by a combination of condition effects and a residual component. The random effects are taken into account in the estimation of the size of the condition effects.

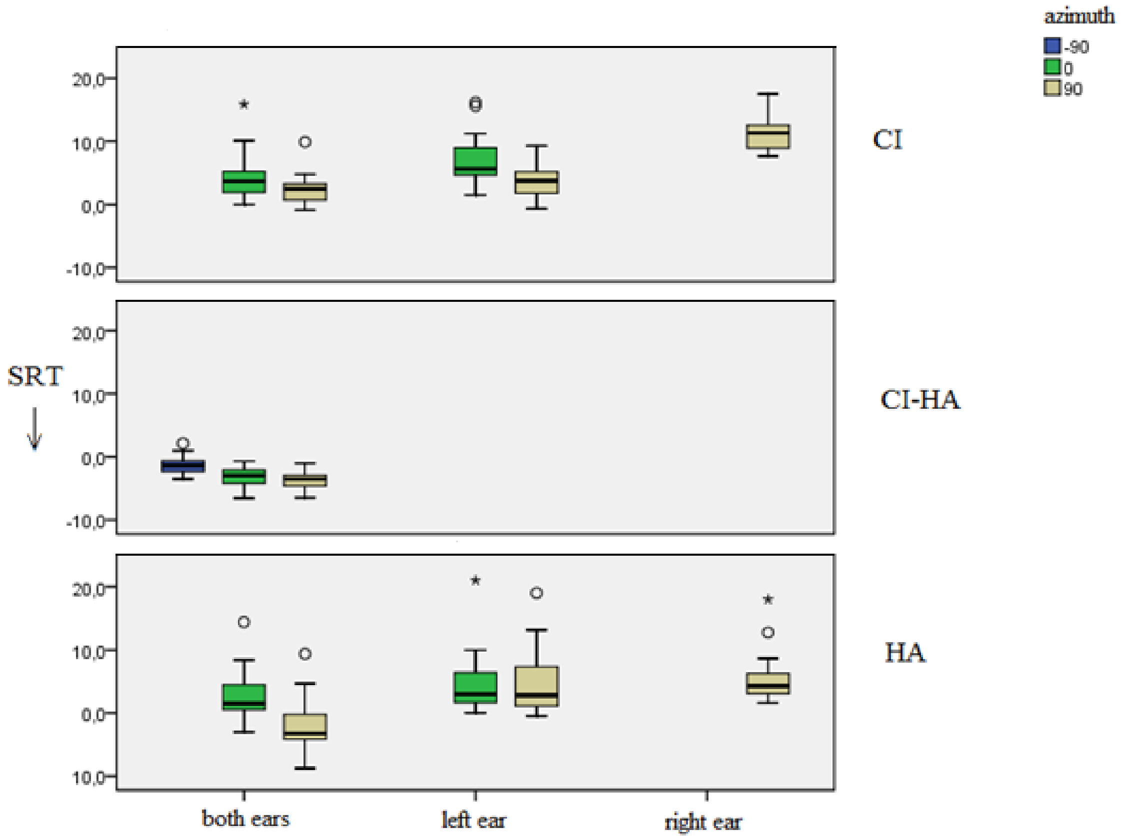

Fig. 10. Average SRT scores (in dB) are plotted for the HA (top), bimodal (CI-HA; middle) and CI listening conditions (bottom), broken down for presentation mode (binaural 'both ears', monaural 'left ear', monaural 'right ear'; x-axis) and noise azimuth (colour). Green indicates that the noise source is at 0° azimuth (N0), olive green indicates that the noise is at 90° azimuth (N90), and blue indicates that the noise is at -90 (N-90).

Pairwise Comparisons Study I

Contrast	(I)Condition	(J)Condition	Mean Difference (I-J) (dB)	Sig.
1	CI left ear N90	CI right ear N90	**-7.3***	**.000**
2	CI both ears N90	CI both ears N0	**-1.5**	**.064**
3	CI both ears N0	CI left ear N0	**-3.4***	**.000**
4	CI-HA both ears N90	CI-HA both ears N0	**-0.1**	**.953**
5	CI-HA both ears N-90	CI-HA both ears N0	**1.9***	**.016**
6	HA both ears N90	HA both ears N0	**-4.5***	**.000**
7	HA left ear N90	HA right ear N90	**-0.5**	**.540**
8	HA both ears N0	HA left ear N0	**-1.9***	**.019**
9	CI-HA both ears N0	CI both ears N0	**-7.2***	**.000**
10	CI-HA both ears N90	CI both ears N90	**-5.8***	**.000**
11	CI-HA both ears N-90	CI both ears N90	**-3.8***	**.000**

*. The mean difference is significant at the .05 level.

Table 3. Pairwise comparisons of mixed-effects modelling with participants and sentence lists as two crossed random effects. Mean differences in SRT between selected conditions (in dB) are provided, along with the p-value.

This study examined to what extent listeners with simulated bilateral and bimodal fittings can benefit from a binaural advantage. The effects of binaural unmasking, head shadow and summation are tested for all listening conditions by means of pairwise comparisons. Table 3 gives an overview of the contrasts of interest.

Firstly, the largest effect found in the CI condition was a head shadow effect. This effect can be assessed by comparing SRTs for left and right ear only when the speech and noise sources are spatially separated (contrast 1). SRTs for the right ear were 7.3 dB higher than those found for the left ear ($p < 0.0001$). The binaural unmasking effect can be examined by comparing binaural situations where speech and noise are spatially separated or coincide in space (contrast 2). Only a small binaural unmasking effect of 1.5 dB was found, which approached significance ($p = .064$). The effect of summation is computed by comparing monaural and binaural conditions (contrast 3), when speech and noise coincide in space. The results reveal an effect of summation of 3.4 dB ($p < 0.0001$).

Secondly, the effect of binaural unmasking was assessed in the bimodal listening conditions. When the noise was located contralaterally to the CI ear (N90), no effect of unmasking was detected (contrast 4). However, when the noise was located contralaterally to the HA ear (N-90), SRT scores were 1.9 dB higher than when the speech and noise sources coincided in space ($p < 0.05$) (contrast 5). These results suggest an effect of masking, rather than unmasking, when the noise is at the CI ear.

Thirdly, the results for the HA listening condition suggest that the largest advantage is to be had in terms of binaural unmasking. The data reveal an unmasking effect of 4.5 dB ($p < 0.0001$) (contrast 6). Since the HA condition was simulated using a low-pass filter at 1000 Hz, the effect is probably mainly due to ITD cues and small ILDs present at low frequencies. The data indicate that no head shadow effect is present in the HA condition ($p = .540$) (contrast 7). This finding can be explained by the fact that frequencies above 1 kHz were filtered from the signal, whereas the head shadow effect is restricted to high frequencies. A summation effect of 1.9 dB was found ($p < 0.05$) when monaural and binaural conditions with the speech and noise sources at 0° azimuth were compared (contrast 8).

Finally, SRT scores are compared for bilateral and bimodal listening conditions. The results show that listeners performed significantly better with simulated bimodal hearing devices than with simulated bilateral implants. When speech and noise coincide in space (contrast 9), listeners derive an advantage of 7.2 dB with simulated bimodal fittings ($p < 0.0001$). Similarly, when speech and noise are spatially separated, SRTs are significantly lower with bimodal hearing devices than with bilateral implants. There was an advantage of 5.8 dB with bimodal fittings when the noise source was located at 90° azimuth ($p < 0.0001$) (contrast 10). SRTs were 3.8 dB lower for the bimodal listening condition when the noise source was located at -90° azimuth ($p < 0.0001$) (contrast 11).

3.3 Study II

Since in Study I the binaural unmasking effect in the bilateral CI condition only approached significance, a second study was conducted using a different simulation technique. Study I used a noise carrier and a smoothing filter of 400 Hz to simulate the CI. It may be the case, however, that this technique does not accurately preserve ITD and ILD cues. In this second experiment, CIs were therefore also simulated using a noise-vocoder with a smoothing filter of 50 Hz and a tone-vocoder with a smoothing filter of 50 Hz. Timing cues in particular may be more accurately preserved using these simulation techniques[6].

In addition, the experiment examined more closely the relative importance of ITD and ILD information with regard to the binaural unmasking effect in the bilateral CI condition. It is generally assumed that ITD information is not well preserved in signals transmitted by a CI. Hence, it was expected that the small effect of binaural unmasking which was detected in Study I should primarily be attributed to the presence of ILD cues. In order to test this hypothesis, listeners were presented with stimuli that contained either ITD or ILD information alone.

3.3.1 Listeners

A second group of listeners was recruited to participate in this study. Twelve normal hearing listeners (12 females) participated in the experiment. As before, all participants were monolingual native speakers of English. They ranged from 18 to 30 years (mean 22 years). None of the participants had any prior experience with the test materials. All listeners had normal hearing, defined as pure-tone thresholds of 20 dB HL or better at octave frequencies between .5 and 8 kHz. Participants signed a consent form approved by UCL/UCLH Committee on the Ethics of Human Research and were paid for their participation. The experiment consisted of a single two-hour session.

3.3.2 Stimulus materials and conditions

The target sentences were identical to those used in study I. Perceptual abilities were tested simulating CIs by means of noise- and tone-vocoding. The tone-vocoded stimuli were presented both monaurally (left and right ear alone) and binaurally. Noise-vocoded stimuli, on the other hand, were only presented binaurally since the head shadow and binaural summation effects using a noise carrier were already assessed in study I. In addition, the stimuli differed in terms of the smoothing filter used. The tone-vocoded stimuli were only presented with a smoothing filter of 50 Hz. The noise-vocoded stimuli were created using filters of 400 (sm400) and 50 Hz (sm50).

As in Study I, all sentences were presented in a background of twenty-talker babble noise. Speech was always presented at 0° azimuth. The noise source was either presented at the same angle or spatially separated from the target speech, at 90° azimuth.

An additional condition was introduced which concerned the availability of spatial cues. This variable was introduced for binaurally presented stimuli in which speech and noise were spatially separated. Stimuli contained either ITD cues, ILD cues, or both.

Participants were tested in 13 different CI conditions, each with a unique combination of type of vocoding, smoothing filter, listening mode, noise azimuth and the availability of

[6] See section 1.3.4 for a brief overview of various ways to simulate a CI.

spatial cues (Table 4). Stimuli were presented in blocks organised according to the type of vocoding. The organisation of different conditions was as in study I, using a Latin square design.

	Vocoding	Smoothing filter	Listening mode	Noise azimuth	Spatial cues
1	Noise	Sm400	both ears	N0	ITD + ILD
2	Noise	Sm400	both ears	N90	ITD + ILD
3	Noise	Sm50	both ears	N0	ITD + ILD
4	Noise	Sm50	both ears	N90	ITD + ILD
5	Noise	Sm50	both ears	N90	ITD
6	Noise	Sm50	both ears	N90	ILD
7	Tone	Sm50	both ears	N0	ITD + ILD
8	Tone	Sm50	both ears	N90	ITD + ILD
9	Tone	Sm50	both ears	N90	ITD
10	Tone	Sm50	both ears	N90	ILD
11	Tone	Sm50	left ear	N0	ITD + ILD
12	Tone	Sm50	left ear	N90	ITD + ILD
13	Tone	Sm50	right ear	N90	ITD + ILD

Table 4: Test conditions, each with a unique combination of type of vocoding, smoothing filter, listening mode, noise azimuth and the availability of spatial cues.

3.3.3 Signal processing

The processing techniques used to create the stimuli were similar to those in study I. However, some changes were made with respect to the implementation of spatial configurations of the speech and noise sources as well as the techniques recruited for the simulation of the CI.

The spatial configurations were again applied using the spherical head model developed by Brown and Duda (*1998*). As was already mentioned, this model consists of a head shadow model simulating ILDs and a time delay model simulating ITDs. A parameter was added to the program which allowed separate implementation of the head shadow and time delay models. As a result, ILD or ITD alone could be presented in the signal.

Subsequently, the stimuli were simulated to sound as if they were heard through a CI. As before, the signals covered a frequency range of between 100 and 7500 Hz and were bandpass filtered into 6 contiguous channels. In this study, however, the channels were created using a steeper filter, namely a sixth-order Butterworth filter. This results in a frequency response which decreases by 36 dB per octave instead of 12 dB as in the previous study (*c.f. Bianchi and Sorrentino, 2007*). A steeper filter was used to ensure no F0-related modulations would occur. Even when a low cutoff is used in the envelope, a shallow filter does not attenuate F0 components completely. The envelope was extracted from each frequency band by full-wave rectification and smoothing at 400 Hz or 50 Hz. This study used full-wave rectification since it was thought that this might lead to less F0-related modulation in the envelope. In addition, for high cutoff frequencies full-wave rectification was expected to result in a clearer representation of periodicity. The envelopes were then used to modulate a white noise source, which was then filtered to the channel bandwidth, or a sine wave at the centre frequency of the band. All stimuli were presented at 70 dB SPL.

3.3.4 Procedure

The procedure of the experiment was as before. However, changes were made with respect to the training prior to the experiment.

Participants received training on the noise- and tone-vocoded CI conditions. In addition, listeners were presented with unprocessed speech to familiarize themselves with the type of sentences used in the test phase. The training for the two CI conditions was done using target sentences in the form of a short story. Participants trained last on the condition with which they would start in the experiment proper. In addition, they were trained on one BKB list, using unprocessed speech, to familiarize themselves on the type of sentences used in the test phase. As in study I, the target sentences were presented in twenty-talker babble noise and the speech and noise sources were spatially separated. Each participant received 10 minutes of training on both CI conditions and listened to one BKB list in the unprocessed condition, resulting in approximately 20-25 minutes of training in total.

3.3.5 Results

Figure 11 summarises SRTs for all noise-vocoded test conditions. SRTs for the tone-vocoded test conditions are presented in figure 12. SRTs were again obtained by calculating average SNRs over all reversals.

As before, the effects of the different conditions were tested by means of mixed-effects modelling. It turns out that approximately 30% of the variance within the data set can be attributed to variance between participants and 7% is due to variation between the sentence lists. The remaining 63% can be explained by a combination of condition effects and a residual component.

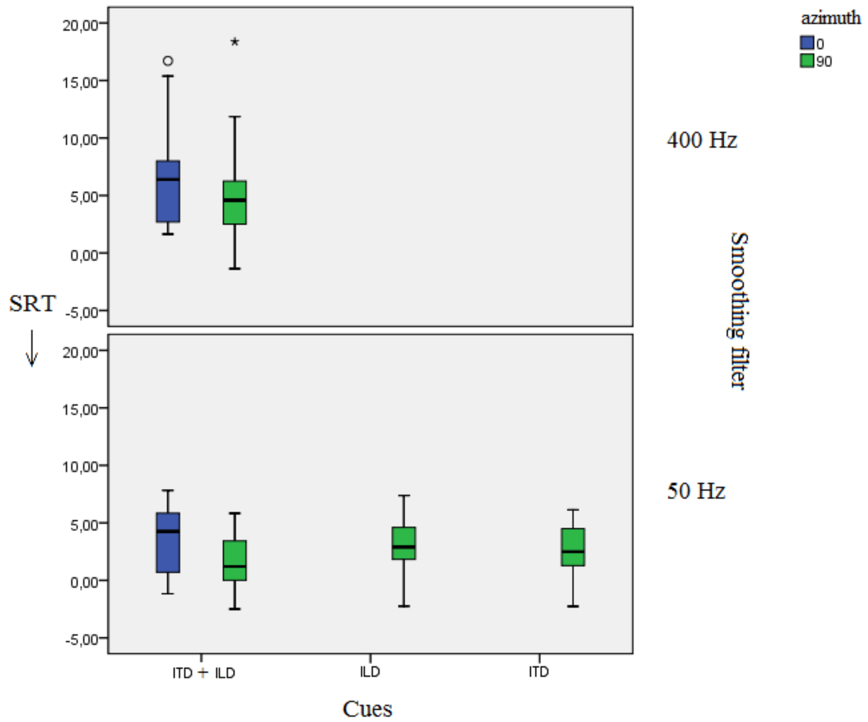

Fig. 11. Average SRT scores (in dB) are plotted for the noise-vocoded conditions, with a smoothing filter of 400 Hz (top) or 50 Hz (bottom), broken down for binaural unmasking cues (ITD and ILD, ITD alone, ILD alone; horizontal axis) and noise azimuth (colour). Blue indicates that the noise source is at 0° azimuth (N0), and green indicates that the noise is at 90° (N90).

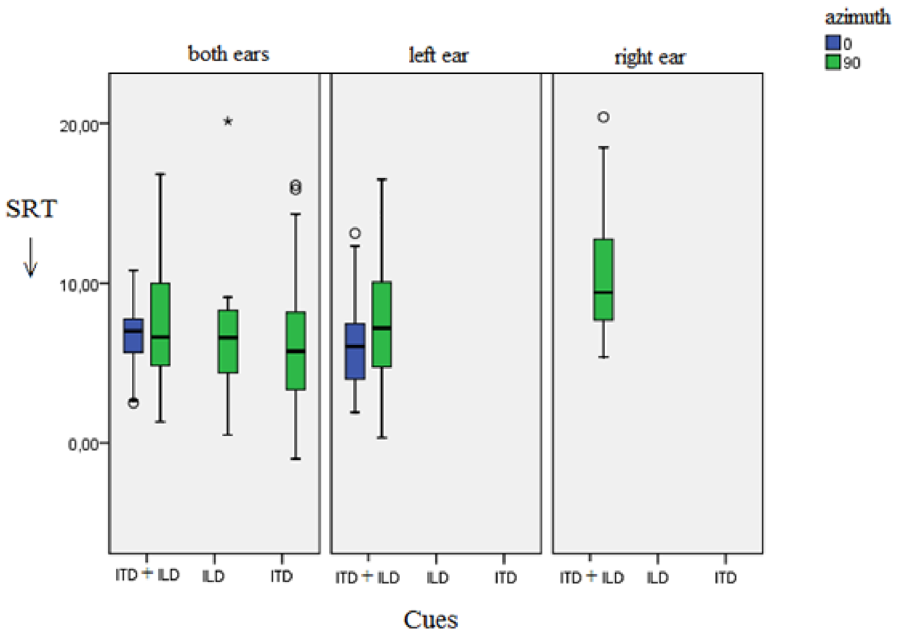

Fig. 12. Average SRT scores (in dB) are plotted for the tone-vocoded conditions, under binaural or monaural (left, right) listening conditions (horizontal axis), broken down for binaural unmasking cues (ITD and ILD, ITD alone, ILD alone) and noise azimuth (colour). Blue indicates that the noise source is at 0° azimuth (N0), and green indicates that the noise is at 90 (N90).

Pairwise Comparisons Study II

Contrast	(I)Condition	(J)Condition	Mean Difference (I-J)	Sig.
1	Noise Sm400 both ears N90 ITD+ILD	Noise Sm400 both ears N0 ITD+ILD	**-1. 9***	**.034**
2	Noise Sm50 both ears N90 ITD+ILD	Noise Sm50 both ears N0 ITD+ILD	**-1.3**	**.124**
3	Noise Sm400 both ears N90 ITD	Noise Sm400 both ears N0 ITD+ILD	**-0.9**	**.331**
4	Noise Sm400 both ears N90 ILD	Noise Sm400 both ears N0 ITD+ILD	**-0.5**	**.553**
5	Tone both ears N90 ITD+ILD	Tone both ears N0 ITD+ILD	**0.1**	**.893**
6	Tone both ears N90 ITD	Tone both ears N0 ITD+ILD	**-0.8**	**.342**
7	Tone both ears N90 ILD	Tone both ears N0 ITD+ILD	**-0.6**	**.500**
8	Tone left ear N90 ITD+ILD	Tone right ear N90 ITD+ILD	**-3.4***	**.000**
9	Tone both ears N0 ITD+ILD	Tone left ear N0 ITD+ILD	**1.0**	**.235**

*. The mean difference is significant at the .05 level.

Table 5. Pairwise comparisons of mixed-effects modelling with participants and sentence lists as two crossed random effects. Mean differences in SRT between selected conditions (in dB) are provided, along with the p-value.

This study focused on the binaural unmasking effect for listeners with simulated bilateral implants. The effect was examined using various techniques to simulate the CI signal. Particular attention was paid to the relative contribution of ITDs and ILDs to the effect of binaural unmasking. In addition, the head shadow and summation effects were examined for conditions that were created using a tone-vocoder. These effects were tested by means of pairwise comparisons. Table 5 gives an overview of the contrasts of interest.

Firstly, the effect of binaural unmasking was assessed for the noise-vocoded conditions with different smoothing filters. When a smoothing filter of 400 Hz was used a binaural unmasking effect of 1.9 dB was found ($p < .05$) (contrast 1). Note that the same simulation technique was used in Study I but failed to reveal a significant binaural unmasking effect ($p = .064$). These findings suggest that a small binaural unmasking effect (1.5 – 1.9 dB) is present, which approximates the 5% criterion. No effect of binaural unmasking is detected for stimuli created with a smoothing filter of 50 Hz ($p = .124$) (contrast 2). The use of a low envelope cutoff, which is thought to minimize the influence of inherent fluctuations in the noise carrier, does not lead to an increased binaural unmasking effect.

Secondly, the binaural unmasking effect was examined for the conditions created by means of tone-vocoding. Contrary to expectation, effects of binaural unmasking were not found for the tone-vocoded conditions when both ITD and ILD cues were presented ($p = .893$) (contrast 5). It was expected that a tone-vocoder would better preserve temporal information in the signal, which would lead to a bigger binaural unmasking effect compared to the noise-vocoded conditions.

Thirdly, the question of whether the presence of ITD or ILD alone could result in a binaural unmasking effect was assessed for the noise-vocoded conditions with a smoothing filter of 400 Hz and the tone-vocoded conditions with a smoothing filter of 50 Hz. No effect of binaural unmasking was detected when ITD alone was presented to the listeners ($p = .331$, Noise, $p = .342$, Tone) (contrasts 3 and 6). It was hypothesized that ILD alone would account for the small binaural unmasking effect in the CI condition in the absence of ITD information. It turns out, however, that no binaural unmasking effect was found when ILD was presented alone ($p = .553$, Noise, $p = .500$, Tone) (contrasts 4 and 7).

Finally, the head shadow and summation effects were examined for the tone-vocoded conditions. In Study I a significant effect of head shadow and summation was detected for the conditions created by means of noise-vocoding. In this study, a head shadow effect of 3.4 dB was detected ($P < 0.0001$) (contrast 8), but no summation effect was found for the tone-vocoded stimuli ($p = .235$) (contrast 9).

3.4 Discussion

The studies described above assessed to what extent listeners with simulated bilateral and bimodal hearing devices can benefit from binaural advantages with regard to the perception of speech in noise. Firstly, the results indicate that listeners with simulated bilateral implants primarily benefit from a head shadow effect. They may also derive an advantage from summation and binaural unmasking effects, although it should be noted that the magnitude of these effects is largely dependent on the technique used to simulate the CIs (section 3.4.1). It was hypothesized that the small binaural unmasking effect could primarily be attributed to the presence of ILDs. It turned out, however, that no effect of binaural unmasking was found when either ILDs or ITDs alone were preserved in the signal (section 3.4.2). Secondly, the results reveal that listeners with simulated bimodal hearing devices do not benefit from binaural interaction effects (section 3.4.3). Nevertheless, it should be pointed out that listeners performed significantly better with simulated bimodal hearing devices than with simulated bilateral implants. This finding suggests that it may be more advantageous to provide unilateral CI users with a contralateral HA than with a second CI (section 3.4.4).

3.4.1 Bilateral implants

Previous studies have shown that the perception of speech in noise improves when unilateral CI users are provided with a second implant (*e.g., Van Hoesel and Tyler, 2003; Mosnier et al., 2009*). The advantage of bilateral implantation can to some extent be explained in terms of binaural advantages listeners can derive when the speech and noise sources are spatially separated.

The current work indicates that the advantage of bilateral implantation should primarily be ascribed to the head shadow effect. As the high frequency components of the noise source present in the CI signal are attenuated by the obstruction of the head, bilateral CI users can selectively attend to the ear with the best SNR. The magnitude of the effect varies depending on the technique used to simulate the CIs. The tone-vocoded conditions indicate an effect size of approximately 3.5 dB, whereas the noise-vocoded conditions reveal a head shadow effect of approximately 7 dB. On average, listeners with simulated bilateral hearing devices derive an advantage which is similar to the effect size of approximately 5 dB previously found for actual CI users (*Tyler et al., 2002; Van Hoesel and Tyler, 2003*). This effect is somewhat smaller than the effect size of approximately 9 dB generally found for NHLs (*Bronkhorst and Plomp, 1988; Arsenault and Punch, 1999*).

Previous studies have suggested that bilateral CI users may also benefit from a small binaural unmasking effect when the speech and noise sources are spatially separated (*Van Hoesel and Tyler, 2003*; *Schleich et al., 2004; Long et al., 2006; Garadat et al., 2009*). On the basis of the current work, it is difficult to predict to what extent actual bilateral CI users can benefit from the binaural unmasking effect in natural listening situations. Whether or not listeners with simulated bilateral CIs could derive a binaural unmasking effect depended largely on the technique used to simulate the implants. A binaural unmasking effect of 1.5 – 1.9 dB was found, which only approached significance, when the stimuli were created with a noise-vocoder with a smoothing filter of 400 Hz. By contrast, when a low cutoff frequency (50 Hz) or a tone-vocoder was used to simulate the CI conditions, listeners were unable to derive an unmasking effect. The size of the binaural unmasking effect is lower than the effect size found for NHLs, which is approximately 5 dB (*Bronkhorst and Plomp, 1988; Arsenault and Punch, 1999*).

Furthermore, there are indications that bilateral CI users may also benefit from a small summation effect. Whether or not listeners were able to benefit from this effect again depended on the simulation technique. Listeners could derive a summation effect of 3.4 dB when the stimuli were created with a noise-vocoder, but were unable to benefit from this effect when a tone-vocoder was used. The size of the effect found in the noise-vocoded condition is somewhat larger than the effect size of 1- 2 dB generally found for NHLs (*Ching et al., 2006*).

The findings of this simulation study cannot necessarily be generalized, however, to actual CI users. Whether or not listeners in fact derive binaural advantages largely depends on the correlation between the signals from the implants. Firstly, the fact that the implants operate independently has an adverse effect on the sensitivity to ITDs. The CIs provide the listener with uncorrelated pulse trains which create an inherent temporal mismatch in the signal across the ears. In addition, a likely mismatch in frequency-to-place maps across the ears may be detrimental for the detection of ITDs, since these cues cannot be derived from non-corresponding frequency channels (*Nuetzel and Hafter, 1981; Magezi and Krumbholz, 2008*). In fact, ITD detection for users of bilateral CIs has been shown to improve when electrode pairs are matched according to pitch[7] (*Long et al., 2003; Poon et al., 2009; Lu et al., 2010*). Secondly, the independent fitting of CIs may also have an adverse effect on the sensitivity to ILDs. Since the implants rely on independently fitted AGC systems, the degree of compression of the signal may be different across the ears. This may lead to unbalanced loudness across the ears. Consequently, when the level of a sound fluctuates, the ILD may

[7] Pitch matching can be explained as follows. Different electrodes across the ears can be paired based on measurements of interaural pitch comparisons. In order to find a match, an electrode pair is stimulated sequentially and the listener is asked to make a pitch judgement. If the perceived pitch is the same in both ears, the electrode pair is matched according to pitch.

alter even in the absence of a change in location. Similarly, one ear may receive the signal at a relatively higher level a priori, which may also affect the reliability of ILD cues.

3.4.2 Relative importance ITD and ILD cues

It still remains largely unclear to what extent bilateral CI users can benefit from a binaural unmasking effect. It seems that, even if people with bilateral implants are able to derive an advantage due to binaural unmasking, the size of the effect is limited compared to the effect size generally found for NHLs. This finding may be explained in terms of the relative contribution of ITDs and ILDs.

The fact that the binaural unmasking effect is limited for bilateral CI users may primarily be explicable in terms of decreased sensitivity to ITDs. Bilateral CI users are sensitive to level differences at the two ears, but are relatively insensitive to ITD cues. On average, listeners can discriminate sounds with an ILD of between 0.5 and 1.4 dB (*Van Hoesel and Tyler, 2003; Laback et al., 2004*). These results are similar in magnitude to performance of NHLs, who have JNDs of between 0.5 and 1 dB (*Mills, 1958, 1960*). Much larger JNDs for ITDs are reported, however, for bilateral CI users in comparison to those found for NHLs. On average, JNDs of approximately 100 – 200 μs have been reported for CI users (*Laback et al., 2004*) whereas NHLs have JNDs as small as 10 μs (*Mills, 1958*). It should be noted, however, that the findings are similar in magnitude to JNDs for envelope ITDs reported for NHLs (*Bernstein and Trahiotis, 2002*). The fact that CI users are relatively insensitive to ITD cues may therefore partly be attributed to the absence of informative TFS in the signal. Moreover, envelope ITD cues in the CI signal are ambiguous, as indicated by a large degree of intra- and inter-individual variation (*Laback et al., 2004*).

Since CI users are relatively more sensitive to ILD cues than to ITD cues, it can be hypothesised that the small benefit that can be derived from the binaural unmasking effect should be attributed primarily to the presence of ILD information. The results of study II indicate, however, that listeners cannot derive a binaural unmasking effect when either ITDs or ILDs alone are presented. This seems to suggest CI users rely on a combination of ITD and ILD cues to benefit from the spatial separation of speech and noise. Moreover, the results seem to indicate that the limited magnitude of the binaural unmasking effect may be the result of decreased sensitivity to ITDs as well as ILDs. NHLs can benefit from a binaural unmasking effect even in the presence of ILDs alone. Given that listeners with simulated bilateral CIs cannot benefit from ILDs when ITDs are filtered from the signal, the detection of ILD cues is not on a par with NHLs.

3.4.3 Bimodal hearing devices

The perception of speech in noise also improves when unilateral CI users are provided with a contralateral HA (*e.g., Kong, 2005; Ching et al., 2006, 2007; Cullington and Zeng, 2010*). The advantage of bimodal fittings, however, cannot be explained in terms a genuine binaural interaction effect resulting from spatial separation of the speech and noise sources.

Users of bimodal fittings are unable to benefit from a binaural unmasking effect, presumably because no comparison across the ears can be made since ITD cues are not well preserved in the CI and since ILD cues are limited at the low frequencies in the HA signal. Bimodal listeners are less sensitive to time and level differences across the ears than bilateral CI users and NHLs. A study by Francart et al. (*2008b*) indicated that bimodal listeners are sensitive to ITD differences of between 100 and 200 µs. These results are similar in magnitude to those found for bilateral CI users and for envelope ITDs for NHLs (*Bernstein and Trahiotis, 2002; Laback et al., 2004*). It should be noted, however, that the stimuli were presented under idealised conditions. The electric signal was delayed, for example, to ensure that the acoustic and electric signals would be presented simultaneously. In addition, the average threshold of hearing of the participants was less than 100 dB HL at 1 and 2 kHz, which indicates that they might have been able to recruit envelope ITDs, which are available above 1 kHz. It is likely that bimodal listeners are not sensitive to ITD cues of this magnitude in more realistic situations. Similarly, users of bimodal fittings are relatively insensitive to ILD cues. Whereas NHLs and bilateral CI users have JNDs of between approximately 0.5 and 1.5 dB, bimodal listeners can detect ILD differences no smaller than 3 dB (*Francart, 2008a*). This may in part be due to the fact that ILDs are only present in the signal above about 1500 – 1800 Hz.

In fact, the results show that, when the noise source was located at the CI ear, performance decreased significantly. SRT scores were 1.9 dB higher than when speech and noise coincided in space. In other words, an effect of masking, rather than unmasking, is found when the noise source is at the CI ear. The effect of masking has also been found in another study. Mok et al. (*2007*) studied the perception of speech in noise for prelingually deafened children with bimodal hearing devices. They found that SRTs were 0.4 dB higher with the noise at the CI ear than when the speech and noise sources were at 0° azimuth. Even though the size of the effect was only small, it was highly significant ($p < .0001$). The masking effect can perhaps be explained in terms of the head shadow effect. As was already mentioned, bilateral CI users benefit primarily from a head shadow effect. By contrast, even though listeners with simulated bilateral HAs can benefit from summation and binaural unmasking effects, they are unable to derive a head shadow effect. When the noise source is presented at the CI ear, contralateral to the HA ear, listeners do not benefit from an improved SNR at the HA ear since low frequency sounds are not attenuated by the head. The effect of masking is discussed in more detail in section 4.4.

It may be the case, however, that the findings of this simulation study cannot be generalized to actual bimodal listeners. Whether or not someone can in fact benefit from the spatial separation of speech and noise depends on, for example, the degree of hearing loss and the synchronisation of the devices. Firstly, the amount of residual hearing can affect the degree to which people can derive binaural advantages. If hearing remains above 1 – 2 kHz, for example, binaural unmasking may be facilitated since envelope ITDs and ILD information may be available to the listener. In addition, cochlear damage may result in abnormalities in the propagation time of the travelling wave along the BM or in the phase of neural spikes (*c.f. Moore, 1998*). The propagation time of a travelling waveform is approximately 5 – 9 ms for NHLs. However, cochlear loss may result in changes in propagation time to as much as 1 ms (*c.f. Moore, 1998*; *Rugerro, 1994*). This may significantly affect the ability to detect timing differences across the ears. Secondly, a lack of synchronisation between the devices may also affect the detectability of ILD and ITD cues. In addition, loudness growth may differ across modalities. This often results in unbalanced loudness across the ears, which may affect the reliability of ILD cues.

3.4.4 Implications for unilateral CI users

The current work has important implications for the debate about whether unilateral CI users with sufficient residual hearing should be provided with a contralateral HA or a second CI. The results of study I indicate that bimodal stimulation may be more beneficial than bilateral implantation. When the speech and noise sources coincided in space, SRTs were approximately 7 dB lower for the bimodal conditions. Similarly, the bimodal simulations resulted in an advantage of 4 – 6 dB when the speech and noise sources were spatially separated (at N-90 and N90 respectively).

The advantage of bimodal fittings cannot be explained, however, in terms of a genuine binaural interaction effect. The advantage is more likely to arise from the provision of additional information by the HA. The added low-frequency information provided by the HA can be beneficial for the intelligibility of speech in noise. It has been proposed that the improved perception by users of bimodal hearing devices can partly be attributed to the availability of pitch cues in the acoustic signal (*Turner et al., 2004; Qin and Oxenham, 2006; Brown and Bacon, 2009a, 2009b*), enabling listeners to segregate the speech and noise sources. As was already mentioned, unilateral CI users have difficulties perceiving pitch, but F0 can often be detected in the amplified acoustic signal. Brown and Bacon (*2009a, 2009b*) examined the effect of replacing the low-frequency target speech, which would normally be provided by the HA, with a tone that was frequency-modulated following the F0-pattern of the target speech and amplitude modulated following the amplitude envelope of the target speech. They showed that a tone carrying F0 and amplitude cues of target speech can improve the perception of speech in noise for CI users. This suggests the importance of pitch cues in the acoustic signal. It should be mentioned, however, that in a similar study Kong and Carlyon (*2007*) found a significant benefit from amplitude-modulation, although their results did not reveal a benefit which could be attributed to F0-modulation. Alternative explanations for the bimodal advantage are related to phonetic information retained in the low-frequency range. Kong and Carlyon (*2007*) suggest that improved performance with a contralateral HA may also be attributed to the availability of, for example, F1 cues below 500 Hz, coarticulation cues such as formant transitions, and low-frequency consonant cues such as nasality and voicing. Furthermore, it has also been suggested that the bimodal advantage can in part be explained in terms of glimpsing (*Kong and Carlyon, 2007*). As was already mentioned, CI users probably cannot listen in the dips, since the TFS of the input signal is discarded. Bimodal listeners, however, may benefit from dips in the masker signal since TFS is preserved in the HA signal. It should be noted, however, that the bimodal advantage due to glimpsing may be limited, as sensorineural hearing loss has been associated with a reduced ability to use TFS (*Festen and Plomp, 1990; Lorenzi et al., 2006*).

4. Bimodal hearing devices: improving perception in noise

4.1 Enhancing the binaural unmasking effect for bimodal listeners

Even though bimodal stimulation is to be favoured over bilateral stimulation, performance of speech in noise is not on a par with that of NHLs (*Ching et al., 2005*). This can in part be explained by the fact that users of bimodal fittings cannot benefit from spatially separating the speech and noise sources. In order to improve speech perception abilities in noise for bimodal listeners, unmasking effects for electro-acoustic hearing should be increased. The experiments described in this section aim to improve the detection of ITD and ILD cues and thereby create a binaural unmasking effect. An attempt is made to improve detectability of spatial cues by introducing ILDs in the frequency range covered by the HA (i.e. below 1 kHz) (section 4.1.1) and by enlarging the size of the ITDs (section 4.1.2). The question that needs to be answered is whether a binaural unmasking effect can be facilitated and the perception of speech in noise can be improved for listeners with simulated bimodal fittings when ITD and ILD cues are enhanced. In addition, the relative contribution of enhanced ITD and ILD cues to the binaural unmasking effect is assessed. It is expected that, even though NHLs can benefit from ITDs as well as ILDs, bimodal listeners will only benefit from enhanced ILD cues (section 4.1.3).

4.1.1 Enhancing ILD cues

Since ILDs may be the most salient cue with respect to the binaural unmasking effect for NHLs (*c.f. Bronkhorst and Plomp, 1988*), improving sensitivity to ILD cues may prove beneficial for bimodal listeners. It is unlikely that users of bimodal hearing devices can use normal head-induced ILD cues to segregate concurrent sound sources, since the HA only amplifies sounds that fall within the frequency range where hearing remains, which is typically up to about 1 kHz. As was already mentioned, however, ILD cues are generally only present above about 1500 – 1800 Hz (*Feddersen et al., 1957*).

One way to improve the detection of ILD cues for bimodal listeners is to introduce these cues in the frequency range covered by the HA. As already mentioned, Brungart and Rabinowitz (*1999*) showed that listeners are sensitive to ILD information at low frequencies when a sound source is located close to the listener’s head (within 1 m). Francart et al. (*2009*) examined localisation abilities of NHLs with simulated bimodal fittings when the level of the HA signal (< 1000 Hz) was increased or decreased to introduce level differences across the ears. The results of the localisation experiment showed that, for broadband noise, performance improved significantly when ILDs were introduced in the low frequencies of the HA signal. It may similarly be expected that bimodal listeners can benefit from the introduction of ILD cues in the low frequencies for the perception of speech in noise. It is worth mentioning that a mismatch of place of stimulation across the ears, which is common for bimodal listeners, only minimally affects the detection of ILD cues.

In this study, ILD cues will be introduced at frequencies within the range of the HA signal (below 1 kHz) by simulating an enlarged head size and thereby increasing the distance between the ears. Since level differences across the ears are the result of a head shadow effect, enlarging the size of the head will result in increased attenuation of frequencies, especially below 1500 Hz. The head shadow effect is normally restricted to high frequencies, since the short wavelengths do not exceed the size of the head and diffraction occurs. When the size of the head is increased, however, frequencies with relatively longer wavelengths are diffracted as they no longer exceed the size of the head. In this study, the size of the (simulated) head will be increased by a factor of two, which will introduce ILD cues below 1 kHz. Whereas ILDs are normally negligible below 1500 Hz, doubling the size of the head will reduce this threshold to 750 Hz. Since this falls within the frequency range covered by the HA signal (<1 kHz), it can be expected that bimodal listeners will derive a binaural unmasking effect when the size of the head is artificially increased.

4.1.2 Enhancing ITD cues

Even though NHLs can obtain a binaural unmasking effect as a result of the presence of ITDs (*Bronkhorst and Plomp, 1988*), it is unlikely that bimodal listeners will derive an advantage from ITD enhancement strategies that can be implemented in current speech processors. The inability of bimodal listeners to recruit ITD cues can be explained by the fact that timing cues are poorly preserved by the CI processor. First, the signal processor discards TFS from the signal, which results in CI users being unable to detect ITDs in the low frequency range. Second, it is likely that actual CI users are also relatively insensitive to envelope ITDs in the high frequency range. Even though Laback et al. (*2004*) suggest that bilateral CI users have JNDs in envelope ITDs of approximately 100 – 200 µs, it is likely that in more realistic situations (i.e. when there is a temporal and frequency mismatch across the ears), CI users have much larger JNDs.

There are several strategies that may result in a facilitation of ITD detection for bimodal listeners. One possibility may be to increase the size of the head, as was also suggested for the enhancement of ILD cues. Increasing the size of the head will result in an enlarged time delay across the ears. Since in this study the size of the head will be increased by a factor of two, the maximum ITD is doubled (from 607 µs to 1214 µs). This may prove beneficial, since actual CI users probably have relatively large JNDs in ITD and may therefore not be able to benefit from normal head-induced interaural delays (600 - 700 µs). It is worth mentioning that cells in the auditory brainstem are able to respond to larger interaural delays (*c.f. Kuwada and Yin, 1983*). In addition, it has been shown that broadband signals can be lateralised with ITDs that exceed the natural range (*Trahiotis and Stern, 1989*). There are several reasons, however, to expect that bimodal listeners will not be able to benefit from enlarged ITDs. It has been shown, for example, that time delays exceeding 500 – 1000 µs (up to 10 ms) yield little or no gain in the perception of speech in noise for NHLs (*Schubert, 1959; Levitt and Rabiner, 1969*). In addition, Zerlin (*1966*) showed that masking level differences (MLD) for NHLs increased with greater ITD until the interaural delay reached 1 ms but decreased again as it exceeded 1 ms. Furthermore, bimodal listeners may be unable to extract ITDs, as this would entail comparison of delays across different frequency channels. Whereas the CI signal only preserves envelope ITDs in the high frequency range, the HA signal generally only provides the listener with TFS ITDs in the low frequency range (up to about 1000 Hz). As was already mentioned, ITD extraction from non-corresponding frequency channels is problematic, since interaural delays are assumed to be determined on the basis of a comparison across neurons with the same CF (*Magezi and Krumbholz, 2008*). In addition, it should be noted that TFS and envelope ITDs are processed in different areas of the auditory brainstem (*Palmer, 1995*). It seems unlikely, therefore, that enlarging ITD cues will be beneficial for the perception of speech in noise by bimodal listeners.

A better alternative would be to introduce informative TFS into the CI signal. As has already been mentioned, the signal processor in the CI replaces the TFS in the input signal by an electrical pulse train with fixed-rate amplitude fluctuations, which are unrelated to the rapid oscillations in the input signal. TFS, however, is important with respect to the perception of speech in noise (*Drennan et al., 2007*). Several attempts have been made to represent TFS in such a way that it can be perceived by CI users (*e.g., Stickney et al., 2002, 2004b; Wilson et al., 2005; Van Hoesel et al., 2008; Chen and Zhang, 2008*). The general tendency seems to be that adding TFS will facilitate the detection of ITD cues in the low-frequency range and lead to improved perception of speech in noise for bilateral CI users (*Waltzman and Roland, 2006; Wilson et al., 2005*). It can similarly be expected that the addition of TFS to the CI signal will be beneficial for the perception of speech in noise by bimodal listeners. The computational demands, however, prevent algorithms introducing informative TFS from being implemented in actual CIs. Therefore, this thesis does not explore the advantage bimodal listeners could possibly derive from implementing informative TFS in the CI signal. Instead, the effect of enlarging ITD cues is examined.

4.1.3 Relative importance of enhanced ITD and ILD cues

It can be expected that, even though NHLs benefit from ITDs as well as ILDs for the perception of speech in noise, bimodal listeners primarily benefit from the presence of enhanced ILD cues. ILDs are probably successfully enhanced by enlarging the size of the head. As this strategy introduces level cues in the frequencies present in the HA signal (<1 kHz), bimodal listeners are likely to obtain a binaural unmasking effect. By contrast, increasing ITDs as an effect of enlarging the size of the head is not expected to result in improved perception of speech in noise, since listeners would have to derive ITDs from non-corresponding frequency channels.

Enlarging the size of the head results in the simultaneous enhancement of ITDs and ILDs. In order to tease the effects of these cues apart, however, listeners were presented with stimuli that contained either ITD or ILD information alone.

4.2 Study III

In order to test whether listeners with bimodal fittings can benefit from enhanced ITD and ILD cues for the perception of speech in noise, a simulation experiment was conducted. The experiment was designed to examine whether listeners with simulated bimodal fittings can derive an advantage from a binaural unmasking effect when ITD cues are enlarged and ILD cues are introduced in the low frequency range (> 750 Hz).

4.2.1 Listeners

Twelve normal-hearing listeners (1 male, 11 female) participated in this study. All participants were native speakers of Dutch. They ranged in age from 18 to 23 years (mean 20 years). All listeners had normal hearing, defined as pure-tone thresholds of 20 dB HL or better at octave frequencies between .5 and 8 kHz. None of the participants had any prior experience with the test materials. They were recruited from a volunteer subject pool at Utrecht University and were paid for their participation. The experiment consisted of a single session of 90 minutes.

4.2.2 Stimulus materials and conditions

The target stimuli used in this experiment consisted of a set of Dutch sentences (*Versfeld et al., 2000*) produced by a male speaker. Three key words were selected for each sentence.

Speech perception abilities were tested in a simulation of bimodal hearing devices, where a CI was simulated in the left ear and a HA at the other ear. As in the previous studies, all sentences were presented in a background of twenty-talker babble noise. Speech was always presented at 0° azimuth. The noise was either spatially separated from the speech signal, at 90° (N90) or -90° azimuth (N-90), or coincided in space with the speech, at 0° azimuth (N0). In the spatially separated conditions the noise was either at the side of the CI (N-90) or at the side of the HA (N90).

Stimuli were created with normal head-induced spatial cues (unenhanced) and enhanced spatial cues (enhanced). ITDs and ILDs were enhanced by simulating an increased head size in the HRTF calculations. This technique leads to the introduction of ILD cues in the low frequency range (> 750 Hz) and the enlargement of ITD cues.

An additional condition concerned the availability of enhanced spatial cues. Stimuli were created to contain either ITD cues, ILD cues, or both ITD and ILD cues. This condition allowed assessment of the relative contribution of the enhanced ITD and ILD cues to the binaural unmasking effect.

A total of 10 different bimodal listening conditions were created, each with a unique combination of the use of enhancement strategies, noise azimuth, and the availability of spatial cues (Table 6).

	Enhancement	Noise azimuth	Spatial cues
1	Unenhanced	N0	ITD + ILD
2	Unenhanced	N90	ITD + ILD
3	Unenhanced	N-90	ITD + ILD
4	Enhanced	N0	ITD + ILD
5	Enhanced	N90	ITD + ILD
6	Enhanced	N-90	ITD + ILD
7	Enhanced	N90	ILD
8	Enhanced	N-90	ILD
9	Enhanced	N90	ITD
10	Enhanced	N-90	ITD

Table 6. Test conditions, each with a unique combination of the use of enhancement strategies, noise azimuth, and the availability of spatial cues.

The stimuli were presented in blocks organised according to head size (unenhanced, enhanced). The order of the blocks was counterbalanced between participants, using a Latin square design. In addition, the order of conditions within blocks was also counterbalanced. As before, participants listened to 20 sentences in each condition. They were again tested on all conditions twice; the second time conditions were presented in reversed order.

4.2.3 Signal processing

The stimuli were created within the MATLAB environment as before. First, the spatial configurations of the speech and noise sources were applied, using a spherical head model (*Brown and Duda, 1998*). Subsequently, the stimuli were processed to simulate bimodal fittings. One channel of the stereo signal was manipulated to sound as if it was heard through a CI and the other as if it was heard through a HA.

The CI signal contained 6 contiguous channels covering frequencies from 100 to 7500 Hz. The channels were created using a sixth-order Butterworth filter. The envelope was extracted from each frequency band by full-wave rectification and smoothing at 400 Hz. Subsequently, the envelopes were used to amplitude-modulate a white noise source, which was then filtered to the channel bandwidth. Independent noise sources were used to simulate CIs in the two ears.

A noise vocoder with a relatively high smoothing filter was used in this study, since the results from study II indicated that listeners with simulated bilateral CIs are more likely to derive a binaural unmasking effect using this simulation technique than when using a noise- or tone-vocoder with a filter of 50 Hz. It can similarly be expected that listeners with simulated bimodal hearing devices are also more likely to derive a binaural unmasking effect when the CI signal is created with a noise-vocoder with a cutoff of 400 Hz.

The HA simulation was again created by means of low-pass filtering at 1000 Hz using a 20^{th} order Butterworth filter. All stimuli were normalized to the same rms level (0.07).

4.2.4 Procedure

The participants were seated in a soundproof booth and listened to the stimuli over headphones (Beyerdynamic DT250). As in the previous studies, the participants' task was to repeat verbatim what they heard and the experimenter scored responses. The scoring screen was not visible to the participants and no feedback was provided.

Before the experiment proper, the participants received training on the bimodal listening conditions in the presence of enhanced and unenhanced spatial cues, with the noise source at 90° and -90° azimuth. This enabled participants to familiarize themselves with the task and to get used to listening to these forms of distorted speech. Participants were trained using Dutch sentences produced by a female speaker (*Versfeld et al., 2000*). They received 10 minutes of training on the bimodal condition with the unenhanced spatial cues, consisting of 5 minutes with the noise source at the side of the HA (N90) and 5 minutes at the side of the CI (N-90). In addition, participants received another 10 minutes of training on the condition with the enhanced spatial cues, consisting of 5 minutes with the noise source at the side of the HA (N90) and 5 minutes at the side of the CI (N-90).

4.2.5 Results

Figure 13 summarises RT scores (in dB) for each test condition. The effects of the different conditions were again tested by means of mixed-effects modelling. It turns out that it is especially useful to include 'participants' as a random effect in the model since approximately 22% of variance within the data set can be attributed to variance between participants. Variance between sentence lists, on the other hand, does not contribute significantly to the overall variance in the data set. The remaining 78% can be explained by a combination of condition effects and a residual component.

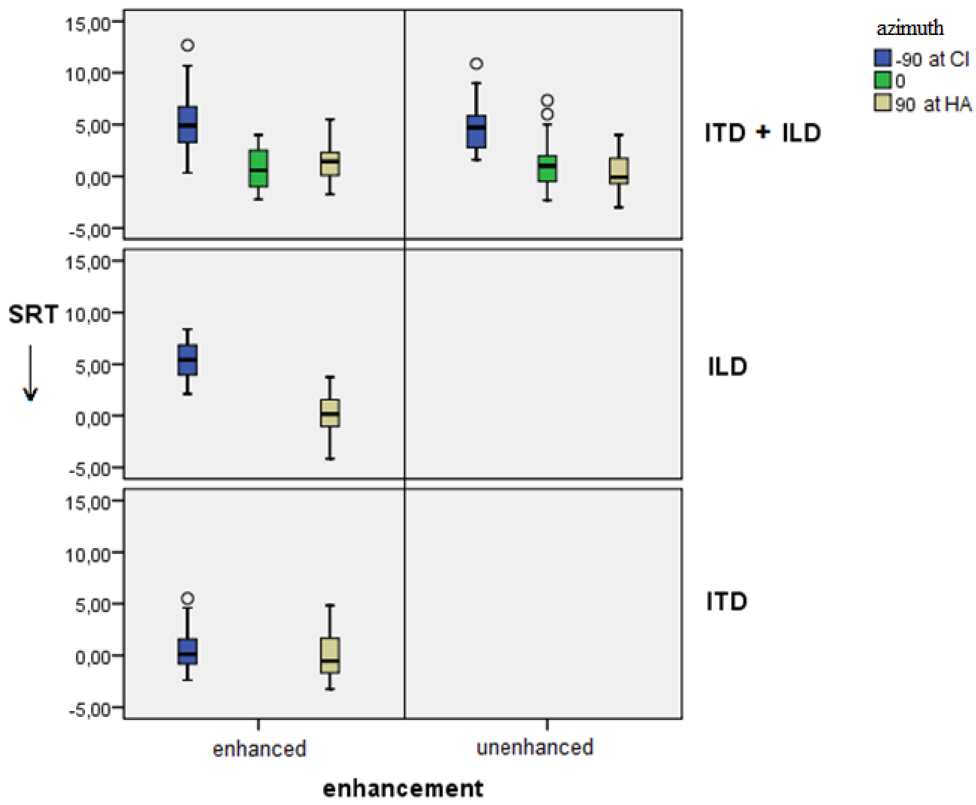

Fig. 13. Average SRT scores (in dB) for the conditions in which ITD and ILD cues are enhanced by enlarging the size of the head (enhanced, left panels) and conditions in which they are not enhanced (unenhanced, right panels). The conditions are further broken down in terms of the spatial cues present in the signal (ILD + ITD, ILD alone, ITD alone) and noise azimuth (colour). Green indicates that the noise source is at 0° azimuth (N0), olive green indicates that the noise is at the side of the HA at 90° azimuth (N90) and blue indicates that the noise is at the side of the CI at -90° (N-90).

Pairwise Comparisons Study III

Contrast	(I)Condition	(J)Condition	Mean Difference (I-J) (dB)	Sig.
1	enhanced N0 ITD+ILD	unenhanced N0 ITD+ILD	-0.4	.442
2	enhanced N90 ITD+ILD	unenhanced N90 ITD+ILD	0.9	.082
3	enhanced N-90 ITD+ILD	unenhanced N-90 ITD+ILD	0.2	.769
4	enhanced N90 ITD+ILD	enhanced N0 ITD+ILD	0.7	.208
5	enhanced N-90 ITD+ILD	enhanced N0 ITD+ILD	4.3*	.000
6	enhanced N90 ILD	enhanced N0 ITD+ILD	-0. 5	.393
7	enhanced N-90 ILD	enhanced N0 ITD +ILD	4.8*	.000
8	enhanced N90 ITD	enhanced N0 ITD+ILD	-0.7	.198
9	enhanced N-90 ITD	enhanced N0 ITD+ILD	-0.2	.722
10	unenhanced N90 ITD+ILD	unenhanced N0 ITD+ILD	-0.7	.212
11	unenhanced N-90 ITD+ILD	unenhanced N0ITD+ILD	3.8*	.000

*. The mean difference is significant at the .05 level.

Table 7. Pairwise comparisons of mixed-effects modelling with participants and sentence lists as two crossed random effects. Mean differences in SRT between selected conditions (in dB) are provided, along with the p-value.

The effect of enhancing ITDs and ILDs on the ability to perceive speech in noise with simulated bimodal hearing devices was assessed. In particular, the binaural unmasking effect was examined in the presence of enhanced ITDs and ILDs and compared to conditions containing normal head-induced ITDs and ILDs. Furthermore, the relative contribution of enhanced ITD and ILD cues to the binaural unmasking effect was studied. The effects were assessed by means of pairwise comparisons. Table 7 gives an overview of the contrasts of interest.

Firstly, the effect of enhancing ITD and ILD cues was examined by comparing SRT scores between enhanced and unenhanced conditions. It turns out that introducing ILDs in the frequency range covered by the HA signal (> 750 Hz) and at the same time enlarging ITDs does not result in an increase or decrease of SRTs. When the speech and noise sources coincided in space (contrast 1), no significant differences were detected between the enhanced and unenhanced conditions ($p = .442$). It should be mentioned, however, that no difference was expected between enhanced and unenhanced conditions when the speech and noise sources coincide in space, since there are no time and level differences across the ears when the sources are presented directly in front of the listener at 0° azimuth. Similarly, when the speech and noise sources were spatially separated, with the noise presented at 90° azimuth at the HA ear or at -°90 azimuth at the CI ear, an enhancement effect was not found (contrasts 2 and 3). The results show no significant differences between the enhanced and unenhanced conditions ($p = .082$ (N90), and $p = .769$ (N-90)).

Secondly, the binaural unmasking effect was examined in the bimodal listening conditions with enhanced and unenhanced spatial cues. A similar pattern of effects arose when the speech and noise sources were enhanced and when they remained unenhanced. Firstly, when the noise source was located contralateral to the CI ear (N90), no effect of binaural unmasking was detected in the enhanced and unenhanced conditions (contrasts 4 and 10) ($p = .208$, enhanced; $p = .212$, unenhanced). Secondly, when the noise source was presented at the CI ear (N-90), a masking effect, rather than an unmasking effect, was detected in both conditions. In the enhanced condition SRT scores were 4.3 dB higher ($p < .0001$) and in the unenhanced condition 3.7 dB higher when the noise was at the CI ear at -90° azimuth (contrasts 5 and 11).

In addition, to shed some light on the relative contributions of enhanced spatial cues to the binaural effects, SRTs were also measured in the presence of either enhanced ITDs or enhanced ILDs alone. When only ILD cues were present in the signal, a pattern arose which was similar to the results found when enhanced or unenhanced ILD and ITD cues were present. No effect of binaural unmasking ($p = .393$) was detected when the noise source was at 90° azimuth, at the HA ear (N90), (contrast 6). In addition, when the noise source was presented at -90° azimuth, at the CI ear (N-90), a masking effect of 4.7 dB ($p < .0001$) was found (contrast 7). However, when listeners were presented with enhanced ITD cues in the absence of ILD cues, a different pattern emerged. When only enhanced ITD cues were preserved in the signal (contrasts 8 and 9), no effects of binaural masking or unmasking were detected ($p = .198$ (N90); $p = .722$ (N-90)).

4.3 Control experiment

The fact that a masking effect, rather than an unmasking effect, was found when the noise was at the CI ear, at -90° azimuth cannot easily be explained. Therefore, a short control experiment was carried out to test whether the technique used to simulate the noise source at -90° azimuth might have introduced any abnormalities into the data. The effect of binaural unmasking was therefore examined in normal hearing conditions. If the processing technique functions properly, the results should reveal a binaural unmasking effect of similar magnitude, when the speech and noise sources are spatially separated with the noise at 90° azimuth or at -90° azimuth.

4.3.1 Listeners

Eight normal-hearing listeners (2 male, 6 female) participated in this study. All participants were native speakers of Dutch and had self-reported normal hearing. They ranged in age from 21 to 28 years (mean 23 years). None of the participants had any prior experience with the test materials. They were volunteers recruited from Utrecht University. The experiment consisted of a single session of 15 minutes.

4.3.2 Stimulus materials and conditions

As in the previous experiment, the target stimuli consisted of pre-recorded Dutch sentences (*Versfeld et al., 2000*) produced by a male speaker and containing three key words each.

Speech perception abilities were tested in three normal-hearing listening conditions. As in the previous experiment, all sentences were presented in a background of twenty-talker babble noise. Speech was always presented at 0° azimuth, coming from directly in front of the listener. The noise was either spatially separated from the target speech, presented at 90° azimuth (N90) or at -90° azimuth (N-90), or came from the same direction as the speech source, presented at 0° azimuth (N0). In total, 3 different normal hearing listening conditions were created (Table 8).

	Noise azimuth
1	N0
2	N90
3	N-90

Table 8. Test conditions control experiment, which differ with regard to degrees of azimuth of the noise source.

The conditions were counterbalanced between participants. As in the previous experiment, participants listened to 20 sentences in each condition. Participants were only tested on each condition once.

4.3.3 Signal processing

The processing techniques used in this experiment were similar to those in the previous experiment. Stimuli were created within the MATLAB environment. Only the spatial configurations were applied, using a spherical head model (*Brown and Duda, 1998*). Since the stimuli represent normal-hearing listening conditions, the signal was subsequently left unprocessed.

4.3.4 Procedure

The experimental procedure was as before. Since the experiment involved normal-hearing conditions, however, participants did not receive any training prior to the experiment proper.

4.3.5 Results

Figure 14 summarises SRT scores (in dB) for each test condition. The effect of spatially separating the speech and noise sources was tested by means of mixed-effects modelling. Approximately 24% of the variance within the data set can be attributed to variance between participants. By contrast, variance between sentence lists does not contribute significantly to the overall variance in the data set. The remaining 76% can be explained by a combination of condition effects and a residual component.

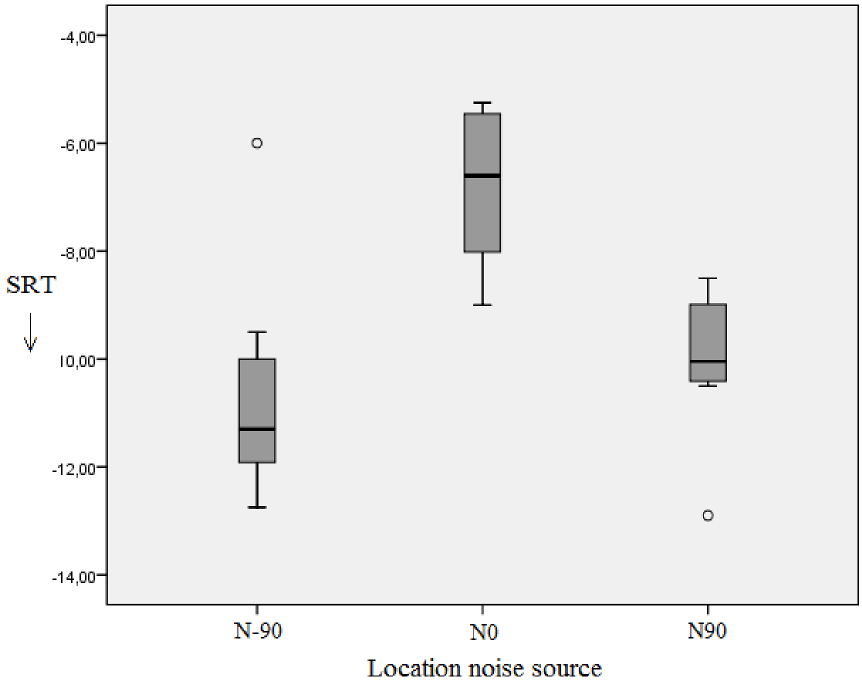

Fig. 14. Average SRT scores (in dB) for the conditions in which the speech and noise sources coincide in space (N0) and in which they are spatially separated with the noise at -90° azimuth (N-90) and at 90° azimuth (N90).

Pairwise Comparisons Study IV

Contrast	(I)Condition	(J)Condition	Mean Difference (I-J)	Sig.
1	N-90	N0	-3.8*	.000
2	N90	N0	-3.2*	.000
3	N-90	N90	-0.6*	.386

*. The mean difference is significant at the .05 level.

Table 9. Pairwise comparisons of mixed-effects modelling with participants and sentence lists as two crossed random effects. Mean differences in SRT between selected conditions (in dB) are provided with the p-value.

The effect of spatially separating the speech and noise sources under normal-hearing listening conditions was examined. In particular, the effect of presenting the noise source at -90° azimuth was assessed since this condition revealed unexpected results in studies I and III. Table 9 gives an overview of the contrasts of interest.

First, the effect of binaural unmasking was examined. It turned out that when the noise was presented at the left ear, at -90° azimuth, a binaural unmasking effect of 3.8 dB was found ($p < .0001$) (N-90 vs. N0). Similarly, when the noise was at the other ear, at 90° azimuth, an effect of binaural unmasking was found of 3.2 dB ($p < .0001$).

Second, the results were compared between the conditions in which the noise was at -90° azimuth and at 90° azimuth. SRT scores did not differ significantly between these two conditions ($p = .386$).

This control experiment showed that there was a binaural unmasking effect of similar magnitude, whether the noise was presented at 90° azimuth or at -90° azimuth. This suggests that the masking effect cannot be attributed to the simulation technique used in studies I and III and should be explained with reference to the combined use of a CI and a contralateral HA.

4.4 Discussion

The central aim in this study was to create a binaural unmasking effect by enhancing ITD and ILD cues and thereby improve the perception of speech in noise for bimodal listeners. The results of the experiment reveal that enlarging ITDs and introducing ILDs in the low frequency range (i.e. > 750 Hz) does not lead to improved perception as the result of a binaural unmasking effect (section 4.4.1). In fact, an effect of masking, rather than unmasking, was found when the noise source was presented at the CI ear, contralateral to the HA ear. This pattern emerges in the presence of enhanced as well as unenhanced spatial cues. By contrast, this effect was not detected when the enhanced ILDs were removed from the signal. Several possible explanations of the masking effect, taking the relative contribution of ITD and ILD cues into account, will be provided in section 4.4.2. Furthermore, the implications for actual bimodal listeners will be described (section 4.4.3) and suggestions for future research will be given (section 4.4.4).

4.4.1 The effect of enhancing ITD and ILD cues

This study aimed to improve the perception of speech in noise for listeners with simulated bimodal fittings, by enhancing spatial cues. The results indicate, however, that enlarging ITDs and introducing ILDs in the frequency range covered by the HA signal (i.e. > 750 Hz) does not improve the perception of speech in noise. SRTs of a similar magnitude were found in the presence of enhanced and unenhanced ITDs and ILDs. The fact that there is no advantage of enhancing spatial cues can be explained with reference to the binaural unmasking effect. It was hypothesized that a binaural unmasking effect could be created by enhancing ITDs and ILDs. The results of this experiment suggest, however, that the presence of enhanced spatial cues does not facilitate this effect.

There are two possible explanations for the absence of the binaural unmasking effect in the enhanced conditions. Firstly, it may be the case that listeners with simulated bimodal fittings are unable to detect the enhanced spatial cues. In particular, listeners may be relatively insensitive to ITDs. This may be due to the fact that ITD extraction involves comparison across non-corresponding frequency channels in the CI and HA ear (*c.f. Magezi and Krumbholz, 2008*). Listeners with simulated bimodal fittings may also be relatively insensitive to enhanced ILDs. This may be argued on the basis of the assumption that bimodal listeners are unable to detect normal head-induced ILDs and the fact that a binaural unmasking is absent in the presence of enhanced as well as unenhanced ILDs. It should be noted, however, that bimodal listeners may in fact be sensitive even to normal head-induced ILDs, which are present, although negligible, even below 500 Hz (*Feddersen et al., 1957*). An alternative explanation for the absence of the binaural unmasking effect is that listeners do detect the enhanced spatial cues but cannot use them to segregate concurrent sound sources. It may be true that listeners with simulated bimodal hearing devices are sensitive to ITDs, but do not depend on this cue to separate the speech and noise sources in the input signal. This would be in accordance with the suggestion that NHLs do not depend on ITDs to group simultaneous frequency components and thereby segregate sound sources (*Culling and Summerfield, 1995; Darwin and Hukin, 1999*). Similarly, bimodal listeners may be sensitive to ILDs but unable to derive an advantage from these cues. In fact, a masking effect was found in the presence of enhanced ILDs. This finding seems to suggest that the presence of ILDs leads to a disadvantage, rather than an advantage, with regard to the perception of speech in noise.

4.4.2 Masking effect

Perhaps the most interesting finding is the fact that there was an effect of masking, rather than unmasking, when the speech and noise sources were spatially separated with the noise at the CI ear. It turns out that the effect of masking was absent when the enhanced ILDs were removed from the signal. This suggests that enhanced ILDs, rather than ITDs, play an important role with respect to the masking effect.

The fact that no effect of masking was found when the enhanced ILD cues were removed from the signal does not necessarily mean, however, that masking occurs as a result of ILD detection. As was already mentioned, it may be assumed that bimodal listeners are unable to detect enhanced and unenhanced ILDs. Even if bimodal listeners were able to detect ILD cues, it is unlikely that this could contribute to the masking effect. If the masking effect were due to ILD detection, the effect should also have been found with the noise at the HA ear. Instead, bimodal listeners may be sensitive to fluctuations in the level of the noise at the CI ear resulting from a head shadow effect, rather than to level differences across the ears.

Several possible explanations can be provided for the masking effect in terms of the perceived loudness of the noise. Firstly, fluctuations in the level of the noise source, and consequently also of the SNR, are especially likely to arise at the CI ear as a result of the head shadow effect. Improvements in SNR may be limited at the HA ear, since users of bilateral HAs generally do not benefit from a head shadow effect (see study I). It has therefore been suggested that improvements in SNR at the HA ear may not be sufficient to compensate for the decrease in SNRs at the CI ear (*Mok et al., 2007*). According to this view, performance in noise is determined by the combined detectability of the speech signal in the CI and HA ears. Secondly, it may be the case that listeners predominantly focus on the input provided to the CI ear. Mok et al. (*2007*) tested pre-lingually deafened children, who may have had difficulties using the acoustic input provided through the HA. This led them to suggest that these children may pay more attention to the CI signal. In the current study, however, all participants were NHLs and were therefore used to hearing acoustic input. Nevertheless, it is possible that they focused more on the input provided at the CI ear. Future research will need to determine, however, what exactly causes the masking effect.

4.4.3 Implications for actual bimodal listeners

There are several important implications that can be derived from this simulation study for actual bimodal listeners and for strategies aiming to improve perception in noise for these listeners. Firstly, the results indicate that the perception of speech in noise for actual bimodal listeners is not likely to improve when ITD cues are enlarged and ILD cues are introduced in the frequency range covered by the HA. This suggests that developing processing strategies which enhance ITDs and ILDs in this way are unlikely to improve perceptual abilities of bimodal listeners. The results furthermore suggest that actual bimodal listeners are likely to experience an effect of masking when the noise is at the CI ear. Since this effect can perhaps be linked to level fluctuations of the noise source at the CI ear in combination with the absence of the head shadow effect at the HA ear, it may be more beneficial to improve the SNR in the HA ear when the noise is at the CI ear, than to improve the detectability of spatial cues.

The findings of this simulation study may not necessarily be generalised, however, to actual users of bimodal hearing devices. Firstly, the effects of long term hearing loss may influence the ability to perceive speech in noise of actual bimodal listeners. Two factors that play an important role with regard to performance in noise are loudness recruitment and reduced frequency selectivity (*Villchur, 1974; Florentine et al., 1980; Nejime and Moore, 1997*). Loudness recruitment distorts the perceived amplitude relationships among the acoustic elements of the signal (*c.f. Villchur, 1974*). Similarly, the amplitude relationships between the speech and noise sources may be distorted, possibly leading to an increase or decrease of SNR in the HA ear. This may be of particular interest with regard to the masking effect, as it was suggested that the SNR at the HA ear may not be sufficient to compensate for the decrease in SNR at the CI ear. In addition, reduced frequency selectivity also affects the detectability of the speech signal in the presence of background noise. The perception of speech in noise becomes increasingly difficult as the bandwidths of the auditory filters are broadened. Secondly, the lack of auditory stimulation in profoundly deaf individuals before the fitting of hearing devices may lead to deterioration of speech perception abilities (*c.f. Silman et al., 1984; Neuman, 1996*). Auditory deprivation, especially in the case of pre-lingual deafness, often leads to a reorganization of other sensory modalities in the brain *(Dye et al., 2007)*. In the case of profound deafness auditory deprivation is met by the enhancement of visual cognition, for example, as an effect of compensatory plasticity (*Bavelier et al., 2006*). Consequently, auditory processing may be impaired in actual bimodal listeners; this may result in decreased performance for the perception of speech in noise. Thirdly, a mismatch in terms of the synchronization of the hearing devices may impair performance by bimodal listeners. The fact that the implant and the hearing aid present the signal to the listener with a different delay may, for example, impair ITD detection. Since the masking effect has been linked to level fluctuations of the noise source, the unbalanced loudness that arises across the ears when combining an implant and a hearing aid is perhaps more important. The loudness mismatch may primarily be detrimental to the perception of ILDs.

4.4.4 Suggestions for future research

The aim of this thesis was to create a binaural unmasking effect and thereby improve the perception of speech in noise for bimodal listeners. The enhancement strategies used in this experiment, however, did not lead to a perceptual advantage. Future research should therefore be directed at exploring other strategies to improve performance in noise. It may be worthwhile to examine whether bimodal listeners can benefit from a binaural unmasking effect when TFS is introduced into the CI signal. Introducing TFS is likely to facilitate the detection of ITDs, since interaural delays can be determined by comparing corresponding frequency channels in the CI and HA ear. When trying to improve the perception of speech in noise, however, it may be more beneficial to develop strategies that increase the SNR in the HA ear, rather than to improve the detectability of spatial cues. Increasing the SNR at the HA ear may decrease the effect of masking when the noise is at the CI ear.

Further research is required to determine what factors contribute to the masking effect. Firstly, it was suggested that the effect may primarily be attributed to level fluctuations of the noise source at the CI ear in combination with the absence of the head shadow effect at the HA ear. Since this study only assessed the relative contribution of enhanced ITDs and ILDs, it may be worth examining to what extent normal head-induced timing and level fluctuations in the masker signal contribute to the masking effect. If the masking effect can indeed be attributed to the level of the masker signal, it may be worthwhile to examine whether performance improves when the SNR at the HA ear is artificially increased. Secondly, it was also suggested that the masking effect arises because listeners selectively attend to the CI ear, even when this is detrimental to the perception of speech in noise. Future research needs to determine whether bimodal listeners do indeed predominantly focus on the CI signal. If this is indeed the case, the perception of speech in noise may be improved by training bimodal listeners to pay more attention to the information provided by the HA ear. Last but not least, it is important to examine whether the results of this simulation study are replicated when actual bimodal listeners are tested, since the findings have important implications for the development of future processing techniques that aim to improve the perception of speech in noise.

5. Summary and conclusions

The central aim of this thesis was to assess to what extent users of bilateral and bimodal fittings can benefit from binaural stimulation when listening to speech in noise. In order to shed some light on this issue, several simulation experiments were conducted. The main conclusions of these experiments are summarised below.

- The results of the simulation studies indicated that bilateral CI users primarily benefit from a head shadow effect. It is difficult to predict whether actual CI users can also benefit from summation and binaural unmasking effects, since the results were largely dependent on the technique used to simulate the CIs. No effect of binaural unmasking was found when ILDs or ITDs alone were preserved in the signal. This seems to suggest that the limited effect of unmasking should be attributed to decreased sensitivity to ILDs and ITDs.

- Listeners performed significantly better with simulated bimodal hearing devices than with simulated bilateral implants. This suggests that actual unilateral CI users may derive a larger advantage from a contralateral HA than from a second CI. The advantage of bimodal stimulation cannot be attributed, however, to a genuine binaural interaction effect. Listeners with simulated bimodal fittings could not benefit from a binaural unmasking effect. This was explained by the fact that no comparison across the ears could be made, since ITD cues are not well preserved in the CI signal and ILD cues are limited at the low frequencies in the HA signal.

- The perception of speech in noise by listeners with simulated bimodal fittings did not improve when ITDs were enlarged and ILDs were introduced in the frequency range covered by the HA (i.e. > 750 Hz). Listeners did not obtain a binaural unmasking effect when spatial cues were enhanced. In fact, an effect of masking, rather than unmasking, was found when the noise source was presented at the CI ear. The masking effect was absent, however, when the enhanced ILDs were removed from the signal and only enhanced ITDs remained. It was hypothesized that the masking effect arises as a result of an increase in the level of the noise at the CI ear, combined with the absence of the head shadow effect at the HA ear.

References

American Speech-Language Hearing. Association (ASHA). *Type, Degree, and Configuration of Hearing Loss.* Retrieved March 6, 2010, from www.asha.org/public/hearing/disorders/types

Arsenault, M.D., and Punch, J.L. (1999). Nonsense-syllable recognition in noise using monaural and binaural listening strategies. *Journal of the Acoustical Society of America*, 105.3, 1821 – 1830.

Armstrong, M., Pegg, P., James, C., and Blamey, P. (1997). Speech perception in noise with implant and hearing aid. *The American Journal of Otology*, 18, S140 – S141.

Baayen, R.H., Davidson, D.J., and Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390 – 412.

Bavelier, D., Dye, M.W.G., Hauser, P.C. (2006). Do deaf individuals see better? *TRENDS in Cognitive Sciences,* 10.11, 512-518.

Bench, J., Kowal, Å and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) Sentence List for Partially-Hearing Children. *British Journal of Audiology*, 13.3, 108 – 112.

Bernstein, L.R., and Trahiotis, C. (1985a). Lateralization of low-frequency, complex waveforms: The use of envelope-based temporal disparities. *Journal of the Acoustical Society of America*, 77.5, 1868 – 1880.

Bernstein, L.R., and Trahiotis, C. (1985b). Lateralization of sinusoidally amplitude-modulated tones: Effects of spectral locus and temporal variation. *Journal of the Acoustical Society of America*, 78.2, 514 – 523.

Bernstein, L.R., and Trahiotis, C. (2002) Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". *Journal of the Acoustical Society of America*, 112.3, 1026-1036.

Bianchi, G., and Sorrentino, R. (2007). *Electronic filter simulation and design*. New York: McGraw-Hill.

Blamey, P. J., Dooley, G. J., James, C. J., and Parisi, E. S. (2000). Monaural and binaural loudness measures in cochlear implant users with contralateral residual hearing. *Ear and Hearing*, *21*, 6–17.

Blauert, J. (1983). *Spatial hearing: The psychophysics of human sound localization.* Cambridge, MA: MIT Press.

Bonding, P. (1979). Frequency Selectivity and Speech Discrimination in Sensorineural Hearing Loss. *Scandinavian Audiology*, 8. 4, 205 — 215.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: Bradford Books/MIT Press.

Brokx, J. P. L., & Nooteboom, S. G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics, 10,* 23-36.

Bronkhorst, A.W., and Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, 83, 1508 – 1516.

Brown, C.A., and Bacon, S.P. (2009a). Low-frequency speech cues and simulated electric-acoustic hearing. *Journal of the Acoustical Society of America*, 125.3, 1658 – 1665.

Brown, C.A., and Bacon, S.P. (2009b). Achieving electric-acoustic benefit with a modulated tone. *Ear and Hearing*, 30.5, 1 – 5.

Brown, C.P., and Duda, R.O. (1998). A structural model for binaural sound synthesis. *IEEE transactions on speech and audio processing*, vol. 6, No. 5, 476 – 488.

Brungart, D.S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, 109.3, 1101 – 1109.

Brungart, D.S., and Rabinowitz, W.M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *Journal of the Acoustical Society of America*, 106.3, 1465 – 1479.

Butterworth. S. (1930). On the theory of filter amplifiers. *Experimental wireless & the wireless engineer*, 7, 536 – 541.

Cherry, E.C. (1953). Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*, 25.5, 975 – 979.

Chen, F., and Zhang, Y. (2008). A novel temporal fine structure-based speech synthesis model for cochlear implant. *Signal Processing*, 88, 2693 – 2699.

Ching, T.Y.C., Hill, M., Brew, J., Incerti, P, Prioli, S., Rushbrook, E., Forsythe, L. (2005). The effect of auditory experience on speech perception, localization, and functional performance of children who use a cochlear implant and a hearing aid in opposite ear. *International Journal of Audiology*, 44.12, 677 – 690.

Ching, T.Y.C., Incerti, P., Hill, M., and van Wanrooy, E. (2006). An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices. *Audiology and Neurotology*, 11 (supplement 1), 6 – 11.

Ching, T.Y.C., Van Wanrooy, E., and Dillon, H. (2007). Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. *Trends Amplification,* 11.3,161-192.

Colburn, H.S. (1973). Theory of binaurali nteractionb asedo n auditory-nerved ata. I. General strategy and preliminary results on interaural discrimination. *Journal of the Acoustical Society of America*, 54.6, 1458 – 1470.

Crystal, D. (2008). *A dictionary of linguistics and phonetics*. Oxford: Blackwell Publishing.

Culling, J. F, and Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America, 98,* 785-797.

Cullington, H.E., and Zeng, F. (2010). Bimodal hearing benefit for speech recognition with competing voice in cochlear implant subject with normal hearing in contralateral ear. *Ear and Hearing*, 31.1, 70 – 73.

CVZ uitspraak 29096895(29131579) (2009). College voor zorgverzekeringen. Meerwaarde bilaterale cochleaire implantatie bij kinderen en volwassenen.

Darwin, C.J., and Hukin, R.W. (1999). Auditory Objects of Attention: The Role of Interaural Time Differences. *Journal of Experimental Psychology: Human Perception and Performance*, 25.3, 617-629.

Djourno A, Eyries C, Vallancien B. (1957). De l'excitation electrique du nerf cochleaire chez l'homme, par induction a distance, a l'aide d'un micro-bobinage inclus a demeure. *La Presse Médicale,* 65, 612.

Dorman, M.F., Loizou, P.C., and Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America*, 102.4, 2403-2411.

Dorman, M.F., Loizou, P.C., Fitzke, J., and Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6 – 20 channels. *Journal of the Acoustical Society of America, 104.6, 3583 – 3585.*

Dorman, M. F., Gifford, R. H., Spahr, A. J., and McKarns, S.A. (2008). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiology Neurotology*, *13*, 105–112.

Drennan, W.R., Won, J.H., Dasika, V.K. and Rubinstein, J.T. (2007). Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. *Journal of the Association of Research in Otolaryngology*, 8, 373 – 383.

Durlach, N.I. (1963). Equalization and cancellation theory of binaural masking-level differences. *Journal of the Acoustical Society of America*, 35.8, 1206 – 1218.

Durlach, N.I., Mason, C.R., Shinn-Cunningham, B.G., Arbogast, T.L., Colburn, H.S., and Kidd, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*, 114.1, 368 – 379.

Dye, M.W.G., Baril, D.E., Bavelier, D. (2007). Which aspects of visual attention are changed by deafness? The case of the Attentional Network Test. *Neuropsychologia*, 45, 1801-1811.

Faulkner, A., Rosen, S., and Moore, B.C.J. (1990). Residual frequency selectivity in the profoundly hearing-impaired listener. *British Journal of Audiology*, 24, 381 – 392.

Faulkner, A. (2006). Adaptations to distorted frequency-to-place maps: Implications of simulations in normal listeners for cochlear implants and electroacoustic stimulation. *Audiology and Neurotology,* 11(suppl 1), 21 – 26.

Favre, E., and Pelizzone, M. (1993). Channel interactions in patients using the Ineraid multichannel cochlear implant. *Hearing Research*, 66, 150 – 156.

Feddersen, W.E., Sandel, T.T., Teas, D.C., & Jeffress, L.A. (1957). Localization of high-frequency tones. *Journal of the Acoustical Society of America*, 29, 988 – 991.

Festen, J.M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88, 1725 – 1736.

Florentine, M., Buus, S., Scharf, B., and Zwicker, E. (1980). Frequency selectivity in normally-hearing and hearing-impaired observers. *Journal of Speech and Hearing Research,* 23, 646 – 669.

Francart, T. (2008a) Perception of binaural localization cues with combined electric and *acoustic hearing*. PhD dissertation, Katholieke Universiteit Leuven.

Francart, T., Brokx, J., and Wouters, J. (2008b). Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation. *Journal of the Association of Research in Otolaryngology*, 10, 131–141.

Francart, T., Bogaert, T., Moonen , M., and Wouters, J. (2009). Amplification of interaural level differences improves sound localisation in acoustic simulations of bimodal hearing. *Journal of the Acoustical Society of America*, 126.6, 3209 – 3213.

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. *Journal of the Acoustical Society of America*, 106.6, 3578 – 3588.

Friesen, L.M., Shannon, R.V., Baskent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparisons of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America*, 110.2, 1150 – 1163.

Fu, Q, Shannon, R.V., and Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *Journal of the Acoustical Society of America, 104.6, 3586 – 3596.*

Fu, Q., and Shannon, R.V. (1999). Effects on electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant. *Ear and Hearing*, 20.4, 332 – 344.

Garadat, S.N., Litovsky, R., Yu, G., and Zeng, F. (2009). Role of binaural hearing in speech intelligibility and spatial release from masking using vocoded speech. *Journal of the Acoustical Society of America*, 126.5, 2522 – 2535.

Green, T., Faulkner, A., Rosen, S., and Macherey, O. (2005). Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification. *Journal of the Acoustical Society of America*, 118.1, 375 – 385.

Greenwood, D. (1990). A cochlear frequency-position function for several species – 29 years later. *Journal of the Acoustical Society of America*, 87.6, 2592 – 2605.

Henning, G.B. (1974). Detectability of interaural delay in high-frequency complex waveforms. *Journal of the Acoustical Society of America*, 55, 84 – 90.

Henning, G.B. (1983). Lateralization of low-frequency transients. *Hearing Research*, 9, 153 – 172.

Hirsh, I.J. (1950). The Relation between Localization and Intelligibility. *Journal of the Acoustical Society of America*, 22.2, 196 – 200.

Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology*, *41*, 35-39.

Kong, Y.Y., Stickney, G.S., and Zeng, F.G. (2005). Speech and melody recognition in binaurally combined acoustic and electric hearing. *Journal of the Acoustical Society of America*, 117, 1351–1361.

Kong, Y., and Carlyon, R. P. (2007). Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *Journal of the Acoustical Society of America* **,** 121, 3717–3727.

Kuhn, G.F. (1977). Model for the interaural time differences in the azimuthal plane. *Journal of the Acoustical Society of America*, 62.1, 157 – 167.

Kuwada, S., and Yin, T.C.T (1983). Binaural interaction in low-frequency neurons in inferior colliculus of the cat: Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *Journal of Neurophysiology*, 50, 981 – 999.

Laback, B., Pok, S.M., Baumgartner, W.D., Deutsch, W.A., and K. Schmid, K. (2004). Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear and Hearing*, 25.5, 488-500.

Levitt, H., and Rabiner, L.R. (1969). Binaural release from masking for speech and gain in intelligibility. *Journal of the Acoustical Society of America*, 42.3, 601 – 608.

Loizou, P.C. (1998). Introduction to cochlear implants. *IEEE Signal Processing Magazine*, 101 – 130.

Long, C.J., Eddington, D.K., Colburn, H.S., and Rabinowitz, W.M. (2003). Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user. *Journal of the Acoustical Society of America*, 114.3, 1565 – 1574.

Long, C., Carlyon, R. P., Litovsky, R., and Downs, D. (2006). Binaural unmasking with bilateral cochlear implants. *Journal of the Association of Research in Otolaryngology*, 7.4, 352-360.

Lord Rayleigh, J.W.S. (1907). *The theory of sound.* New York: Dover.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B.C.J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 18866 – 18869.

Lu, T., Litovsky, R., and Zeng, F. (2010). Binaural masking level differences in actual and simulated bilateral cochlear implant listeners. *Journal of the Acoustical Society of America*, 127.3, 1479 – 1490.

MacLeod A., and Summerfield Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology*, 24, 29–43.

Magezi, D.A., and Krumbholz, K. (2008). Can the binaural system extract fine-structure interaural time differences from noncorresponding frequency channels? *Journal of the Acoustical Society of America*, 124.5, 3095 – 3107.

McDermott, H.J., and Sucher, C.M. (2006). Perceptual dissimilarities among acoustic stimuli and ipsilateral electric stimuli. *Hearing Research*, 218 (1-2), 81 - 88.

McFadden, D., and Pasanen, E. G. (1976). Lateralization at high frequencies based on interaural time differences. *Journal of the Acoustical Society of America*, 59.3, 634 – 639.

McKay, C.M, McDermott, H.J., and Clark, G.M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *Journal of the Acoustical Society of America*, 96.5, 2264 – 2673.

Mills, A.W. (1958). On the Minimum Audible Angle. *Journal of the Acoustical Society of America*, 30.4, 237 – 246.

Mills, A.W. (1960). Lateralization of High-Frequency Tones. *Journal of the Acoustical Society of America*, 32.1, 132 – 234.

Mok, M., Galvin, K.L., Dowell, R.C., and McKay, C.M. (2007). Spatial unmasking and binaural advantage for children with normal hearing, a cochlear implant and a hearing aid, and bilateral implants. *Audiology and Neurotology*, 12, 295 – 306.

Moore, B.C.J. (1998). *Cochlear hearing loss*. London: Whurr Publishers.

Moore, B.C.J. (2008a). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, 9, 399 -406.

Moore, B.C.J. (2008b). *An introduction to the psychology of hearing*. Bingley: Emerald.

Mosnier, I., Sterkers, O., Bebear, J., Godey, B., Robier, A., Deguine, O., Fraysse, B., Bordure, P., Borel, S., Ambert-Dahan, E., and Ferrary, E. (2009). Speech performance and sound localization in a complex noisy environment in bilaterally implanted adult patients. *Audiology and Neurotology*, 14, 106 – 114.

Nejime, Y., and Moore, B.C.J. (1997). Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise. *Journal of the Acoustical Society of America*, 102.1, 603 – 615.

Neuman, A.C. Late-onset auditory deprivation: A review of past research and an assessment of future research needs. *Ear and Hearing*, 17, 3S – 13S.

NICE technology appraisal guidance 166 (2009). National Institute for Health and Clinical Excellence (NHS). Cochlear implants for children and adults with severe to profound hearing loss.

Nuetzel, J.M, and Hafter, E.R. (1981). Discrimination of interaural delays in complex waveforms: Spectral effects. *Journal of the Acoustical Society of America*, 69.4, 1112 – 1118.

Offeciers, E., Morera, C., Müller, J., Huarte, A., Shallop, J., and Cavallé, L. (2005). International consensus on bilateral cochlear implants and bimodal stimulation. Second meeting consensus on auditory implants, 19 – 21 February 2004, Valencia, Spain. *Acta Oto-Laryngologica*, 125, 918 – 919.

Palmer, A.R. (1995). Neural signal processing, in *Hearing*, edited by Moore, B.C.J. (Academic, New York).

Plomp, R., and Mimpen, A.M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *International Journal of Audiology*, 18.1, 43 – 52.

Poon, B.B., Eddington, D.K., Noel, V., and Colburn, H.S. (2009). Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing. *Journal of the Acoustical Society of America*, 126.2, 806 – 815.

Qin, M. K., and Oxenham, A. J. (2003). Effects of simulated cochlear implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114, 446–454.

Qin, M. K., and Oxenham, A. J. (2006). Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *Journal of the Acoustical Society of America*, 119, 2417–2426.

Quené, H., and Van den Berg, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*, 43, 103 – 121.

Rose, J.E., Gross, N.B., Geisler, C.D., and Hind, J.E. (1966). Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *Journal of Neurophysiology*, 29, 288-314.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences*, Vol. 336, No. 1278, Processing of Complex Sounds by the Auditory System, 367-373.

Rosen,S., Faulkner,A., Wilkinson,L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *Journal of the Acoustical Society of America*, 106, 3629 - 3636.

Rugerro, M.A. (1994). Cochlear delays and travelling waves: Comments on ‘Experimental look at cochlear mechanics’. *Audiology*, 33, 131 – 142.

Sayles, M., and Winter, I.M. (2008). Reverberation challenges the temporal representation of the pitch of complex sounds. *Neuron*, 58, 789 – 801.

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 and Hearing*, 25.3, 197 – 204.

Schubert, E.D. (1959). Some preliminary experiments on binaural time delay and intelligibility. *Journal of the Acoustical Society of America*, 28.5, 895 – 901.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., and Ekelid M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303-304.

Silman, S., Gelfand, S.A., and Silverman, C.A. (1984). Late-onset auditory deprivation: Effects of monaural versus binaural hearing aids. *Journal of the Acoustical Society of America*, 76.5, 1357 – 1362.

Smith, Z.M., Delgutte, B., and Oxenham, A.J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87 – 90.

Souza, P., and Rosen, S. (2009). Effects of envelope bandwidth on the intelligibility of sine- and noise-vocoded speech. *Journal of the Acoustical Society of America*, 126.2, 792 – 805.

Stickney, G.S., Nie, K., and Zeng, F. (2002). Realistic listening improved by adding fine structure. *Journal of the Acoustical Society of America, 112.5, 2355.*

Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (2004a). Cochlear implant recognition with speech maskers. *Journal of the Acoustical Society of America*, 116, 1081–1091.

Stickney, G.S., Nie, K., Kong, Y.Y., Chen, H., and Zeng, F.G. (2004b). Temporal fine structure: the missing component in speech processing algorithms. *International Congress Series*, 1273, 23 - 26.

Summerfield, Q., Marshall, D.H., Barton, G.R., and Bloor, K.E. (2002). A cost-utility scenario analysis of bilateral cochlear implantation. *Archives of otolaryngology - head and neck surgery*, 128, 1255 – 1262.

Trahiotis, C., and Stern, R.M. (1989). Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase. *Journal of the Acoustical Society of America*, 86.4, 1285 – 1293.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *Journal of the Acoustical Society of America*, 115, 1729–1735.

Tyler, R.S., Parkinson, A.J., Wilson, B.S., Witt, S., Preece, J.P., and Noble, W. (2002). Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. *Ear and Hearing*, 23.2, 98 – 105.

Van de Par, S. (1998). A comparison of binaural detection at low and high frequencies. Dissertation Eindhoven University.

Van Hoesel, R.J., and Tyler, R.S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of the Acoustical Society of America*, 113.3, 1617-1630.

Van Hoesel, R.J.M., Böhm, M., Pesch, J., Vandali, A., Battmer, R.D., and Lenarz, T. (2008). Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *Journal of the Acoustical Society of America*, 123.4, 2249 – 2263.

Versfeld, N.J., Daalder, L., Festen, J.M., and Houtgast, T. (2000). Method for the selection of sentences materials for efficient measurement of the speech reception threshold. *Journal of the Acoustical Society of America*, 107.3, 1671 – 1684.

Villchur, E. (1974). Simulation of the effect of recruitment on loudness relationships in speech. *Journal of the Acoustical Society of America*, 56.5, 1601 – 1611.

Waltzman, S.B., and Roland, J.T. (2006). *Cochlear Implants*. Thieme medical publishers: New York.

Whitmal, N.A., Poissant, S.F., Freyman, R.L., and Helfer, K.S. (2007). Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience. *Journal of the Acoustical Society of America*, 12.4, 2376 – 2388.

Whitworth, R.H., and Jeffress, L.A. (1961). Time vs intensity in the localization of tones. *Journal of the Acoustical Society of America*, 33.7, 925 – 929.

Wightman, F.L., and Kistler, D.J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America*, 91.3, 1648-1661.

Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D.K., and Rabinowitz, W.M. (1991). Better speech recognition with cochlear implants. *Nature,* 352, 236 – 238.

Wilson, B.S., Schatzer, R., Lopez-Poveda, E.A., Sun, X., Lawson, D.T., and Wolford, R.D. (2005). Two new directions in speech processor design for cochlear implants. *Ear and Hearing*, 26, 73S – 81S.

Wilson, B.S., and Dorman, M.F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hearing Research*, 242, 3 – 21.

Xu, L., and Pfingst, B.E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception (L). *Journal of the Acoustical Society of America*, 114.6, 3024 – 3027.

Yost, W.A. (1974). Discriminations of interaural phase differences. *Journal of the Acoustical Society of America*, 55.6, 1299 – 1303.

Zeng, F.G., Popper, A.N., and Fay R.R (2004). *Cochlear implants*. New York: Springer-Verlag.

Zeng, F.G., and Galvin, J.J. (1999). Amplitude Mapping and Phoneme Recognition in Cochlear Implant Listeners. *Ear and Hearing*, 20.1, 60 – 74.

Zerlin, S. (1966). Interaural time and intensity difference and the MLD. *Journal of the Acoustical Society of America,* 39, 134-137.

Zwislocki, J., and Feldman, R.S. (1956). Just noticeable differences in dichotic phase. *Journal of the Acoustical Society of America*, 28, 860-864.