LATERALIZATION OF SOUNDS BASED ON INTERAURAL TIME DIFFERENCES IN COCHLEAR IMPLANT LISTENERS

Dissertation zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

Institut für Elektronische Musik und Akustik
Universität für Musik und darstellende Kunst
Graz, Österreich

Piotr Majdak

February 2008

Thesis Advisor: Robert Höldrich

Co Advisor: Bernhard Laback

Supported by:

Acoustics Research Institute,
Austrian Academy of Sciences

Revision: 295 (DIN A4, bibus),
Modified: 11. Feb. 2008

Zusammenfassung

Interaurale Zeitdifferenzen (engl.: *interaural time differences*, ITD) entstehen sobald der Schall ein Ohr des Zuhörers früher erreicht, was bei Schallquellen außerhalb der Medianebene des Zuhörers der Fall ist. Modulierte Signale wie Sprache oder Musik beinhalten die ITD sowohl in der schnell fluktuierenden Feinstruktur als auch in der langsam verlaufenden Einhüllenden. Es wird angenommen, dass die Feinstruktur-ITD eine sehr wichtige Rolle bei der Lateralisation von Schallquellen sowie beim Sprachverständnis im Störgeräusch spielt. Moderne Cochleaimplantate (CI), neuronale Hörprothesen für Schwerhörige, werden bilateral implantiert, um den Patienten die Vorteile des binauralen Hörens bieten zu können. Die bei klinischen CI-Systemen eingesetzten pulsatilen Stimulationsstrategien übertragen ITDs nur über die Einhüllende des Signals – die Feinstruktur des Signals und damit die Feinstruktur-ITD wird nicht kodiert. Einer der Gründe hierfür ist der geringe Wissensstand über die Effekte der Feinstruktur-ITD bei den eingesetzten Pulsraten von bis zu 1600 Pulsen pro Sekunde (pps).

In dieser Arbeit wurden die Effekte der Feinstruktur-ITD auf die Lateralization von Schallquellen bei CI-Trägern untersucht. Drei aufeinander aufbauende Studien zeigen, dass CI-Träger die Feinstruktur-ITD wahrnehmen und bis zu einer Pulsrate von 800 pps auswerten können. In der vierten Studie wurde eine Methode zur Verbesserung der Feinstruktur-ITD-Wahrnehmung vorgestellt und untersucht. Diese Methode verwendet binaural synchronisierten Jitter, um die Feinstruktur-ITD-Wahrnehmung bei hohen Pulsraten zu verbessern. Mit binaural synchronisiertem Jitter konnte die Lateralisation von Schallquellen bei CI-Trägern bei Pulsraten von bis zu 1515 pps verbessert werden.

Abstract

Interaural time differences (ITD) occur when a binaural signal is delayed at one ear. They arise when a sound is presented outside of the median plane of a subject. In modulated sounds like speech, ITD is present in both the rapidly varying fine structure and the slowly varying envelope. There is evidence that ITD in the fine structure of a sound is most important for sound localization and for understanding speech in noise. Cochlear implants (CI), neural prosthetic devices that restore hearing in the profoundly deaf, are increasingly implanted to both ears to provide implantees with the advantages of binaural hearing. Bilateral CI listeners currently use stimulation strategies that encode ITD in the temporal envelope but do not transmit ITD in the fine structure. The reason for this is that it is not clear if CI listeners are sensitive to the fine structure ITD for stimulation pulse rates which are required for a satisfactory perception of speech.

This thesis investigates the fine structure ITD sensitivity in CI listeners in three different studies. The results show that CI listeners are sensitive to the fine structure ITD up to a pulse rate of 800 pps. In a fourth study, a new method is investigated which improves the lateralization based on the fine structure ITD for pulse rate up to 1600 pps. The method introduces a binaurally-synchronized jitter to the regular pulse train and thus, reduces the binaural adaptation effect. With this method the CI listeners were able to lateralize sounds for high pulse rates as commonly used in current stimulation strategies.

Preamble

During the last six years I have been working on many projects for the Acoustics Research Institute (ARI, *Institut für Schallforschung*), which is a part of the Austrian Academy of Sciences (*Österreichische Akademie der Wissenschaften*). In 2003, Bernhard Laback and I started a working program called "Interaural time differences in the fine structure with cochlear implants", which has been funded by the Austrian Academy of Sciences. This thesis comprises portions of the outcome of this program.

This thesis has been designed and written by myself. The novel outcome presented in this thesis is based on already published or submitted articles, which were prepared together with Bernhard Laback. According to common standards in cases of shared work, statements about my contribution to the shared work are as follows:

- Majdak, P., Laback, B., and Baumgartner, W-D. (**2006**). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," *J Acoust Soc Am*, **120**, 2190-201.

  B.L. and P.M. discussed and developed the main ideas originally proposed by P.M.; P.M. reviewed the relevant literature; P.M. and B.L. designed the experiments; P.M. implemented the experimental procedures; P.M. and B.L. conducted the experiments; P.M. analyzed the results; B.L. and P.M. discussed the outcome; P.M. wrote the article; W-D.B. provided medical support for the subjects.

  Contribution: P.M. 70 %, B.L. 30 %.

- Laback, B., Majdak, P., and Baumgartner, W. (**2007**). "Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing," *J Acoust Soc Am*, **121**, 2182-91.

  B.L. and P.M. discussed and developed the main ideas originally proposed by B.L.; B.L. reviewed the relevant literature; P.M. and B.L. designed the experiments; P.M. implemented the experimental procedures; P.M. and B.L. conducted the experiments; B.L. analyzed the results; B.L. and P.M. discussed the outcome; B.L. wrote the article; W-D.B. provided medical support for the subjects.

  Contribution: P.M. 30 %, B.L. 70 %.

- Majdak, P., and Laback, B. "Effect of center frequency on the sensitivity to interaural delay in high-frequency clicks," submitted to *J Acoust Am Soc*.

B.L. and P.M. discussed and developed the main ideas; P.M. reviewed the relevant literature; B.L. and P.M. designed the experiments; P.M. implemented the experiment; P.M. and B.L. conducted the experiments; P.M. analyzed the results; B.L. and P.M. discussed the outcome; P.M. wrote the article.

Contribution: P.M. 60 %, B.L. 40 %.

- Laback, B., and Majdak, P. **(2008)**. "From the cover: Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," *Proceedings of the National Academy of Sciences U S A* **105**, 814-817.

  B.L. and P.M. discussed and developed the main ideas; P.M. and B. L. reviewed the relevant literature; P.M. and B.L. designed the experiments; P.M. and B.L. implemented the experimental procedures; P.M. and B.L. conducted the experiments; B.L. analyzed the results; B.L. and P.M. discussed the outcome; B.L. wrote the article.

  Contribution: P.M. 40 %, B.L. 60 %.

Piotr Majdak

Vienna, February 2008

Acknowledgments

I am very grateful to the many people who have helped me in every aspect of the thesis and encouraged me to begin, continue, and finish this work. My sincere thanks go:

- To my friend and mentor, Bernhard Laback, for all his help and support in completing this thesis. I totally appreciate the time being under his supervision and look forward to many more years of working with him. Don't give up on me, Bernhard, our (sometimes long-winded) discussions are not in vein!
- To Robert Höldrich, for trusting in my work and providing me with six years of protected freedom.
- To Jens Blauert, for using his valuable time to review my work.
- To Werner A. Deutsch for backing the ARI experimental audiology group.
- To Matt Goupell, for fruitful discussions, a lot of patience while going through my manuscripts, and being the best English teacher I ever had.
- To Michael Mihocic, for keeping the lab running. *Fahrma mess'n?*
- To Peter Balazs, Konrad Antonicek, Wolfgang Kreuzer, Jonnie White, Toni Noll, Timo Becker and other members of the ARI family, for their discussions on almost every aspect of science and life. I'm proud to be a part of you.
- To the test persons, in particular the cochlear implant listeners, for their patience while performing the longsome tests.
- To Wolf-Dieter Baumgartner, for his enthusiasm for cochlear implants.
- To MED-EL Corporation, especially Peter Nopp, for providing the equipment for bilateral direct electric stimulation, and for their interest in our work.

- To my friends and family, for encouraging and supporting me during the long period of my academic education.

- To Chiara and Sarah, for putting a smile on my face and keeping me awake, even when tired (especially during the night, Sarah ;-)

- To Berni, the best au-pair we ever had, for giving me the extra time to write not only these words.

- And most importantly, to my best friend, Ali, for her undying and unconditional love and support, and for being a wonderful companion in my life. I love you!

Piotr Majdak

Vienna, February 2008

Contents

I. Introduction

Blindness separates us from things but
deafness separates us from people.

Helen Adams Keller (1880 – 1968)

1. Cochlear implants

Over millions of years, mammals have developed five senses: sight, touch, taste, hearing, and smell. They provide the information about our environment and in the past have been essential for the survival of an individual: missing one sense sometimes determined the further existence.

Even in our “sophisticated” human civilization within the last century the presence of all senses is important. Severe hearing impairment or profound deafness is a serious handicap. Hearing impaired people frequently have the feeling of being socially restricted and being separated from the majority of society. Fortunately, beginning with the first direct electric stimulation of an auditory nerve in 1957 (Djourno and Eyries, 1957), they have a chance to partially retrieve their lost sense. Since the first successful implantation of an electrode in the cochlea in the years 1964 to 1967 (House, 1995), cochlear implants (CI) have gained popularity and as of 2008, over 100.000 people worldwide have received cochlear implants.

Current studies like Harris *et al.* (1995) show that recovery of hearing can improve both the quality of life and psychological well-being of CI users. They can achieve speech understanding, which is comparable with the level of normal hearing listeners in quiet (Friesen *et al.*, 2001). Since the speech understanding performance, associated with the question “what”, is satisfactory for most users, the question “where” is raised very soon after. Hence, the localization of sounds became an important topic in the electric hearing research as a consequence of the recent trend of implanting CIs to both ears. Naturally, with bilateral CI systems bilateral stimulation strategies have to be developed. Currently, several laboratories work on different aspects of such stimulation strategies. They address issues related to binaural hearing like lateralization of sounds and speech understanding in noise. In this thesis the focus was on the lateralization of sounds, in particular with respect to interaural time differences in the fine structure of a sound.

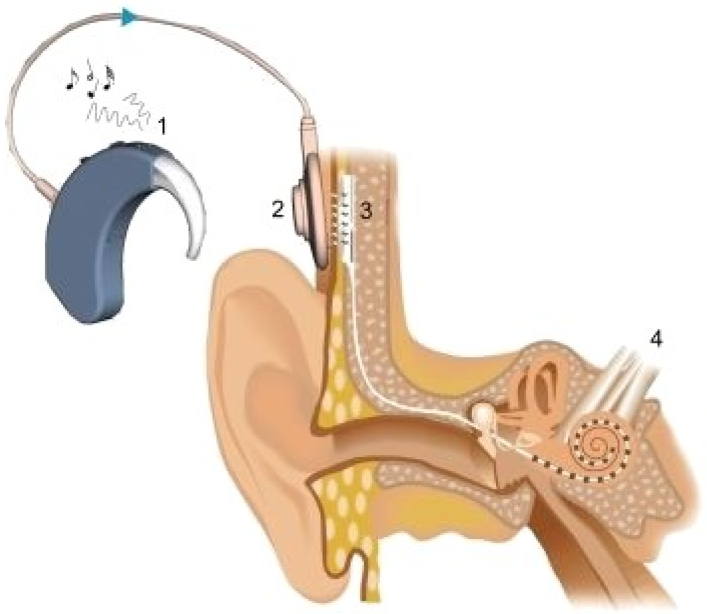

Figure 1: Cochlear implant system consisting of speech processor (1), transmitter coil (2), implant (3), and electrode bundle (4). From http://recorlsa.club.fr/implantcochleaire/descripgenerale.html (29.11.2007).

1.1. General

A cochlear implant is a partially-implanted electronic device that provides a sense of sound to a severely hard of hearing or profoundly deaf person. Unlike conventional hearing aids, a CI does not amplify sound. The incoming sound is captured and processed. Then, the information is converted to electric current pulses which directly stimulate parts of the cochlea.

Nowadays, the CI, or rather the CI system, consists of external and implanted parts (see Fig. 1). The speech processor (1) includes a microphone, a signal-processing device, and a power supply. The speech processor is usually placed behind the ear, although earlier devices were worn on the body. The processing of the acoustic information is based on a stimulation strategy and includes gain control and frequency dependent acoustic-to-electric mapping. The general concept of a stimulation strategy is presented in the next section. The output of the speech processor is connected to a transmitter coil (2), which transmits power and the processed information to the internal device by electromagnetic induction.

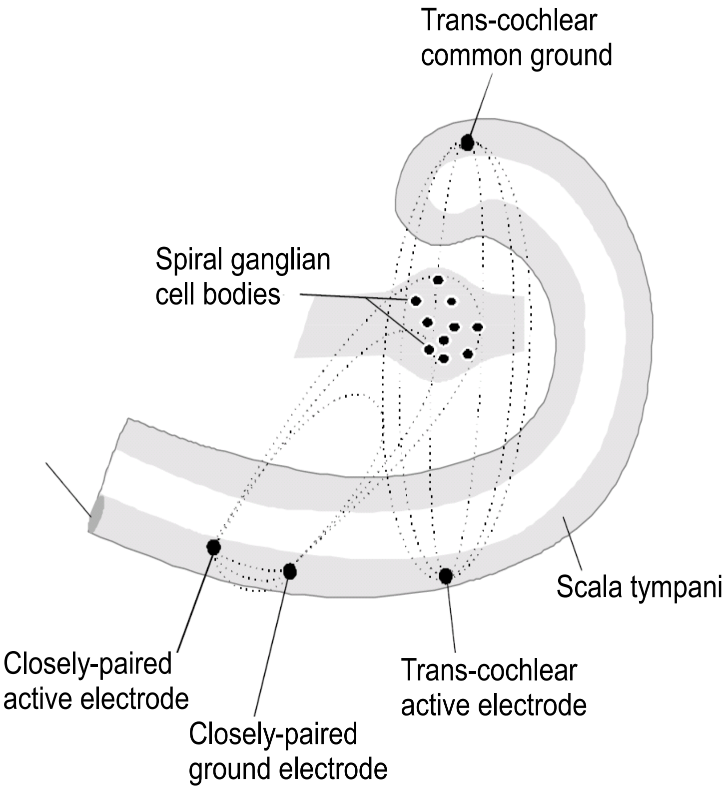

Figure 3: Comparison of the electrode configuration. Monopolar configuration (trans-cochlear electrodes) and bipolar configuration (closely-paired electrodes). From House (1995).

The internal receiver, the actual implant (3), is surgically placed under the skin behind the ear. It includes a receiver coil, a decoder unit and a stimulator. The stimulator is a configurable current source, which converts the received signal to electric currents and drives the electrodes (4). Additionally, the implant contains a strong magnet, which allows good adherence of the transmitter coil to the implant and efficient coupling of the electromagnetic field.

Up to 24 electrode pads are bundled and coated in silicon. Then, that electrode bundle is inserted via a small hole near to the round window and typically placed in the scala tympani (see Fig. 2). The typical insertion depth is 20 to 30 mm, however insertions as deep as 32 mm are used as well (Hamzavi and Arnoldner, 2006). Depending on the electrode configuration and manufacturer, 12 to 24 active electrodes are available.

There are two concepts for the electrode configuration, which are shown in Fig. 3. The monopolar configuration uses one of the electrodes located in the scala tympani as the active electrode. A common ground electrode is usually located outside the cochlea so that the current is directed across the modiolus and the spiral ganglion cells. In the bipolar configuration a common ground electrode is not needed. The current is applied between two closely-paired electrodes in the scala tympani. Compared to the monopolar configuration, the electric field is more focused (Shannon, 1983b); however, it requires higher currents (Pfingst *et al.*, 1997) to achieve hearing thresholds. Despite differences in the electric field patterns, there is no consistent effect of the electrode configuration on speech perception or pitch discrimination (Zwolan *et al.*, 1996; Franck *et al.*, 2003).

1.2. Stimulation strategies

The stimulation strategy has a large influence on the performance of a CI. The first CIs used only a single channel for stimulation with the advantage of simplicity of the processing strategy. Even with such simple systems, some subjects could obtain high scores on speech recognition tests (e.g. Tyler, 1988). The general speech understanding improved dramatically after introduction of multi-channel strategies in the early 1980s (Hinderink *et al.*, 1995; Cohen *et al.*, 1993). Nowadays, most stimulation strategies use multi-channel pulsatile stimulation. The reason for using pulses instead of a continuous stimulation is the broad excitation spread of the electric field in the cochlea. Simultaneous stimulation on more than one electrode results in an uncontrolled interaction through vector summation of the electric fields from each of the electrodes (Pelizzone *et al.*, 1999; de Balthasar *et al.*, 2003). This is avoided when stimulating one electrode at a time. Present pulsatile stimulation strategies include the continuous interleaved sampling (CIS; Wilson *et al.*, 1991), spectral peak (SPEAK; Skinner *et al.*, 1994); advanced combination encoder (ACE; Kiefer *et al.*, 2001), and "*n*-of-*m*" (corresponding to selection of *n* among *m* channels; Wilson *et al.*, 1988) strategies. Each of these strategies can achieve high performance in terms of speech perception as supported by many studies with between-strategy comparisons. All these strategies are monaural; any potential advantage of bilateral stimulation are automatically dismissed.

The CIS strategy is the most straight-forward strategy and is presented in Fig.4. The microphone signal is preprocessed by an automatic gain control (AGC) and a pre-emphasis filter, and then converted to a digital audio stream (ADC). The preprocessed data are directed to a

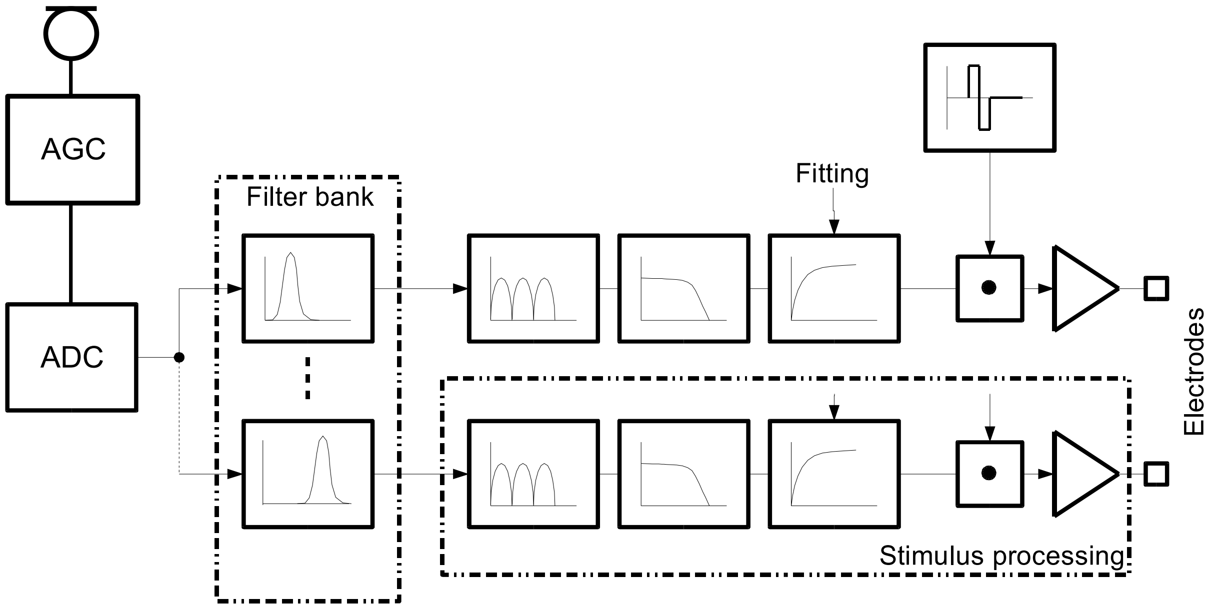

Figure 4: Block diagram of the continuous interleaved sampling (CIS) strategy. From left to right: analog-to-digital converter; filter bank with bandpass filters; rectifier; low-pass filter; signal-to-current mapping according to subjects fitting; modulation with pulse trains as carrier; amplifier.

bandpass filter bank and split into channels. Each channel is separately processed, which includes envelope detection, compression and modulation.

Envelope detection is performed either by rectification of the signal followed by a low-pass filtering or by means of Hilbert transform. The envelope detector usually has a low-pass cut-off frequency of 200 to 400 Hz. This range includes the fundamental frequency of voiced speech and rapid transitions of the consonants in the signal. Furthermore, for most subjects, this cut-off frequency is in the region of the upper limit for temporal pitch discrimination (Shannon, 1983a; Zeng, 2002).

In the next stage, the envelope signal is mapped to fit the stimulation range. I. e., the relatively wide dynamic range of the envelope signal is compressed to the narrow electric stimulation range. This mapping is based on parameters determined individually for each subject and each electrode. The most important parameters are: the threshold (THR), the most comfortable level (MCL), and the compression factor, which determines the shape of the mapping curve. Stimulation currents cover a range of a few hundreds of µA, rarely exceeding one or two mA.

The mapped-envelope signal is used to modulate a pulse train. The resulting modulated pulse train steers a current source connected to an electrode. In most systems, the pulses are biphasic, beginning with the cathodic phase. This results in a balanced charge transmission and is necessary to reduce the depolarization of the neurons in the cochlea. The pulse trains are interleaved across different channels to avoid simultaneous stimulation and separate the electric fields along the tonotopy. The rate of the pulse trains must be at least two times higher than the cut-off frequency of the envelope detector to avoid aliasing effects. However, results from recordings of auditory nerve responses indicate that the pulse rate should be even higher to achieve a good neural representation of the modulation waveforms (Wilson, 1997). Thus, CIS processors use pulse rates of 1000 pulses per second (pps) or higher for each electrode. For example, C40 and C40+ implants, which were used in this study, are typically driven with a pulse rate of 1515 pps per electrode.

Figure 5 illustrates the results of the stimulation for a simplified implementation of a four-channel CIS processor. The left panels show a 40-ms long vowel "a", the right panels show a 90-ms long unvoiced consonant "p". The uppermost panels show the acoustic input signals. The input signal was low-pass filtered with a cut-off frequency of 8.5 kHz and processed by a pre-emphasis filter (6 dB/octave attenuation below 1.2 kHz). The result was split into four channels by a filter bank and the envelopes were calculated. They were mapped to electric representation – the map was logarithmic with THR set to zero and MCL set to the maximum of the dynamic range. The results were used to modulate biphasic pulse trains with the pulse rate of 1515 pps and 50 µs pulse duration for each phase. The lower four panels show the output stimulation signal at the four electrodes. The electrodes are arranged in an apex-to-basal order; electrode four is the most-basal electrode. This arrangement mimics the tonotopic organization of the cochlea in normal hearing listeners.

The CIS, SPEAK, ACE, and “*n*-of-*m*” share the principle of interleaving the stimulation pulses across electrodes. The latter three strategies are based on CIS and include some additional features. The ACE and “*n*-of-*m*” strategies perform an additional scan to select the *n* highest envelope signals in each frame. Then, the stimulus pulses are delivered only to the subset *n* of *m* electrodes which correspond to the *n* selected channels. This is applied to reduce the density of stimulation across channels.

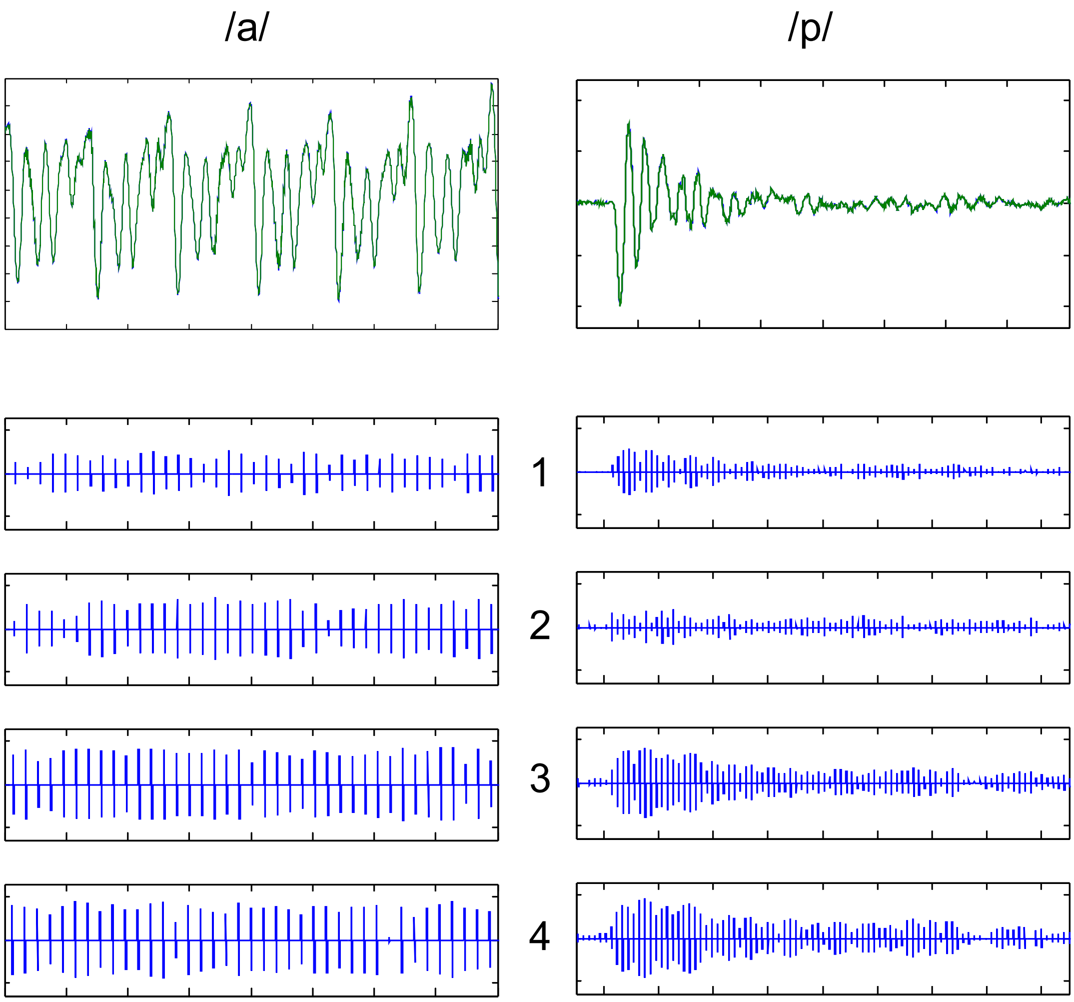

Figure 5: Signals produced by an implementation of a continuous interleaved sampling (CIS) processor. The left panels correspond to the 40-ms long vowel "a", the right panels correspond to the consonant "p" with the length of 90 ms. The uppermost panels show the acoustic input signals. The lower four panels show the output stimulation signal at the four electrodes. The lowest number corresponds to the most-apical electrode, the highest to the most-basal electrode. The pulse rate of 1515 pps per electrode was used.

The SPEAK strategy is an extension of the "n-of-m" approach: the number of stimulated channels n is adapted in each frame. The choice of n depends on the number of envelope signals above a threshold, on the distribution of energy across frequencies, and is mostly between six and ten channels. With this approach, the average stimulation rate can be lowered to 300 pps, which is necessary for systems with a slow transcutaneous link between the processor and implant. Results from comparisons of the SPEAK and ACE strategies show generally a lower performance of the SPEAK strategy (Kiefer *et al.*, 2001). Thus, CIS or ACE are the most preferred strategies by many clinicians.

However, all the multichannel strategies described here share the disadvantage of discarding the fine structure information. The pulse trains are generated in a predefined scheme, which does not consider the fine structure in the acoustic signal at all. This is particularly important for the low-frequency channels, where the fine structure transmits the temporal information important for pitch perception and sound localization (Wightman and Kistler, 1992; Smith *et al.*, 2002). The investigation of the effects of the fine structure for lateralization of sounds in electric hearing is the main topic of this thesis.

It is noteworthy that recently a new strategy has been presented by Med-El Corporation which considers the fine structure in the signal (Hochmair *et al.*, 2006). Their fine structure processing (FSP) strategy alters the fixed stimulation scheme of the CIS strategy to allow a stimulation at a zero crossing of the band-passed acoustic signals. This results in a better representation of the fine structure than that for the CIS strategy. However, currently FSP is performed for the most-apical electrodes only and no comparisons of speech perception and pitch discrimination scores between the FSP and CIS strategies have been published yet.

Another stimulation strategy, which considers the fine structure in the signal is the peak derived timing (PDT) strategy developed by van Hoesel and Tyler (2003). The PDT strategy was mainly investigated with respect to binaural effects and, thus, it is elaborately described in the section 3.2 of this chapter.

2. Interaural time differences (ITD)

2.1. Lateralization of sounds

A sound field produced by a lateral source arrives first at the ipsilateral ear and then at the contralateral ear (Fig. 6). The difference in time between the arrival of the wavefront at the two ears is called the interaural time difference (ITD). Additionally, the sound field is altered by the reflections of torso, head and pinna. The broadband difference in intensity between the two incoming sound fields is called the interaural level difference (ILD). Both ITD and ILD are used by the binaural auditory system to estimate the lateral position of sounds.

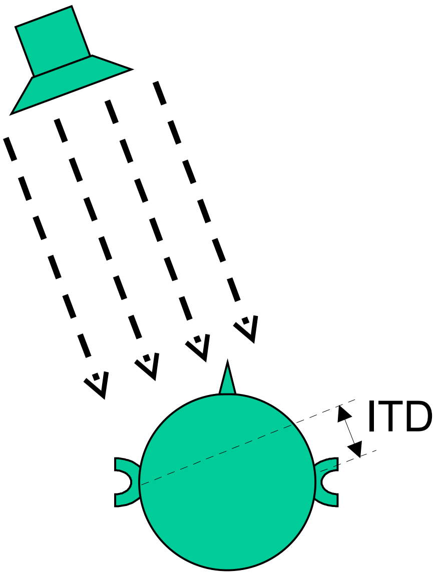

Figure 6: Origin of the interaural time differences (ITD).

Interaural differences were explored with respect to the localization of sounds more than hundred years ago by Strutt (1907) (also known as Lord Rayleigh), among others. His understanding of the localization of sounds in the lateral dimension has come to be known as the “Duplex Theory”. Lord Rayleigh's studies on the Duplex Theory were performed with pure tones presented via pipes. The pure tones evoked a shift in the perceived horizontal position of the sound event. This is referred to as lateralization. It is possible to change the perceived lateral position of sounds binaurally presented via headphones by applying ILD and ITD to signals.

The research in the past few decades shows, in addition to the Duplex Theory, the importance of spectral-shape cues. Spectral cues result from the direction-dependent filtering of broadband sounds by the torso, head and external ear (Wightman and Kistler, 1989) and can be described by the head-related transfer functions in 3-dimensional space (Møller *et al.*, 1995). Usually, a polar coordinate system is used to describe the position of a sound: azimuth describes the horizontal position and elevation describes the vertical position. However, using the polar coordinate system, the interaural and the spectral cues do not directly correspond to the azimuth and elevation, respectively. Morimoto and Aokata (1984) introduced a double-pole coordinate system (sometimes called horizontal-polar system), which represents the spatial position by the lateral angle for the horizontal direction and the polar angle for the vertical direction. They showed that only interaural cues are required for the lateral angle judgment. Spectral cues correspond to the perception along the polar angle. Hartmann and Wittenberg (1996) showed that both interaural and spectral cues are indispensable for localization of sounds accompanied by the out-of-head perception (externalization).

Macpherson and Middlebrooks (2002) revisited the Duplex Theory and investigated the contribution of ITD, ILD, monaural, and interaural spectral cues to the perception of the sound position along the lateral dimension. They found that the interaural cues are sufficient to achieve a proper lateralization in agreement with the Duplex Theory. Furthermore, they could show that the spectral cues, monaural and interaural, have only a little impact on the perceived lateral angle. Additionally, a comparison between ILD and ITD showed a higher or equal weighting for ITD than for ILD. This demonstrates the significance of the ITD as an important cue in the lateralization of sounds.

2.2. ITD in normal hearing listeners

Lord Rayleigh showed that human listeners are sensitive to interaural differences in the ongoing phase of low-frequency pure tones and that ITD alone could provide cues to the perception of the lateral positions of sounds. For ITDs in pure tones, the lower detection limit is approximately 10 µs (Blauert, 1997). It is strongly subject-dependent and training is mostly required to obtain such low limits. Generally, the sensitivity to ITD increases as the stimulus level or stimulus duration increases. This indicates an advantage for binaural tasks of having more responding neural units in the auditory system (Dye and Hafter, 1984). The auditory system detects the ITD in low-frequency pure tones with the help of coincidence detectors, which receive neural input from the left and right ear. This mechanism requires a sufficient phase locking to the phase of the stimulus. However, the phase locking is limited in frequency. In most mammals it becomes progressively less precise for stimulus frequencies above 1 kHz, and it disappears completely at about 5 kHz, although the exact upper limit varies across species (Palmer and Russell, 1986). These effects limit ITD perception on the neural level (Kuwada *et al.*, 1987). Additionally, for frequencies higher than approximately 1.5 kHz, the acoustic wave length is shorter than the interaural distance. Thus, interaural time differences yield ambiguous information about the actual interaural delays and an unique lateral position cannot be deter-

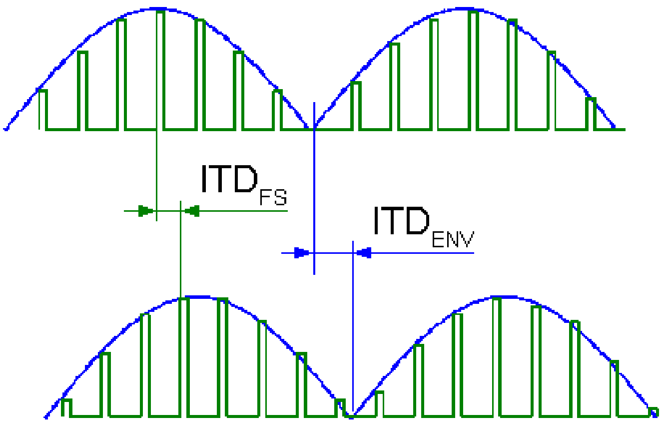

Figure 7: Fine-structure (FS) ITD (ITD_{FS}) and envelope (ENV) ITD (ITD_{ENV}) in a modulated pulsatile stimulus.

mined. In conclusion, because of both neurological and acoustic effects, ITD sensitivity to low-frequency pure tones rapidly decreases for stimuli exceeding 1.5 kHz (Zwislocki and Feldman, 1956; Blauert, 1997).

However, Klump and Eady (1956) showed that listeners are sensitive to ITDs in high-frequency complex sounds. Further investigations by others (e.g. Henning, 1974; Nuetzel and Hafter, 1976; Dye and Hafter, 1984; Buell and Hafter, 1988; Dye *et al.*, 1994; Bernstein and Trahiotis, 1994) showed sensitivity to ITD in the slowly-varying envelope of broadband signals. Physiological studies support these finding on the neural level (e.g. Yin *et al.*, 1984; Skottun *et al.*, 2001; Shackleton *et al.*, 2003). It seems that the auditory system is able to extract the timing information from the envelopes of high-frequency portions of sounds and detect the ITD in the envelope. Thus, following two types of ITD can be defined (see Fig. 7):

- Envelope ITD (ENV ITD), which is the interaural delay in the envelope of a signal. The envelope in a signal represents the slowly-varying fluctuations of the signal.
- Fine-structure ITD (FS ITD), which is the interaural delay in the fast-varying fine structure of a signal.

The ability to transmit both envelope and fine structure information in separate paths seems to be a general property of the auditory system (Liang *et al.*, 2002) and was also found in other sensory modalities like the visual and electrosensory systems (Middleton *et al.*, 2006).

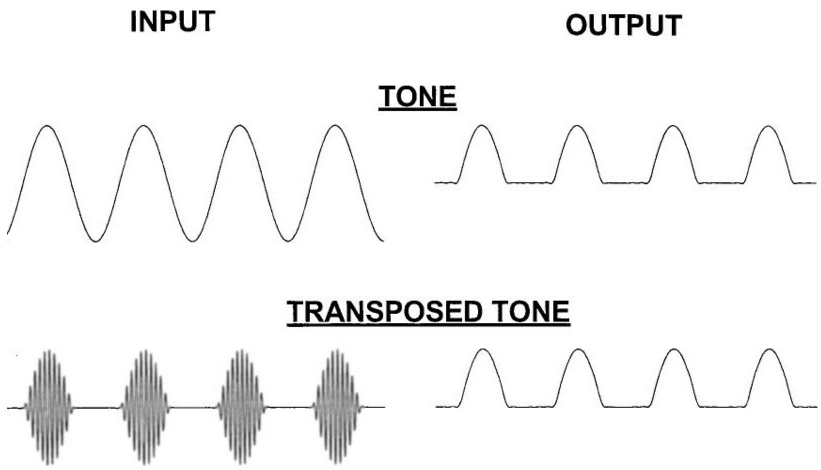

Figure 8: Left column: a 250-Hz tone (top) and a 250-Hz tone transposed to 4 kHz. Right column: the same stimuli after a simulation of the auditory filters (bandpass filtering, halfway-rectification, and low-pass filtering). From Bernstein and Trahiotis (2002).

Most studies show that ITD sensitivity to amplitude-modulated high-frequency sounds is lower than to low-frequency pure tones (e.g. Klump and Eady, 1956). Recently, Bernstein and Trahiotis (2002) measured ITD sensitivity to the transposed tones (van de Par and Kohlrausch, 1997). Transposed tones attempt to mimic the neural response of the auditory filters to low-frequency pure tones. They are composed by modulating a high-frequency carrier with a halfwave-rectified and subsequently lowpass-filtered low-frequency tone (see Fig. 8). Bernstein and Trahiotis (2002) showed that for lower modulation rates (< 256 Hz), the ITD sensitivity was comparable between conditions with transposed and pure tones. However, by increasing the modulation frequency, the performance for the transposed tones decreased rapidly while the performance for pure tones remained approximately constant. Bernstein and Trahiotis (2002) discussed these findings in terms of a possible effect of the peripheral processing, namely, auditory filtering. This effect smears the envelope fluctuations, which results in a lower modulation depth of the internal representation of the acoustic signal. This may yield lower ITD sensitivity. Bernstein and Trahiotis (2002) repeated the experiment for higher center frequencies (CF). If auditory filtering limits the performance, the shorter responses of the filters at higher CFs, producing less smearing, should result in a higher ITD sensitivity. They found a generally worse performance at higher CFs, which rejected the hypothesis of auditory filtering as the limiting factor. They concluded that it must be a mechanism “that serves to limit the ability to 'follow' or to encode high rates of fluctuation of the envelope of high-frequency, complex waveforms” (Bernstein and Trahiotis, 2002, pp. 1033). Although this effect has been shown for transposed tones, it is still unclear if this effect can be generalized to other types of stimuli. The investigation of the impact of auditory filtering on ITD perception for bandpass-filtered pulse trains is presented in Chapter IV.

Besides amplitude modulation (AM) of stimuli, investigations on the ITD sensitivity to stimuli that have frequency modulation (FM) have been reported (Henning, 1974; Blauert, 1981; Saberi, 1998). These stimuli are of special interest because they do not have amplitude fluctuations in the envelope and they transmit ITD information in the phase of the carrier only. Blauert (1981)[1] tested ITD sensitivity using frequency modulation in the stimulus by jittering the phase of a 4-kHz pure tone. The jitter was binaurally synchronized, which preserved the ITD information in the stimulus. The modulated stimuli evoked higher ITD sensitivity than unmodulated pure tones. He explained the performance improvements by an FM-to-AM-mapping of the signal on the basilar membrane, which results from applying an FM signal to a very steep bandpass filter, such as an auditory filter. Hence, the auditory system still detects the ITD information in the envelope fluctuations because the auditory filters map the acoustic FM signal to the amplitude-modulated representation. Further systematic investigations by Saberi (1998) confirmed the FM-to-AM conversion from the cochlear filtering.

2.3. Binaural adaptation

Hafter and Dye (1983) used bandpass-filtered click trains to demonstrate that ITD information is integrated over time. They assumed that ITD perception in one single click is limited by internal noise in the auditory system of the listener. Assuming a statistical independence between the subsequent pulses, each pulse contributes to the perception and lowers the standard error of the internal noise. Thus, the thresholds should decline by the square root of the total number of clicks n:

1 Nordmark (1976) published a study about the ITD sensitivity to jittered pulse trains before the study of Blauert (1981). However, the methods are unclear and the results could not be confirmed, so we consider this study as questionable. The study of Blauert (1981) was a response to the study of Nordmark (1976).

$$JND_n = \frac{JND_1}{\sqrt{n}} \quad or \quad \log JND_n = \log JND_1 - 0.5 \cdot n \tag{1}$$

where JND_1 and JND_n refer to the thresholds for one and n pulses, respectively. Although this integration model could explain the data of Hafter and Dye (1983) for lower pulse rates (≤ 100 pps), for higher pulse rates the JND_ns were higher than predicted.

Five years later, by varying the duration of the bandpass-filtered pulse trains Buell and Hafter (1988) found that the integration over the number of pulses n follows the inter-pulse interval (IPI). Hence, they generalized Eq. 1 to:

$$\log JND_n = \log JND_1 - 0.5\, k \cdot n \ , \tag{2}$$

where

$$0.0 \leq k \leq 1.0 \quad and \quad k = f(IPI) \ . \tag{3}$$

They showed that for lower pulse rates (approx. 100 pps), the integration of ITD information is optimal ($k=1$) and each pulse pair contributes to the ITD perception according to Eq. 1. However, for higher pulse rates, there is less integration ($k<1$): by adding new pulse pairs to the stimulus, ITD perception improves less than for lower pulse rates. Buell and Hafter (1988) referred this effect to as *binaural adaptation* and they used the factor k to describe the extent of binaural adaptation. They showed that the factor k decreases with increasing pulse rate.

For higher pulse rates, binaural adaptation has such a strong effect on ITD perception that primarily the onset of the sound is processed by the binaural system (Saberi, 1996). Stecker and Hafter (2002) showed that the onset, offset, and ongoing parts of a signal have generally different weighting in ITD perception at higher rates. Hence, the envelope ITD information is transmitted in different parts of the stimulus, which can be referred to as:

- Onset ITD, which is the ITD in the onset of a signal.
- Offset ITD, which is the ITD in the offset of a signal.
- Ongoing ITD, which is the ITD in the ongoing part of the signal. The ongoing ITD is not necessarily related to the interaural delay in the fine structure of a signal. It relates also to the slowly-varying envelope of a signal.

The effect of binaural adaptation seems to be ubiquitous covering a variety of stimuli including low-frequency pure tones, noise, and high-frequency AM stimuli (Yost and Hafter, 1987).

Going one step further, studies with normal-hearing (NH) listeners showed that by introducing a change in the ongoing signal (a trigger) the portion of the signal following the trigger becomes more important. This is supported by an improved ITD sensitivity (Hafter and Buell, 1990; Stecker and Hafter, 2002). They explained this effect as a *recovery from binaural adaptation* (RBA). The experiments were performed with NH listeners and thus could not rule out spectral effects because introducing a trigger in the time domain results in a concomitant spectral change. Thus, it is still unclear if the RBA effect is caused by temporal or spectral trigger.

Fortunately, a purely-temporal trigger can be realized by testing the ITD sensitivity in CI listeners when a single binaural electrode pair is used. The RBA effect was investigated in this thesis using a purely-temporal trigger in CI listeners. The study and the findings are described in Chapter V.

3. Bilateral CI systems

3.1. General

In electric hearing, the stimulus is directly presented via an electrode at a given tonotopic position. The timing of pulses in an electric pulse train corresponds to the fine structure of the signal. Thus, fine-structure ITD is the interaural delay between two binaurally-corresponding pulses in the trains. In modulated pulse trains, which for example result from processing of a speech signal with the CIS strategy, envelope ITD is the interaural difference between the envelopes.

A single electric pulse presented at one electrode can be approximated in acoustic hearing by a bandpass-filtered pulse. This is demodulated in the cochlea and stimulates a tonotopic position corresponding to the electrode position. An unmodulated electric pulse train can be represented in acoustic hearing as a bandpass-filtered unmodulated pulse train. As a result of a bandpass filtering, every filtered pulse has an “envelope”, which is the envelope of the impulse response of the bandpass filter. In such a case, the “fine structure” would refer to the fast fluctuations of the impulse responses. An envelope of the electric pulse train, e.g., as a result of the CIS strategy, would appear in the acoustic stimulus as a slowly-varying second-order envelope. Thus, the terms “envelope” and “fine structure” have different meaning in the NH and CI literature. Therefore, there is an inconsistency in nomenclature, which may cause confusion.

Thus, the definitions from the CI literature have been adopted for the acoustic stimuli. The term “fine structure” defines the total impulse response of the bandpass filter, not the carrier only. The term “envelope” refers to the slowly-varying amplitude modulation of the filtered pulse trains. Thus, the terminology now describes the same effects in electric and acoustic stimulation.

3.2. Binaural cues available in current CI systems

Currently, most stimulation strategies have been designed for monaural use. They do not include any special binaural coordination: electric stimulation is controlled by two independently running speech processors. However, the effect of wave propagation between the two ears acts as a kind of acoustic synchronization.

For lateral sounds, the head shadow causes interaural level differences (ILD). Access to ILD cues is easily available provided that the speech processors have the same configuration, even if they are not binaurally synchronized. However, the implemented AGC changes the operation level of incoming sound. If this is performed differently at each ear, the ILD is modified and this can lead to higher lateralization errors (Schleich and Nopp, 2002). However, if slow time constants are used for the AGC, which is the more frequently used configuration for speech compression, the ILD cues in the onset of the signals are not affected by the AGC. This reduces larger lateralization errors in onset sensitive CI listeners (Noel and Eddington, 2007).

The other interaural cue is the ITD, which can be transmitted in the fine structure and in the envelope of the stimulus. However, the CIS-based strategies do not consider the timing of the fine structure – only the envelope is encoded. As a consequence, the ITD information is coded in the envelope only and FS ITD is not available.

The stimulation strategies have one other aspect in common: according to the specification, they use a constant stimulation pulse rate at both ears. Due to the lack of synchronization between the two ears, the stimulation pulses have an uncontrolled interaural delay, which can be regarded as FS ITD. This depends on the switch-on delay between the processors and has a random value between 0 µs and the IPI. If CI listeners are sensitive to FS ITD, it might interact with other lateralization cues like ENV ITD or ILD. It is yet unclear if the CI listeners are sensitive to FS ITD.

One strategy which encodes timing information in the fine structure is peak derived timing (PDT) introduced by van Hoesel and Tyler (2003). PDT takes into consideration the fine structure of acoustic signals and provides electric signals that do transmit FS ITD. In the PDT strategy, the temporal position of an acoustic peak in a subband is determined and an electric pulse is applied to the corresponding electrode at the corresponding time. As a consequence, the pulse rate varies according to the temporal properties of the acoustic signal at each channel and was limited to a maximum of 1400 pps. Van Hoesel and Tyler (2003) could not find any clear difference between the PDT strategy and the standard clinical stimulation strategy with respect to sound localization and speech perception in noise. Unfortunately, the comparison between the two strategies was confounded by differences in the experimental setup such as automatic gain control (AGC), dynamic range, and number of electrodes. The average pulse rate was about 700 pps, which might have resulted in a low FS ITD sensitivity. Hence, more detailed investigations into the basic effects of fine-structure ITD are essential to determine the extent of lateralization improvement for CI listeners.

4. Outline of the current thesis

The main purpose of this thesis is to investigate fine-structure ITD sensitivity in CI listeners. Additionally, hypotheses about how to improve that sensitivity are presented and tested.

In **Chapter II** very basic experiments investigate if CI listeners are sensitive to fine-structure ITD. The main goal of that study is to separate the effects of the fine-structure (ongoing) ITD and portions of the envelope (onset and offset) ITD. Special stimuli, consisting of only four pulses, are used in lateralization discrimination experiments with CI listeners. Additional experiments with NH listeners are performed by using an acoustic simulation of the electric stimulation. The results show that all subjects are sensitive to ITD cues in the ongoing parts of the stimuli, which indicates that the fine-structure ITD is an important cue in lateralization.

However, the effects of using such artificially-short stimuli may be not generalizable to the real-world conditions.

Thus, in **Chapter III**, more realistic stimuli are used to test sensitivity to fine-structure and envelope ITD. The stimuli are 300-ms amplitude-modulated pulse trains. Again, both subject groups are tested. The results show that the fine-structure ITD has a larger impact on lateralization discrimination than the envelope ITDs tested. However, the ITD sensitivity decreases with increasing pulse rate, showing almost no sensitivity at rates that are important for transmission of speech information. Additionally, the results for ITD values larger than ¼ IPI reveal an ambiguity in the left/right representation for stimuli with the fine-structure ITD only. A method to reduce this problem in a potential stimulation strategy is presented.

In both studies, some CI listeners show a better FS ITD sensitivity at higher pulse rates than the NH subjects. This is unexpected and thus, it is conjectured that the comparison of the effects between the subject groups may be affected by the specific type of acoustic simulation of electric stimulation. One potential problem comparing the subject groups is auditory filtering in the acoustic hearing. Auditory filtering may smear the temporal information in the stimuli in such a way that it forms a rate limit in the ITD perception. In electric stimulation, the auditory filters are bypassed.

Therefore, **Chapter IV** presents a study, which addresses the effect of auditory filtering at ITD perception in acoustic hearing. The effects of center frequency and rate in acoustic stimulations are investigated and discussed. Even though a significant effect of the center frequency was found, it is concluded that it is unlikely that the auditory filtering is the rate limiting factor in ITD perception. The most feasible explanation for the rate limitation is some mechanism in the higher stages of the auditory system. The origin of differences in the performance between the subject groups is assumed to be due to differences in the phase locking mechanisms.

Typically, clinical stimulation strategies are configured to pulse rates of 1000 pps and higher. For such high pulse rates, the fine-structure ITD has essentially no effect on lateralization because the sensitivity to this cue degrades rapidly for pulse rates above a few hundred pps. Hence, in **Chapter V**, a method to reduce the effect of the rate limit is presented. Based on findings in the NH literature about the phenomena of binaural adaptation and methods to evoke a recovery from binaural adaptation, a new method to improve ITD perception at higher rates is presented and evaluated. Temporal triggers in form of binaurally-synchronized jitter in pulse trains are tested in CI listeners for pulse rates up to 1515 pps. For the jittered stimuli, the results show a drastic improvement of ITD sensitivity compared to the results for regular pulse trains. This promises improvements in the localization of sound sources and in the understanding of speech in noise with future stimulation strategies.

Chapter VI summarizes the output of the studies presented and provides a short outlook into the future.

II. Interaural time differences in different portions of four-pulse sequences

Based on *Journal of the Acoustical Society of America* **121**, 2182-2191 **(2007)**

As mentioned in the introduction of this thesis, ITD perception in CI listeners was already topic of several psychophysical studies (van Hoesel and Clark, 1997; van Hoesel and Tyler, 2003; Lawson *et al.*, 1998; Lawson *et al.*, 2001; Long *et al.*, 2003; Laback *et al.*, 2004; Senn *et al.*, 2005). These studies showed that CI listeners are sensitive to ITD, although there is a large inter-individual variability in performance. Most of these studies used, among other stimuli, unmodulated, rectangularly-gated pulse trains. For this type of stimulus, the results do not reveal to what extent listeners exploit ITD information in the ongoing signal as opposed to the information in the onset and offset portions (gating) of the stimulus. Motivated by those studies, the present study used a lateralization discrimination task to investigate if CI listeners are sensitive to ITD information presented in different portions of a pulse train. For that purpose, a specific stimulus is required, which allows the strict separation and independent control of ITD in the ongoing and in the gating portions. Thus, a stimulus was chosen, which consists of a sequence of four pulses with constant amplitude. In this stimulus, the first and last pulse represent the gating portions and the two pulses in the middle represent the ongoing portion. Therefore, the ITD information could be presented either in the ongoing signal, onset, offset, or in the combinations of them, i.e. both onset and offset, or in the entire pulse train. While this kind of stimulus has several advantages, which are described later in this chapter, its artificially-short duration may complicate the generalization of the outcomes of this study to longer stimulus durations. Thus, in the follow-up study, which is presented in the next chapter, we used more realistic stimuli. However, this artificial stimulus allowed to investigate the relative contributions of the ongoing, onset, and offset portions to the ITD sensitivity.

Considering an unmodulated electrical pulse train, ongoing ITD is present solely in the fine timing of the individual pulses. This can be referred to as the "fine structure". This term is used in the psychoacoustic literature to define the rapidly varying carrier frequency of an acoustic waveform. Referring to the "fine structure" of an electrical pulse train without envelope modulation is uncommon, however it appears appropriate in this context, as explained in Chapter I. Thus, in this case, the lateralization discrimination performance for ongoing ITD can be interpreted as an estimation of fine structure ITD sensitivity.

Findings from the NH literature show that the relative importance of ongoing and gating ITD depends on the rate of the stimulus. As long as the frequency of the signal component transmitting the ITD, the carrier or the envelope, does not exceed a certain limit, the ongoing

ITD is the primary lateralization cue. If the ITD is transmitted in the carrier, the frequency limit is around 1500 Hz (Klump and Eady, 1956; Zwislocki and Feldman, 1956). In the case of envelope ITD imposed on a high frequency carrier, the frequency limit depends on the temporal characteristics of the stimulus and appears to be lower than for the carrier ITD (Henning, 1974; Hafter and Dye, 1983; Bernstein and Trahiotis, 1994; Bernstein and Trahiotis, 2002). Generally, ITD in the gating portions has more impact at higher signal frequencies (Saberi, 1996), particularly when ambiguous ongoing ITD cues are involved (Freyman *et al.*, 1997).

To examine these effects in CI listeners, the current study examined lateralization discrimination for different ITD conditions as a function of pulse rate. The measurements obtained from the CI listeners were compared to the measurements performed on NH subjects who listened to an acoustic simulation of electric stimulation. McKay and Carlyon (1999) and Carlyon *et al.* (2002) have shown that some aspects of temporal pitch perception of NH subjects listening to such an acoustic simulation correspond to the perception of CI listeners. We hoped that the simulation technique used in these studies will mimic some aspects of ITD perception in electric hearing because both temporal pitch perception and ITD perception are based on the temporal properties of the stimulus. However, it has to be kept in mind that using that simulations, NH listeners can discriminate rate pitch up to much higher pulse rates than CI listeners.

In NH listeners, it could be demonstrated that ITD information in the temporal fine structure is important for the lateralization of sound sources (Wightman and Kistler, 1992; Smith *et al.*, 2002) and for speech perception in noise (Nie *et al.*, 2005; Zeng *et al.*, 2005). Thus, if CI listeners are sensitive to ITD in the ongoing signal and thus the fine structure, then transmission of fine structure ITD in future CI stimulation strategies will be a promising approach for improving the ability to lateralize sound sources and to understand speech in noise.

1. Method

1.1. Subjects

We tested four postlingually deafened CI listeners, which were implanted bilaterally at the Vienna University Hospital (CI1, CI3, and CI8) and at the University Clinic Würzburg (CI12).

Subject	Participating in the experiments	Etiology	Age	Age at implantation (yr)		Duration of deafness		Binaural electrical stimulation experience	Performance for baseline condition (% correct)	Performance for 300 ms version (% correct)
			(yr)	L	R	L	R			
CI1	yes	Meningitis	20	14	14	5.5 mo	1.5 mo	6 yr	80,0	98,3
CI3	yes	Meningitis	21	21	21	2 mo	2 mo	1 mo	96,0	99,0
CI8	yes	Osteogenesis imperf.	41	41	39	3 yr	12 yr	2 mo	73,0	77,0
CI12	yes	Sudden hearing loss	40	35	34	8 yr	3 yr	5 yr	95,0	99,0
CI2	no	Skull trauma	58	54	48	21 yr	25 yr	4 yr	58.3	70,0
CI6	no	Progressive	42	41	39	8 yr	8 yr	1 yr	60,0	65,0
CI5	no	Otosclerosis	44	35	42	2 yr	9 yr	2 yr	59,4	73,7
CI9	no	Progressive	58	50	51	5 yr	5 yr	7 yr	60,6	75,9

Table 1: Data of the four CI listeners completing the experiments (CI1, CI3, CI8, and CI12) and of the four CI listeners who showed too poor sensitivity thus were not included for participation in the experiments. The rightmost columns show the lateralization discrimination scores for the baseline condition (four pulses; waveform ITD of 600 µs; 100 pps) and the 300-ms version of the baseline condition. Data reprinted from Laback et al. (2007).

They were selected from a total of eight CI listeners invited for participation in the study. These four listeners were selected because they were able to reproducibly perform left/right discrimination on the basis of 600-µs waveform ITD in a sequence of four pulses at a pulse rate of 100 pps. The remaining four listeners showed very low discrimination scores for this baseline condition, even after many hours of training. Table 1 shows data on all eight listeners' etiology. Additionally, it shows the percent correct scores achieved for the baseline condition in a final test after training. The performance for a 300-ms version of the same stimulus is included to allow comparison with other studies which used this stimulus duration. All these scores are based on at least 180 item repetitions. The data of the patients not participating in the experiments are included in the table to make them available for future analysis.

All eight subjects had been supplied with the C40+ system by MED-EL Corp. With this implants, non-simultaneous biphasic current pulses (cathodic phase first) on up to 12 electrodes (2.4-mm spacing) can be generated. They provide stimulation in monopolar configuration with an extracochlear ground electrode. The numbering order of the electrodes is from apical to basal electrodes' position in the cochlea.

The four selected listeners had normal hearing before the onset of deafness. The time period between the beginning of deafness at the first ear and the activation of the second CI (duration of binaural deafness) was 2 months (CI3), 5.5 months (CI1), 8 years (CI12), and 12 years

(CI8). Subject CI3 was supplied with both CIs in one operation. Subjects CI1, CI12, and CI8 were successively supplied with CIs at the two ears with a temporal gap of four months, one year, and two years, respectively. The binaural electric stimulation experience was 6 years (CI1), 1 month (CI3), 2 months (CI8), and 4 years (CI12). The short binaural experience of CI3 and CI8 may be considered as a potential problem. However, we had the opportunity to repeat tests with these two listeners two years (CI3) and one year (CI8) after the main tests. We observed no change in ITD sensitivity. The results of these tests are presented in Chapter III[2]. Notice that clinical CI systems use constant pulse rates and thus discard fine binaural timing cues. Thus, they provide no stable fine structure ITD cues in everyday listening. This indicates that the short binaural CI experience of the listeners CI3 and CI8 did not influence their ITD sensitivity.

Five NH listeners participated in this study. They were 25 to 35 years old. None of them had any indication of present or past hearing disorder. Subjects NH2 and NH4 were the first two authors of the study Laback *et al.* (2007). Only NH6 had previous no experience with psychoacoustic experiments.

1.2. Apparatus

A personal computer system was used to control electric and acoustic stimulation. Each implant was controlled by a Research Interface Box (RIB), manufactured at the University of Innsbruck, Austria. The two RIBs were synchronized, providing an interaural accuracy of stimulation timing of 2.5 μs. Both RIBs were connected to the personal computer system via serial interfaces. The stimuli were verified using a pair of dummy implants (Detektorbox, MED-EL), which were connected to a two-channel storage oscilloscope (softDSP, SDS 200).

The stimuli for acoustic stimulation were output via a 24-bit stereo A/D-D/A converter (ADDA 2402, Digital Audio Denmark) using a sampling rate of 96 kHz per channel. The converter received the data from the computer via a digital audio interface (DIGI 96/8 PRO, RME). The analog signals were sent through a headphone amplifier (HB6, TDT) and an attenuator (PA4, TDT) and presented to the subjects via a circumaural headphone (K501, AKG). Calibration of the headphone signals was performed using a sound level meter (2260, Brüel & Kjær) connected to an artificial ear (4153, Bruel & Kjær). The headphone signals were further

2 see also Majdak et al. (2006)

Subject	Test electrodes	Right electrode higher	Current levels (in µA) used in exp. I and II
	L / R	(in %)	L / R
CI1	4 / 1	50.0	261 / 248
CI3	4 / 3	42.3	265 / 283
CI8	7 / 5	45.0	376 / 358
CI12	2 / 2	53.8	547 / 601

Table 2: Parameters of the electric stimulation. The column "Right electrode higher" shows the results of the pitch discrimination pretest. Values between 32% and 68% indicate not significant pitch discriminability ($p > 0.05$). Data reprinted from Laback et al. (2007).

inspected by digital signal analysis software after digitizing them through the A/D-D/A converter.

1.3. Stimuli

Trains of four biphasic current pulses with constant amplitude and a phase duration of 26.7 µs were used for the electric stimulation. The pulse rate varied between 100 and 800 pps. The amplitude of the stimuli was set to a comfortable loudness, as shown in Table 2. It was constant for all pulse rates.

To determine the comfortable loudness, the threshold, the comfortable level (CL), and the most comfortable level for each electrode were determined. The perceived loudness was indicated by the subjects by pointing to the appropriate position on a continuous scale, ranging from "not audible" to "just uncomfortably loud". The CL corresponded to the subject's response "comfortable". The same procedure was then applied to determine the binaural CL, i.e., the comfortable level when both ears were stimulated simultaneously. Starting at 80% of the monaural CLs, levels were varied simultaneously in equal steps at the two ears. Subjects were instructed to attend to the overall loudness in the binaural case rather than to "hear out" a left-ear or right-ear contribution. Following the initial adjustment of the binaural CL, centralization of the perceived stimulus was checked and monaural levels were adjusted if necessary. All CI listeners required a reduction of current levels in the binaural condition relative to the monaural conditions to achieve the same loudness.

Acoustic stimuli were used to simulate electric stimulation at a single electrode. They were similar to stimuli used by McKay and Carlyon (1999) and later by Carlyon *et al.* (2002). Pulse trains consisting of monophasic pulses with a duration of 10 μs were band-pass filtered with -3 dB cutoff frequencies at 3900 and 5400 Hz. The filter was a digital eighth-order Butterworth filter and had slopes of 48 dB/octave. The bandwidth of the filter was broad enough to preserve the modulation in the stimuli. The stimuli were windowed with the Tukey window (duration of the onset and offset tapers: 0.6 ms) to avoid truncation of the impulse response, which could cause detection of transient cues. The sound pressure level re 20μ Pa (SPL) corresponded to 78 dB of a continuously presented pulse train with a pulse rate of 100 pps.

The choice of the band-pass filter center frequency was a compromise between two effects. On one hand, the auditory filter bandwidth increases with center frequency. Therefore, its smearing effect on the stimulus temporal envelope decreases with increasing center frequency, which favors the choice of a high frequency. On the other hand, for carrier frequencies exceeding 4 to 6 kHz, the sensitivity to ITD in amplitude modulated tones decreases (Henning, 1974; Bernstein and Trahiotis, 2002), which favors the choice of a low frequency. Hence, the center frequency of 4589 Hz was chosen, which was also the choice of McKay and Carlyon (1999).

The main idea of the acoustic simulation of the electrical stimulation is using band-pass filtered pulse trains, where each pulse corresponds to the impulse response of the band-pass filter. Thus, each electric pulse is represented by an acoustic complex waveform, which has a fine structure and an envelope. This differs from electrical pulses, which have no fine structure of their own. It is assumed that information in the envelope of band-pass filtered acoustic pulse trains corresponds to information in the “fine structure” of electric pulse trains, even though the fine structure is not effectively represented in the neural response to high acoustic frequencies. Thus, we consider the band-pass filtered pulse trains as an appropriate simulation of electrical stimulation.

The band-pass filtered click trains produce stimulation outside of the pass-band. Thus, interaurally uncorrelated pink noise was presented continuously at both ears to mask the portions outside of the desired frequency band. The frequency range of the noises was from 50 to 10050 Hz. The spectrum SPL was 15.2 dB at 4.6 kHz. The noises were generated and mixed with the pulse trains in real-time. Furthermore, presenting the noise to the subject, any nonlinear distortion products were masked. The continuous background noise may cause an overall decrease of ITD sensitivity. Thus, the ITD sensitivity was measured for a reference condition (300 ms pulse trains, waveform ITD, pulse rate of 100 pps). The just noticeable differences (JND) measured in five NH listeners were as low as 40 µs with a standard deviation of 5 µs, which is close to the minimum detectable ITD (e.g. Blauert, 1997). Hence, the noise did not have substantial impact on ITD sensitivity in this study.

2. Pretests

Interaural electrodes with similar pitches seem to be more likely to show ITD sensitivity, although the effect of increasing interaural place difference can be small (van Hoesel, 2004; Long *et al.*, 2003). Thus, for the CI listeners, pretests were required in order to locate an interaural pair of electrodes eliciting the same pitch percept. These pretests were performed using the 300-ms version of the baseline condition (pulse rate of 100 pps; duration of 300 ms).

The procedure to find a pitch-matched electrode pair involved the following steps:

- Determination of electric dynamic range and comfortable level for electrodes 1-8 on each ear;
- Estimation of monaural pitch sensation across the electrode arrays to determine pitch-matched interaural electrode pair candidates;
- Interaural loudness balancing for each interaurally pitch-matched pair candidate;
- Measurement of pitch discriminability for each interaurally pitch-matched pair candidate and final selection of one pitch-matched pair.

A magnitude estimation procedure was applied to obtain an estimate of the perceived pitch across the electrodes at both ears, similar to the procedures applied by Busby *et al.* (1994) and Collins *et al.* (1997). Stimuli were presented randomly between both ears and at each of the electrodes 1 to 8, using the binaural CLs determined before. Subjects were instructed to assign numbers according to the perceived pitch of each stimulus. No restrictions on the range and type of numbers were given. Each stimulus was presented ten times. The distribution of pitch judgments across the electrodes and the two ears allowed selection of about 16 interaural electrode pairs supposed to elicit similar pitch sensation at the two ears. These pairs were evaluated further in the pitch-ranking task.

An automated procedure was applied to obtain interaurally loudness-balanced levels for each of the electrode pairs used further in the pitch-ranking task. The members of each electrode pair were presented in two subsequent intervals. By pressing one of two buttons the subjects adjusted the relative level of the signals between the two ears in steps corresponding to the smallest amplitude changes realizable by the implants to arrive at an interaurally matched loudness. The sum of the two levels within a trial was held constant and corresponded to the sum of the binaural CLs determined for the respective electrodes. The level difference at the beginning of each run was randomly roved. The mean value resulting from four runs was defined as the loudness-balanced levels for the members of the respective electrode pair.

In the pitch-ranking procedure the members of each of the electrode pairs were directly compared with respect to the perceived pitch difference, using a two-interval, two-alternative forced-choice (2-AFC) procedure. The pair members were presented randomly either in the first or second observation interval. Subjects were required to indicate which of the two stimuli sounded higher in pitch while concentrating on pitch rather than on other attributes such as timbre or loudness. Electrode pairs with an average discriminability across 25 repetitions within the range of chance ($50 \pm 18\%$) were considered as pitch-matched. For subjects with more than one pitch-matched electrode pair, the pair at medial tonotopic position was chosen.

For all CI listeners at least two pitch-matched electrode pairs ($p < 0.01$) were identified. Table 2 indicates, for each subject, the electrode pair members finally selected for presenting stimuli in the ITD studies and the corresponding percentage of trials in which the right electrode was judged to be higher in pitch.

3. Experiment I: Lateralization discrimination

In this experiment the effects of ITD in different signal portions on lateralization discrimination was investigated as a function of pulse rate.

3.1. Procedure

A “lateralization discrimination” task was used to measure the ITD sensitivity in the subjects. The task required left/right judgments of target's positions relative to the position of a comparison stimulus. The comparison stimulus with zero ITD was always presented in the first interval and evoked a centralized auditory image. The target differed from the comparison stimulus in that the pulses at one ear were delayed relative to the other ear. The target was always presented in the second interval. The subjects were requested to indicate whether the second stimulus was perceived to the left or to the right of the first stimulus by pressing the left or right button. The separation between the intervals was 300 ms. Visual indication of the stimulus intervals was provided on a computer screen. After each trial, visual feedback was provided, which indicated “correct” or “incorrect”.

To determine the JNDs in ITD with respect to lateralization discrimination the method of constant stimuli was applied. JNDs were estimated from a maximum-likelihood cumulative Gaussian fit to the percent correct data. The package “psignifit”[3] version 2.5.41 was used to fit psychometric functions to the percent correct data. Psignifit is described in Wichmann and Hill (2001a) and Wichmann and Hill (2001b). The data were collected at four to six ITD values. The ITD values depended on the sensitivity of each subject for the individual conditions. Generally, the largest presented ITD was 800 µs. However, in case of ongoing ITD at the pulse rate of 800 pps, the largest ITD value was restricted to 500 µs, because ongoing ITD approaching half of the inter-pulse interval can introduce ambiguous cues[4]. Generally, each stimulus was presented 60 times. However, in cases where the psychometric function did not exceed 66 percent correct, at least 60 further item repetitions were presented in order to reduce the randomness in the data. This occurred for some conditions in case of listener CI8. For the ordering of the conditions, we used a completely randomized design, in which all levels of the independent variables and their repetitions were pooled in one list which was then randomized. Depending on the constitution and motivation of the subject, 6 to 10 blocks were completed in one testing day, where each block was approximately 30 minutes of testing. Prior to the main experiment, the subjects were trained in three stages using the same procedure as in

3 Available from http://bootstrap-software.org/psignifit/

4 As this study was performed, the idea of ambiguous cues was just a speculation. Later, as shown in Chapter III, this ambiguity could be confirmed.

the main experiment. The first stage used a 300-ms version of the baseline condition, the second stage used the baseline condition (4 pulses), and the third stage used a list containing all stimulus conditions of the main experiment. The training was continued until the subjects showed stable performance in each stage. The four CI listeners, which were selected for the main experiment, required about two hours of training. The other four CI listeners showed poor performance even after a training for one day. Thus, it was decided decided to exclude them from this study.

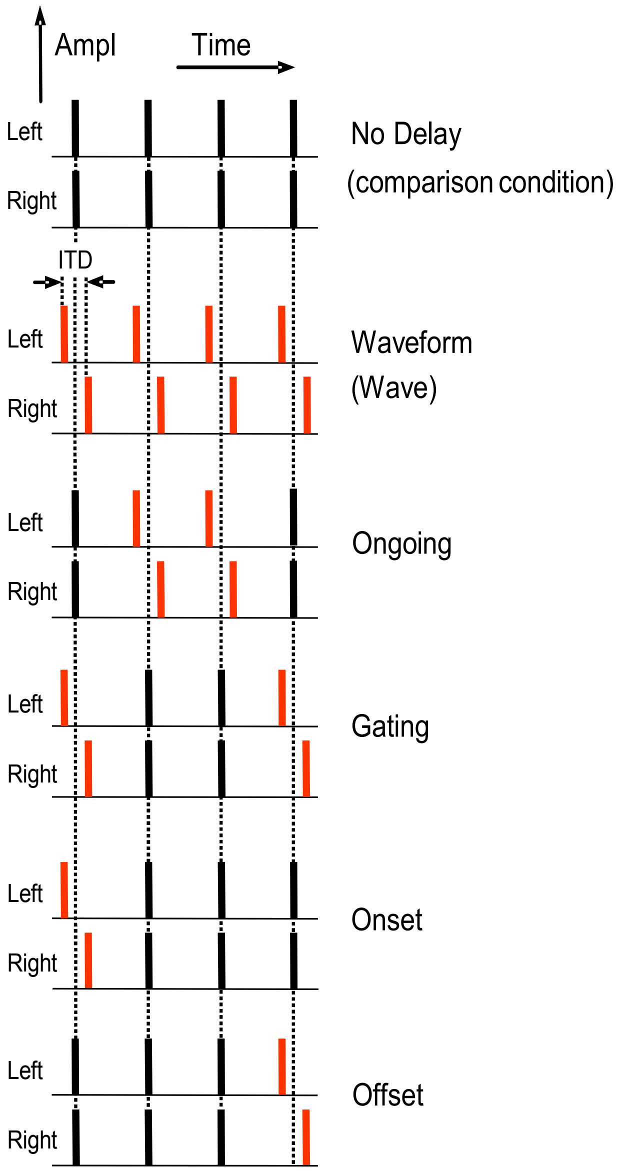

Figure 9: Schematic diagram of the pulse trains at the two ears for different ITD types. Notice that electric pulses were actually biphasic and for reasons of clarity the positive phase is shown only. Figure from Laback et al. (2007).

3.2. Stimulus conditions

In this study very short pulse trains were used, which consisted of four pulses only. There were two reasons for using a constant number of four pulses at each pulse rate. First, the number of information units (in terms of pulses, which contain ITD information) remained constant across pulse rates. This allows a comparison of the contribution of different stimulus portions across the pulse rates. Second, the amplitude could be held constant across pulse rates, which reduced confounding effects of different loudness across the pulse rates. This was verified by an informal loudness estimation task.

The different types of ITD tested in this experiment are schematically illustrated in Figure 9. This figure shows the amplitude versus time representations of the pulse trains at the two ears. The comparison condition with zero ITD is shown on the top of the figure. The stimulus shown beneath contains ITD in each of the interaural pulse pairs and is referred to as waveform delay (Wave). The third stimulus from the top shows the condition containing the interaural delay in the ongoing portion of the signal (Ongoing). This means that the ITD is present in the two pulse pairs in the middle of the train and the first and last pulse pairs have a zero ITD. In the gating delay condition (Gating) the first and last pulse pairs contain ITD, whereas the two pairs in the middle have zero ITD. The condition onset delay (Onset) contains ITD in the first (onset) pulse pair only. The condition offset delay (Offset) contains ITD in the last (offset) pulse pair only. In each condition, the ITD is always divided between the two ears, i.e., at one ear, the leading pulse already starts at half the ITD before the reference position and, at the opposite ear, the lagging pulse starts at half the ITD after the reference position. This reduces the temporal irregularity and was intended to reduce a potential monaural discrimination cue.

All conditions were tested at pulse rates of 100, 200, 400, and 800 pps and, thus, the duration of the stimuli ranged from 5 ms (800 pps) to 40 ms (100 pps). The subject CI8 could not attend the tests with the pulse rate of 200 pps because of his/her limited availability.

3.3. Results

The collected percentage correct scores were used to estimate the just noticeable differences (JND). In order to allow the estimation of JNDs for all subject a threshold criterion of 65% was used. Such a low threshold criterion was required because of large inter-individual differences in the overall lateralization discrimination performance of the CI listeners. In cases where the psychometric function does not exceed the threshold, no JND could be determined. In all other cases, the psychometric functions are monotonic.

All tested subjects reported hearing fused images in all conditions. The distribution of the left/right judgments for each listener shows sufficient symmetry, thus, it was not necessary to remove response bias by adjusting the percent correct scores.

CI listeners

The JNDs for each of the four CI listeners are shown in Figure 10 as a function of pulse rate for the various ITD types. Error bars indicate the 95% confidence intervals, which were calculated by the BCa bootstrap method based on 1999 simulations (implemented in psisignifit; described in Wichmann and Hill, 2001b). The difference between two JNDs was tested for significance with a test based on Monte Carlo simulations of the fits to the underlying psychometric functions. Because the error bars shown in Fig. 10 are based on the bootstrap method and not Monte Carlo simulations, their overlap between two conditions does not exclude a significant difference between the mean values. JNDs, which could not be determined are plotted at “ND”.

For the condition Wave and two CI listeners, we found a significant increase of the JNDs with the pulse rate as supported by the significant difference between JNDs at 100 and 800 pps ($p = 0.003$ for CI1; JND of 398 μs at 100 pps and indeterminable JND at 800 pps for CI8). The remaining two CI listeners (CI3 and CI12) show approximately constant JNDs across pulse rates.

The CI listeners, who showed sensitivity to waveform ITD at specific pulse rates, are also sensitive to gating ITD at these pulse rates. In these cases, the JNDs do not vary significantly

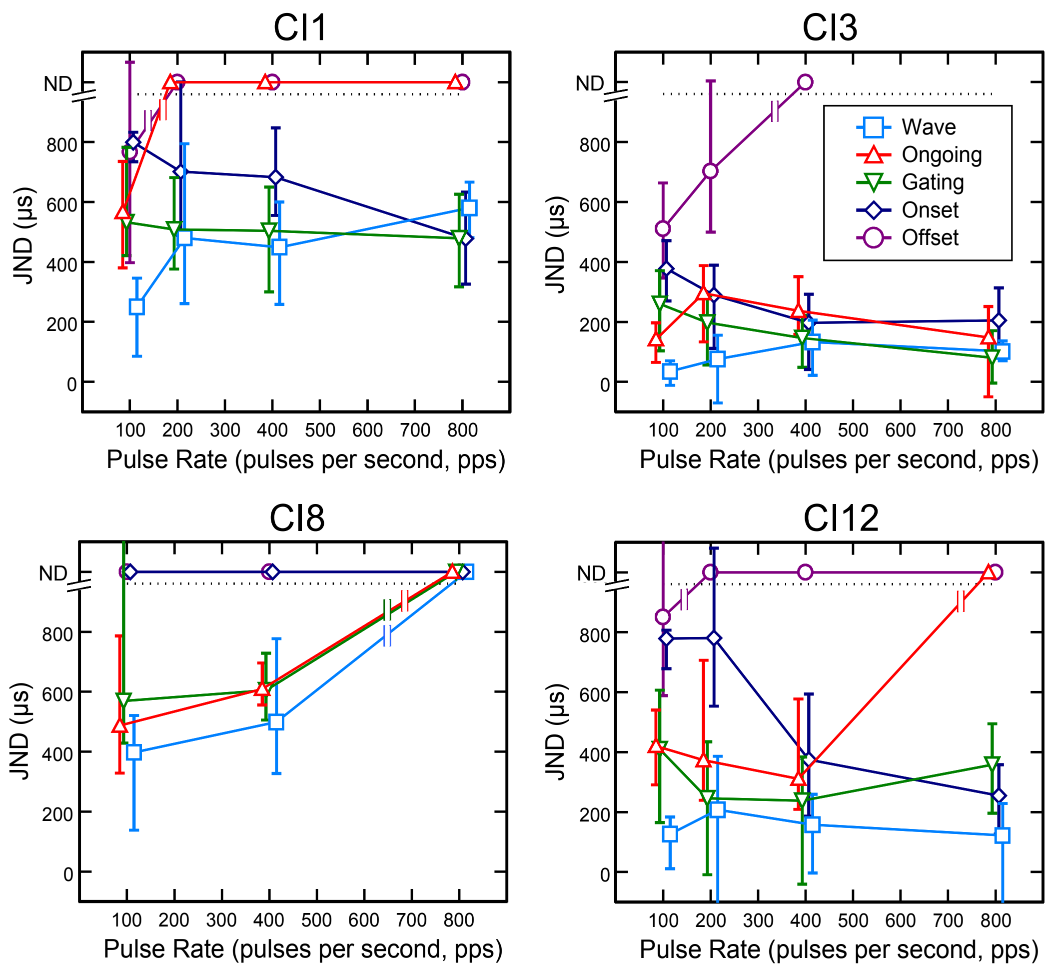

Figure 10: JNDs from Experiment I as a function of pulse rate for each CI listeners. The parameter is the type of ITD. Error bars indicate the bootstrap 95% confidence intervals. Not determinable JNDs are shown at "ND". For some conditions the error bars are smaller than the symbols and were omitted. Figure from Laback et al. (2007).

across different pulse rates. Especially for listener CI3, the apparent decrease of JNDs with increasing rate is not statistically significant.

At the lowest pulse rate (100 pps), all CI listeners show higher JNDs for the conditions Ongoing and Gating compared to the condition Wave. These differences are significant for all listeners ($p < 0.043$), except for CI8 ($p > 0.05$). Thus, compared to condition Wave, the absent ITD in either the ongoing or the gating signal portion causes significant degradation in performance. This implies that both ongoing and gating ITD substantially contribute to lateralization discrimination.

With increasing pulse rate, the CI listeners show different performance with respect to their sensitivity to ongoing ITD. Listener CI3 shows sensitivity up to 800 pps in terms of the highest pulse rate with determinable JND. Listeners CI8 and CI12 show sensitivity up to 400 pps, and listener CI1 shows sensitivity at 100 pps only. Since the observation of sensitivity up to 800 pps for CI3 was rather unexpected, the measurements for condition Ongoing at 100, 400, and 800 pps were repeated at another day. The data show exactly the same results indicating that CI3 is sensitive to the ongoing ITD at pulse rates as high as 800 pps.

Only three CI listeners reveal sensitivity to onset ITD (CI1, CI3, and CI12) showing increasing sensitivity with increasing pulse rate. The significance of this effect is supported by the significant differences between the JNDs at 100 and 800 pps ($p = 0.001$, $p = 0.04$, and $p = 0.034$, respectively).

For the condition Offset, the JNDs are determinable for three subjects at the pulse rate of 100 pps (CI1, CI2, and CI3) and for one subject at the pulse rate of 200 pps (CI3).

NH listeners

The results of the NH listeners are presented in Figure 11. The results are sufficiently homogeneous, which justifies averaging of the JNDs over subjects. Thus, in Figure 11, group results are presented only. The error bars indicate ± 1 standard deviation of the mean values across the listeners. The left panel of Figure 11 shows the JNDs determined using the 65% threshold criterion, which was used for the JND estimation for the CI listeners. The right panel shows the JNDs determined for the same data but using the 80% threshold criterion. Notice the different scaling of the ordinates. The JNDs obtained for the two threshold criteria reveal similar effects of the stimulus conditions. This indicates that the choice of the threshold criterion of 65% has no substantial impact on our investigation. In the following, the JNDs between different conditions were compared with two-tailed *t* tests.

In condition Wave, the performance decreases with the pulse rate. This is supported by the significantly higher JNDs at 800 pps than at 100 pps ($p = 0.008$).

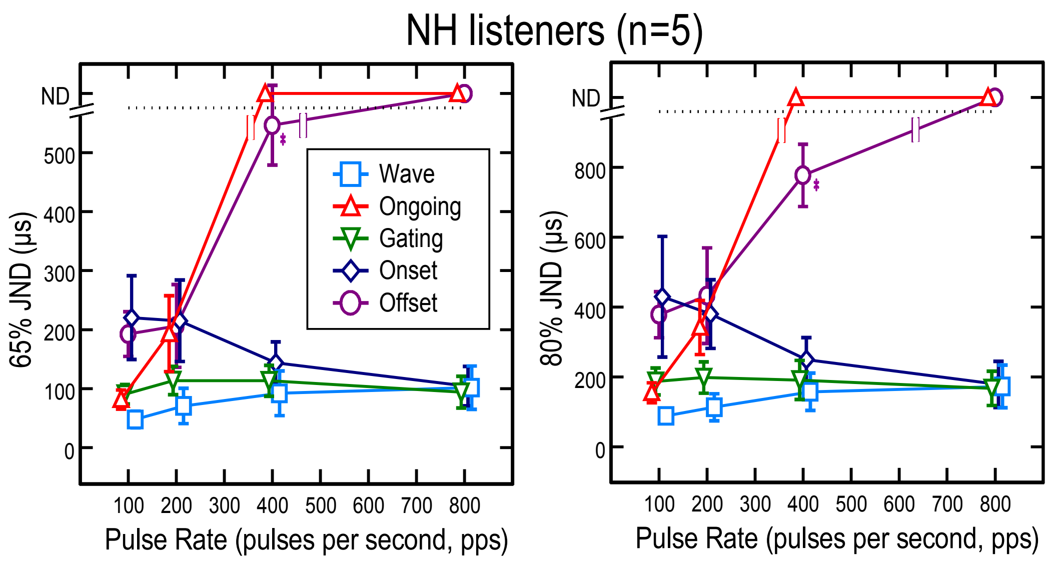

Figure 11: Average JNDs of five NH listeners. The left panel shows the JNDs determined using to the 65% threshold criterion. The right panel shows the JNDs using the 80% threshold criterion. Note the different scaling of the ordinates in the two panels. Not determinable JNDs are shown at "ND". Error bars indicate ± 1 standard deviation. The data point marked with asterisk (400 pps, condition Offset) is based on data from two listeners only. Figure from Laback et al. (2007).

In condition Gating, the listeners show sensitivity at all pulse rates. The performance does not vary significant across different pulse rates ($p = 0.64$ for JND difference between 100 pps and 800 pps).

In the condition Ongoing, the sensitivity decreases monotonically with increasing pulse rate. In four listeners, the highest rate with determinable JNDs is 200 pps. In one listeners, the highest rate with determinable JND is 400 pps. This JND amounts to 486 µs.

In the condition Onset, the sensitivity increases monotonically with the pulse rate, as supported by the significant difference for the pulse rate 100 and 800 pps ($p = 0.007$). At 800 pps, the JND in condition Onset (104 µs) is not significantly different from the conditions Gating (94 µs) and Wave (102 µs; $p = 0.22$ and $p = 0.84$, respectively).

In the condition Offset, the sensitivity decreases with the pulse rate. Already at 400 pps, the JNDs are indeterminable for the three listeners and can be determined for two listeners only.

The determinable JNDs are high (mean of 547 µs) and are showed in Figure 11. At 800 pps, the JNDs are indeterminable for all NH listeners.

As we are interested in the relative contributions of the ongoing and gating portions of the signals, the following analysis compares the effects of conditions Ongoing and Gating with respect to the condition Wave. At the lowest pulse rate (100 pps), the mean JND for the condition Ongoing (82 µs) is significantly higher than that for condition Wave (48 µs; $p = 0.01$). The mean JND for condition Gating (91 µs) is also significantly higher than that for condition Wave ($p = 0.001$).

For the contribution of the onset and offset portions of the signal, the results are compared with respect to the Gating condition. The mean JNDs for conditions Onset (220 µs) and Offset (193 µs) are significantly higher than those for condition Gating (91 µs; $p = 0.02$ and $p = 0.01$, respectively). Thus, these results indicate that, at 100 pps, each interaural pulse pair significantly contributes to the ITD perception. However, as a function of pulse rate, the JNDs for conditions Onset and Offset with respect to the condition Gating reveal complementary contributions of onset and offset ITD.

3.4. Discussion

CI listeners

Generally, the results from Experiment I show large inter-individual differences. However, three main effects can be observed.

1. The three CI listeners, who reveal sensitivity to onset ITD show increasing contribution of onset ITD with increasing pulse rate.
2. Gating ITD contributes to lateralization discrimination at all pulse rates for which sensitivity to waveform ITD was observed.
3. All subjects show sensitivity in condition Ongoing, i.e. sensitivity to ITD in the two pulses in the middle of the train: one listener up to 800 pps, two listeners up to 400 pps, and one listener at 100 pps only.

Sensitivity in condition Ongoing implies that ITD in the temporal fine structure of the short pulse trains contributes to lateralization discrimination. These stimuli present “pure” fine structure ITD, containing no other ITD cues in the onset, offset, or ongoing envelope. This is in contrast to Majdak *et al.* (2006), the study presented in Chapter III, where we studied fine structure ITD sensitivity in four CI listeners using amplitude modulated pulse trains with a 300-ms duration. In that study, fine structure ITD was created by delaying the pulses at one ear and subsequently applying a trapezoidal envelope with zero ITD. This could potentially involve confounding effects, namely, the first audible pulse could be that on the lagging side, since the first pulse on the leading side was at the absolute threshold. However, the finding of this study, which uses “pure” fine structure stimuli, indicates that conflicting ITD cues during the onset in the study Majdak *et al.* (2006) did not confound substantially its outcome.

At 800 pps, some ITD values approached half of the inter-pulse interval. This may have introduced ambiguous ITD cues, which would be reflected by a decline of the psychometric function in that ITD range. Actually, the psychometric functions obtained showed no such decline even at the largest ITDs tested (500 µs and 600 µs in conditions Ongoing and Wave, respectively). Thus, the ambiguity probably did not substantially affected the results. This is in contrast to the results of Majdak *et al.* (2006), the study presented in Chapter III, where a pronounced decline of the psychometric functions was observed for ITDs between approximately one quarter and one half of the inter-pulse period. This effect was observed for 300-ms pulse trains with ITD in the fine structure only. One explanation for the absence of non-monotonicity of the psychometric functions in the current study is the short duration of the stimuli: we assume that for such short stimuli (four pulses) the onset is more dominant in resolving the ambiguity than for longer stimuli (e.g. 300 ms).

The results for the listeners CI1, CI3, and CI12 at the different pulse rates show that the sensitivity to gating ITD is qualitatively consistent with the corresponding contributions of onset and offset ITD. This indicates that the contributions of the onset and offset ITD are combined in the condition Gating. For CI8, however, the JND in condition Gating could be determined, although the JNDs in conditions Onset and Offset were indeterminable. This indicates that the presence of ITD in either onset or offset alone is just too weak to be evaluated by the auditory system.

NH listeners

In condition Ongoing, the decrease in sensitivity with increasing pulse rate is qualitatively consistent with other studies. Hafter and Dye (1983) and Saberi (1996) studied the the perceptual contribution of the ongoing signal as a function of pulse rate. They used band-pass filtered clicks. For rates higher than 200 pps, they showed a decreasing contribution of ITD transmitted in clicks which follow the onset. In our study, the maximum pulse rate where all NH listeners were sensitive to ITD in the two center pulses (condition Ongoing ITD) was 200 pps. The results of a study by Bernstein and Trahiotis (2002) show a qualitatively comparable rate limit. They used sinusoidally amplitude-modulated and transposed 4-kHz tones and found sensitivity to ongoing ITD for modulation frequencies up to 256 Hz.

The results of our study show a high sensitivity to onset ITD at 400 pps and the complete dominance of onset ITD at 800 pps. This is qualitatively consistent with the results of Freyman *et al.* (1997) who showed a similar dominance of the onset in lateralization of high-pulse rate click trains. In particular, they showed that click trains with rates $\geq$ 500 pps are lateralized toward the ear favored by ITD in the onset click. This happens even if the remaining clicks in the stimulus have an ITD favoring the opposite ear.

In condition Gating, the ITD sensitivity seems to be constant across different pulse rates. This effect has also been observed for three of the CI listeners.

4. Experiment II: Monaural Detection

In all conditions but Wave, the lateralization judgments could have been based on monaural cues such as periodicity pitch or timbre. This is because of temporal irregularities in the stimuli we used. In the condition Ongoing, the changes in the inter-pulse interval (IPI) from the first to the second and from the second to the third IPI yield such irregularities. Such cues could theoretically have been exploited by the listeners. Thus, we performed Experiment II to falsify this hypothesis and to prove that the data from Experiment I are based on binaural cues only. Experiment II tested if the subjects exceed chance performance in detecting monaural versions of the stimuli used in Experiment I. If the subjects do not exceed chance performance then the performance in Experiment I was not based on monaural cues. In case the subjects exceed chance performance, the data obtained in Experiment I could have been based on monaural cues.

In this experiment, a three-interval, two-alternative forced-choice procedure was used. The listener had to find the target stimulus, namely, to chose a stimulus different from the stimuli in the other two intervals. The target stimulus was a monaural version of the target stimulus used for lateralization discrimination and had irregular IPIs. The comparison stimulus was a monaural version of the respective reference stimulus and had regular IPIs. The magnitude of the deviation from the regular IPI corresponded to half the ITD values tested in Experiment I. The chosen ITDs were always just a little bit higher than the corresponding JNDs for each condition from Experiment I. Visual feedback about the correctness of the response was provided after each trial. The conditions were all combinations of pulse rates and ITD types except for Wave. For each condition, a total of 36 stimulus presentations were tested. The stimuli were presented in one session in completely randomized order. A training consisting of 20 presentations per condition was completed before collecting data. All other details were the same as in Experiment I.

The the performance achieved in all conditions was within the range of chance rating ($p > 0.05$). In particular, the performance of the best condition for individual listeners was: 47.2% (CI1), 39% (CI3), 44.4% (CI8), 47.2% (NH2), 47.2% (NH3), 36% (NH4), 42% (NH5), and 47.2% (NH6). These results indicate that the data obtained in Experiment I are based on binaural cues only.

5. General Discussion

The study presented in this chapter investigated the contribution of ITD in various portions of four pulse sequences to lateralization discrimination. The tests were performed for different pulse rates with NH and CI listeners. As already mentioned, four of the eight invited CI listeners showed low and unstable sensitivity for a baseline condition and were therefore not included for participation in the main experiments. Because only four CI listeners participated in the main experiments, no general conclusions can be drawn from this study for the population of bilateral CI listeners. The results should rather be considered as case studies.

Four cochlear implant listeners were sensitive to ITD in the signal portion referred to as ongoing signal, which consisted of the two pulses in the middle of the train. Even though the results of the CI listeners differed with respect to the pulse rate, all CI listeners showed sensitivity to ITD in the temporal fine structure of the pulse sequences at least at 100 pps. Furthermore, all listeners (NH and CI) showed sensitivity to ITD in the onset and offset pulses at all those pulse rates for which they showed sensitivity to waveform ITD. For the three listeners sensitive to the onset ITD, the contribution of the onset ITD increased with the pulse rate.

The design of our stimuli intended to avoid the influence of confounding parameters like different number of pulses and different amplitudes in the comparison across different pulse rates. Using the same number of pulses, which have the same constant amplitude at each pulse rate, the same information units could be presented at each pulse rate. This allowed us to compare the effects across pulse rates. However, the four-pulse sequences are quite short, in particular at the higher pulse rates. Thus, it is possible that for higher rates the neurons were in refractory state immediately after the onset pulse and did not respond to the following pulses. In real-life, such sustained timing cues mostly continue for more than just two electrical pulses. Thus, the weighting of different types of ITD cues may differ for more realistic stimuli.

To investigate the effects of fine structure ITD on lateralization discrimination in more realistic situations a follow-up study was performed. This study is presented in the next chapter. Additionally, within the current study, a preliminary test in one CI listener was performed. Laback *et al.* (2007) write:

“To check the outcomes for longer stimuli with a constant duration across pulse rates, additional data have been collected. The stimuli had the same duration of 300 ms at all pulse rates (100, 400, and 800 pps). The stimulus amplitude at 100 pps was the same as in Experiment I and for higher pulse rates it was adjusted to elicit equal loudness. The conditions Wave, Ongoing, and Gating were tested. The methodology was the same as in Experiment I. Since this experiment was done with just one CI listener (CI3, who was available for further testing), the interpretation of the results has to be considered as preliminary. The results confirmed two main findings from Experiment I. First, the listener was able to lateralize upon ongoing ITD. Second, gating ITD contributed to lateralization discrimination at all three pulse rates tested. Furthermore, up to 400 pps, the sensitivity to ongoing ITD was higher for the 300-ms stimuli than for the four-pulse stimuli. This improvement is most likely due to temporal integration of ITD information. At 800 pps, however, no sensitivity to ongoing ITD was observed, which is in contrast to the results for the four-pulse stimuli. This may be related to the lower amplitude compared to the four-pulse stimuli. Lowering of the amplitude was necessary to obtain equal loudness increasing the stimulus duration. In summary, these results reveal an interaction between the effects of the parameter pulse rate and the parameters pulse number and amplitude. To use a constant number of information units containing ITD at each pulse rate in Experiment I circumvented these interactions."

We observed large inter-individual differences of the upper rate limit, especially in condition Ongoing. These findings suggest that some unknown factors (besides the controlled interaural pitch and loudness matching) limit the perception of ongoing ITD in electric hearing. Some studies showed that the recovery from forward masking can vary considerably between CI listeners (Chatterjee, 1999; Nelson and Donaldson, 2002). This may be related to the decay of excitation, which reduces the internal representation of the ongoing ITD cues and, thus, may be one potential candidate for an uncontrolled factor in our study.

In condition Ongoing, the upper rate limit for the CI listeners is 800 pps, whereas for the NH listeners the upper rate limit is 200 pps. We think that this difference is related to the specific properties of electric and acoustic stimulation. For example, the phase locking is known to be stronger in electric hearing than in acoustic hearing because of bypassing the synaptic mechanism at the hair cell (Abbas, 1993). Additionally, in acoustic hearing, the auditory filtering on the basilar membrane may be a limiting factor. This is bypassed in electric hearing. In acoustic stimulation, the auditory filtering effectively reduces the modulation depth. Thus, this may complicate extraction of the ITD information from the pulses following the onset at higher pulse rates. Of course, the JNDs obtained for the NH listeners might depend on the frequency region of the bandpass filter applied on the stimulus. If the auditory filtering were the limiting factor, higher sensitivity would be expected at higher frequency regions because the duration of the impulse responses of the auditory filters decreases with their center frequency. However, Bernstein and Trahiotis (2002) found that the upper rate limit in the sensitivity to ongoing ITD does not increase with the center frequency (they measured center frequencies from 4 to 10 kHz). This does not support the assumption that auditory filtering is the limiting factor. On the other hand, the stimuli used by Bernstein and Trahiotis (2002) had a constant bandwidth in Hz. Using an increasing bandwidth with the center frequency, e.g. constant bandwidth in the ERB scale, would stimulate a larger number of neurons. Additionally, the modulation depth in the internal representation of the stimulus would be higher because of the larger bandwidth of the auditory filter. Thus, with increasing bandwidth, the performance at higher center frequencies would be higher compared to that at the lower frequencies. This qualifies the statement of Bernstein and Trahiotis (2002). To clarify this issue an additional study was performed with normal hearing listeners where the effect of ITD as a function of the center frequency was investigated. This study and its results are presented in Chapter IV of this thesis.

A recent study on monaural rate discrimination by Chen and Zeng (2004) supports our finding that temporal fine structure cues can be exploited in electric hearing for pulse rates as high as 800 pps. They investigated the detection of sinusoidal frequency modulation and found that three CI listeners were able to do that for rates up to 1000 pps. This finding differs from previous studies (e.g. Zeng, 2002), which have shown an upper rate limit of 300 to 500 pps in pitch difference detection tasks. In 2003, van Hoesel and Tyler presented a new stimulation strategy designed to encode fine structure ITD cues for CIs (van Hoesel and Tyler; 2003). The results show no clear advantage in sound source localization compared to conventional strategies like CIS or ACE. Thus, the potentials of stimulation strategies encoding ITD information in the fine structure should be determined in future studies, which consider the complexity of the parameters and effects involved. For example, one potential effect, namely, the channel interactions due to current spread, may disrupt low-frequency ITD cues in the fine structure.

6. Conclusions

The conclusion from the data collected in this study is that bilateral CI listeners may benefit from encoding fine structure ITD information in future CI stimulation strategies with respect to the lateralization of sound sources. This conclusion is based on the results from left/right discrimination tests performed with four bilateral CI and five NH listeners as a function of pulse rate. The subjects were listening to four-pulse sequences, which contained ITD in different pulses. Specifically, the ITD information was presented in the two middle pulses (ongoing), in the gating portions (onset and offset pulses), or in the entire train (waveform). The

NH subjects were listening to acoustic simulations of electrical stimulation. One of the CI listeners showed sensitivity to ITD in the two middle pulses up to 800 pps, two CI listeners up to 400 pps, and one CI listener up to 100 pps. Four NH listeners showed sensitivity to ITD in the two middle pulses up to 200 pps, and one up to 400 pps. For all listeners, gating ITD contributed at all pulse rates. For all NH and three CI listeners, the sensitivity to onset ITD increased with the pulse rate. A monaural detection experiment verified that the listeners did not use monaural cues in the lateralization discrimination task.

III. Interaural time differences in fine structure and envelope of 300-ms stimuli

Based on *Journal of the Acoustical Society of America* **120**, 2190-2201 **(2006)**

It is well established that ITD information in unmodulated signals can only be processed up to about 1500 Hz (Zwislocki and Feldman, 1956; for reviews see Blauert, 1997; and Wightman and Kistler, 1997). At higher frequencies a slow modulation of the carrier transmits the ITD information (e.g. Bernstein, 2001). Using modulated signals, like speech, at least two different types of ITD can be defined: ITD in the envelope (ITD ENV) and ITD in the fine structure (ITD FS). Signals with equal ITD ENV and ITD FS can be considered as a special case in that the whole waveform of one channel is delayed relative to the other channel. This case is most often found in natural signals and is referred to as waveform delay (WD).

Several studies have examined ITD perception in cochlear implant (CI) listeners. Lawson *et al.* (1998) showed that lateralization discrimination using ITD only is possible. Using a pulse rate of 480 pulses per second (pps) they obtained a just noticeable difference (JND) of 150 µs. More detailed studies were performed by van Hoesel and Clark (1997), van Hoesel *et al.* (2002), and van Hoesel and Tyler (2003). In general, the performance was much worse than that of normal hearing (NH) listeners and had high inter-subject variability. For unmodulated stimuli the JNDs increased for higher pulse rates and could not be determined at a pulse rate of 800 pps. However, when using low-frequency amplitude-modulated stimuli at this pulse rate JNDs were on the order of JNDs for unmodulated stimuli with carrier pulse rate equal to that of the low-frequency modulation. Unfortunately, they did not separate the relative contribution of ITD ENV and ITD FS, which may be important for amplitude-modulated stimuli like speech. Laback *et al.* (2004) investigated the effects of ITD ENV manipulation (the ITD FS was random and uncontrolled) in electric hearing by presenting acoustic stimuli via unsynchronized speech processors. They showed that JNDs differed between NH subjects (19 µs) and CI listeners (259 µs and 384 µs, best JNDs for CI listener S2 and S1, respectively) and depended on the type of stimulus (lowest for click trains, highest for speech or noise bursts).

Current cochlear implant systems use a variety of stimulation strategies to transmit the acoustic information, an overview of which is given in Wilson (2004). Almost all strategies were designed for monaural use and do not include any binaural synchronization: the electric stimulation is controlled by two independently running speech processors. As a result, the ITD information is coded in the envelope only. The strategies have one other aspect in common: according to the specification, they use a constant stimulation pulse rate at both ears. Due to the lack of synchronization between the two ears, the stimulation pulses have an interaural delay, which can be regarded as an ITD FS. This depends on the switch-on delay between the processors and has a random value between 0 μs and the interpulse interval (IPI). If CI listeners are sensitive to ITD FS, it will interact with other lateralization cues like ITD ENV or interaural level differences.

Due to the manufacturing tolerances, the time bases deviate between the speech processors at the two ears, resulting in different IPI. Therefore, the pulse rates can not be assumed to be equal at both ears. This leads to a dynamically changing ITD FS, which varies between 0 and the IPI. The period of this "ITD beat" increases with decreasing deviation in the pulse rates. If subjects are sensitive to ITD FS, the dynamically changing ITD FS will result in a movement of the auditory image.

In the previous chapter, it was shown that ITD FS contributes to lateralization discrimination for lower pulse rates (Laback *et al.*, 2007). In this case, a controlled ITD FS may support the effect of ITD ENV and improve the lateralization of sound sources. Coding ITD FS information also may be advantageous for speech perception in noise (Licklider, 1948; Hirsch, 1950; Dirks and Wilson, 1969; Bronkhorst and Plomp, 1988; Hawley *et al.*, 1999) or speech segregation (Drennan *et al.*, 2003; Culling *et al.*, 2004). One study, performed with anesthetized cats, indicates that ITD FS in a low-frequency carrier may be a much stronger cue than ITD ENV in an amplitude-modulated high-frequency carrier: Smith and Delgutte (2005) showed that the neuronal tuning curves in the inferior colliculus are sharper for ITD FS in a low-frequency stimulus (tested up to 320 pps) than for ITD ENV in a high- frequency carrier (tested 1000 pps carrier and modulation frequencies up to 160 Hz).

The goal of this study is to systematically investigate the effects of fine structure ITD manipulation on lateralization discrimination in electric stimulation using amplitude-modulated stimuli. It was expected that CI listeners would be sensitive to ITD FS at lower pulse rates. In addition, the same experiments were performed with normal hearing (NH) subjects using a simulation of electric stimulation to compare their performance with that of the CI listeners. The results allow the assessment of the need for the synchronization of speech processors, taking some synchronization methods into account.

1. Methods

1.1. Subjects and Apparatus

Four NH subjects participated in this study, of whom one (NH3) was female. All subjects were between the ages of 25 and 35 years old and had no indication of hearing abnormalities. One of them was the author of this thesis (NH4).

Four cochlear implant (CI) listeners were tested. Three of them were implanted bilaterally with the C40+ implant system manufactured by MED-EL Corp. This system provides pulsatile, nonsimultanous biphasic current pulses on up to 12 electrodes with a minimum phase duration of 26.7 µs. One CI listener (CI2) used the C40+ in the left ear and an older implant, the C40, in his right ear. The C40 provides current pulses on up to 8 electrodes with a minimum phase duration of 40 µs. Clinical data of CI listeners can be found in Table 3. The subjects were selected from a total of seven CI listeners invited for participation in the study. These four listeners fulfilled the selection criterion, as defined by the ability to reproducibly perform left/right discrimination on the basis of waveform ITD in a pulse train with a pulse rate of 100 pps in a reasonable amount of time.

The apparatus was identical to the one used in the study presented in Chapter II. A personal computer system was used to control electric and acoustic stimulation. Each implant was controlled by a Research Interface Box (RIB), manufactured at the University of Technology Innsbruck, Austria. The two RIBs were synchronized, providing an interaural accuracy of stimulation timing better than 2.5 µs. Prior to the experiment, the stimuli were verified using a pair of dummy implants (Detektorbox, MED-EL). The stimuli for acoustic stimulation were output via a 24-bit stereo A/D-D/A converter (ADDA 2402, Digital Audio Denmark) using a sampling rate of 96 kHz per channel. The analog signals were sent through a headphone amplifier (HB6, TDT) and an attenuator (PA4, TDT) and presented to the subjects via a circumaural headphone (K501, AKG). Calibration of the headphone signals was performed using a sound level meter (2260, Brüel & Kjær) connected to an artificial ear (4153, Brüel & Kjær).

Subject	Aetiology	Age at implant		Deafness duration		Binaural electric stimulation experience
		L	R	L	R	
CI1	Meningitis	14 yr	14 yr	5.5 mo	1.5 mo	6 yr
CI2	Skull trauma	54 yr	48 yr	21 yr	25 yr	4 yr
CI3	Meningitis	21 yr	21 yr	2 mo	2 mo	1 mo
CI8	Osteogenesis imperf.	41 yr	39 yr	3 yr	12 yr	2 mo

Table 3: Clinical data of CI listeners.

1.2. Stimuli

The stimuli were amplitude-modulated pulse trains which were designed as pulse trains multiplied by a predefined envelope. ITD FS and ITD ENV were introduced by delaying the temporal position of the pulses and of the envelope, respectively, at one ear relative to the other ear. The following ITD conditions were specified: ITD in envelope only (ENV), ITD in fine structure only (FS), no ITD at all, which is the reference condition (REF), and the identical ITD in both the envelope and the fine structure, referred to as waveform delay (WD).

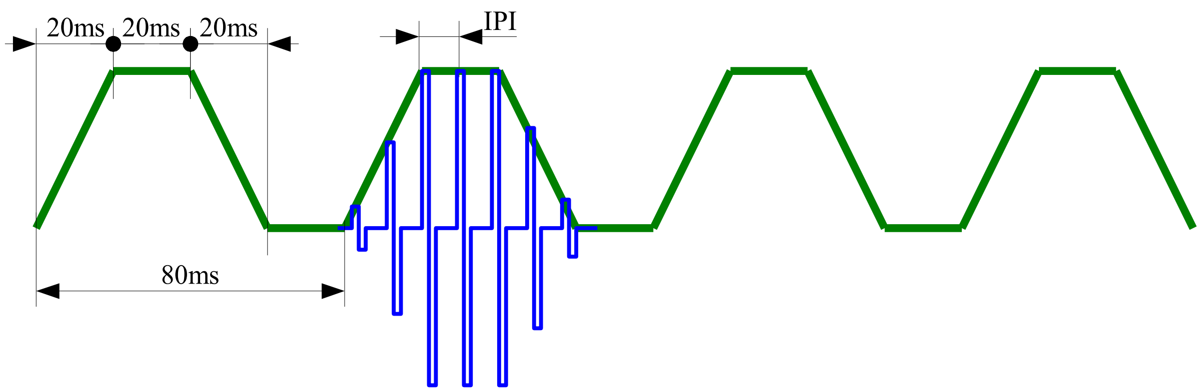

Figure 12: A schematic representation of the stimulus used in this study. For readability purposes the fine structure characteristics was shown in one trapezoid only. The ramps slope down to the absolute threshold of each subject. Between the trapezoids the amplitude was set to zero. In acoustic stimuli, pulses with positive amplitude are applied.

The envelope consisted of 4 trapezoids with durations of 60 ms, each repeated at a period of 80 ms, resulting in 20 ms gaps between two successive trapezoids and a total stimulus duration of 300 ms (Figure 12). The trapezoid period of 80 ms yields an amplitude modulation frequency of 12.5 Hz. Since the envelope modulation is trapezoidal, the modulation spectrum contains multiples of the 12.5 Hz as well. There were several reasons for selecting a relatively slow modulation rate. Although sensitivity to ITD increases with growing modulation frequency up to approximately 125 Hz (Henning, 1974; Bernstein, 2001) values in that range interfere with the lower limit of the pulse rates used here (100 pps). Even modulation rates as low as tens of Hertz reduce the information available in the fine structure. Furthermore, pulse trains with modulation rates on the order of 12.5 Hz more closely resemble real-world signals than pulse trains modulated with higher rates. In particular, speech has a modulation spectrum peak of approximately 5 Hz (Greenberg *et al.*, 2003). Considering these aspects, these values of modulation frequency lead to signals providing sufficient ITD information in both the fine structure and the envelope. Finally, the rise and release time of each trapezoid was set to 20 ms. This value was chosen to emphasize the onset and the offset effects, which are assumed to enhance sensitivity to ITD ENV. The level of the first and last pulse of each trapezoid was set at the subject's absolute threshold, which was determined in pretests (see section “Pretests”). Between the trapezoids, the amplitude was set to zero. The acoustic amplitudes were interpolated logarithmically and the electric currents were interpolated linearly.

Van Hoesel and Tyler (2003) found that subjects differed strongly in their sensitivity to ITD as a function of pulse rate. Furthermore, this was a major finding of our study presented in Chapter II. The two studies the sensitivity was generally highest at lowest pulse rates tested, which were 50 pps and 100 pps, respectively. On the other hand, current stimulation strategies use pulse rates up to about 1600 pps (Wilson, 2004), thus, testing these pulse rates is important for real world applications. Consequently, the pulse rates to be tested must be selected individually for each subject: three to four pulse rates between 100 pps and 1600 pps, corresponding to IPI between 10 ms and 625 µs, were chosen for each subject on the basis of lateralization discrimination pretests described in section 2.3.

In the case of electric stimulation, the pulse trains were composed of biphasic current pulses. Each phase of a pulse had a duration of 26.7 µs and 40 µs for the C40+ and the C40 devices, respectively. An interaurally pitch-matched electrode pair, selected in pretests (see section “Pretests”), was used for all experiments.

To allow a direct comparison of the results from NH subjects with those from CI subjects, the electric stimulation was simulated in the NH subjects using a method developed by McKay and Carlyon (1999) and further successfully applied by Carlyon *et al.* (2002). Pulse trains were composed of monophasic pulses with a duration of 10.4 µs, corresponding to one sampling interval at a sampling rate of 96 kHz. The pulse trains were passed through a digital eighth-order Butterworth filter with a geometric center frequency of 4590 Hz and -3 dB bandwidth of 1500 Hz. The acoustic simulation was identical to the simulation used in the study presented in Chapter II.

Due to the filtering of amplitude-modulated pulse trains, a possible naming clash might have been introduced. In the NH literature, the "fine structure" of the acoustic stimulus refers to the carrier frequency, which is 4590 Hz in our case. Following this definition, every filtered pulse has an "envelope", which is the envelope of the impulse response of the bandpass filter. Furthermore, the envelope of the pulse train appears as a second order envelope. The carrier frequency arises from the filtering procedure, which it is not the object of interest in this study. Thus, the definitions from the CI literature have been adopted to the acoustic signals. In reference to acoustic signals, the term "fine structure" defines the total impulse response of the bandpass filter, not the carrier only, and "envelope" refers to the slow trapezoidal amplitude modulation of the filtered pulse trains. Keeping in mind that the acoustic stimuli represent a simulation of electric stimulation, the same terms can be used to describe electric and acoustic stimulation effects.

Given that the sound pressure level (SPL) depends on the pulse rate, stimulation amplitudes were adjusted to maintain a constant SPL of 59 dB, measured at the headphones, at all rates for all NH subjects. Despite the filtering of the pulse trains, some artifacts like harmonic distortions or intermodulation at the basilar membrane can cause stimulation outside the desired frequency band. To prevent these artifacts from being heard, a binaurally uncorrelated pink noise with a spectrum level of 15.2 dB SPL at 4.6 kHz was continuously played throughout the testing.

Eight ITD FS values were chosen for each pulse rate, which corresponded to values from 0µs up to seven-eighths of the IPI in steps of eighth IPI. These values covered the range of ITD FS which would occur in a setup of unsynchronized speech processors and included ITDs exceeding the natural head-width delay for lower pulse rates. The investigations on effects of ITD ENV were secondary in this study; thus, only two values were used. The intended values of 400 µs and 625 µs represent large ITD values with respect to the head size, and correspond to ITD ENV cues as they occur in real-world situations. Unfortunately, in the lateralization pretests the CI listeners showed no sensitivity at 400 μs and very low sensitivity at 625 μs. Thus, intending to produce as much effect as possible, larger ITD ENV values were chosen for the CI listeners: 625 μs and 800 μs.

1.3. Pretests

In the experiments with CI users, pretests were performed to determine a binaurally loudness balanced, pitch-matched electrode pair for each listener. The pretests used pulse trains of 300 ms duration with zero ITD, 100 pps pulse rate, no amplitude modulation, and consisted of a manual up/down procedure to estimate each listener's threshold (THR), comfortable level (CL), and maximum comfortable level (MCL); a balancing procedure to iteratively determine levels of binaurally equal loudness for each electrode pair; a monaural pitch estimation procedure to reduce the number of candidates for pitch matching for both ears; and a pitch ranking procedure to determine the pitch discriminability for the pair candidates and finally select one pitch matched pair. Generally, the pretest procedure is an extension of the procedure presented in Chapter II.

To determine the THR, CL, and MCL for each electrode the perceived loudness was indicated by the subjects by pointing to the appropriate position on a continuous scale, ranging from “not audible” to “just uncomfortably loud”. The CL corresponded to the subject's response “comfortable”. The same procedure was then applied to determine the binaural CL, i.e., the comfortable level when both ears were stimulated simultaneously. Starting at 80% of the monaural CLs, levels were varied simultaneously in equal steps at the two ears. Subjects were instructed to attend to the overall loudness in the binaural case rather than to “hear out” a left-ear or right-ear contribution. Following the initial adjustment of the binaural CL, centralization of the perceived stimulus was checked and monaural levels were adjusted if necessary. All subjects required a reduction of current levels in the binaural condition relative to the monaural conditions to achieve the same loudness.

A magnitude estimation procedure was applied to obtain an estimate of the perceived pitch across the electrodes at both ears, similar to the procedures applied by Busby *et al.* (1994) and Collins *et al.* (1997). Stimuli were presented randomly between both ears and at each of the

Pulse rate (pps)	Stimulation currents left/right (µA)			
	CI1	CI2	CI3	CI8
100	-	358/1045	-	-
150	-	355/1045	-	-
200	478/486	362/1031	-	-
400	470/401	393/909	478/524	376/586
600	-	-	-	-
800	-	-	440/470	376/586
938	-	-	-	371/532
1600	501/424	-	347/370	-

Table 4: Stimulation levels for each CI listener as parameter of pulse rate. "-" shows not tested pulse rates.

electrodes 1-8, using the binaural CLs determined before. Subjects were instructed to assign numbers according to the perceived pitch of each stimulus. No restrictions on the range and type of numbers were given. Each stimulus was presented ten times. The distribution of pitch judgments across the electrodes and the two ears allowed selection of about 16 interaural electrode pairs supposed to elicit similar pitch sensation at the two ears. These pairs were evaluated further in the pitch-ranking task.

An automated procedure was applied to obtain interaurally loudness-balanced levels for each of the electrode pairs used further in the pitch-ranking task. The members of each electrode pair were presented in two subsequent intervals. By pressing one of two buttons the subjects adjusted the relative level of the signals between the two ears in steps corresponding to the smallest amplitude changes realizable by the implants to arrive at an interaurally matched loudness. The sum of the two levels within a trial was held constant and corresponded to the sum of the binaural CLs determined for the respective electrodes. The level difference at the beginning of each run was randomly roved. The mean value resulting from four runs was defined as the loudness-balanced levels for the members of the respective electrode pair.

In the pitch-ranking procedure the members of each of the electrode pairs were directly compared with respect to the perceived pitch difference, using a two-interval, two-alternative forced-choice (2-AFC) procedure. The pair members were presented randomly either in the first or second observation interval. Subjects were required to indicate which of the two stimuli sounded higher in pitch while concentrating on pitch rather than on other attributes such as timbre or loudness. Electrode pairs with an average discriminability across 25 repetitions within the range of chance (50 ± 18%) were considered as pitch-matched. For subjects with more than one pitch-matched electrode pair, the pair at medial tonotopic position was chosen. The selected electrode pairs were (left/right): 4/1 (CI1), 2/3 (CI2), 4/3 (CI3), and 7/5 (CI8).

Using the automated loudness balancing procedure described above, the levels for each pulse rate were determined with the goal of obtaining a binaurally balanced, comfortable loudness level for the selected pitch-matched electrode pair. Table 4 depicts the subject-dependent stimulation currents for each pulse rate.

1.4. Procedure

A two-interval, 2-AFC procedure was used in the lateralization discrimination tests. The first interval contained a reference stimulus with zero ITD, evoking a centralized auditory image. The second interval contained the target stimulus with the ITD tested. The subjects were requested to indicate whether the second stimulus was perceived to the left or to the right of the first one by pressing an appropriate button. All stimuli were repeated at least 60 times, in a balanced format with 30 targets on the left and 30 targets on the right. Thus, a subject with no ITD sensitivity could get 50 % responses correct by guessing. A score of 100 % correct responses would indicate that all stimuli were discriminated, with lateralization corresponding to the ear receiving the leading signal. In contrast, a score of 0 % implies perfect discrimination as well, but with lateralization at the ear receiving the delayed signal[5]. To avoid biasing the subjects towards a particular manner of responding, no feedback was given. To simplify the interpretation of the results, scores ranging from 0 % to 100 % were mapped to a range from -100 % to +100 %, referred to as “lateralization discrimination”. Lateralization discrimination of 0 % means that the target could not be discriminated from the reference stimulus with respect to the lateral position and represents 50 % correct responses.

5 Lateralization to the “wrong” side was possible due to ambiguous ITD information in the fine structure in cases where the ITD exceeded 0.5 IPI

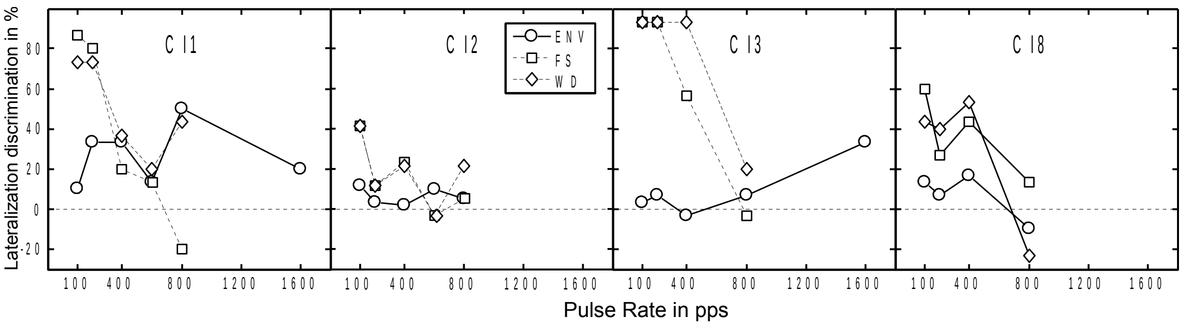

Figure 13: Pretest results as lateralization discrimination (LD) for different pulse rates and four CI listeners. Conditions: ENV: ITD ENV=625µs; FS: ITD FS=625µs; WD: ITD FS=ITD ENV=625µs

Lateralization discrimination pretests were performed to select, for each listener, the pulse rates to be used in the main experiment. Discarding conditions with very low sensitivity to ITD kept the test time as short as possible. In the pretests, one ITD value of 625 µs was presented in three different ITD conditions: ENV, FS and WD. The results for each CI listener are shown in Figure 13. Based upon these results and the availability of the subjects three to four pulse rates were chosen for each subject to be tested in main experiments (CI1: 200, 400, 1600 pps; CI2: 100, 150, 200, 400 pps; CI3: 400, 800, 1600 pps; CI8: 400, 800, 938 pps). The NH subjects were tested at 400, 600, 800 and 938 pps.

2. Results

The lateralization discrimination (LD) data of the individual CI listeners are shown in Figure 14 (CI1), Figure 16 (CI2), Figure 15 (CI3) and Figure 17 (CI8). The results of the NH subjects were more homogeneous; as a result, the mean scores of all four NH listeners are provided in Figure 18. For all listeners, there was a common pattern of LD as a function of ITD FS. At the lowest pulse rates (different for each subject) in the conditions ITD ENV = 0 µs, LD increased monotonically with ITD FS for ITD FS ≤ 0.25 IPI with a maximum at about 0.25 IPI. For ITD at approximately 0.5 IPI, LD was at chance (= 0 %), confirming the ambiguity in the lateralization task using ITD FS only. As ITD FS exceeded 0.5 IPI, the magnitude of LD as a function of ITD FS was similar to that for ITD FS < 0.5 IPI but with the opposite sign. This indicates that LD upon ITD FS is periodic and that stimuli with ITD FS > 0.5 IPI effectively represent stimuli with negative ITD FS. At the highest pulse rates tested (different for each sub-

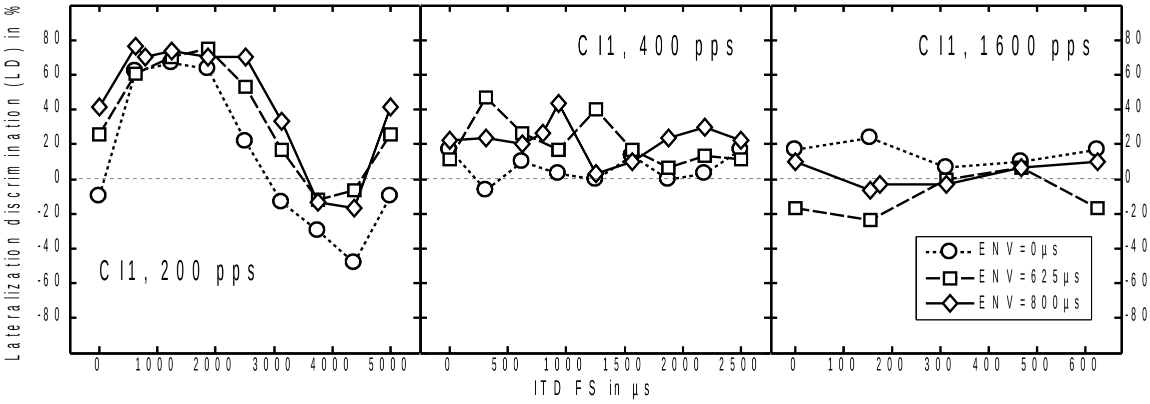

Figure 14: Lateralization discrimination for CI1 and different pulse rates. To point out the periodicity of the ITD FS the data points for the ITD FS = IPI are copies of the data points for ITD FS = 0µs. Note the different scaling of the X axes.

ject) the dependence of LD on ITD FS disappeared. Introducing a non-zero ITD ENV resulted in a lateralization shift towards the ear receiving the stimulus with the leading envelope. This effect seems to increase with increasing pulse rate.

Although most trends were easily distinguishable by visual inspection, a statistical analysis was used to determine the significance of the trends. The statistical method employed was multidimensional contingency table analysis (Lienert, 1978; Agresti, 1984; Agresti, 1996) implemented in “stats” package of R (R Development Core Team, 2004). This is a useful method for inter-subject comparisons of results obtained by a 2-AFC task, for which the variance analysis is not available[6], although it is an unusual method in psychoacoustics. A general description of this method would exceed the scope of this thesis. Thus, only a summary of the tests and models applied in the context of the data analysis is provided.

6 Note that the lateralization data could not be modeled by means of JNDs since the functions are non-monotonic.

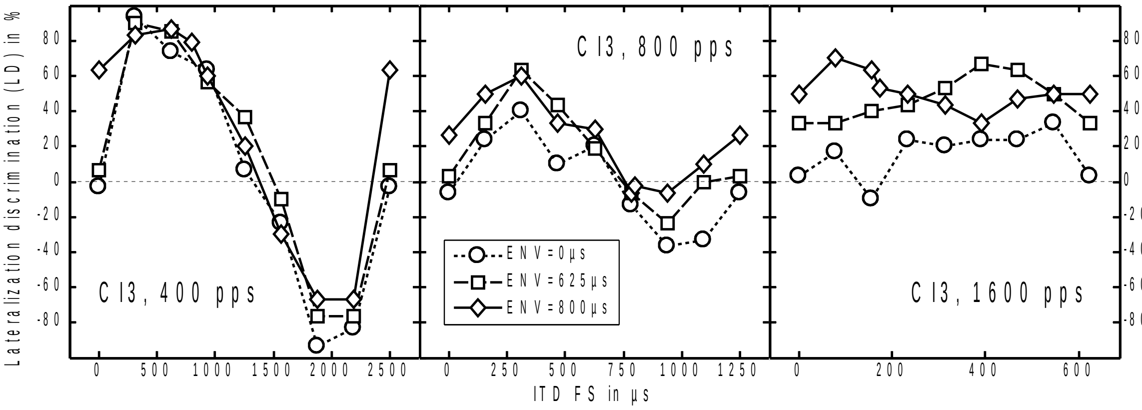

Figure 15: Lateralization discrimination for CI3 and different pulse rates. All other conventions are as in Fig. 14.

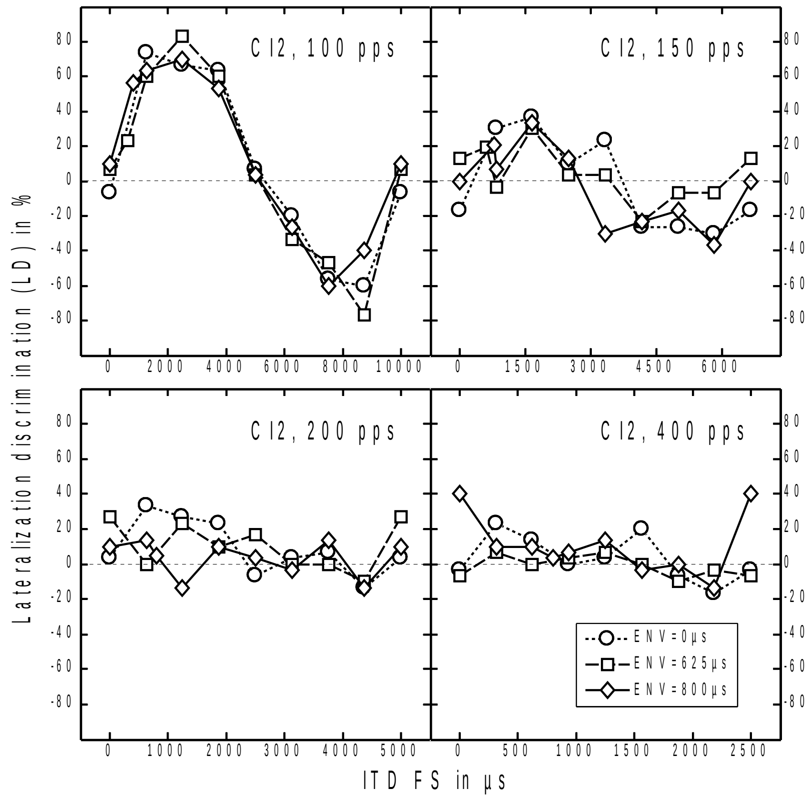

Figure 16: Lateralization discrimination for CI2 and different pulse rates. All other conventions are as in Fig. 14.

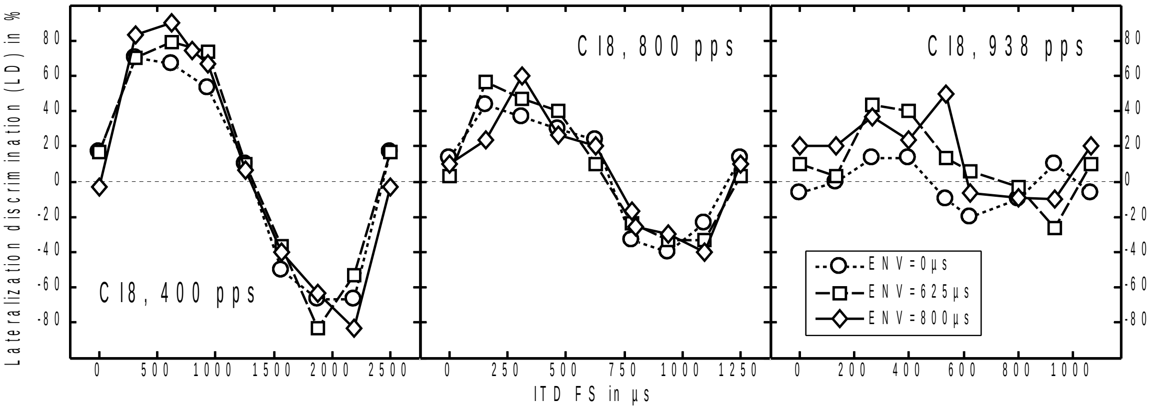

Figure 17: Lateralization discrimination for CI8 and different pulse rates. All other conventions are as in Fig. 14.

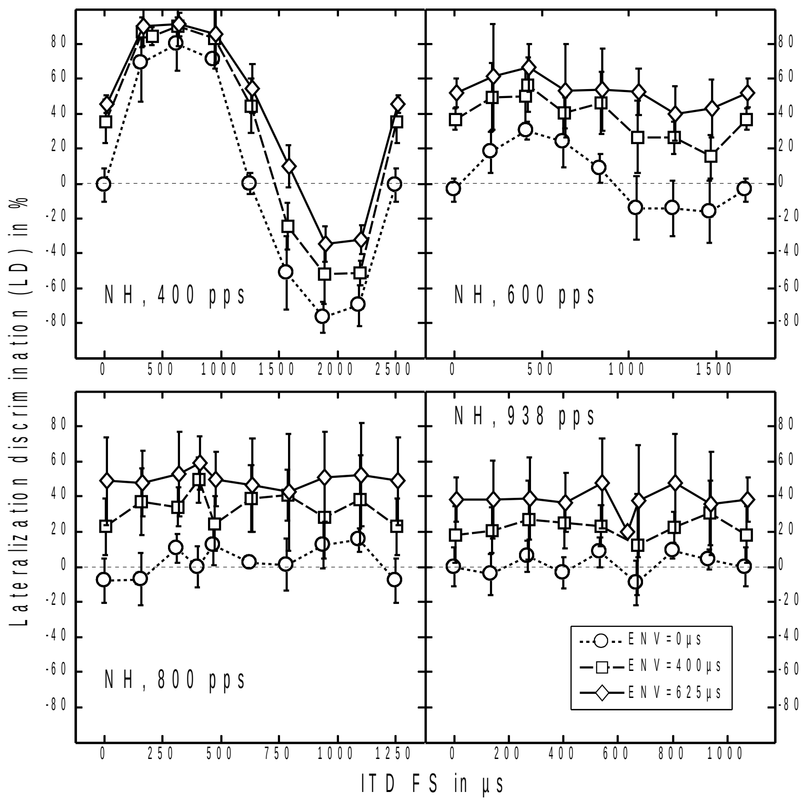

Figure 18: Average lateralization discrimination for all NH subjects and different pulse rates. To point out the periodicity of the ITD FS the data points for the ITD FS = IPI are copies of the data points for ITD FS = 0µs. The bars shows the standard deviation. Note the different scaling of the X axes.

The significance of the differences between two conditions was tested by obtaining the two-tailed probability p of the Pearson χ^2 statistic for the null hypothesis that the logarithmic odds ratios[7] for both conditions are equivalent (Agresti, 1984). Log-linear models were fitted to determine the interactions between different factors. The goodness-of-fit of a model to the data is described by the G^2, df and p-values. In these cases, the significance of an effect is given by the significance of the corresponding model parameter. In cases of fitted data, the calculation of odds ratios was done using estimated response frequencies, according to Agresti (1984) (pp. 47-69). Some conditions were tested with a higher number of repetitions (> 60) due to differences in the availability of subjects. Thus, the investigations of marginal associations, collapsing[8] the data over the variable tested with different number of repetitions, were done using regularization of the data to the smallest common number of repetitions (= 60). To increase the test power, pulse rate and ITD ENV were treated as ordinal factors.

The statistical analysis is structured as follows: in the first section (Groups of Subjects) the differences between subjects will be analyzed and a classification will be done to show the homogeneity in the results of the CI listeners and how it compares to the results of the NH listeners. The next section (Interaural Synchronization) considers effects of interaural synchronization on lateralization discrimination and shows some improvements which can be achieved by specific coding of the ITD. In the last section (Effects of ITD ENV) the effects of ITD ENV will be analyzed in relation to the factors pulse rate and subject.

2.1. *Groups of Subjects*

The pulse rate of 400 pps and ITD ENV of 0 µs and 625 µs were used for the analysis of differences between subjects, since they were the only values available for all subjects. Using these data, the odds ratios for the categorical factor subject were calculated and analyzed. The subjects were clustered according to the type of stimulation and the analysis was performed on

7 The odds ratio is the ratio of the probabilities of obtaining a correct response for one condition compared to another condition. An odds ratio of 1 shows that there is no difference in the correct responses between the two conditions. The significance of the difference between the two conditions can be calculated using the logarithmic odds ratios and their corresponding confidence intervals. In multi-factorial design, logarithmic odds ratios directly correspond to the interaction terms in the log-linear models allowing investigation of interactions between factors.

8 Agresti (1984) uses the term “to collapse” to describe the process of averaging a factor in a data pool. This process is also known as marginalization by a factor.

each group separately. The results showed that the group of NH subjects was homogeneous and the group of CI listeners was heterogeneous $(\chi_3^2=20.0, p<0.001)$. Therefore, further analyses were performed on the NH listeners as a group and for each CI listener individually. However, the better performing CI listeners (CI3 and CI8) performed sufficiently similar to the NH subjects that they could have been clustered with the NH subjects to form a larger homogeneous group $(\chi_5^2=1.34, p=0.93)$.

2.2. *Interaural Synchronization*

To address the question of the need for interaural pulse synchronization, the dependence of LD on ITD FS was investigated. If a subject lateralizes the stimuli for ITD FS < 0.5 IPI to one side and for ITD FS > 0.5 IPI to the opposite side, that implies that LD depends on ITD FS. Therefore, the data for all conditions fulfilling ITD ENV = 0 µs were grouped as follows: the first group contained all LDs for 0 < ITD FS < 0.5 IPI and the second group all LDs for 0.5 IPI < ITD FS < IPI. Results for ITD FS = 0 and 0.5 IPI were discarded as there is no lateralization information, and they should be at chance rate. It was hypothesized that if there is a significant difference between the direction of LD for both ITD FS groups, then LD depends on ITD FS, thus indicating the necessity of interaural pulse synchronization.

For the NH listeners the data pool (subject × ITD FS group × pulse rate × response) was collapsed across subjects because of the homogeneity of their performance. A log-linear model was fitted to this data pool with the result that only the saturated model[9] could give an accurate fit, showing a strong interaction between the factors pulse rate and ITD FS group. Thus, the dependency on ITD FS was investigated for each pulse rate separately, analyzing the odds ratios in partial contingency tables with fixed pulse rate. For the NH subjects 600 pps was the highest pulse rate with significant sensitivity on ITD FS. For the CI listeners the data were analyzed separately for each subject in the same way as for the NH listeners: odds ratios in partial contingency tables were analyzed for each pulse rate. The better performing CI listeners (CI3, CI8) showed significant sensitivity to ITD FS for pulse rates up to 800 pps, while for the

9 According to Agresti (1984) a saturated model contains all interactions of all factors. In this model, no factor may be averaged and the data must be analyzed for each level of each factor separately.

Pulse rate (pps)	NH	CI1	CI2	CI3	CI8
100	-		<0.001	-	-
150	-		<0.001	-	-
200	-	**<0.001**	**0.006**	-	-
400	<0.001	0.752	0.21	<0.001	<0.001
600	**<0.001**	-	-	-	-
800	0.139	-	-	**<0.001**	**<0.001**
938	0.712	-	-	-	0.45
1600	-	0.459	-	0.108	-

Table 5: Probability (p-value) of no dependence of LD on ITD FS for ITD ENV = 0µs conditions. Conditions with the highest pulse rate with significant sensitivity on ITD FS are shown bold; df was 1 for all results.

poorer performing CI listeners (CI1, CI2) a significant sensitivity could be found only for pulse rates up to 200 pps. Detailed results of the analysis are shown in Table 5.

It was further hypothesized that for conditions showing a dependence of LD on ITD FS, the synchronization of the ITD FS to the ITD ENV would result in a better LD than synchronizing ITD FS to zero. This was evaluated by keeping the ITD constant and comparing the LD between two synchronization conditions: one in which the ITD is carried both by the envelope and the fine structure (WD) or alternatively by the envelope only, keeping ITD FS at zero (ENV). In the statistical analysis, the synchronization conditions were regarded as a factor with two levels (WD, ENV) and the ITD as a factor with two levels, which depended on the subject group (NH: 400 µs and 625 µs; CI: 625 µs and 800 µs). The analysis was performed for the NH group and for each CI listener separately. As before, log-linear models were fitted to investigate the interactions. The hypothesis of no interaction between the factors ITD and synchronization condition could not be rejected for CI1 ($G^2=8.3867$, df=10, p=0.591) nor CI8 ($G^2=8.4617$, df=10 , p=0.584). Thus, the data were collapsed over ITD for these subjects. For all other subjects, separate partial tables were used for each ITD value. The probabilities for the hypothesis of equal LDs in both synchronization conditions for each subject and pulse rate showing dependence on ITD FS (p<0.05 in Tab. 5), are given in Table 6.

The NH subjects showed an improvement using the WD condition for pulse rates up to 600 pps for an ITD of 400 µs (p=0.04). Increasing the ITD to 625 µs, the improvement due to

Pulse rate	NH		CI1	CI2		CI3		CI8
(pps)	400µs	625µs	-[a]	625µs	800µs	625µs	800µs	-[a]
100	-	-	-	0.356	**0.007**	-	-	-
150	-	-	-	0.681	0.143	-	-	-
200	-	-	**<0.001**	0.141	0.729	-	-	-
400	<0.001	<0.001	-	-	-	**<0.001**	0.091	**<0.001**
600	**0.04**	0.696	-	-	-	-	-	-
800	-	-	-	-	-	0.266	0.043[b]	0.17

Table 6: Probability for equal LD in conditions ENV and WD. Conditions with the highest pulse rate with significantly higher LD for WD than for ENV are shown in bold. a) ITD was marginalized in these cases. b) LD for ENV condition was higher than for WD condition. χ^2 values have been omitted for readability purposes.

synchronization decreased, and could be found for 400 pps only ($p<0.001$). This revealed an interesting effect of combining ITD FS and ITD ENV: assuming a dependence of LD on ITD FS, it can be expected that increasing the ITD in both the envelope and fine structure (WD) improves LD up to about 0.25 IPI. Above this point, up to ITD = 0.5 IPI, LD is expected to decrease because at ITD = 0.5 IPI the ITD FS cue provides ambiguous information. Increasing the ITD further, depending on the relative perceptual contribution of ITD ENV, the stimulus may even be lateralized towards the opposite side. This actually happened for CI8 (800 pps, ITD FS=800 µs, see Figure 17). Thus, the synchronization of the fine structure to the envelope gives an improvement for ITD values smaller than half IPI only.

For the CI listeners, improvements due to synchronization were observed for the following pulse rates: 200 pps (CI1), 100 pps (CI2), and 400 pps (CI3 and CI8). For CI3, at a pulse rate of 800 pps and ITD of 800 µs, there was a significant difference ($p=0.043$), but the LD was higher for ENV than for WD. Since the ITD exceeded 0.5 IPI, this is an example of the effect of combining conflicting ITD FS and ITD ENV, which was also seen in the NH subjects' results. This effect seems to reduce the positive effects of synchronizing ITD FS to ITD ENV, but it allows an optimization of coding ITD FS: for ITD values greater than 0.25 IPI the WD condition was modified, diminishing the ITD FS to 0.25 IPI. This condition is termed *diminished waveform delay* (WD_{DIM}) and a new formula for ITD FS coding is proposed:

$$ITD\ FS = min\left(ITD\ ENV, \frac{1}{4}\ IPI\right) \qquad (4)$$

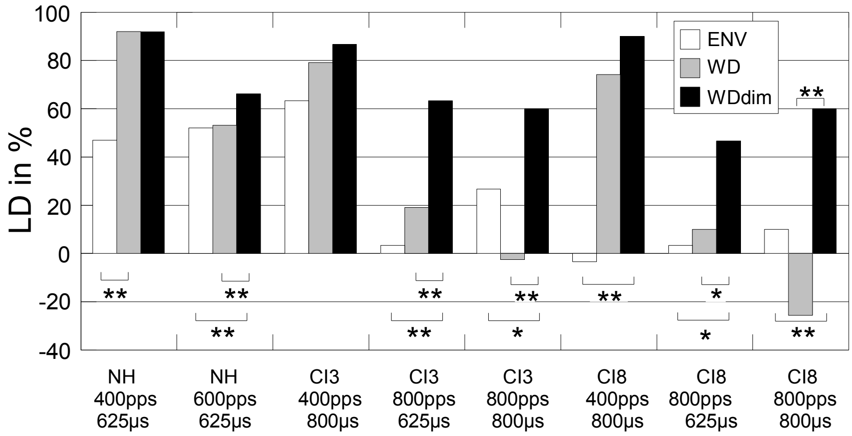

*Figure 19: Comparison of lateralization discrimination for conditions ENV, WD and WD_{DIM}. Significance codes: * p<0.05; ** p<0.01*

To obtain an improvement using WD_{DIM}, the ITD must be greater than 0.25 IPI. To fulfill this requirement for low ITDs the pulse rate must be as high as possible. On the other hand, the effect of ITD FS is weak for higher pulse rates. Thus, the WD_{DIM} optimization is efficient only for medium pulse rates showing sensitivity to ITD FS. Figure 19 compares lateralization discrimination between the conditions ENV, WD, and $WD_{DIM,}$ for each of the ITDs and pulse rates measured and fulfilling the specified requirements. In most cases LD increased using WD_{DIM} optimization; in one case (CI8, 800 pps, 800 µs) even a reversal of lateralization into the correct direction could be achieved.

2.3. Effects of ITD ENV

To determine the effects of envelope delay, a comparison of the sensitivity to ITD ENV between the subjects was performed for stimuli with 400 pps and ITD ENV values of 0 µs and 625 µs. For subjects showing no effect of ITD FS at this pulse rate and ITD ENV held constant at either value, all ITD FS values can be averaged to increase the power of the test. Even for subjects showing dependence of ITD FS on LD at this pulse rate, the marginalization of ITD FS is justified, based on the finding of independence between ITD ENV and ITD FS. Thus, log-linear models including the factors subject, response and ITD ENV for results collapsed over ITD FS were fitted to the data. For both CI and NH listeners only the saturated model could be fitted, showing strong inter-subject variability (p<0.001 for the 3-way interaction term). Thus, the NH and CI listener groups were analyzed separately. For the NH listener group the model with the marginalized factor subject fits well (G^2=3.432, df=12, p=0.992), revealing homogeneity of subjects within this group and high sensitivity to ITD ENV (p<0.001). For the CI listener group the model including interactions [subject × response] and [ITD ENV × response] gave the best fit (G^2=6.121, df=12, p=0.41) showing a significant overall sensitivity to ITD ENV of the group (p=0.031); however, a significant deviation of the performance of listeners CI1 (p=0.039) and CI2 (p=0.032) from the group of CI subjects could be found. Therefore the ITD ENV sensitivity of the CI listeners was analyzed separately in a contingency table analysis, revealing a significant sensitivity for listener CI1 (p=0.007) and no sensitivity for the rest (p>0.1 for CI2, CI3, CI8) at 400 pps.

Finally, the effect of ITD ENV (0, 625, 800 µs) on LD was investigated with pulse rate as parameter. The three-way interaction parameters of saturated log-linear models including the factors pulse rate, ITD ENV, and subjects' response for each subject were examined. An interaction between ITD ENV and pulse rate could be found for CI3: increasing the pulse rate within the range of values tested or raising the ITD ENV increased the odds ratio by the factor of 1.129 (p=0.011), indicating a greater sensitivity to ITD ENV with increasing pulse rate. For the NH subjects, sensitivity was independent of the pulse rate (p=0.66) but strongly associated with ITD ENV (raising the ITD ENV from 0 to 400 µs or from 400 to 625 µs increased the odds ratio by the factor of 1.558, p<0.001).

3. Discussion

The results of this study show that the tested CI listeners are sensitive to ITD in the fine structure for pulse rates up to 800 pps, depending on the individual listener. The NH subjects listening to a simulation of electric stimulation showed sensitivity up to 600 pps for the same conditions. This is qualitatively consistent with the results of the study presented in Chapter II, where we used unmodulated trains with four pulses (Laback *et al.,* 2007). We found sensitivity to ITD FS up to 800 pps in two out of three CI listeners and up to 400 pps in the NH listeners, depending on the individual listener. Furthermore, in both studies the data of the NH subjects show little inter-subject variability, as opposed to the results of the CI listeners, who in this and most other studies show a wide range of performance. Such strong inter-subject variability implies that at least one factor influencing sensitivity to fine structure ITD has not been taken into account in this study. Thus, a statistical evaluation of CI listeners as a group may yield misleading conclusions.

Recovery from forward masking (Chatterjee, 1999; Nelson and Donaldson, 2001; Nelson and Donaldson, 2002) could be an explanation for the large subject dependence among CI listeners. Forward masking in CI listeners may imply longer decay in the excitation pattern, which would smear the fine structure resulting in a lower sensitivity to fine structure ITD.

The strong inter-subject variability, combined with the small number of subjects, limits the interpretation of the results to case studies. It is interesting that the better performing CI listeners (CI3, CI8) had 1-2 months of bilateral experience, while the worse performers (CI1, CI2) had years of bilateral experience. One may be tempted to hypothesize that CI1 and CI2 lost their ability to utilize fine structure cues because of the longer period of stimulation with signals that are uncorrelated in the fine structure. Actually, one reviewer of the paper corresponding to this study (Majdak *et al.*, 2006) mentioned that issue, too. We had the opportunity to test CI8 one year after the main tests under similar conditions. In contrast to the hypothesis, the performance of CI8 was not significantly different than the previous results, which suggests, that either the time constant of the unlearning process is much longer than one year or the origins underlying the individual differences in performance are much more complex. Actually, testing one subject to validate a hypothesis based on evidence derived from two groups of two subjects appears to be very speculative. Thus, more extensive tests with more subjects are required to validate this hypothesis.

There are at least two possible explanations for the higher maximum rate showing significant effects of ITD FS on LD in the better performing CI listeners compared to the NH listeners. First, one of the limiting parameters in acoustic hearing could be the smearing of the temporal information by auditory filtering in the cochlea. Filtering the acoustic signals with a simulation of the auditory filter (center frequency: 4590 Hz) shows that by increasing the pulse rate, the modulation depth of the stimuli decreases, but is still present for pulse rate as high as 938 pps, which was the highest pulse rate used in the experiments with the NH listeners. In electric hearing, the auditory filters are bypassed. Second, in electric hearing, the degree of neural phase locking is known to be stronger than in acoustic hearing due to bypassing the synaptic mechanism at the hair cell (Abbas, 1993). An appropriate characteristic of both effects with respect to the pulse rate might lead to higher ITD FS sensitivity in electric hearing at higher pulse rates, and may account for these results.

ITD FS was varied in the range between zero and IPI, which corresponds to a setup of unsynchronized speech processors, in which ITD FS varies periodically between 0 and IPI. The results obtained for these stimuli show that ITD FS can cause a lateral shift in the perceived position up to a pulse rate of 800 pps (CI listeners) and 600 pps (NH subjects). Therefore, to control the lateral position of the auditory image, the fine structure of the stimulus should be encoded for stimulation pulse rates up to 800 pps.

Comparing the subjects' sensitivity to ITD ENV at 400 pps, three out of four CI listeners showed no sensitivity, as opposed to the results of CI1 and the NH listeners. It appears to be contradictory that CI listeners, who performed comparably to NH listeners with respect to ITD FS, showed much worse sensitivity to ITD ENV. One possible explanation for that is the effect of the amplitude modulation shape. In our study the trapezoidal modulation was a compromise of providing strong envelope and fine structure cues. As an example, the rectangular modulation is expected to provide a stronger ITD ENV cue, but, it allows ITD ENV values in integer multiples of the IPI only and therefore, it is not adequate for this study. On the other hand, it is expected that extending ramps beyond 20 ms results in, besides a higher ITD ENV resolution, an attenuation of the onset effects in each trapezoid. Furthermore, by applying different ramps or changing the duty factor, the amount of information in the fine structure changes, which has an effect on ITD FS sensitivity. A "nice" alternative may be a speech-shaped pulse train, providing information on lateralization discrimination sensitivity to ITD ENV for more realistic stimuli.

Considering all pulse rates tested, one subject (CI3) showed a consistent improvement of sensitivity to ITD ENV with increasing pulse rate. The NH listeners showed a positive effect of ITD ENV but no significant effect of pulse rate. This is in agreement with the results of Henning (1974), who showed no monotonous effect of rate. In general, the sensitivity to fine structure ITD was higher than to envelope ITD for all subjects in this study.

In the real world, most stimuli carry coherent ITD information in both the fine structure and envelope corresponding to the waveform delay condition (WD) tested in this study. A comparison of the WD and ENV conditions is important with respect to practical applications.

These results show that the WD condition results in better LD for pulse rates up to 400 pps (CI listeners) and 600 pps (NH listeners) relative to condition ENV. It was also shown that for the combination of higher pulse rates and higher ITD values, the WD condition leads to a deterioration of LD as a result of ITD FS cues pointing to the wrong side and ITD ENV cues pointing to the correct side. To avoid this negative effect an optimized WD condition called *diminished waveform delay* ($WD_{DIM.}$) was introduced, in which ITD FS was limited to 0.25 IPI. Using WD_{DIM} resulted in an improvement of LD relative to WD for pulse rates up to 800 pps for CI listeners CI3 and CI8. There are practical constraints with regard to implementing the WD_{DIM} rule in bilateral CI systems. Whether to use WD_{DIM} or not should be based on the pulse rate applied in the stimulation strategy: if ITD values greater than 0.25 IPI are expected, WD_{DIM} will improve lateralization discrimination. The efficacy of WD_{DIM} is restricted to pulse rates with sensitivity to ITD FS, i.e., to pulse rates up to 800 pps only. Furthermore, WD_{DIM} requires a bilateral processor with the ability to extract and control ITD cues in the envelope and fine structure, which may be difficult to implement.

There is also another way to provide ITD FS cues to CI listeners: coding the temporal information in the envelope of a very high pulse rate carrier (several thousands of pps) such as “HiRes” (Wilson, 2004). HiRes was investigated in several monaural studies (e.g. Frijns *et al.*, 2002, Filipo *et al.*, 2004, Bosco *et al.*, 2005), showing some improvements, particularly a better speech recognition in noise, compared to pulse rates in the region of 1500 pps. However, it is difficult to interpret these results in the context of bilateral stimulation and effects on fine structure ITD sensitivity.

Improving ITD FS perception requires using lower pulse rates, which may influence the performance with respect to monaural speech perception in quiet. Several studies have compared speech intelligibility performance by varying the pulse rate. For example Fu and Shannon (2000) and van Hoesel *et al.* (2002) tested different pulse rates, however they did not consider listener accommodation to a new stimulation rate. In contrast, Vandali *et al.* (2000) and Holden *et al.* (2002) tested different pulse rates and did consider listener accommodation. In general these studies provide no contraindication to using pulse rates as low as 250 pps with respect to speech intelligibility in stimulation strategies optimized for ITD FS coding.

One strategy which encodes timing information in fine structure is Peak Derived Timing (PDT) introduced by van Hoesel and Tyler (2003). PDT takes into consideration the fine structure of acoustic signals and provides electric signals similar to the WD condition in this paper. In the PDT strategy, the temporal position of an acoustic peak in a subband is determined and an electric pulse is applied to the corresponding electrode at the corresponding time. As a consequence, the pulse rate varies according to the temporal properties of the acoustic signal at each channel and was limited to a maximum of 1400 pps. Van Hoesel and Tyler could not find any clear difference between the PDT strategy and the standard clinical stimulation strategy with respect to sound localization and speech perception in noise. Unfortunately, the comparison between the two strategies was confounded by differences in the experimental setup such as automatic gain control (AGC), dynamic range, and number of electrodes. Hence, more detailed investigations into the efficiency of encoding fine structure timing information with various strategies are required to determine the actual extent of lateralization improvement for CI listeners.

4. Summary

This study shows that CI listeners are able to lateralize stimuli using interaural time differences in the fine structure only, up to pulse rates as high as 800 pps. This may affect the lateralization of sounds using speech processors which do not consider the synchronization of the fine structure. In contrast to the study presented in the previous chapter, more realistic stimuli with the duration of 300 ms were used in this study. Three different synchronization conditions were introduced and tested with four CI listeners, indicating some possible constraints for future stimulation strategies to take more advantage of the interaural time difference information in the fine structure and envelope.

IV. Interaural time differences as function of center frequency

Submitted to *Journal of Acoustic Society of America*

The study presented in this chapter has a strong relation to the studies presented in Chapters II and III. Both studies have been already published in professional journals. Thus, for reasons of simplicity, the studies presented in Chapters II and III are referred to as Laback *et al.* (2007) and Majdak *et al.* (2006), respectively.

As already mentioned, ITD information in unmodulated signals can only be processed up to about 1500 Hz (see Chapter I or for review Blauert, 1997). At higher frequencies, ITD information can be processed from the slowly-varying temporal envelope only. This is supported by various psychoacoustic (Henning, 1974; Nuetzel and Hafter, 1976; Nuetzel and Hafter, 1981; Bernstein, 2001; Bernstein and Trahiotis, 2002) and physiologic studies (Yin *et al.*, 1984; Skottun *et al.*, 2001; Shackleton *et al.*, 2003). The ability to transmit both envelope and carrier information in separate paths seems to be a general property of the auditory system (Liang *et al.*, 2002) and was also found in other sensory modalities like visual and electrosensory systems (Middleton *et al.*, 2006).

There is strong evidence that envelope ITD sensitivity depends on the temporal properties of the fast-varying carrier signal. This was shown for sinusoidally amplitude-modulated (SAM) tones (Henning, 1974; Stellmack *et al.*, 2005; Nuetzel and Hafter, 1976), two-tone complexes (McFadden and Pasanen, 1976; Bernstein and Trahiotis, 1994), transposed tones (Bernstein and Trahiotis, 2002), and bandpass-filtered click trains (Hafter and De Maio, 1975; Hafter and Dye, 1983). All these studies found that envelope ITD sensitivity is limited with respect to the rate of the envelope fluctuations. However, ITD sensitivity depends not only on the rate of envelope fluctuations. For example, Bernstein and Trahiotis (2002) found that transposed tones yield higher performance than SAM tones for comparable rates. On the physiological basis, it was shown that, under certain conditions, the response to the envelope ITD of transposed tones is comparable with the response to the ITD of low-frequency pure tones (Griffin *et al.*, 2005; Dreyer and Delgutte, 2006).

In cochlear implant (CI) listeners, Majdak *et al.* (2006) and Laback *et al.* (2007) investigated the sensitivity to the ongoing ITD presented at one binaural electrode pair. At least one CI listener tested showed sensitivity up to 800 pulses per second (pps). Further, in that studies, we compared the results to the ongoing *envelope* ITD sensitivity of normal hearing (NH) subjects. In NH experiments, we used bandpass-filtered pulse trains to simulate pulsatile stimulation with CIs, similar to McKay and Carlyon (1999). The best NH listeners showed sensitivity up to 600 pps (Majdak *et al.*, 2006) and 400 pps (Laback *et al.*, 2007) only. So, compared to the best performance in the electric stimulation, the performance of the best NH listeners was lower. We concluded that the comparison between these two groups may be affected by different properties of electric and acoustic stimulation. In acoustic stimulation the bandpass-filtered pulse train passes through the basilar membrane, where the modulation depth can be reduced depending on the stimulation place. This filtering process is bypassed in electric stimulation. So, assuming that a large modulation depth is required for ITD sensitivity (Stellmack *et al.*, 2005), the performance of the NH listeners may have been limited by auditory filtering in the cochlea. The investigation of the possible limitation in the comparison between NH and CI listeners' results was the main motivation for this study.

If the auditory filtering limits ITD perception, then the increase of the stimulus center frequency (CF) should improve the performance. This is because the auditory filter bandwidth increases with the CF, causing relatively less reduction in the modulation depth (e.g. Moore, 1978), especially at high modulation rates. However, the inversion of that argument is not valid: if the performance does not improve with CF, it does not mean that the auditory filtering has no effect. The effect of auditory filtering may be superimposed by other effects associated with the stimulation of different tonotopic places. Those are:

- Changing hearing thresholds, which may result in a change of ITD sensitivity (Nuetzel and Hafter, 1976).
- Stimulating different amounts of peripheral neurons when constant bandwidth stimuli are applied (Buell and Hafter, 1991).
- Differences in the higher processing stages (e.g., changing number of responding cells in the central nervous system as indicated in Bernstein and Trahiotis, 1994, or differences in the tonotopic specialization of the coincidence detectors as shown in chicks by Kuba *et al.*, 2005).

In NH listeners, Hafter and Dye (1983) investigated the detection of ITD in bandpass-filtered click trains as a function of the inter-pulse interval (IPI) and the number of clicks. Although they did not collect data for different CFs, their results indicate that the modulation depth reduction due to auditory filtering does not explain the decrease in performance for increasing pulse rate. However, additional data on the ITD sensitivity at different CFs are required to substantiate their conclusions.

For other types of stimuli, there are strong indications that auditory filtering is not the limiting factor in ITD perception, as supported by the results of two studies: Bernstein and Trahiotis (1994) and Bernstein and Trahiotis (2002). For 100% SAM tones and two-tone complexes, Bernstein and Trahiotis (1994) found a general decrease of ITD sensitivity when the CF was increased. They varied the modulation rate as well, showing that the sensitivity to ITD decreased as the rate of envelope fluctuation increased. The comparison of the modulation rate effect at different CFs led to the conclusion that the change in ITD sensitivity is unlikely affected by auditory filtering. For transposed tones, Bernstein and Trahiotis (2002) investigated ongoing ITD sensitivity and compared their results to those obtained with SAM and pure tones. For the lower modulation frequencies (< 256 Hz), listeners showed even higher ITD sensitivity to transposed tones than to pure tones with comparable frequency. However, using higher modulation frequencies, JNDs could not be determined for transposed tones. Furthermore, increasing the center frequency of the transposed tones resulted in lower sensitivity as well. From their results, Bernstein and Trahiotis (2002) concluded that for SAM and transposed tones, the ITD perception is not primarily limited by the effects of auditory filtering.

Nevertheless, Bernstein and Trahiotis (2002) found large differences in the ITD sensitivity between the SAM and transposed tones. Those findings indicate that the relative decrease in ITD sensitivity usually observed for high-frequency stimuli (compared to low-frequency stimuli) is associated with the type of used stimuli, as specified by center frequency, rate, and shape of envelope fluctuations. Hence, conclusions derived for SAM and transposed tones do not automatically apply to another type of stimuli like bandpass-filtered click trains. These stimuli have different envelope fluctuation shapes even for the same CF and rate. Thus the impact of the auditory filtering on ITD perception may be different and require a separate investigation for click trains.

In Majdak *et al.* (2006) and Laback *et al.* (2007), we used bandpass-filtered click trains because they appear to be a good approximation of the electric stimulation in CIs when investigating temporal effects (McKay and Carlyon, 1999). Electric pulse trains presented at one electrode in the cochlea have a very steep onset, which may be best approximated by bandpass-filtered click trains in acoustic stimulation. Of course, when compared to electric pulse trains, the response of an auditory filter to bandpass-filtered click trains results in temporal fluctuations with a clearly reduced modulation depth and shallower onsets. However, the resulting modulation depth is still larger and the onsets are still steeper than for SAM or transposed tones. Dreyer and Delgutte (2006) showed that phase locking to transposed tones is better than for SAM tones and comparable to pure tones under certain restrictions. Thus, stimuli with even steeper onsets and larger modulation depths like bandpass-filtered pulse trains promise further improvements in the phase locking effects. An additional advantage of bandpass-filtered pulse trains is that the onset slopes do not change with the pulse rate as it is the case for SAM and transposed tones.

The amplitude spectrum of a click train is a harmonic series, which can be limited in frequency by applying a bandpass filter. Thus, the bandwidth and the CF are independent, allowing to systematically stimulate different amounts of neurons at different tonotopic places. For both SAM and transposed tones the bandwidth is directly related to the modulation rate and is constant for all CFs. It is well-known that the logarithmic scaling is a good approximation of the frequency-to-place mapping (Greenwood, 1961). Thus, applying SAM and transposed tones at higher CFs may stimulate less neurons than at low CFs because the bandwidth is constant. Even though the effect of bandwidth on the sensitivity to envelope ITD is still unclear, there are some indications that increasing bandwidth leads to improved envelope ITD perception (Dye *et al.*, 1994). Using stimuli with a logarithmically-scaled bandwidth may result in relatively higher sensitivity at higher CFs.

The main focus of this study is to reconsider the comparison between CI and NH listeners we performed in Majdak *et al.* (2006) and Laback *et al.* (2007) under the hypothesis that the NH performance was not underestimated because of auditory filtering. In this study, the effect of auditory filtering was investigated by systematically varying the center frequency and the pulse rate of bandpass-filtered pulse trains and testing the sensitivity to ongoing envelope ITDs in NH listeners. Additionally, the results are discussed in the context of the general effects of CF and pulse rate on ITD perception.

1. Methods

1.1. Subjects and Apparatus

Five NH subjects participated in this study. All subjects were aged between 25 and 35 years and had no indication of hearing abnormalities.

A personal computer system was used to control the experimental task. The stimuli were output via a 24-bit stereo D/A converter (ADDA 2402, Digital Audio Denmark) using a sampling rate of 96 kHz per channel. The analog signals were sent through a headphone amplifier (HB6, TDT) and an attenuator (PA4, TDT), and presented to the subjects via a circumaural headphone (K501, AKG). Calibration of the headphone signals was performed using a sound level meter (2260, Brüel & Kjær) connected to an artificial ear (4153, Bruel & Kjær).

1.2. Stimuli

The stimuli were 300-ms pulse trains composed of monophasic pulses with a duration of 10.4 µs, corresponding to one sampling interval at a sampling rate of 96 kHz. The pulse rate was varied from 200 pps to 588 pps. The smallest inter-pulse interval (IPI) was 1700 µs. Both studies, Majdak *et al.* (2006) and Laback *et al.* (2007) showed that for pulse rates above 600 pps the ITD sensitivity degrades to chance rate. Hence, higher pulse rates were not required.

The ITD was introduced by delaying the temporal position of the pulses at one ear relative to the other ear. To restrict the ITD to the ongoing part of the stimulus only, the position of the first pulse pair (onset) and the last pulse pair (offset) was fixed to ITD of zero (Laback *et al.*, 2007). To minimize the monaural perception of irregularities, half of the ITD was applied to the leading ear and the other half to the lagging ear. Laback *et al.* (2007) showed that these irregularities are not perceptible. The ITD values varied from 25 to 400 µs. The ITD was adapted to account for the subject's sensitivity. According to Majdak *et al.* (2006), ITDs higher than 0.25 IPI can lead to lower performance compared to ITDs lower than 0.25 IPI. This is an effect of the ambiguity in ongoing ITD information provided by the stimulus for ITD = 0.5 IPI and results in a nonmonotonic psychometric function. Thus, the combinations of ITD and pulse rate in this study resulting in ITD values higher than 0.25 IPI were not tested.

The pulse trains were passed through a digital eight-order Butterworth filter. Three different center frequencies (CF) with the corresponding bandwidths were tested: 4589 (CF1), 6490 (CF2), and 9178 Hz (CF3). The lowest CF, CF1, was chosen to be identical with the CF used in Majdak *et al.* (2006) and Laback *et al.* (2007). The highest CF, CF3, is double the frequency of CF1. CF2 is the geometrical average of CF1 and CF3. Using CFs higher than 10 kHz would lead to even shorter impulse responses of the auditory filtering. However, we did not test CFs above 10 kHz since Bernstein and Trahiotis (1994) showed that it is hard to retrieve valid data for very high CFs. Also in electric hearing, the most-basally implanted electrode usually corresponds to frequencies below 10 kHz. The details about the filter configuration can be found in Table 7. It is important to notice that the bandwidth (in Hz) changes with the CF to achieve a constant stimulation range (in mm) on the basilar membrane.

Code	Geometrical Center Frequency (Hz)	Bandwidth (Hz)	Lower edge (Hz)	Upper edge (Hz)
CF1	4589	1500	3900	5400
CF2	6490	2121	5515	7637
CF3	9178	3000	7800	10800

Table 7: Filter configuration for the conditions CF1, CF2, and CF3.

Given that the sound pressure level re 20 µPa (SPL) depends on the pulse rate, the amplitudes of the stimuli were adjusted to maintain a constant SPL of 66 dB, measured at the headphones, for all rates and CFs. We did not compensate for differences in hearing thresholds at different CFs. Despite the filtering of the pulse trains, some artifacts like harmonic distortions or intermodulation at the basilar membrane can cause stimulation outside the desired frequency band. To prevent these artifacts from being heard, a binaurally uncorrelated white noise with a spectrum SPL of 9.2 dB, ranging from 50 Hz to 20 kHz, was continuously played throughout the testing.

1.3. Procedure

A two-interval, two-alternative forced-choice (2-AFC) procedure was used in lateralization discrimination test. The first interval contained a reference stimulus with zero ITD evoking a centralized auditory image. The second interval contained the target stimulus with the ITD tested. The subjects were requested to indicate whether the second stimulus was perceived to the left or to the right of the first one by pressing an appropriate button.

Each combination of CF and pulse rate was tested in a separate block. Each block contained 70 presentations of four ITD values in a randomized order with 35 targets leading to the left and 35 targets leading to the right. At least two blocks were completed for each condition and the order of blocks was randomized for each subject. Visual response feedback was provided after each trial. The chance rate was 50%. A score of 100% correct responses indicates that all stimuli were discriminated, with lateralization corresponding to the ear receiving the leading signal.

Stimuli with the pulse rate of 200 pps and ITD of 600 µs at CF1 were used to train subjects before the main test started. The subjects were trained until they showed a stable performance which was achieved within a few hours.

2. Results

Fig. 20 shows the percent correct (P_c) scores as function of the ITD with the pulse rate as parameter. Each panel shows data for one subject and one CF. The dashed horizontal lines show the threshold used for the JND estimation.

The comparison of the data between the three CFs (columns) suggests a difference in performance. Especially comparing CF3 and CF1 for subjects NH2, NH7, NH8, and NH9, the performance appears to decrease. The effect of the pulse rate seems to be more salient: the performance decreases with increasing pulse rate, independent of subject, CF, and ITD. It appears that with increasing pulse rate the differences between the CFs become larger, particularly for subject NH8.

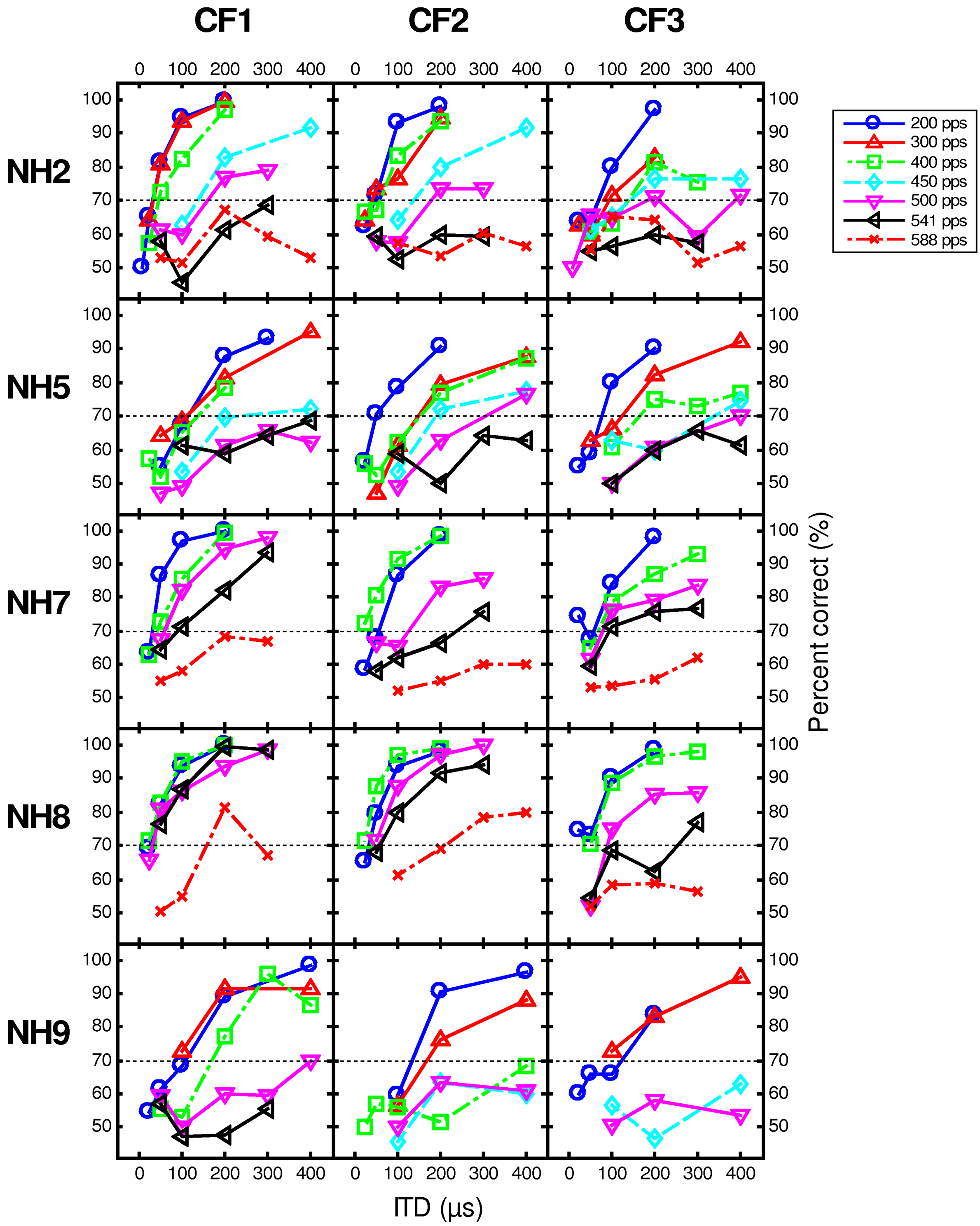

Figure 20: Performance in percent correct (P_c) as function of the ITD. The parameter is the pulse rate in pulses per second (pps). Each column shows one center frequency (CF) and each row shows results for one subject. The dashed lines show the threshold used for the JND estimation.

The data were statistically analyzed by calculating a multiway repeated-measures analysis of variance (RM ANOVA) with the factors CF, pulse rate, and ITD. To not violate the assumption of homogeneity of variance required for ANOVA, the P_c scores were transformed using the rationalized arcsine transform (Sherbecoe and Studebaker, 2004). The main effect of CF is highly significant ($p<0.0001$) with decreasing performance for increasing CF. The main effect of the pulse rate is highly significant ($p<0.0001$); the performance decreases with increasing pulse rate. As expected, the effect of ITD is highly significant ($p<0.0001$) showing increasing performance for increasing ITD values.

Due to the individual performance of the subjects, the analyzed data matrix is not full rank. This means that for some subjects and conditions there are more extreme P_c scores like 100% or 50% than for others. These extreme values, which represent the tails of the psychometric functions, have an impact on the analysis producing ceiling or floor effects. For example, the Tukey-Kramer post-hoc test revealed that the results for the pulse rates of 200, 300, and 400 pps are not significantly different ($p>0.05$), which indicates a possible ceiling effect for pulse rates lower than 400 pps. To reduce this problem, the RM ANOVA was performed again. However, this time we excluded all conditions where the subjects performed below 60% and above or equal to 90%. The exclusion of conditions with extreme performance resulted in a well populated condition matrix and the interactions could be included in the analysis. Hence, the main factors were CF, ITD, and pulse rate and the interactions were CF × ITD and CF × pulse rate. The interactions of pulse rate × ITD and pulse rate × ITD × CF were not included.

The results for the main effects did not change ($p=0.002$ for CF; $p<0.0001$ for the effects of ITD and pulse rate). Both interactions are not significant ($p=0.1188$ for CF × ITD; $p=0.4247$ for CF × pulse rate) showing the independence of the effects of CF, ITD, and pulse rate.

The significant effect of CF results from a decreasing performance with increasing CF (CF1: 72.7%, CF2: 71.5%, CF3: 69.2%). However the difference is small: only 3.6% between CF1 and CF3. For all rates, except for 588 pps, the Tukey-Kramer post-hoc tests revealed a decrease in performance between CF1 and CF3 ($p<0.05$). For example, at 200 pps, the performance is 82.1% (CF1), 78.7% (CF2), and 79.1% (CF3) resulting in a decrease in performance of 3% between CF1 and CF3. Only 588 pps showed no significant effect of CF ($p>0.05$; with

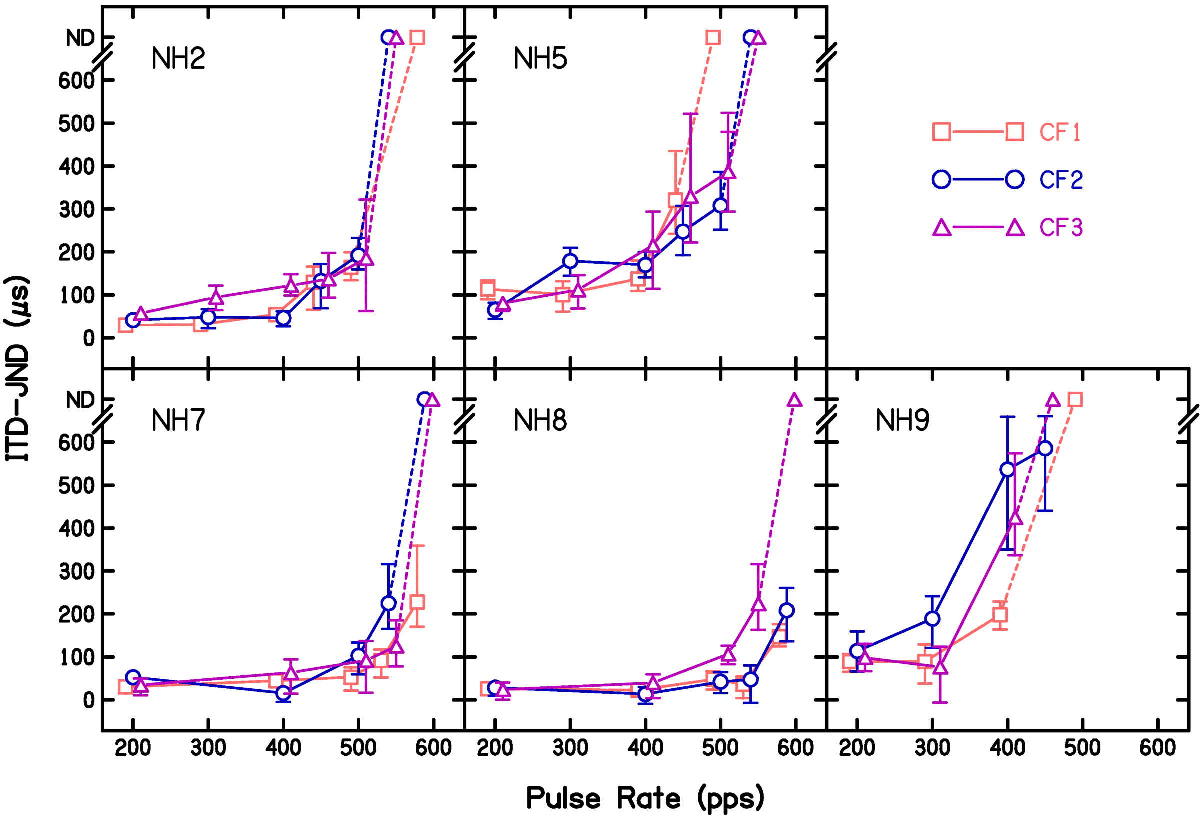

Figure 21: ITD-JNDs as function of the pulse rate in pulses per second (pps). The parameter is the center frequency (CF). The vertical bars indicate the 95% confidence intervals. The non-determinable JNDs are drawn at "ND" and are connected with dotted lines.

an *increase* in performance of 0.98% between CF1 and CF3). However, for this pulse rate the data represent results from three subjects only.

To compare the results to the literature the ITD-JNDs were calculated. Psychometric functions were estimated from the maximum-likelihood cumulative Gaussian fits to the raw data (Wichmann and Hill, 2001a)[10]. The ITD value yielding 70% of the P_c score along the psychometric function was defined as JND. The estimated JNDs are shown as functions of the pulse rate in Fig. 21, one panel per subject. The parameter is CF. In some cases, mostly for the higher pulse rates, the psychometric functions did not reach 70% and the JNDs could not be determined. These not determinable (ND) JNDs are plotted at "ND". The vertical bars show the

10 Using psignifit version 2.5.41 (see http://bootstrap-software.org/psignifit/), a software package for fitting psychometric functions to psychophysical data (Wichmann and Hill, 2001a).

95% confidence intervals of the JNDs and were determined applying a bootstrapping method (Wichmann and Hill, 2001b).

The idea of using JNDs for the statistical analysis has been rejected because of the problem with the treatment of the ND JNDs. Thus, the JNDs are used for the comparison with the literature. The issue of performing statistical tests on JNDs is discussed in the appendix of this chapter.

The JNDs are highly individual over subjects. It seems that subjects NH2, NH7, and NH8 are better performers than the others. So, it is not useful to present the data averaged across the subjects. Comparing the performance for 200 pps only, the average JNDs are 57.3 µs (CF1), 59.9 µs (CF2), and 58.7 µs (CF3). Thus, for this pulse rate, the CF appears to have no effect. By increasing the pulse rate, some differences in JNDs as a function of CF can be found, especially for subjects NH7, NH8, and NH9. However, no general trend can be revealed just by visual inspection. Unfortunately, many JNDs could not be determined for the higher pulse rates. The consequences of this problem for the statistical analysis based on JNDs are discussed in the Appendix. However, the estimated JNDs are used to facilitate the interpretation of the results and for a comparison to the literature in the next section.

3. Discussion

Bernstein and Trahiotis (2002) studied the sensitivity to ongoing ITDs using transposed tones. For 4 kHz and modulation rates of 128 and 256 Hz, they found JNDs of 79 and 100 µs, respectively. These conditions are comparable with our condition CF1 (4.6 kHz) and 200 pps, where we found an average JND of 57 µs. This indicates that bandpass-filtered pulse trains may evoke a higher sensitivity to ITD than transposed tones at this rate. One explanation may be the steeper onsets and larger modulation depth of the stimuli after passing the auditory filtering. For CF1 and 200 pps, we found in Laback *et al.* (2007) a JND[11] of approximately 200 µs. However, we used four-pulse stimuli, resulting in a stimulus duration of 20 ms. The integration effect of the ITD information for our much longer stimuli is the most likely reason for such a large difference. This is supported by findings of Hafter and Dye (1983), who

11 In that study, we calculated the JNDs for 65% threshold, which is different from our threshold of 70% used in this study. However, with the 70% threshold the difference in JNDs would become even larger.

showed that, for 200 pps, appending more and more pulses to the stimulus improves ITD sensitivity.

Our results show a small but significant decrease in ongoing envelope ITD sensitivity as CF increases. For the particular pulse rate of 200 pps, the JNDs are about 60 µs and show no effect of CF. This is contrary to the results of Bernstein and Trahiotis (2002), who found a considerable decrease in ITD sensitivity as CF increased for transposed tones; for the modulation rate of 128 Hz, the JNDs increased from 79 to 170 µs when CF was increased from 4 to 10 kHz. Their results for 128-Hz SAM tones showed an even larger effect of CF: the JNDs increased from approximately 150 µs (for 4 kHz) to 350 µs (for 10 kHz). The larger CF effect for SAM and transposed tones can be explained by comparing the envelopes of the auditory filter response to the stimuli: by increasing the CF, the envelope of SAM or transposed tone response has only a slight increase of the onset steepness because of the constant bandwidth. In contrast, the envelope of our bandpass-filtered click train response becomes steeper because of increasing bandwidth. This results in longer dead times. Additionally, our stimuli were designed to yield a constant stimulation range on the basilar membrane. By doing this we hoped to stimulate approximately constant amounts of neurons for different CFs. It seems that, stimuli with such a bandwidth almost compensated for the effect of CF observed with SAM and transposed tones. Interestingly, a constant number of stimulated neurons is assumed to be the case in electric stimulation, where the place (electrode) has no systematic effect on the width of excitation patterns (Cohen *et al.*, 2003).[12]

The results show a strong effect of pulse rate. The rate limit, which is the highest rate yielding a valid JND, is 500 pps (median). Bernstein and Trahiotis (2002) found a rate limit of 256 Hz for a CF less than 6 kHz, which decreased to 128 Hz at 10 kHz. The lower limit for 10 kHz probably results from constant stimulus bandwidth for all CFs for transposed tones. For CF1, we found in Laback *et al.* (2007) a limit of 200 pps, which is, again, lower than in this study. This is probably an effect of overall lower performance of the subjects resulting from using much shorter stimuli in Laback *et al.* (2007). In Majdak *et al.* (2006), we tested the subjects for pulse rates of 400, 600, 800, and 938 pps using stimuli of 300 ms duration, which is the same as in the current study. Assuming a threshold of 70%, the estimated highest rate

12 Contrary to this idea, in a recent study, Litovsky *et al.* (2005) showed an effect of place on ITD sensitivity in electric stimulation. However, the effect was not consistent across all subjects and the most plausible explanation seems to be the place-dependent survival of neurons.

yielding a JND was 400 pps in that study, thus, the rate limit is in the range between 400 and 600 pps. This is in agreement with the results of the present study. According to Majdak *et al.* (2006), the rate limitation in CI listeners varied between 100 and 800 pps, depending on the subject. Also, in Laback *et al.* (2007), we found similar rate limits in CI listeners, again varying between 100 and 800 pps across the subjects. The origin for the strong subject dependence is still unclear. Generally, the rate limit may be explained in terms of a binaural adaptation effect, which appears to affect the ITD sensitivity for higher pulse rates (for NH listeners: Hafter and Dye, 1983; for CI listeners: Laback and Majdak, 2007).

Let us return to the main question posited in this study: do we have an effect of auditory filtering? Considering CF as only parameter, an increase in ITD sensitivity with increasing CF would argue for an effect of auditory filtering. However, this can be overshadowed by the effects of changing bandwidths, changing absolute thresholds, or even differences in neural processing when the CF of the stimulus changes. Thus, a decreasing performance with increasing CF does not rule out a possible effect of auditory filtering. Fortunately, the impact of the auditory filtering can also be revealed by varying the pulse rate: by increasing the pulse rate the auditory filtering smears the timing information more and more and reduces the modulation depth. Thus, a significant interaction between pulse rate and CF would indicate that auditory filtering affects the ITD sensitivity. Indeed, the statistical analysis showed no interaction between pulse rate and CF suggesting no effect of auditory filtering[13]. However, this is not proof that auditory filtering does not contribute at all. It still could be the case that the test power was too small to reveal an effect. Nevertheless, the test power was strong enough to reveal a difference in P_c of 3.6% between CF1 and CF3 as a significant main effect. Thus, if the auditory filtering affects ITD sensitivity at all, then the effect is small and it appears not to be the limiting factor in the perception of ITDs.

Our results suggest that the comparisons between the electric stimulation and the acoustic CI simulation in Majdak *et al.* (2006) and Laback *et al.* (2007) were not overly affected by auditory filtering. Thus, the origin of the performance differences between the two subject groups is still unclear. One possible explanation may be the higher degree of phase locking in

13 An extreme impact of the auditory filtering would lead to a change in performance direction with increasing CF with respect to the rate, namely, a higher rate limit at higher CFs. Indeed, a rudimentary form of that effect could be found. For 200 pps, the performance significantly decreased by 3% when the CF was increased from CF1 to CF3. For 588 pps, the performance non-significantly increased by 0.98% for the same increase of the CF. However, this change in the performance direction is not significant.

electric hearing (Abbas, 1993). It seems to be reasonable considering the following across-study comparison. First, the results of the study by van den Honert and Stypulkowski (1987) indicate that the neural response to electric *pulse train* is better synchronized than the response to electric *sinusoidal* stimulation. Using such sinusoidal stimuli, Hartmann *et al.* (1984) demonstrated that the phase locking is stronger for *electric* stimulation than for *acoustic* stimulation. Subsequently, Dreyer and Delgutte (2006) showed that, for acoustic stimulation, the response to *sinusoidal stimuli* evokes in most cases a higher degree of phase locking than the response to *transposed tones*. Thus, assuming a similar degree of phase locking for transposed tones and bandpass-filtered pulse trains, pulsatile electric stimulation seems to yield stronger phase locking than acoustic stimulation with bandpass-filtered pulse trains. This may be a reason for the higher sensitivity to ongoing ITDs in CI listeners compared to NH subjects listening to the acoustic CI simulation.

4. Conclusions

This study showed a small but significant decrease in sensitivity to the ongoing *envelope* ITD when the center frequency is increased from 4.6 to 9.2 kHz. Also, the ITD sensitivity decreased with increasing rate showing an average limit at 500 pps. No interaction between the effects of the parameter pulse rate and center frequency could be found. This indicates that the effects of rate and center frequency are not caused by the ringing of the auditory filtering. It it conjectured that ITD sensitivity is limited by mechanisms at a more central stage of the auditory system. Furthermore, the results indicate that a constant stimulation range on the basilar membrane when the center frequency is increased reduces the center frequency effect and yields an almost constant ITD performance. Thus, it is concluded that the comparison between electric stimulation and acoustic CI simulation presented in Chapters II and III (Majdak *et al.,* 2006 and Laback *et al.,* 2007) was not importantly affected by the effects of the auditory filtering. The difference in the performance between both subject groups in those previous studies is hypothesized to be associated with differences in the degree of phase locking.

5. Appendix

In some cases, mostly for the higher pulse rates, the psychometric functions did not reach 70% for all tested ITD values and the JNDs could not be determined. This appendix shows the problematic of including conditions with non-determinable JNDs (ND JND) in an RM ANOVA.

The RM ANOVA is usually performed on all data. However, conditions with ND JNDs can not be included in the analysis. In the ANOVA model, those conditions are treated as “no information given”, leading to a condition matrix not of full rank. This results in a lower test power of the ANOVA. To compensate for this problem, the ND JND conditions can be replaced by a hypothetical JND value, forcing the ANOVA to treat these conditions as if they were measured, even though no sensitivity was found. Thus, the ND JND replacement should be of a sufficiently high magnitude.

We analyzed the JND data using the RM ANOVA under different conditions. In the first ANOVA we excluded the ND JND conditions, taking into account having a condition matrix not of full rank. In the other ANOVAs, we included the ND JND conditions and each ANOVA was performed with a different ND JND replacement. In the second ANOVA, the ND JNDs were set to 630 µs, which corresponds to the natural ITD for lateral sounds for an average subject (Blauert, 1997). In the third ANOVA, the ND JNDs were set to ¼ IPI, assuming that this yields maximal sensitivity for the ongoing ITD cue (Majdak *et al.*, 2006). In the forth ANOVA, the ND JNDs were set to the highest JND derived for each subject and CF. These values ranged from 125 to 585 µs. In the fifth and last ANOVA, the ND JNDs were set to 10 ms (an approximation of infinity, which could not be used due to technical limitations). All ANOVA results are presented in Table 8. The effect of CF is just significant and not significant ($0.0373 \leq p \leq 0.2327$), depending on the choice for the ND JND replacement. The main effect of the pulse rate is always significant ($p<0.0001$) and the interaction between CF and pulse rate is not significant for all analyses ($p>0.2776$). The comparison of different ND JND replacements clearly shows that inclusion of the ND JND conditions in the ANOVA is problematic.

ND JND	p (CF)	p (pulse rate)	p (CF x pulse rate)
excluded	0.0859	<0.0001	0.9826
630 µs	0.0414	<0.0001	0.5045
¼ IPI	0.0373	<0.0001	0.805
max(JND)	0.0545	<0.0001	0.9722
10 ms	0.2327	<0.0001	0.2776

Table 8: Results from separate ANOVAs applying different assumptions for ND JND replacements. Significant effects are shown in bold.

The most consequent choice is to exclude ND JND conditions from the analysis. However, this leads to a low test power for at least two reasons. First, for the higher pulse rates, only a small amount of data is available for all subjects. This is because of the highly individual results. Thus, the condition matrix is rather sparse. Unfortunately, this is the range of pulse rates where we expect an effect of CF and needs to be well represented by the data. Second, at lower pulse rates, although the data for all subjects are available, the test power of the ANOVA is reduced using the concept of the JND. Performing the ANOVA on JNDs implies comparing only one sample of data per condition and subject. This reduces the test power compared to the analysis based on P_c scores: they are represented by three to five samples per condition and subject.

Furthermore, the concept of the sigmoidal psychometric function is questionable in our study. This model requires a stimulus variable which, under certain conditions, can always be detected (a small observer lapse is allowed; Wichmann and Hill, 2001a). Said another way, the function range should be between 50% and 100% in our discrimination task. This is not always the case: the performance in the ND JND conditions never reached 70%. Furthermore, for conditions with determinable JNDs, it can not be ensured that increasing ITD would lead to $P_c = 100\%$ (e.g. NH8, CF1, 588 pps). For example, in Majdak *et al.* (2006), we showed more general results for ongoing ITD, which were above 70%-threshold but never reached 100% (Fig. 18 in Chapter III or Fig. 7 in Majdak *et al.*, 2006; left upper panel). These results show that sigmoidal psychometric functions are not an appropriate model for the statistical analysis of such an effect. Thus, in the absence of an adequate model for our JND data, only the statistical analyses for the P_c data are appropriate for discussion. Nevertheless, the estimated JNDs remain important to facilitate the interpretation of the results and for a comparison to the literature.

V. Effect of binaural jitter on ITD sensitivity at high pulse rates

Based on *Proceedings of the National Academy of Sciences U S A* **105**, 814-817 **(2008)**.

Applied for patent in EU and USA.

This chapter presents a method to improve sensitivity to fine structure ITD at higher pulse rates in electric hearing. This method is based on the findings presented in Chapters II and III where we demonstrated that bilateral CI listeners are sensitive to fine structure ITD. However, as already presented, their sensitivity degrades dramatically at a pulse rate of a few hundred pps. This finding is also supported by van Hoesel and Tyler (2003) and van Hoesel (2007). In contrast, NH listeners are able to detect ITD in the fine structure of pure tones up to much higher frequencies (Zwislocki and Feldman, 1956). The limitation in electric hearing appears to be disadvantageous, especially with respect to speech coding where high rates are required. Thus, the goal was to develop a method which eludes the rate limitation and improves the ITD sensitivity at such high rates. As already stated in Chapter IV, for reasons of simplicity, we refer to the published articles Laback *et al.* (2007) and Majdak *et al.* (2006) and not to Chapters II and III, respectively.

For modulated high-frequency stimuli, investigations with NH listeners showed that the sensitivity to ITD information degrades with increasing modulation rate (Hafter and Dye, 1983; Bernstein and Trahiotis, 2002). Hafter and Dye (1983) used filtered pulse trains in investigations of the ITD sensitivity as a function of pulse rate and stimulus duration. They compared the ITD sensitivity to results obtained from a model. The model was an optimum integration of ITD information across time and was described in Section 2.3 of Chapter I in this thesis. They showed that as the pulse rate increases, increasing the stimulus duration results in a smaller improvement of ITD sensitivity they predicted by the model. The decline of ITD perception improvement at higher rates has been referred to as *binaural adaptation*. It has a very strong effect on ITD perception: at higher pulse rates, the onset of a sound receives maximum perceptual weight (Saberi, 1996; Stecker and Hafter, 2002). As a consequence, the ongoing signal contributes only little to ITD perception. Further studies (Hafter and Buell, 1990; Stecker and Hafter, 2002), showed that binaural system can be restarted by introducing a change in the ongoing signal. Then, the portion of the signal following the trigger becomes more important and this results in improved ITD sensitivity. The restarting of the adaptation process has been called the *recovery from binaural adaptation* (RBA).

Based on the results of these studies, we assumed that the decreasing ITD sensitivity with increasing pulse rate in CI listeners may be a form of binaural adaptation. Thus, an appropriate way of introducing a trigger in the signal would cause an RBA and improve the ITD sensitivi-

Subject	Etiology	Age (yr)	Age at implantation (yr)		Duration of deafness		Binaural electrical stimulation experience	Test electrodes
			L	R	L	R		L / R
CI3	Meningitis	24	21	21	2 mo	2 mo	3 yr	4 / 3
CI8	Osteogenesis imperfekta	44	41	39	3 yr	12 yr	3 yr	7 / 5
CI10	Sudden hearing loss	54	44	48	43 yr	43 yr	6 yr	7 / 8
CI11	Temporal bone fracture	28	22	22	2 yr	2 yr	6 yr	2 / 3
CI12	Sudden hearing loss	40	35	34	8 yr	3 yr	5 yr	2 / 2

Table 9: Subjects' data. The rightmost column shows the electrodes on the left and right side evoking equal pitch perception. The electrodes are numbered from apex to base. Data reprinted from Laback et al. (2008).

ty at higher pulse rates. In contrast to Hafter and Buell (1990), who used acoustic stimulation, we used electric stimulation in CI listeners. Thus, by testing at a single interaural electrode pair, we were able to change the temporal and not the spectral properties of the stimulus. We attempted to trigger on every pulse by randomly varying (jittering) the inter-pulse interval (IPI). The purpose was to integrate the trigger effect from each pulse and thus multiply the RBA effects. An important requirement was that the ITD information in the fine structure had to be preserved. Thus, the jitter was synchronized between the two ears and is referred to as binaurally-synchronized jitter or shortly *binaural jitter*. In this study, the effect of binaural jitter on ITD-based lateralization discrimination was tested in CI listeners at different pulse rates and with different amounts of jitter.

1. Methods

1.1. Subjects

The subjects were five bilaterally implanted listeners, all wearing the Combi 40+ implant (MED-EL, Austria). They were postlingually deafened, had high speech recognition scores with the help of implants, and had more than three years of binaural stimulation experience at the time of the tests. Their data are presented in Tab. 9.

1.2. Stimuli

We used 300-ms pulse trains of biphasic electric pulses with trapezoidal amplitude modulation (see Fig. 22a). This kind of signal is an approximation of the characteristics of real-world

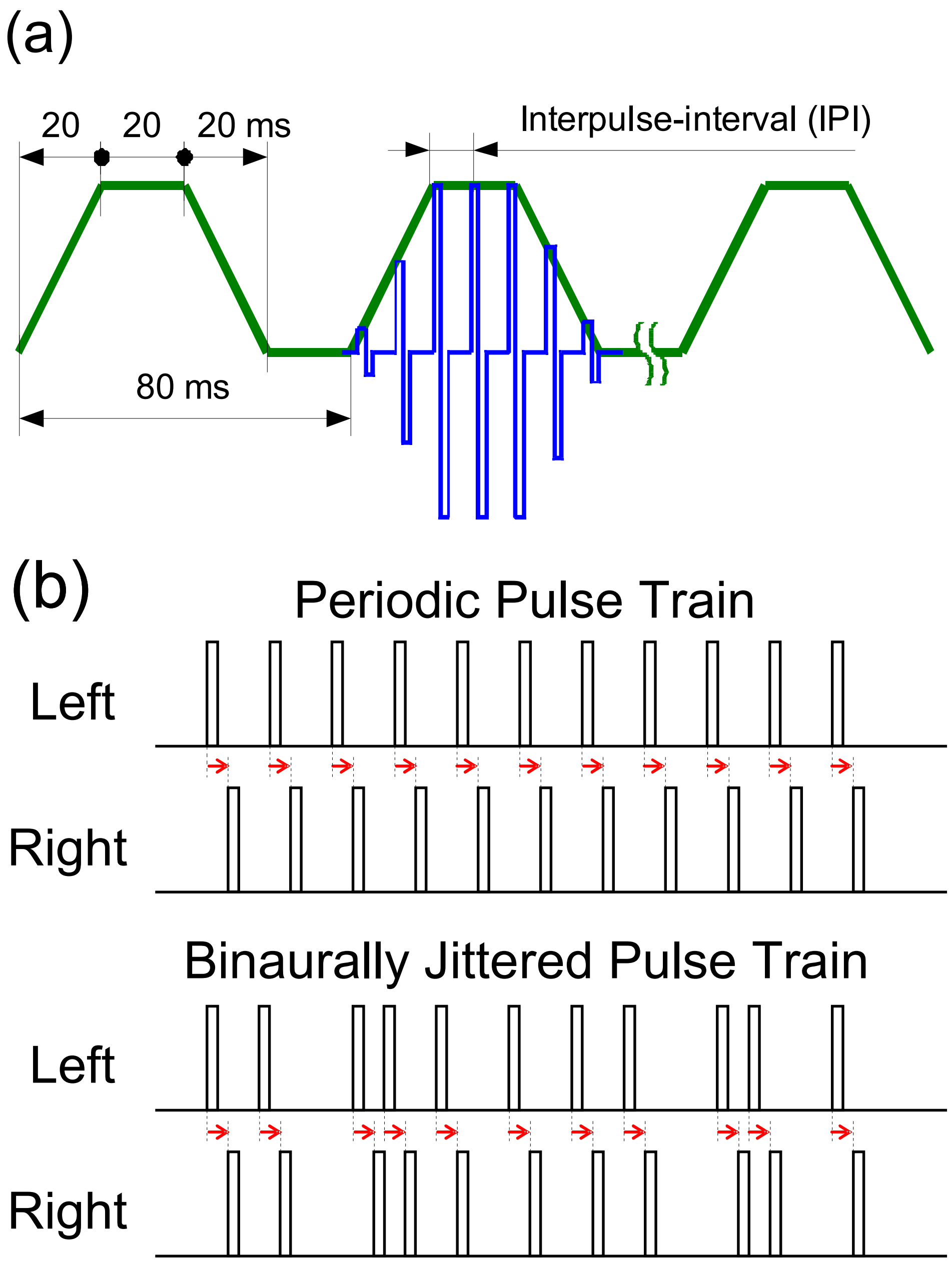

Figure 22: Schematics of the stimuli presented to the subjects. (a) Details on the envelope and fine structure of the stimulus. For clarity, only three of the four trapezoids are shown and the fine structure characteristic is shown in one trapezoid only. (b) Stationary portion of a periodic pulse train (upper) and of a binaurally-jittered pulse train (lower). For clarity, in this part of the figure only the positive phase is shown. Notice the preserved ITD despite of jitter (marked with arrows of constant length). Figure from Laback and Majdak. (2008).

signals like speech and it was already applied in one of our previous studies (Majdak *et al.*, 2006). The stimuli were presented binaurally and had an ITD in the entire waveform. Inclusion of the ITD in the envelope is motivated by the purpose of this study: the main focus is on jitter effects and, thus, any improvement in ITD sensitivity due to the jitter must occur despite the availability of envelope ITD information. Furthermore, the stimuli with waveform ITD promise a higher overall performance, which is more motivating for the subjects than testing near the chance rate.

The periodic pulse trains had a constant IPI, which is the nominal IPI (upper part of Fig. 22b). Jittered pulse trains had randomly varied IPIs (lower part of Fig. 22b), however, the IPI values were chosen so that the average value over the stimulus duration exactly represented the nominal IPI. The jitter was synchronized between the two ears, which was required to preserve the ITD in the fine structure (notice the constant length of the arrows in Fig. 22b). The distribution of the IPI irregularity was rectangular. The distribution width is defined relative to the nominal IPI and is represented by the parameter *k*. Thus, the parameter *k* represents the amount of applied jitter and ranges from 0 (periodic condition, no jitter) to 1 (maximum jitter). The pulse trains were "constructed" pulse by pulse, e.g. for each pulse added, its temporal position was varied within the interval [IPI·(1-k), IPI·(1+k)]. Hence, for $k = 1$, the IPI would vary between zero and 2·IPI. Each stimulus repetition had a new jitter manifestation. Because of technical limitations a k of 1 can not be achieved and a maximum value of $k = 0.9$ was used in our tests.

The stimuli were presented at an interaural electrode pair, which was chosen to evoke equal pitch on both sides. The pretest procedure was identical to the procedure presented in the previous studies Majdak *et al.*(2006) and Laback *et al* (2007). The electrode pairs are specified in the last column of Tab. 9. For each pulse rate, currents were determined to evoke a centralized auditory image at a comfortable level. The details about methods to determine comfortable levels were, again, described in the previous studies. Also, the apparatus was the same: the stimuli were created with the help of a computer and directly transmitted to the implants via an interaurally synchronized research interface.

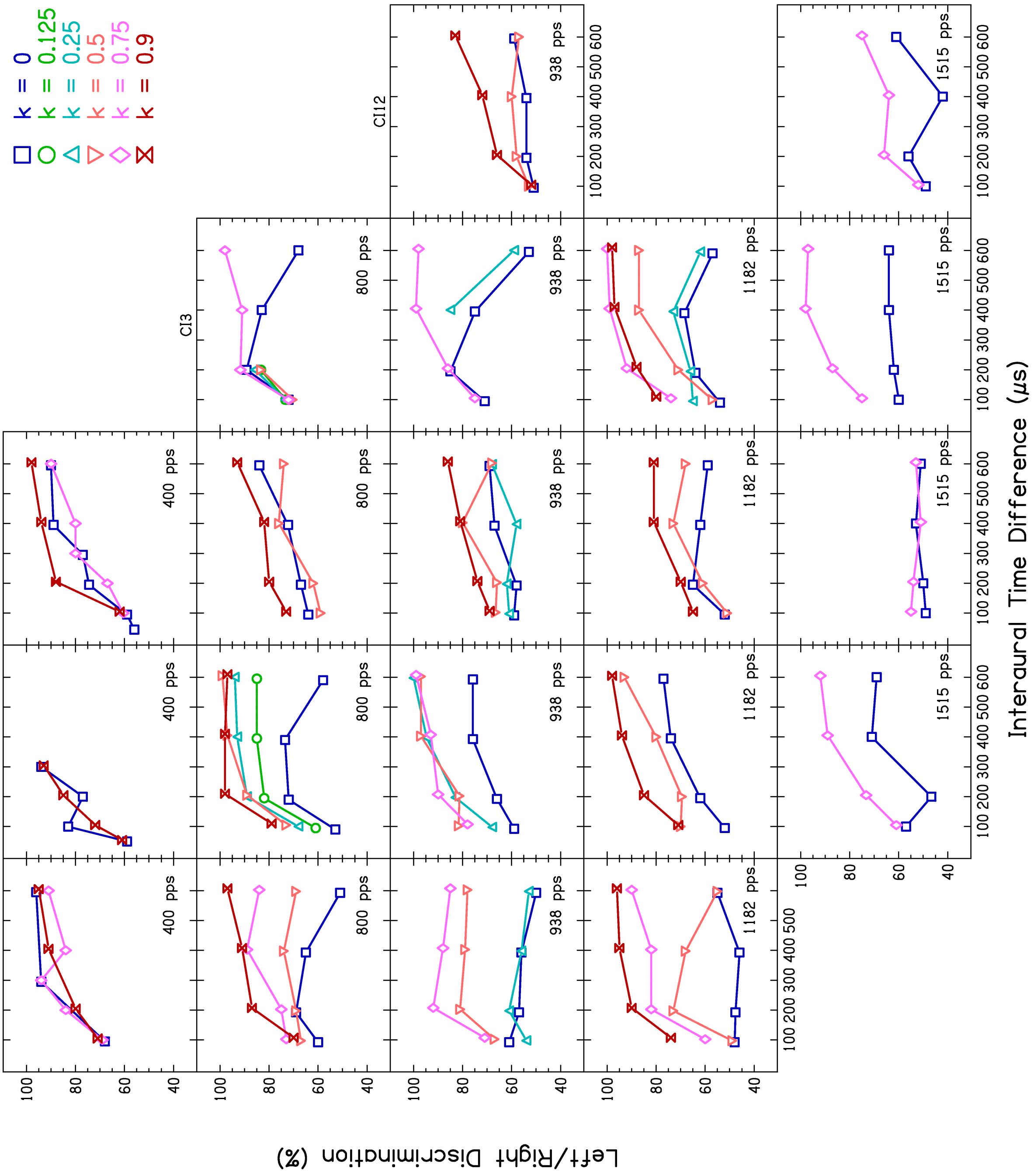

Figure 23: Percent correct (P_c) left/right discrimination scores as a function of the ITD. The parameter is amount of binaural jitter (k). The results for the individual subjects are presented along the rows and the results for the different pulse rates are presented along the columns. Figure from Laback and Majdak (2008).

1.3. Procedure

Lateralization discrimination was tested in a left/right discrimination task. The subject had to compare a target stimulus containing ITD with a preceding reference stimulus with zero ITD. The reference stimulus was always a periodic pulse train (k = 0). The independent variables were *k* (0, 0.125, 0.25, 0.5, 0.75, and 0.9), the ITD (100, 200, 400, and 600 µs), and the pulse rate (400, 800, 938, 1182, and 1515 pps). The conditions corresponded to combinations of the independent variables and were presented in a balanced order. Because of limited availability of the subjects, only the most interesting combinations of *k* and pulse rate were tested for each subject. The presentation of each condition was repeated 100 times. Visual response feedback was provided after each trial. To reduce potential learning effects, the subjects were trained for a few hours before starting the main experiment. For each listener, the distribution of the responses showed sufficient symmetry, thus, the percent correct scores did not require any adjustments with respect to removing response bias.

1.4. Statistical Analysis

The statistical analysis was performed using repeated measures analyses of variance (RM ANOVA). The effects of the parameters *k* and pulse rate were tested on the transformed percentage of correct left/right discrimination (P_c). The P_c scores were transformed using the rationalized arcsine transform to not violate the assumption of homogeneity of variance required for ANOVA (Sherbecoe and Studebaker, 2004). To compare the factor levels of the parameter *k*, Tukey's post hoc tests were applied to the results of the ANOVA.

2. Results

2.1. Individual results

The complete set of results for the individual listeners is presented in Fig. 23. The individual panels present percent correct left/right discrimination scores (P_c) as a function of the ITD. The parameter is the amount of binaural jitter *k*. The results for the different pulse rates are presented along the columns. Comparing across subjects, for each pulse rate and *k*, the results appear to be sufficiently homogeneous, which allows averaging over subjects. Thus, in the following, the results are presented as averaged data only.

Furthermore, the effect of jitter appears to be monotonous: the performance increases with increasing the parameter *k*. Thus, for the purpose of simplicity, the results for the parameter *k* were pooled into large jitter (k = 0.75 and 0.9), small jitter (k = 0.125, 0.25, and 0.5), and no jitter ($k = 0$).

2.2. Sensitivity as a function of pulse rate

Fig. 24 shows P_c averaged over subjects as a function of the pulse rate. The results for different ITDs are presented in separate panels. The effect of jitter is very similar for the different ITD values, despite shifts in overall performance (see the results for ITD = 100 and 200 µs). Thus, as we are interested in the effect of jitter and not ITD, the analysis was further simplified by averaging the data across the ITD values 200, 400, and 600 µs. Fig. 25 shows the summarized results. At the lowest pulse rate tested (400 pps), P_c is at a high level and does not differ significantly between the conditions with and without jitter ($p = 0.98$). At the higher pulse rates (> 400 pps), there is a large and significant difference between the results for the conditions with and without jitter ($p < 0.001$). For the condition without jitter, P_c decreases dramatically with increasing pulse rate ($p < 0.001$) and approaches chance performance at approximately 938 pps. In contrast, for the conditions with binaural jitter, the performance remains constantly high up to 1182 pps and declines at 1515 pps. Even at the highest pulse rate tested (1515 pps), the performance difference between the conditions with and without jitter is highly significant ($p < 0.001$). For small jitter, the improvements with respect to the no-jitter condition are smaller, however, the differences are still significant ($p < 0.001$).

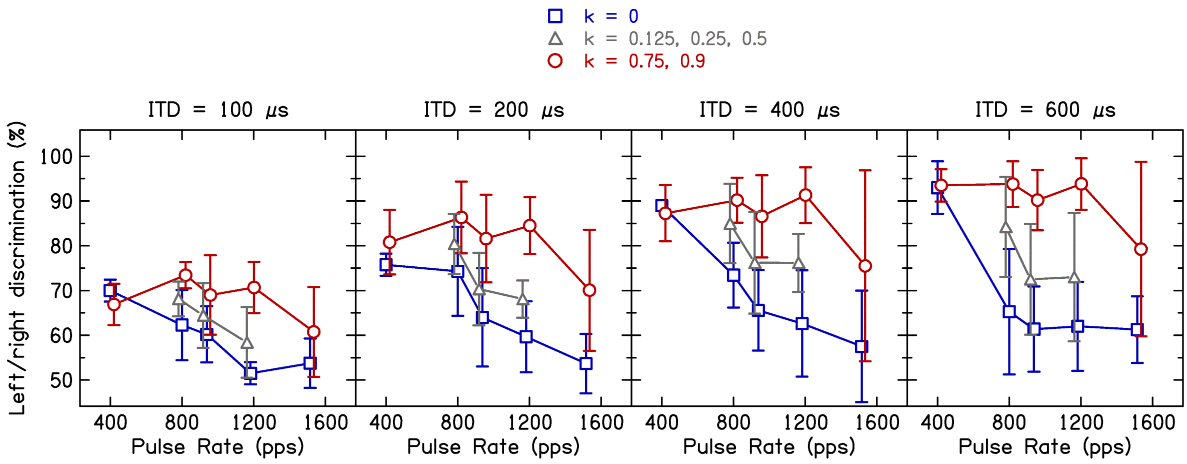

Figure 24: Percent correct scores for left/right discrimination as a function of the pulse rate, averaged over the five subjects. The results for different ITD values are presented in separate panels. The condition without jitter (k = 0) is depicted by the blue squares, the condition with small jitter (k = 0.125, 0.25, and 0.5) is depicted by the grey triangles, and the condition with large jitter (k = 0.75 and 0.9) is depicted by the red circles. The error bars represent 95% confidence intervals. Figure from Laback and Majdak (2008).

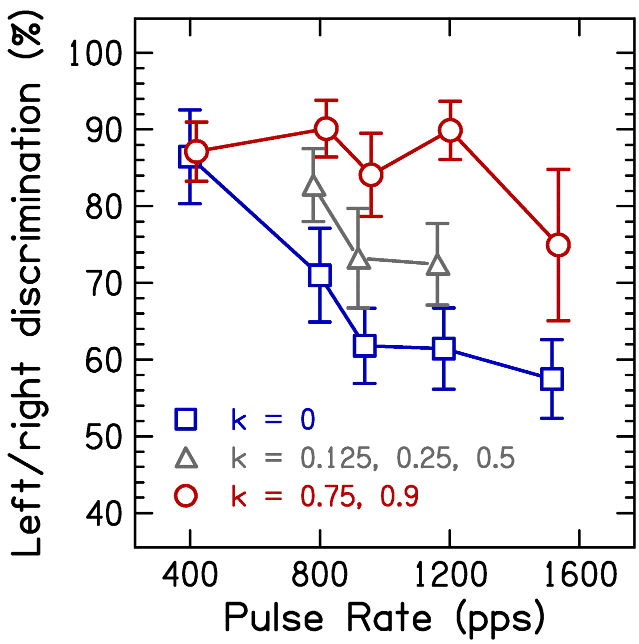

Figure 25: Percent correct scores for left/right discrimination as a function of the pulse rate. The data are averaged over the ITD (values 200, 400, and 600 µs). The condition without jitter (k = 0) is depicted by the blue squares, the condition with small jitter (k = 0.125, 0.25, and 0.5) is depicted by the grey triangles, and the condition with large jitter (k = 0.75 and 0.9) is depicted by the red circles. The error bars represent 95% confidence intervals. Figure from Laback and Majdak (2008).

2.3. Sensitivity as a function of ITD

In the previous section, the data were presented as a function of the pulse rate to reveal the effects of the pulse rate. In this section, we present the results as a function of ITD to analyze the effect of binaural jitter with respect to the ITD. Thus, the data were replotted in Fig. 26 showing the effect of binaural jitter as a function of the ITD. Each panel shows the results for a different pulse rate. For pulse rates > 400 pps, the condition without jitter shows low performance and the conditions with jitter show a monotonous characteristics. Thus, the results were averaged over pulse rates > 400 pps and are presented in Fig. 27. For the condition without jitter, P_c is low at all ITD values. For the conditions with binaural jitter, the performance increases monotonically with the ITD and is significantly different from the no-jitter condition even at the smallest ITD (100 µs; $p < 0.001$ for both large and small jitter). The improvements reach a maximum of 28% at 600 µs for large jitter ($p < 0.001$) and 14% at 400 µs for small jitter ($p < 0.001$). As shown in Majdak *et al.* (2006), ITDs that approach or exceed half of the IPI contain ambiguous ongoing fine structure ITD cues. The results presented in Fig. 27 indicate that binaural jitter resolves that ambiguity. For example, the ITD of 400 µs contains ambiguous cues at all pulse rates from 800 to 1515 pps as it is within ¼ to ¾ of the IPI. However, even in this case, the binaural jitter improves the performance significantly ($p < 0.001$ for both large and small jitter).

3. Discussion

Consistent with previous studies (Majdak *et al.*, 2006; Laback *et al.*, 2007; van Hoesel and Tyler, 2003; van Hoesel, 2007), the ITD sensitivity declines with increasing pulse rate for the periodic condition without jitter. At 400 pps, almost no improvement of binaural jitter was observed. At pulse rates ≥ 800 pps, the performance was severely reduced. This is consistent with the expectations that the performance is less affected by rate limitation mechanisms at 400 pps compared to higher pulse rates. As expected, ongoing envelope ITD appears to have contributed little to ITD sensitivity, which is supported by the low performance in the no-jitter conditions.

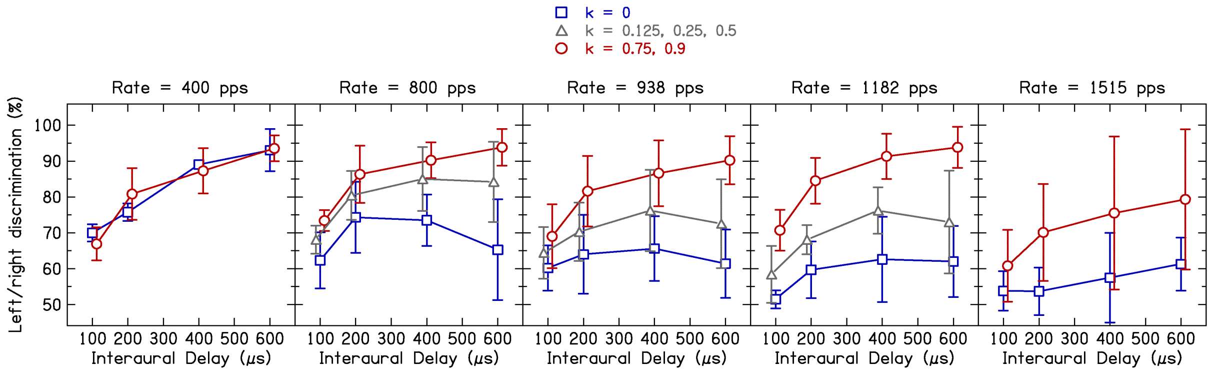

Figure 26: Percent correct scores for left/right discrimination as a function of ITD. The results for different pulse rates are presented in separate panels. The condition without jitter (k = 0) is depicted by the blue squares, the condition with small jitter (k = 0.125, 0.25, and 0.5) is depicted by the grey triangles, and the condition with large jitter (k = 0.75 and 0.9) is depicted by the red circles. The error bars represent 95% confidence intervals. Figure from Laback and Majdak (2008).

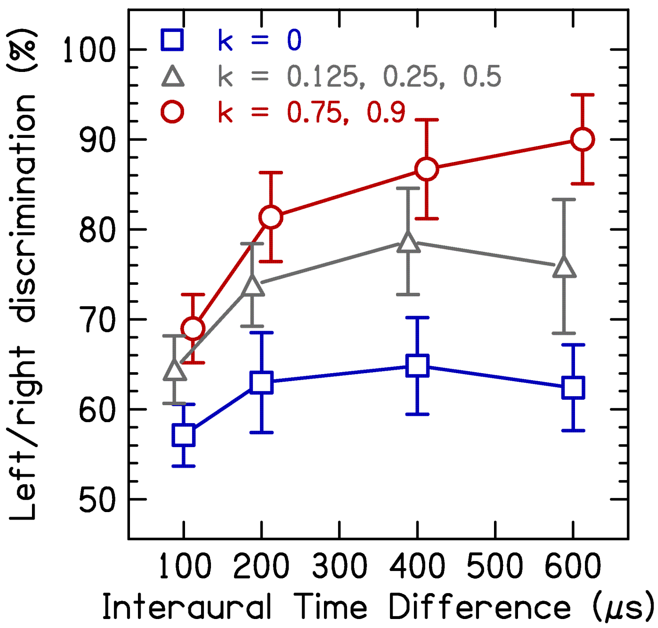

Figure 27: Percent correct scores for left/right discrimination as a function of ITD. The data are averaged over the pulse rates 800, 938, 1182, and 1515 pps. The condition without jitter (k = 0) is depicted by the blue squares, the condition with small jitter (k = 0.125, 0.25, and 0.5) is depicted by the grey triangles, and the condition with large jitter (k = 0.75 and 0.9) is depicted by the red circles. The error bars represent 95% confidence intervals. Figure from Laback and Majdak (2008).

The results clearly show that binaurally-jittered stimulation improves CI listeners' sensitivity to fine structure ITD at rates where NH listeners show sensitivity to ITD in pure tones (Zwislocki and Feldman, 1956; Klump and Eady, 1956). Thus, binaural jitter appears to resolve the discrepancy in the rate limitation between CI and NH listeners. Nevertheless, notice that the absolute performance of the CI listeners is considerably lower.

The large improvements of binaural jitter at higher pulse rates (≥ 800 pps), but not at lower rates, is consistent with the hypothesis that an excessive form of binaural adaptation limits fine structure ITD sensitivity at higher pulse rates. Thus, it seems that introducing ongoing temporal changes in the stimulus causes an RBA effect. Because of using a single binaural electrode pair in the CI listeners, the RBA in our experiments is based on temporal trigger only. This finding extends the conclusion of Hafter and Buell (1990) on the RBA effect in acoustic hearing. They investigated the RBA effect by inserting a trigger, which was a change of the IPI or a brief sound, into a pulse train. They attributed the RBA effect to the spectral changes induced by the trigger. Our results with electric stimulation show that a recovery is possible without spectral changes.

Why are the CI listeners subject to such excessive form of the binaural adaptation effect? Possible reasons are the high degree of phase locking and across-fiber synchrony in the neural response to electric stimulation (Abbas, 1993; Dynes and Delgutte, 1992; Hartmann *et al.*, 1984; Litvak *et al.*, 2001; Wilson *et al.*, 1997). Introducing artificial randomness into the stimulus may reduce the amount of periodicity in the neural response and consequently reduce binaural adaptation. It seems like the binaural jitter holds the binaural system “awake” over the duration of the stimulus and thus improves access to the ITD information. On the neural level, the effect of binaural jitter may be interpreted in terms of a generally better neural representation of temporal information. Both neural models and experimental results suggest that restoring stochastic responses in electric stimulation enhances the neural representation of stimulus timing (Rubinstein *et al.*, 1999; Litvak *et al.*, 2003). Thus, even a monaural jitter in the stimulus may improve rate pitch perception in electric hearing. Zeng (2002) showed that the rate pitch perception is limited to pulse rates up to about 300 pps. Few years ago, Chen *et al.* studied the effect of jitter on monaural pitch discrimination in three CI listeners (Chen *et al.*, 2005). They found no effect on pitch discrimination besides a deterioration at low pulse rates. However, they tested only small amounts of jitter. They did not test larger amounts of jitter for which we observed the largest improvements in ITD sensitivity. On the other hand, such large amounts of jitter would probably smear the pitch cue, counteracting the potential benefit of jitter. Thus, the findings of our study are strictly limited to binaural phenomena and there is no indication of improvements in the monaural temporal information processing by using jittered stimuli.

Our results show that binaural jitter improves the performance even for ITDs which approach or even exceed half of the IPI. Such combinations of ITD and pulse rate produce ambiguous ongoing fine structure ITD cues (Majdak *et al.*, 2006). Our result can be explained by a so-called multiple looks model (Viemeister and Wakefield, 1991). According to this model, the auditory system stores looks of the signal in memory and processes them selectively. In our case, the auditory system resolves the ambiguity in ongoing fine structure ITD by picking out interaural pulse pairs with a large IPI to adjacent pairs. The interaural pulse pairs with a small IPI to adjacent pairs contain an ambiguous ITD cue. Thus, they do not contribute in the process of determining the lateral stimulus position and are just ignored by the auditory system.

The findings of our study have important implications for new stimulation strategies for bilateral CIs. Nowadays, the most popular strategies use periodic or near periodic stimulation, which limits the perception of fine structure ITD to a few hundred pps. This limitation can be removed by introducing binaurally-synchronized irregularities to the signal. As a consequence, fine structure ITD information can be transmitted at pulse rates, which are important for the coding of speech signals (Wilson *et al.*, 1991) and improve the understanding of speech in noise and localization of sound sources in CI listeners.

VI. General conclusions

The important thing is not to stop questioning.

Albert Einstein (1879 – 1955)

The purpose of this thesis was to investigate the sensitivity of CI listeners to ITD. This topic is of special interest because the ITD is an important cue for localization of sounds (e.g. Macpherson and Middlebrooks, 2002), for perceptual segregation of speech sounds (Drennan *et al.*, 2003), and for binaural unmasking of speech in noise (e.g. Bronkhorst and Plomp, 1988; Hawley *et al.*, 1999).

The first study (**Chapter II**) investigated listeners' sensitivity to ITD when taking into account the separation of ITD information transmitted in different parts of a stimulus. Four bilateral CI listeners' ability to left/right discriminate on the basis of ITD was tested as a function of pulse rate. Four-pulse sequences were used to separate the onset, the offset, and the ongoing ITD cues. The two middle pulses transmitted the ongoing ITD information. The first (onset) and last (offset) pulses transmitted the gating ITD information. The amounts of the gating and ongoing ITD information were varied independently, which allowed to separate their contribution to ITD sensitivity. The main finding of this study was that CI listeners were sensitive to ITD in the ongoing portion of signals, which implies that the CI listeners are sensitive to fine-structure ITD. However, the four-pulse sequences used in this study are relatively short, especially at higher pulse rates. In everyday situations, sustained interaural timing cues are likely to continue for more than just two electrical pulses and thus the weighting of different types of ITD cues may differ in these situations. Therefore, a second study was performed to test ITD sensitivity under more realistic conditions.

The second study (**Chapter III**) used 300-ms pulse trains with trapezoidal-shaped envelopes. The duration of these stimuli was more realistic and their envelope corresponded to the slow modulations in a speech signal. In this study, the ITD was independently varied in the envelope and fine structure. The sensitivity to these cues was tested with four CI listeners. The results show that CI listeners are able to lateralize stimuli on the basis of ITD in the fine structure, some of them for pulse rates as high as 800 pps. This indicates that the lateralization of sounds is affected if unsynchronized speech processors are used with pulse rates up to 800 pps. Therefore, to present all of the ITD information to CI listeners binaural synchronization would be required. A new synchronization method called the *diminished waveform delay* was introduced. Results from test conditions with diminished waveform delay showed improved performance in lateralization discrimination for rates between 400 and 800 pps.

In the same study, NH listeners were tested as well. They were presented with an acoustic simulation, which uses bandpass-filtered click trains with a center frequency of 4.6 kHz. In agreement with the results of the first study, the NH subjects showed a lower sensitivity to ITD in the fine structure than the CI subjects, as indicated by a lower rate limit. Possible explanations for the differences in the rate limit between the two subject groups were discussed: 1) smearing of the temporal information by auditory filtering in the cochlea of NH listeners; 2) a higher degree of phase locking due to the bypassing of the synaptic mechanism at the hair cell in electric hearing.

The first explanation, the effect of auditory filtering, was investigated in the third study (**Chapter IV**). The hypothesis was that if the acoustic simulations were performed at higher center frequencies, then the auditory filtering would affect the temporal structure of the stimulus less and this would result in a higher ITD sensitivity, particularly at higher pulse rates. Therefore, the effects of the center frequency and pulse rate on ongoing ITD sensitivity were tested in acoustic hearing. The results indicate that auditory filtering does not limit ongoing ITD perception and suggest the existence of another limiting mechanism, such as phase locking. Furthermore, it was concluded that the performance of NH listeners in the previous two studies was not underestimated by the choice of the center frequency.

Generally, these findings show a direction for future stimulation strategies, which take better advantage of the ITD information in the fine structure and envelope. A bilateral coordination of the speech processors is the first requirement in the implementation of a binaural stimulation strategy. The synchronization of fine structure and the implementation of the diminished waveform delay synchronization could then help to more-fully transmit the ITD information. However, even with such methods, there is a potential limitation: the sensitivity to ongoing ITD decreases rapidly for rates above a few hundred pps. Clinical stimulation strategies like CIS, ACE, and "*n*-of-*m*" usually use rates higher than 1000 pps. Therefore, even with binaural synchronization, fine-structure ITD would have only a small impact on lateralization if a standard clinical strategy were used.

To circumvent this problem, methods to improve the ITD sensitivity at higher pulse rates in CI listeners are required. One promising method was investigated in the last study (**Chapter V**) of this thesis. The hypothesis was that ITD sensitivity at higher pulse rates is limited by the effect of binaural adaptation. This effect is well-known from the NH literature and it was showed that a recovery from binaural adaptation can be achieved by using a trigger in the ongoing part of a signal. In electric hearing, signals are usually presented as regular pulse trains. This may also cause binaural adaptation effect, possibly in an even stronger form than in acoustic hearing due to the higher degree of phase locking. Thus, *binaurally-synchronized jitter* in the stimulation timing of the pulses was applied with the hope of achieving a recovery from binaural adaptation. ITD sensitivity to jittered stimuli was tested for pulse rates up to 1515 pps in five CI listeners. Results showed large improvements in ITD sensitivity for the jittered stimuli compared to the regular pulse trains. This supports the hypothesis that CI listeners are suffering from binaural adaptation. A purely-temporal trigger was able to evoke a recovery from binaural adaptation. Therefore, the results of this study indicate that the method of binaurally-synchronized jitter may improve lateralization of sounds in future stimulation strategies using higher pulse rates.

As of 2008, only two stimulation strategies considering the fine-structure of the acoustic signals have been presented.

In the fine-structure processing (FSP) strategy, two to three of the most apical electrodes stimulate at the zero-crossings of the bandpass-filtered signals (Hochmair *et al.*, 2006). Therefore, for these electrodes, the FSP strategy results in a kind of binaural synchronization of the fine structure when two speech processors are used. The other electrodes are driven according to the CIS strategy. The FSP-driven electrodes stimulate at a low pulse rate derived from the signal timing. However, the exact timing of the pulses underlies some restrictions, which may limit the transmission of precise ITD information to the auditory system.

Additionally, van Hoesel and Tyler (2003) presented a binaural stimulation strategy, which considers the fine-structure of the acoustic signal. Similar to the FSP strategy, the peak derived timing (PDT) strategy derives the temporal position of an electric pulse from the bandpass-filtered acoustic signal. However, in the PDT strategy, the peaks of the waveform determine the temporal positions of the pulses. Thus, the pulse rate varies according to the temporal properties of the acoustic signal in each channel. For CI listeners, van Hoesel and Tyler (2003) could not find any clear difference between the PDT strategy and the standard clinical strategy with respect to sound localization and speech perception in noise. However, the comparison was confounded by several differences in the experimental setup. Additionally, the stimulation pulse rate was rather high: it was 700 pps on average and limited to 1400 pps. For such high pulse rates, the sensitivity to fine-structure ITD is rather low, as supported by results of this thesis.

Both fine-structure strategies presented so far do not include synchronization of the timing between the two ears. Additionally, both strategies share the problem of the rate limitation in ITD perception. Using the binaurally-synchronized jitter may solve this problem and lead to an improvement of the localization performance. However, control over the binaural timing is required to implement binaurally-synchronized jitter.

In conclusion, the presented results show the importance of fine-structure ITD in binaural electric hearing. However, one caveat is that these results were obtained using only one bilaterally pitch-matched electrode pair. A feasible implementation of a multi-electrode fine-structure binaural stimulation strategy may prove to be quite a challenge because of different factors such as channel interaction. In any case, the successful development of such a strategy would be an important milestone in bilateral cochlear implant research.

VII. References

Abbas, P. J. (**1993**). "Electrophysiology," in *Cochlear Implants: Audiological Foundations*, edited by R. S. Tyler (Singular, San Diego)

Agresti, A. (**1984**). *Analysis of ordinal categorical data*, (John Viley & Sons, Inc., New York)

Agresti, A. (**1996**). *Introduction to analysis of categorical Data*, (A Wiley-Interscience Publication, John Viley & Sons, Inc., New York)

Bernstein, L. R. (**2001**). "Auditory processing of interaural timing information: new insights," *J Neurosci Res*, **66**, 1035-46.

Bernstein, L. R., and Trahiotis, C. (**1994**). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise," *J Acoust Soc Am*, **95**, 3561-7.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," *J Acoust Soc Am*, **112**, 1026-36.

Blauert, J. (**1981**). "Lateralization of jittered tones," *J Acoust Soc Am*, **70**, 694-8.

Blauert, J. (**1997**). *Spatial Hearing, 2nd edition*, (MIT, Cambridge, MA)

Bosco, E., D'Agosta, L., Mancini, P., Traisci, G., D'Elia, C., and Filipo, R. (**2005**). "Speech perception results in children implanted with Clarion devices: Hi-Resolution and Standard Resolution modes," *Acta Otolaryngol*, **125**, 148-58.

Bronkhorst, A. W., and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J Acoust Soc Am*, **83**, 1508-16.

Buell, T. N., and Hafter, E. R. (**1988**). "Discrimination of interaural differences of time in the envelopes of high-frequency signals: integration times," *J Acoust Soc Am*, **84**, 2063-6.

Buell, T. N., and Hafter, E. R. (**1991**). "Combination of binaural information across frequency bands," *J Acoust Soc Am*, **90**, 1894-900.

Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., and Clark, G. M. (**1994**). "Pitch perception for different modes of stimulation using the cochlear multiple-electrode prosthesis," *J Acoust Soc Am*, **95**, 2658-69.

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., and Wouters, J. (**2002**). "Temporal pitch mechanisms in acoustic and electric hearing," *J Acoust Soc Am*, **112**, 621-33.

Chatterjee, M. (**1999**). "Temporal mechanisms underlying recovery from forward masking in multielectrode-implant listeners," *J Acoust Soc Am*, **105**, 1853-63.

Chen, H., and Zeng, F. (**2004**). "Frequency modulation detection in cochlear implant subjects," *J Acoust Soc Am*, **116**, 2269-77.

Chen, H., Ishihara, Y. C., and Zeng, F. (**2005**). "Pitch discrimination of patterned electric stimulation," *J Acoust Soc Am*, **118**, 338-45.

Cohen, L. T., Richardson, L. M., Saunders, E., and Cowan, R. S. C. (**2003**). "Spatial spread of neural excitation in cochlear implant recipients: comparison of improved ECAP method and psychophysical forward masking," *Hear Res*, **179**, 72-87.

Cohen, N. L., Waltzman, S. B., and Fisher, S. G. (**1993**). "A prospective, randomized study of cochlear implants. The Department of Veterans Affairs Cochlear Implant Study Group," *N Engl J Med*, **328**, 233-7.

Collins, L. M., Zwolan, T. A., and Wakefield, G. H. (**1997**). "Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects," *J Acoust Soc Am*, **101**, 440-55.

Culling, J. F., Hawley, M. L., and Litovsky, R. Y. (**2004**). "The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources," *J Acoust Soc Am*, **116**, 1057-65.

de Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., and Pelizzone, M. (**2003**). "Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects," *Hear Res*, **182**, 77-87.

Dirks, D. D., and Wilson, R. H. (**1969**). "The effect of spatially separated sound sources on speech intelligibility," *J Speech Hear Res*, **12**, 5-38.

Djourno, A., and Eyries, C. (**1957**). "Prothèse auditive par excitation électruque à distance du nerf sensoriel à l'aide d'un bobinage inclus à demeure," *Presse Méd*, **65**, 1417-1423.

Drennan, W. R., Gatehouse, S., and Lever, C. (**2003**). "Perceptual segregation of competing speech sounds: the role of spatial location," *J Acoust Soc Am*, **114**, 2178-89.

Dreyer, A., and Delgutte, B. (**2006**). "Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization," *J Neurophysiol*, **96**, 2327-41.

Dye, R. H. J., and Hafter, E. R. (**1984**). "The effects of intensity on the detection of interaural differences of time in high-frequency trains of clicks," *J Acoust Soc Am*, **75**, 1593-8.

Dye, R. H. J., Niemiec, A. J., and Stellmack, M. A. (**1994**). "Discrimination of interaural envelope delays: the effect of randomizing component starting phase," *J Acoust Soc Am*, **95**, 463-70.

Dynes, S. B., and Delgutte, B. (**1992**). "Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea," *Hear Res*, **58**, 79-90.

Eddington, D. K., and Pierschalla, M. L. (**1994**). *Cochlear Implants - Restoring Hearing to Deaf*, (http://www.med.harvard.edu/publications/On_The_Brain/Volume03/Number4/Cochlear.html). Retrieved on 29.11.2007

Filipo, R., Mancini, P., Ballantyne, D., Bosco, E., and D'Elia, C. (**2004**). "Short-term study of the effect of speech coding strategy on the auditory performance of pre- and post-lingually deafened adults implanted with the Clarion CII," *Acta Oto-laryngol*, **124**, 368-70.

Franck, K. H., Xu, L., and Pfingst, B. E. (**2003**). "Effects of stimulus level on speech perception with cochlear prostheses," *J Assoc Res Otolaryngol*, **4**, 49-59.

Freyman, R. L., Zurek, P. M., Balakrishnan, U., and Chiang, Y. C. (**1997**). "Onset dominance in lateralization," *J Acoust Soc Am*, **101**, 1649-59.

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: comparison of acoustic hearing and cochlear implants," *J Acoust Soc Am*, **110**, 1150-63.

Frijns, J. H. M., Briaire, J. J., de Laat, J. A. P. M., and Grote, J. J. (**2002**). "Initial evaluation of the Clarion CII cochlear implant: speech perception and neural response imaging," *Ear Hear*, **23**, 184-97.

Fu, Q. J., and Shannon, R. V. (**2000**). "Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners," *J Acoust Soc Am*, **107**, 589-97.

Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (**2003**). "Temporal properties of spontaneous speech - a syllable-centric perspective," *J Phonetics*, **31**, 465-485.

Greenwood, D. D. (**1961**). "Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane," *J Acoust Soc Am*, **33**, 1344-56.

Griffin, S. J., Bernstein, L. R., Ingham, N. J., and McAlpine, D. (**2005**). "Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig," *J Neurophysiol*, **93**, 3463-78.

Hafter, E. R., and Buell, T. N. (**1990**). "Restarting the adapted binaural system," *J Acoust Soc Am*, **88**, 806-12.

Hafter, E. R., and De Maio, J. (**1975**). "Difference thresholds for interaural delay," *J Acoust Soc Am*, **57**, 181-7.

Hafter, E. R., and Dye, R. H. J. (**1983**). "Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number," *J Acoust Soc Am*, **73**, 644-51.

Hamzavi, J., and Arnoldner, C. (**2006**). "Effect of deep insertion of the cochlear implant electrode array on pitch estimation and speech perception," *Acta Otolaryngol*, **126**, 1182-7.

Harris, J. P., Anderson, J. P., and Novak, R. (**1995**). "An outcomes study of cochlear implants in deaf patients. Audiologic, economic, and quality-of-life changes," *Arch Otolaryngol Head Neck Surg*, **121**, 398-404.

Hartmann, R., Topp, G., and Klinke, R. (**1984**). "Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea," *Hear Res*, **13**, 47-62.

Hartmann, W. M., and Wittenberg, A. (**1996**). "On the externalization of sound images," *J Acoust Soc Am*, **99**, 3678-88.

Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (**1999**). "Speech intelligibility and localization in a multi-source environment," *J Acoust Soc Am*, **105**, 3436-48.

Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," *J Acoust Soc Am*, **55**, 84-90.

Hinderink, J. B., Mens, L. H., Brokx, J. P., and van den Broek, P. (**1995**). "Performance of prelingually and postlingually deaf patients using single-channel or multichannel cochlear implants," *Laryngoscope*, **105**, 618-22.

Hirsch, I. J. (**1950**). "The relation between localization and intelligibility," *J Acoust Soc Am*, **22**, 196-200.

Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schösser, H., Garnham, C., and Anderson, I. (**2006**). "MED-EL Cochlear implants: state of the art and a glimpse into the future," *Trends Amplif*, **10**, 201-19.

Holden, L. K., Skinner, M. W., Holden, T. A., and Demorest, M. E. (**2002**). "Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy," *Ear Hear*, **23**, 463-76.

House, W. F. (**1995**). *Cochlear Implants: My Perspective*, (AllHear, Inc.)

Kiefer, J., Hohl, S., Stürzebecher, E., Pfennigdorff, T., and Gstöettner, W. (**2001**). "Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the nucleus CI 24M cochlear implant system," *Audiology*, **40**, 32-42.

Klump, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," *J. Acoust. Soc. Am.*, **28**, 859-860.

Kuba, H., Yamada, R., Fukui, I., and Ohmori, H. (**2005**). "Tonotopic specialization of auditory coincidence detection in nucleus laminaris of the chick," *J Neurosci*, **25**, 1924-34.

Kuwada, S., Stanford, T. R., and Batra, R. (**1987**). "Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit: effects of changing frequency," *J Neurophysiol*, **57**, 1338-60.

Laback, B., and Majdak, P. (**2008**). "From the Cover: Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," *Proc Natl Acad Sci*, **105**, 814-817.

Laback, B., Deutsch, W. A., and Baumgartner, W. (**2004**). "Coding of vowellike signals in cochlear implant listeners," *J Acoust Soc Am*, **116**, 1208-23.

Laback, B., Majdak, P., and Baumgartner, W. (**2007**). "Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing," *J Acoust Soc Am*, **121**, 2182-91.

Lawson, D. T., Wilson, B. S., Zerbi, M., van den Honert, C., Finley, C. C., Farmer, J. C. J., McElveen, J. T. J., and Roush, P. A. (**1998**). "Bilateral cochlear implants controlled by a single speech processor," *Am J Otol*, **19**, 758-61.

Lawson, D. T., Wolford, R., Brill, S., Schatzer, R., and Wilson, B. S. (**2001**). *Twelfth quarterly progress report: Speech processors for auditory prostheses, Center of Auditory Prosthesis Research, Research Triangle Institute, NIH project N01-DC-8–2105. Bethesda.*, (http://www.nidcd.nih.gov/staticresources/funding/programs/npp/pdf/archived/N01-DC-8-2105QPR12.pdf). Retrieved on 10.11.2006

Liang, L., Lu, T., and Wang, X. (**2002**). "Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates," *J Neurophysiol*, **87**, 2237-61.

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

Lienert, G. A. (**1978**). *Verteilungsfreie Methoden in der Biostatistik (Non-parametric methods in biostatistics)*, (2nd edition, Verlag Anton Hain, Meisenheim am Glan)

Litovsky, R., Agrawal, S., Jones, G., Henry, B., and Van Hoesel, R. (**2005**). "Effect of Interaural Electrode Pairing on Binaural Sensitivity on Bilateral Cochlear Implant Users," presented at the 28th MidWinter Meeting of the *Association for Research in Otolaryngology*, New Orleans

Litvak, L. M., Delgutte, B., and Eddington, D. K. (**2003**). "Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains," *J Acoust Soc Am*, **114**, 2079-98.

Litvak, L., Delgutte, B., and Eddington, D. (**2001**). "Auditory nerve fiber responses to electric stimulation: modulated and unmodulated pulse trains," *J Acoust Soc Am*, **110**, 368-79.

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," *J Acoust Soc Am*, **114**, 1565-74.

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited," *J Acoust Soc Am*, **111**, 2219-36.

Majdak, P., and Laback, B. "Effect of center frequency on the sensitivity to interaural delay in high-frequency clicks," submitted to *J Acoust Am Soc*.

Majdak, P., Laback, B., and Baumgartner, W. (**2006**). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," *J Acoust Soc Am*, **120**, 2190-201.

McFadden, D., and Pasanen, E. G. (**1976**). "Lateralization of high frequencies based on interaural time differences," *J Acoust Soc Am*, **59**, 634-9.

McKay, C. M., and Carlyon, R. P. (**1999**). "Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains," *J Acoust Soc Am*, **105**, 347-57.

Middleton, J. W., Longtin, A., Benda, J., and Maler, L. (**2006**). "The cellular basis for parallel neural transmission of a high-frequency stimulus and its low-frequency envelope," *Proc Natl Acad Sci U S A*, **103**, 14596-601.

Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B. (**1995**). "Head-Related Transfer Functions of Human Subjects," *J. Audio Eng. Soc*, **43**, 300-321.

Moore, B. C. (**1978**). "Psychophysical tuning curves measured in simultaneous and forward masking," *J Acoust Soc Am*, **63**, 524-32.

Morimoto, M., and Aokata, H. (**1984**). "Localization cues in the upper hemisphere," *J. Acoust. Soc. Jpn. (E)*, **5**, 165-173.

Nelson, D. A., and Donaldson, G. S. (**2001**). "Psychophysical recovery from single-pulse forward masking in electric hearing," *J Acoust Soc Am*, **109**, 2921-33.

Nelson, D. A., and Donaldson, G. S. (**2002**). "Psychophysical recovery from pulse-train forward masking in electric hearing," *J Acoust Soc Am*, **112**, 2932-47.

Nie, K., Stickney, G., and Zeng, F. (**2005**). "Encoding frequency modulation to improve cochlear implant performance in noise," *IEEE Trans Biomed Eng*, **52**, 64-73.

Noel, V. A., and Eddington, D. K. (**2007**). "Binaural localization ability can be predicted from basic measure of ILD and is robust to effects of AGC," presented at the 9th of the *Conference on Implantable Auditory Prostheses*, Lake Tahoe, California

Nuetzel, J. M., and Hafter, E. R. (**1976**). "Lateralization of complex waveforms: effects of fine structure, amplitude, and duration," *J Acoust Soc Am*, **60**, 1339-46.

Nuetzel, J. M., and Hafter, E. R. (**1981**). "Discrimination of interaural delays in complex waveforms: Spectral effects," *J Acoust Soc Am*, **69**, 1112-8.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," *Hear Res*, **24**, 1-15.

Pelizzone, M., Cosendai, G., and Tinembart, J. (**1999**). "Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for ineraid implanted subjects," *Ear Hear*, **20**, 228-37.

Pfingst, B. E., Zwolan, T. A., and Holloway, L. A. (**1997**). "Effects of stimulus configuration on psychophysical operating levels and on speech recognition with cochlear implants," *Hear Res*, **112**, 247-60.

Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J. (**1999**). "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation," *Hear Res*, **127**, 108-18.

Saberi, K. (**1996**). "Observer weighting of interaural delays in filtered impulses," *Percept Psychophys*, **58**, 1037-46.

Saberi, K. (**1998**). "Modeling interaural-delay sensitivity to frequency modulation at high frequencies," *J Acoust Soc Am*, **103**, 2551-64.

Schleich, P., and Nopp, P. (**2002**). "The Effect Of The AGC On Sound Localisation," presented at the 3rd of the *Wullstein Symposium on Bilateral Cochlear Implants and Binaural Signal Processing*, Würzburg, Germany

Senn, P., Kompis, M., Vischer, M., and Haeusler, R. (**2005**). "Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors," *Audiol Neurootol*, **10**, 342-52.

Shackleton, T. M., Skottun, B. C., Arnott, R. H., and Palmer, A. R. (**2003**). "Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of Guinea pigs," *J Neurosci*, **23**, 716-24.

Shannon, R. V. (**1983a**). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," *Hear Res*, **11**, 157-89.

Shannon, R. V. (**1983b**). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," *Hear Res*, **12**, 1-16.

Sherbecoe, R. L., and Studebaker, G. A. (**2004**). "Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units," *Int J Audiology*, **43**, 442-448.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P. M., Staller, S. J., Shipp, D. B., Shallop, J. K., Everingham, C., Menapace, C. M., Arndt, P. L., and *et al.* (**1994**). "Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System," *Am J Otol*, **15 Suppl 2**, 15-27.

Skottun, B. C., Shackleton, T. M., Arnott, R. H., and Palmer, A. R. (**2001**). "The ability of inferior colliculus neurons to signal differences in interaural delay," *Proc Natl Acad Sci U S A*, **98**, 14050-4.

Smith, Z. M., and Delgutte, B. (**2005**). "Binaural interactions in the auditory midbrains with bilateral cochlear implants," presented at the of the *28th Midwinter Research Meeting of the Association for Research in Otolaryngology*, New Orleans

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

Stecker, G. C., and Hafter, E. R. (**2002**). "Temporal weighting in sound localization," *J Acoust Soc Am*, **112**, 1046-57.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2005**). "Discrimination of interaural phase differences in the envelopes of sinusoidally amplitude-modulated 4-kHz tones as a function of modulation depth," *J Acoust Soc Am*, **118**, 346-52.

Strutt, J. W. (**1907**). "On out perception of sound direction," *Philos. Mag.*, **13**, 214-232.

Tyler, R. S. (**1988**). "Open-set word recognition with the 3M/Vienna single-channel cochlear implant," *Arch Otolaryngol Head Neck Surg*, **114**, 1123-6.

van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," *J Acoust Soc Am*, **101**, 1671-80.

van den Honert, C., and Stypulkowski, P. H. (**1987**). "Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli," *Hear Res*, **29**, 207-22.

van Hoesel, R. J. M. (**2004**). "Exploring the benefits of bilateral cochlear implants," *Audiol Neurootol*, **9**, 234-46.

van Hoesel, R. J. M. (**2007**). "Sensitivity to binaural timing in bilateral cochlear implant users," *J Acoust Soc Am*, **121**, 2192-206.

van Hoesel, R. J. M., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," *J Acoust Soc Am*, **113**, 1617-30.

van Hoesel, R. J., and Clark, G. M. (**1997**). "Psychophysical studies with two binaural cochlear implant subjects," *J Acoust Soc Am*, **102**, 495-507.

van Hoesel, R. J., Ramsden, R., and Odriscoll, M. (**2002**). "Sound-direction identification, interaural time delay discrimination, and speech intelligibility advantages in noise for a bilateral cochlear implant user," *Ear Hear*, **23**, 137-149.

Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (**2000**). "Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system," *Ear Hear*, **21**, 608-24.

Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," *J Acoust Soc Am*, **90**, 858-65.

Wichmann, F. A., and Hill, N. J. (**2001a**). "The psychometric function: I. Fitting, sampling, and goodness of fit," *Percept Psychophys*, **63**, 1293-313.

Wichmann, F. A., and Hill, N. J. (**2001b**). "The psychometric function: II. Bootstrap-based confidence intervals and sampling," *Percept Psychophys*, **63**, 1314-29.

Wightman, F. L. and Kistler, D. J. (**1997**). "Factors affecting the relative salience of sound localization cues," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey, and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ)

Wightman, F. L., and Kistler, D. J. (**1989**). "Headphone simulation of free-field listening. I: Stimulus synthesis," *J Acoust Soc Am*, **85**, 858-67.

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," *J Acoust Soc Am*, **91**, 1648-61.

Wilson, B. (**2004**). "Engineering design of cochlear implants," in *Cochlear Implants*, edited by F.G. Zeng, A.N. Popper, and R. R. Fay (Springer, New York)

Wilson, B. S. (**1997**). "The future of cochlear implants," *Br J Audiol*, **31**, 205-25.

Wilson, B. S., Finley, C. C., Farmer, J. C. J., Lawson, D. T., Weber, B. A., Wolford, R. D., Kenan, P. D., White, M. W., Merzenich, M. M., and Schindler, R. A. (**1988**). "Comparative studies of speech processing strategies for cochlear implants," *Laryngoscope*, **98**, 1069-77.

Wilson, B. S., Finley, C. C., Lawson, D. T., and Zerbi, M. (**1997**). "Temporal representations with cochlear implants," *Am J Otol*, **18**, S30-4.

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**, 236-8.

Yin, T. C., Kuwada, S., and Sujaku, Y. (**1984**). "Interaural time sensitivity of high-frequency neurons in the inferior colliculus," *J Acoust Soc Am*, **76**, 1401-10.

Yost, W. A. and Hafter, E. R. (**1987**). "Lateralization," in *Directional Hearing*, edited by W. Yost, and G. Gourevitch (Springer-Verlag, New York)

Zeng, F. G. (**2002**). "Temporal pitch in electric hearing," *Hear Res*, **174**, 101-6.

Zeng, F., Nie, K., Stickney, G. S., Kong, Y., Vongphoe, M., Bhargave, A., Wei, C., and Cao, K. (**2005**). "Speech recognition with amplitude and frequency modulations," *Proc Natl Acad Sci U S A*, **102**, 2293-8.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just Noticeable Differences in Dichotic Phase," *J Acoust Soc Am*, **28**, 860-4.

Zwolan, T. A., Kileny, P. R., Ashbaugh, C., and Telian, S. A. (**1996**). "Patient performance with the Cochlear Corporation '20 + 2' implant: bipolar versus monopolar activation," *Am J Otol*, **17**, 717-23.