The Psychophysical Bases of Spatial Hearing in Acoustic and Electric Stimulation

Bernhard Laback

Universität Wien, Fakultät für Psychologie

Wien, 25. Februar 2013

Table of Contents

I. PREFACE

1. Motivation and Overview

An important trigger for the work described in this habilitation treatise was a hint I got from a hearing scientist in about 2001. While I was at that time mainly interested in auditory masking and channel interaction effects in cochlear implants, he attempted to convince me that hearing with bilateral cochlear implants (CIs), one implanted to the left and one to the right ear, is going to become an important and "hot" topic in both basic and applied psychophysical science. Indeed, ENT-clinics at that time started to implant more and more bilaterally, seeking to provide the implantees with the advantages of hearing with two ears, i.e. binaural hearing. Binaural hearing is known to be essential for localizing sound sources in the left/right dimension and to understand speech in noise. By shortly reviewing the state of research, I realized how little was known at that time about binaural hearing with CIs. Therefore, in consensus with Werner A. Deutsch, the director of the Acoustics Research Institute (ARI) of the Austrian Academy of Sciences, I decided to start a research program on binaural psychophysics, aiming to better understand the differences in spatial hearing between normal and electrically-stimulated ears. The fruits of the decision to work in this field are described in this treatise. Section I.2. summarizes chronologically the scientific achievements and impact of the individual studies. Section I.3. describes my personal development in this field of science. Section II. first provides an introduction into psychophysics and spatial hearing, and then summarizes the scientific background and context of the individual studies as well as the general conclusions that can be drawn by combining their outcomes. Finally, section III. lists the peer-reviewed publications included in this habilitation treatise.

2. Scientific Achievements and Impact

A focus of my work is on exploring the basic mechanisms underlying spatial hearing with CIs and particularly on the differences compared to the mechanisms in normal (acoustic) hearing. Therefore, an important aspect of the research has been to develop the appropriate psychophysical methodology and to design psychophysical experiments for both acoustic and electric hearing. The comparison between the two hearing modes has the potentials to identify some of the origins for perceptual deficits in electric hearing, to devise new method to compensate for such deficits, and to contribute to the better understanding of spatial hearing mechanisms in normal hearing.

The first study (Laback *et al.*, 2004) addressed the sensitivity of CI listeners to the basic binaural cues, interaural time differences (ITDs) and interaural level differences (ILDs), with experimental stimuli sent directly to clinical CI processors. The study, revealing good ILD sensitivity but poor ITD sensitivity relative to the control group of normal hearing (NH) listeners, has been cited many times since then. The study has been considered as evidence that state-of-the-art CI systems are not capable of providing bilateral CI listeners with sufficient ITD information while they do a good job in providing ILD information. The study contributed to the interpretation of numerous CI sound localization studies in terms of the binaural cues used by the listeners.

The next step was to look closer into the potential reasons for the apparently poor ITD sensitivity. In order to accurately control the detailed stimulus properties and interaural timing, a CI research system is required. Fortunately, the University of Innsbruck had developed such a system that was kindly provided to us by the Austrian CI manufacturer MED-EL Corp. Integration of the research system into our experimental hardware and software framework for psychophysical experiments gave us the possibility to perform binaural CI experiments with an accuracy of binaural stimulation timing of 2.5 μs (more accurate than required even for NH listeners).

In the two studies that followed we investigated the particular contribution of ITD in the carrier signal (referred to as *fine structure* in case of acoustic stimuli and *fine timing* in case of electric stimuli), the ongoing envelope modulation, the onset, and the offset (Majdak, Laback and Baumgartner, 2006; Laback, Majdak and Baumgartner, 2007). These studies provided for the first time evidence that CI listeners are sensitive to ITD in the fine timing, if its rate does not exceed a certain limit. Furthermore, they showed a relatively weak contribution of ongoing envelope ITD and gating ITD. These results suggested the need to explicitly encode fine-timing ITD information with bilateral CIs. Consequently, our studies have been cited many times since then and received considerable attention both from the basic-science community and from the clinical and manufacturer communities. These studies also contributed to motivating CI manufacturers to develop pertinent stimulation strategies.

A severe limitation of the approach to encode fine-timing ITD information with CIs is that fine-timing ITD sensitivity is limited to quite low pulse rates, while higher pulse rates are required to appropriately encode speech information and to achieve sufficiently loud auditory sensations without exceeding maximally tolerable electric power consumption. Inspired by findings reported in the normal-hearing literature on the so-called binaural adaptation phenomenon, referring to a rate limitation in the perception of binaural cues, we hypothesized that the rate limitation in electric hearing might have related origins. Further, we hypothesized that recovery from binaural adaptation, as reported in the literature for NH listeners, might as well be achievable in CI listeners by applying appropriate stimuli. We tested these hypotheses with CI listeners using electric pulse trains with binaurally-synchronized timing jitter. Using these stimuli we were able to demonstrate large improvements in ITD sensitivity compared to conventional periodic pulse trains and published the results in the prestigious journal PNAS (Laback and Majdak, 2008c). This publication received world-wide attention in the pertinent communities and triggered both psychophysical and physiological follow-up research in different labs. The scientific achievement also led to a world-wide patent for a method to improve the access to ITD cues with binaural CI or hearing aid systems (Laback and Majdak, 2008b; Laback and Majdak, 2009).

The relatively weak sensitivity of CI listeners to ongoing envelope modulation observed in our and others' previous studies motivated me to study the particular role of the envelope shape both in electric and acoustic hearing. By systematically varying different features of the ongoing envelope shape, we found that in CI listeners the sensitivity improves with increasing off time (the silent period in each modulation cycle), similar to NH listeners (Laback *et al.*, 2011b; Laback, Zimmermann and Majdak, 2010). The effect of the envelope slope, however, was found to differ substantially in CI listeners compared to NH listeners. These results allowed us to suggest strategies for CI signal processing aiming to improve the CI listener's sensitivity to envelope ITD cues.

The studies conducted so far involved electric stimuli presented at a single interaural electrode pair. This corresponds to the practical situation of a narrowband acoustic sound activating a single CI channel at each ear. However, natural sounds usually cover broader spectral regions, thus involving the simultaneous activation of multiple CI electrodes which are associated with the corresponding frequency channels. Furthermore, acoustic environments often contain multiple sound sources at different locations which results in ITD and ILD cues differing across channels. NH listeners are able to combine conflicting ITD and ILD cues across spectral channels in a meaningful manner. They integrate binaural information across spectral components to form a single auditory object if the monaural signal properties indicate that these components correspond to the same acoustic source, an effect called binaural inference. In contrast, they segregate components corresponding to different sources into separate auditory objects. We addressed the question whether similar grouping mechanisms are operating in bilateral CI listeners. In electric stimulation, channel interactions – a result of current spread from the individual electrodes – may disrupt the monaural cues available within the individual channels and, thus, disturb the grouping mechanisms. In addition, channel interactions may also disrupt the binaural cues. Using combinations of widely spaced electrode pairs to avoid such interactions, we found that grouping mechanisms are overall operating in most of the CI listeners tested (Best, Laback and Majdak, 2011). Although channel interactions would likely disturb grouping mechanisms for more closely spaced electrode pairs to some extent, our results demonstrated that CI listeners might be able to deal with complex listening environments with future CI systems that limit channel interactions.

The acoustic-hearing studies included in this habilitation treatise not only served to provide reference data for the electric-hearing studies, but were designed to provide also new insight and better understanding of auditory processes in normal hearing. For example, the studies on the effect of binaurally-synchronized jitter in acoustic hearing (Goupell, Laback and Majdak, 2009), in which we tested a physiology-based model of auditory nerve and brainstem, provided evidence that recovery from binaural adaptation can at least partly be explained by peripheral effects, in contrast to more central explanations put forward in the literature.

Another example is a study on the effects of center frequency and rate on the sensitivity to ITD in high-frequency click trains (Majdak and Laback, 2009). This study was primarily intended to clarify if the comparison of pulse rate limitations in ITD perception between CI and NH listeners in our earlier studies may have been confounded by the choice of the center frequency of stimulation in NH listeners. In addition to demonstrating that such confounding effects can be ruled out, the study revealed the interesting result that the decrease in ITD sensitivity with increasing center frequency for stimuli with a constant *relative* bandwidth is much smaller compared to previous studies that used stimuli with a constant *absolute* bandwidth and, thus, less salient envelope ITD cues at high center frequencies. This in turn has implications for the understanding of ITD-sensitive mechanisms at high frequencies.

The study on envelope-ITD perception (Laback *et al.*, 2011b) addressed a highly relevant open question in acoustic hearing research, namely, which aspects of the ongoing envelope shape are important for ITD perception. Previous studies on envelope-ITD perception that varied the peakedness of the envelope co-varied several envelope parameters, thus allowing no conclusions on their specific contribution. By independently varying the off time, the slope steepness, and the peak level we were able to isolate the perceptual contributions of these parameters. The general effects observed (Laback *et al.*, 2011b) were in line with an acoustic-hearing study that was performed independently and at the same time at the University of Oldenburg (Klein-Hennig *et al.*, 2011). Discussions with the Oldenburg group about differences between the two studies in the detailed dependencies of the observed effects on the varied parameters led to a collaborative follow-up study (Dietz *et al.*, 2013). In that study we demonstrated that the differences in outcomes can be explained by the differences in the experimental stimuli and proposed a perception-based metric for the off time.

Around 2004, we started thinking if it may be possible to provide CI listeners with localization information exploited by the normal auditory system to localize sounds in the vertical planes (along the front/back/up/down dimensions), namely, the direction-dependent spectral colouring of the incoming sound by the pinnae, head, and torso. Having access to both binaural and spectral localization cues would allow CI listeners to exploit the advantages of spatial hearing in 2-D space. Based on these considerations and following discussions with my colleagues, most notably Piotr Majdak, I wrote a proposal for a project titled "Spectral Cues in Auditory Localization with Cochlear Implants (CI-HRTF)". The project was funded by the FWF (Project No. P18401-B15), starting in Oct. 2005 and ending in Dec. 2010. The parallel work on this project and on the described projects on binaural hearing promised a large potential for strong synergies.

In the first project step we developed a 2-D localization paradigm that allows to train listeners with new spatial information. In the experiments that followed we showed that the inclusion of a visual environment that provides feedback to the listener substantially improves vertical-plane localization. We finally derived a reliable method, based on a virtual audio-visual environment and involving a hand-held pointing device (Majdak, Goupell and Laback, 2010). The method has applications in both basic and applied spatial hearing research, whenever the task is to learn new localization information via visual feedback.

Based on this method, the next step was to determine the ability of bilateral CI listeners, using state-of-the-art speech processors, to localize in the vertical planes in order to evaluate the extent to what they may already use spectral localization cues. Vertical-plane localization performance was found to be much worse than in NH listeners and turned out to rely solely on absolute level cues instead of spectral cues (Majdak, Goupell and Laback, 2011).

Next, we performed a set of basic studies addressing the crucial question if or under what conditions CI listeners have access to spectral localization cues. A preceding study on monaural perception of vowel-like stimuli in CI listeners, showing that both the spectral peaks and the between-peak information of vowels are somewhat resolvable (Laback, Deutsch and Baumgartner, 2004), suggested that the spectral sensitivity might be sufficient. In the following project step, we found that for a 12-channel implant the eight apical-most channels with matched spectral content are sufficient for accurate speech intelligibility in quiet and moderate noise (Goupell *et al.*, 2008b). This suggested that about four channels are available for independent coding of spectral localization cues. Next, in a CI simulation study with NH listeners, we found that 12 frequency channels are sufficient for robust vertical-plane localization in NH listeners (Goupell, Majdak and Laback, 2010). Remarkably, there were only as few as four channels within the range of spectral localization cues (> 3 kHz). Given the aim to encode spectral localization information with a very limited number of channels that is available in electric hearing, this is a promising result for CI listeners. Lastly, we addressed the crucial question to what extent and under which conditions CI listeners are able to discriminate between different spectral shapes, often called profile analysis. The results (Goupell *et al.*, 2008a) showed that while all tested CI listeners were sensitive to spectral peaks and notches imposed on a flat, i.e., constant-loudness background, random trial-to-trial level roving rendered the performance to chance level. This is in contrast to NH listeners who show much less deterioration in performance with level roving in comparable tasks. These results raised important general questions about coding of spectral information in acoustic and electric hearing, triggered two follow-up studies of research groups in the US and UK, and inspired me to develop new hypotheses that I will address experimentally in a future project.

The last step of the project tackled the general question if listeners are able to adapt to a mapping (warping) of the high-frequency spectral localization cues to the stimulation range available in electric hearing. In a long-time localization training experiment with NH listeners (Majdak, Walder and Laback, conditionally accepted) we showed that listeners are largely able to adapt to such a warping. More surprisingly, the listeners adapted completely to a control condition without warping where high frequencies were completely removed.

Taken together, the results of the project CI-HRTF provided new insights into basic mechanisms related to vertical-plane localization. Furthermore, it showed possible strategies for future CI systems to encode spectral information in order to provide CI listeners with the ability to localize in the vertical planes and possibly also to improve their ability to understand speech in noise. A particular challenge will be to find ways to enable listeners to perform spectral profile analysis. The project largely benefited from synergistic effects resulting from our parallel studies on binaural hearing. The ultimate goals of both lines of research, to better understand the psychophysical basics underlying spatial hearing and to improve spatial hearing in CI listeners, overlapped between the two projects. For example, in the stimulus design of the study "Loca#Channels", we took into account not only spectral but also temporal properties of the stimulus pulses, given that they may impact both vertical-plane aspects (spectral cues) and horizontal-plane aspects (ITD cues) of sound localization.

In summary, 16 journal articles[1], one book chapter, one invited editorial focus and one reply-letter in journals, as well as one international patent are included in this treatise, all on the psychophysics of spatial hearing and the spectral sensitivity in normal and electric hearing. For the updated state of the publication list, please see: .http://www.kfs.oeaw.ac.at/laback

3. Personal Development

Based on my research experience in the field of psychophysics in normal hearing and impaired hearing I gathered during my dissertation, mainly at the Free University Hospital in Amsterdam, the ARI in Vienna, and the ENT-department of the Medical University of Vienna, I entered the research field of cochlear implant psychophysics in about 2000. I was immediately fascinated by this technique that not only restores hearing in the deaf but also represents a direct interface to the auditory

1 One of those articles (Majdak, Walder and Laback, conditionally accepted) is not published yet, but is in the state of conditional acceptance for publication in J. Acoust. Soc. Am.

nerve and thus allows to study hearing mechanisms in general. Within the scope of a cooperation with the ENT-department of the University Clinic Salzburg, where I received kind support from Alois Mair, the head of the audiology department, I was acquainted with the audiological techniques underlying CI fitting and testing. I received further training in CI-fitting and experience in CI technology by means of a cooperation with MED-EL Corp., an Austrian CI manufacturer, and on medical aspects of cochlear implantation by Wolf-Dieter Baumgartner, our co-operation partner from the Medical University of Vienna.

Until 2001, I built up laboratory facilities at ARI for performing psychophysical experiments, allowing to present calibrated stimuli to NH listeners via headphones and to CI listeners via freely programmable clinical processors. Based on this setup, we performed monaural studies on the coding of vowellike stimuli and channel interactions (Laback, Deutsch and Baumgartner, 2004) and on auditory stream segregation (Laback and Deutsch, 2001). Following my growing interest on binaural hearing with CIs, in 2002, I was able to obtain a CI research interface (designed by the University of Innsbruck) that allows direct and interaurally-synchronized stimulation with bilateral CIs, which was kindly provided by MED-EL Corporation. Thanks to Piotr Majdak, who joined my group in 2002 and implemented a completely new software framework for electric and acoustic psychophysical experiments, the CI research interface was integrated into our new experimental infrastructure. From about 2004 on, we were one of the very few laboratories worldwide having the possibility to study binaural hearing with CIs under precise control of stimuli presented to the two ears. This allowed me and my group to establish internationally in binaural psychophysics with a focus on CIs within the years that followed. The peer-reviewed publications resulting from these studies are included in this treatise.

In 2005, Piotr Majdak, our laboratory technician Michael Mihocic, and I started building up the experimental infrastructure required for 2-D localization experiments (in both the horizontal and vertical dimensions) and allowing localization training by means of a virtual visual environment. This infrastructure was then completed within the scope of my FWF-project CI-HRTF. Based on this infrastructure, in the following years my group (complemented by Matthew Goupell, whom I hired as a post-doc for CI-HRTF) performed extensive studies on the perception of spectral cues and vertical-plane localization in electric and acoustic hearing. An additional study of intermodal (audio-visual) training effects with manipulated spatial cues allowed to explore the plasticity of localization mechanisms. The peer-reviewed journal papers resulting from that line of research are included in this treatise.

Besides the journal papers, the book chapter, and the patent included in this treatise, I have published 5 more peer-reviewed journal papers, 16 proceedings papers and 3 book chapters on other topics of psychophysics, and presented a total of 78 talks and posters at international conferences. At 14 of those conferences I was invited speaker. I organized scientific sessions at several conferences, including the International Congress on Acoustics (ICA) in Rome (2001), the 9th International Conference on Cochlear implants (CI-2006) in Vienna (2006), and the Acoustics '08 conference in Paris (2008). I was organizer and co-organizer of two small international symposia on binaural hearing with hearing aids and cochlear implants that were held in Sept. 2011 in Vienna (www.kfs.oeaw.ac.at/bihoci).

Normal hearing and cochlear implant research is a highly interdisciplinary field, involving the disciplines psychology, in particular psychophysics, medicine, acoustics, signal processing, and physiology. Because of this interdisciplinarity, cooperation of scientists from the individual fields is an integral part of the research work. This has the consequence that the publications are published together with co-authors. I was, however, deeply involved in all research stages for all the studies contained in this treatise. I was the main researcher in those studies where I am the first author of the corresponding papers. My contributions to the other studies included providing the motivating ideas, applying for respective research grants, planning, programming, and conducting experiments with NH and CI listeners, analyzing data, presenting results at conferences, and writing and revising journal papers. Only for the paper Dietz *et al.* (2013) my role was restricted to providing motivating ideas, discussing the experimental design and contributing to the data analysis and writing of the paper.

Besides my work in the field of spatial hearing, I am also active in other research fields. Most notably, I am working on masking effects in the normal auditory system, with a special focus on time-frequency masking. Together with the Mathematics and Signal Processing in Acoustics Group of ARI and international cooperation partners (the Laboratory of Mechanics and Acoustics, LMA, at CNRS, Marseille), I also work on the incorporation of these results in models of auditory processing and perceptual audio codecs. My particular focus within this research line is on psychophysical measurements and modeling of the spread and additivity of time-frequency masking for spectro-temporally compact sounds (Laback *et al.*, 2011a; Balazs, P., Laback, B., Eckel, G., and Deutsch, W. A., 2010; Laback *et al.*, conditionally accepted; Necciari *et al.*, submitted, 2011). Further areas of my interest are auditory grouping phenomena, perceptual consequences of cochlear hearing impairment, and plasticity and learning effects.

During my research work I have established many international and national co-operations. Amongst others, I co-operated with Virginia Best from the Dept. of Cognitive & Neural Systems (Boston University) and the Computing and Audio Research Lab (University of Sydney), with Bertrand Delgutte and Kenneth Hancock from the Massachusetts Eye and Ear Infirmary (Harvard Medical School and MIT), with Sophie Savel, Sabine Meunier, Richard Kronland-Martinet, and Solvi Ystad from the Laboratory of Mechanics and Acoustics (LMA, CNRS, Marseille), and with Mathias Dietz, Stephan Ewert, Volker Hohmann, and Birger Kollmeier from the Medical Physics Department (University of Oldenburg). I started establishing co-operations with Steven van de Par from the Acoustics Group (University of Oldenburg) and with Maike Vollmer from the Electrophysiology Laboratory (Comprehensive Hearing Center, University Clinic Würzburg). At the national level, I cooperated, amongst others, with Wolf-Dieter Baumgartner from the ENT-clinic of the Medical University Vienna, with Stephan Marcel Pok from the ENT-clinic of the Landeskrankenhaus St. Pölten, and with Alois Mair from the ENT-clinic of the University Clinic Salzburg.

I am experienced in supervising the work of others. I supervised or co-supervised 2 Post-Docs, 5 master theses, 2 PhD theses, and 1 bachelor thesis. All of the students received positions in science or private industry or continued their academic studies. For my FWF-project CI-HRTF, I hired and supervised a postdoc from the USA (Matt Goupell), after he finished his PhD at William Hartmann's lab. After successfully finishing his work on my project, he received a position at the Binaural Hearing and Speech Laboratory at the University of Wisconsin (leading a NIH-funded project) and now was able to obtain a faculty position at the Department of Hearing and Speech Science (University of Maryland, College Park). For my recently funded FWF project BiPhase I hired a Post-Doc from Japan and I am about to hire two master students.

I have experience in teaching at university level. Since 2004, I am permanent lecturer in "Psychoacoustics" for students of electrical engineering and audio engineering at the Institute of Electronic Music & Acoustics (IEM), University of Music and Performing Arts in Graz and the University of Technology in Graz. In the winter semester 2012/2013 I held the proseminar in "Kognitive

Grundlagen des Erlebens und Verhaltens (Cognitive bases of experience and behavior)“ at the Faculty of Psychology of the University of Vienna. In the winter semester 2012/2013 I also held one unit of the lecture “Einführung in wissenschaftliches Denken: Vom schlussfolgernden Beobachten zur experimentellen Methode (Introduction into scientific thinking: from deductive observation to experimental method)” at the Department of Psychology of the University of Vienna. In the summer semester 2002 I held the lecture “Musik als Testsignal in der Psychoakustik und experimentellen Audiologie (Music as test signal in psychoacoustics and experimental audiology)” at the Institute of Musicology of the University of Vienna. In July 2012 I was invited lecturer at the Auditory Cognition Summer School at Plymouth University, UK. My specific topic was “New approaches in cochlear implants”.

4. Acknowledgments

I want to thank all the co-authors of the journal papers and book chapters included in this treatise, i.e. Piotr Madjak, Matthew Goupell, Wolf-Dieter Baumgartner, Werner A. Deutsch, Marcel Pok, Virginia Best, Inge Zimmermann, Karin Schmid, Bertrand Delgutte, Ken Hancock, Thomas Walder, Mathias Dietz, Volker Hohmann, Stephan Ewert, and Torben Wendt. From those people, I want to especially highlight the intensive and rather fruitful ongoing cooperation with my ARI-colleague and friend Piotr Majdak (I always liked our discussions at the end of which each of us took over the opinion of the other) and the productive cooperation with my former ARI-colleague Matthew Goupell. Special thanks go to my colleague Michael Mihocic who provided invaluable help, particularly in taking care of the lab and preparing and running psychophysical experiments: your ability to motivate test subjects, particularly CI listeners, is unbeatable. I want to thank my other cooperation partners and co-authors of the publications not included in this treatise, who were also a source of inspiration. From those people, I particularly appreciate the numerous discussions with my ARI-colleague Peter Balazs, providing me a better understanding of some mathematical concepts in acoustics, especially in time-frequency analysis. I really enjoyed working together with all of you.

I am particularly grateful to my mentor Werner A. Deutsch, the former director of ARI, who evoked my interest in psychoacoustics and accompanied and supported my scientific career. He trusted in my capability to develop my own psychophysical research program and provided me the possibility to build up my own working group at ARI, offering perfect conditions and infrastructure to conduct basic and application-oriented scientific research.

I thank our main cooperation partner at the ENT-department of the Medical University of Vienna, Wolf-Dieter Baumgartner, for believing in my ability to successfully develop and conduct a research program in basic CI-psychophysics, for providing us access to bilaterally supplied CI listeners, for providing me insight into medical and surgical aspects of cochlear implantation, and for always being very supportive with respect to organizational issues. I appreciate the support of Alois Mair from the ENT-department of the Universitätsklinikum Salzburg, particularly in familiarizing me with the audiological techniques required for CI fitting and clinical testing. I thank MED-EL corporation for providing us with all the technical infrastructure required for psychophysical experiments with CI listeners and for always being supportive with our special technical demands. In particular, I want to thank Inge and Erwin Hochmair, Peter Nopp, Peter Schleich, Ewald Thurner, Gerald Schwarzecker, Bernhard Stoebich, Reinhold Schatzer, Eva Kohl, and several others I surely forgot to mention.

I thank my current and former colleagues from ARI for many interesting and insightful interactions, particularly Wolfgang Kreuzer, Georg Rieckh, Bram Alefs, Herbert Griebel, Stefan Marcel Pok, and Thibaud Necciari. I thank the software development group of ARI for their support in programming and software issues, in particular Anton Noll, Jonnie White, and Christian Gottschall.

I am grateful to many international scientists for their critical and constructive feedback on my/our work, for great discussions, and for their willingness to share ideas, including Niek Versfeld, Tammo Houtgast, Joost Festen, Bob Carlyon, Ruth Litovsky, Bernhard Seeber, Zach Smith, Deniz Baskent, Volker Hohmann, Birger Kollmeier, Jesko Verhey, Maike Vollmer, Torsten Dau, Mathias Dietz, Bertrand Delgutte, Ken Hancock, Fan-Gang Zeng, Erv Hafter, Blake Wilson, Richard van Hoesel, Brian Moore, Tom Francart, Steven van de Par, Stefan Brill, and many others I certainly forgot to mention.

I thank Robert Höldrich and Alois Sontacchi from the Institute of Electronic Music and Acoustics for providing me the possibility to teach students of electrical engineering and audio engineering the basics of psychophysical theory and methodology and the psychoacoustics of normal and impaired hearing. I also thank Germain Weber, Ulrich Ansorge, and Claus Lamm from the Faculty of Psychology of the University of Vienna for offering me the possibility to give the course in cognitive bases of experience and behavior, to cooperate scientifically, and to submit this habilitation treatise at their faculty.

I especially thank the Austrian Academy of Sciences for strongly supporting basic science in the field of acoustics and hearing and the FWF for funding my project "Spectral Cues in Auditory Localization with Cochlear Implants"(P18401-B15).

I thank all those of my friends who were always very interested in my work and asked insightful questions. You supported me in my believe that I am doing something useful and that I am privileged in having the possibility to work in such a fascinating field of science.

Finally, I owe a debt to Claudia and my children Jabob and Niklas for their patience during the many instances when I was brooding over some scientific problem while I should have paid more attention to them.

II. INTRODUCTION, SUMMARY, AND CONCLUSIONS

1. Background of the Scientific Field

A. Psychology, Cognitive Science, and Perception

The work described in this habilitation treatise falls into the broad scientific field of cognitive science, one major branch of psychology, and in particular into the field of perception. Perception can be defined as the ability to capture information via the senses and process it in a certain way. The two dominant senses, vision and hearing, allow us to perceive our environment and to focus on the relevant information. Carterette and Friedman (1974) defined the goals of perception research as follows: “The problem of perception is one of understanding the way in which the organism transforms, organizes, and structures information arising from the world in sense data or memory”. This definition already implies that perception is a rather complex process, involving both the more peripheral mechanisms of transforming acoustic or visual information into neural spike patterns and the more central mechanisms of decomposing the incoming information into auditory or visual objects, comparing them with stored objects, and focusing attention on certain objects based on demands from cognitive processes.

B. Psychophysics

The attempt to quantify the sensation of attributes of physical stimuli, mainly by the senses vision, hearing, touch, and taste marked the beginning of experimental psychology and resulted in the establishment of the discipline called psychophysics. According to Gescheider (1997) psychophysics is defined as "the scientific study of the relation between stimulus and sensation". Bruce, Green and Georgeson (1996) provided a more detailed description of the discipline of psychophysics, namely as "the analysis of perceptual processes by studying the effect on a subject's experience or behaviour of systematically varying the properties of a stimulus along one or more physical dimensions". Psychophysics also describes a theoretical and practical framework of methods that can be used to study perceptual systems.

In the following I provide a short summary of the history of psychophysics. A more detailed description can be found for example in Jones (1974). The beginning of psychophysics is often said to date back to 1860, when G. T. Fechner published his famous book “Elemente der Psychophysik” (Fechner, 1860). In that publication, Fechner elaborates on the idea to demonstrate a mathematical relationship between physical stimuli and aspects of consciousness such as sensations. He had been inspired by Ernst Heinrich Weber (Weber, 1834), who found that the just-noticeable difference (JND) of the change in the magnitude of a stimulus is proportional to the magnitude of the stimulus, rather than being an absolute value. Based on that finding (which Fechner named Weber's law) and on theoretical considerations, Fechner concluded that subjective sensation is proportional to the logarithm of the stimulus magnitude. Later work by S. S. Stevens questioned that relationship and, instead, proposed a power-law relationship between subjective sensation and stimulus magnitude (Stevens, 1957). However, there have been discussions up to date on the question if the current evidence justifies to assume a power law relation (e.g., Ellermeier and Faulhammer, 2000).

Other important contributions of Fechner to the field of psychophysics were the statistical treatment of measurement variability and the development of three basic methods of measurement, namely the method of limits, the method of average error (or production method), and the method of right and wrong cases (now called the method of constant stimuli). Fechner's work had great impact, inspiring extensive follow-up research and causing extensive controversies over decades, which has been documented by Titchener (1905). One important topic in psychophysical research following Fechner was the question what a threshold actually is and how it can be measured. An important theoretical contribution in that area was the law of comparative judgment by Thurstone (1927) that later led to the development of the signal detection theory (SDT, Green and Swets, 1966). SDT addresses the problem of separating factors relating to criterion, motivation, and bias from factors relating to purely sensory capabilities by means of statistical treatment of response variability. SDT has been proven successful in providing a consistent measure of sensitivity for a variety of psychophysical tasks measuring thresholds, independent of the different response criteria involved. Another important line of research in psychophysics concerns the so-called direct methods that attempt to relate subjective sensation along a perceptual dimension (e.g., loudness or pitch) to the corresponding physical parameter (e.g., intensity or frequency). One particularly widespread method, the method of magnitude estimation, has been extensively studied by S. S. Stevens, which is summarized in Stevens (1957).

Finally, because the main focus of my work and this habilitation treatise is on hearing, it is worth mentioning that the branch of psychophysics dealing with hearing is called psychoacoustics. However, it is evident that auditory sound localization under real-life conditions involves visual cues about the environment and, thus, implicitly audio-visual interaction effects. Because the studies described in this treatise involved such visual cues and because in some experiments we explicitly included visual cues to train new localization cues, the work is better classified into the broad field of psychophysics rather than psychoacoustics.

C. Other Disciplines Involved

While the basis of my scientific work is perception and psychophysics, the study of the perceptual processes underlying electric hearing with cochlear implants is clearly interdisciplinary, involving the disciplines medicine, audiology, acoustics, signal processing, and physiology. This interdisciplinarity is reflected by the cooperation partners involved in the studies included in this habilitation treatise.

2. Psychophysics of Spatial Hearing

Spatial hearing is an extremely important aspect of auditory perception. In the animal world, it is a basic requirement for survival, by detecting the attack of natural enemies and by acoustically tracking the location of a potential prey, as well as for the communication between conspecifics (e.g., Gridi-Papp and Narins, 2008). In the world of humans, spatial hearing is also a critical feature of our senses, e.g., when detecting the direction of a car approaching in traffic situations. Spatial hearing further allows us to orient in space, to suppress sound reflections from walls that interfere with the direct sound, and to segregate multiple sound sources in complex acoustic environments. Probably most important for everyday life, it facilitates communication by improving the understanding of speech in competing noise. In order to meet these demanding requirements, the human auditory system has developed highly sophisticated spatial-information-processing strategies. Extensive reviews of spatial hearing mechanisms in humans can be found in Blauert (1997), Gilkey and Anderson (1997), and Middlebrooks and Green (1991) and a review of physiological mechanisms in mammals can be found in Grothe, Pecka and McAlpine (2010). For an extensive review of general auditory perception see Moore (2008).

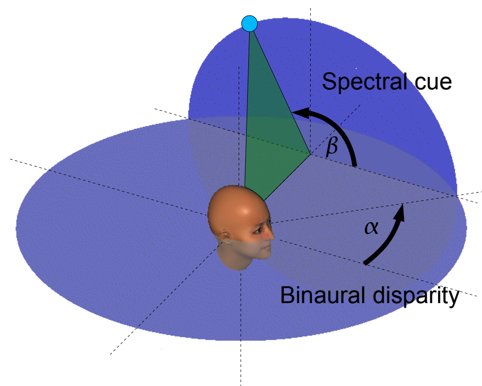

Fig. 1: Horizontal-polar coordinate system.

The spatial position of a sound source relative to the center of the recipient's head can be specified along three dimensions: a) the left/right dimension, b) the vertical dimension (front/back and elevation), and c) the distance dimension. In our work done so far we have focused on the first two of these dimensions, thus, the distance dimension is not described in this treatise. The spatial positions are specified in the horizontal-polar coordinate system (Middlebrooks, 1999b), as shown in Fig. 1. The lateral (left/right) position of a sound source is described by the *lateral angle* α, which is the angle between the median plane and the line connecting the sound source and the center of the listener's head. The vertical-plane position is described by the *polar angle ß*, specifying the angle around the horizontal pole that is given by the listener's interaural axis. According to the results of localization tests (Morimoto and Aokata, 1984; Middlebrooks, 1999b; Morimoto, 2001), the angles α and *ß* are independently determined by binaural cues and spectral-shape cues, respectively. For a description of these types of cues see sections II.2.A. and II.2.B.. The lateral angle ranges from -90° (left) to +90° (right), with 0° corresponding to the position straight ahead (in the median plane). In our studies the polar angle ranges from -30° (front, below eye-level), 0° (straight ahead), +90° (above), to +210° (rear, below eye-level). The following two subsections describe the background of human sound localization along the lateral and vertical dimensions, respectively.

A. Lateralization and Left/Right Localization

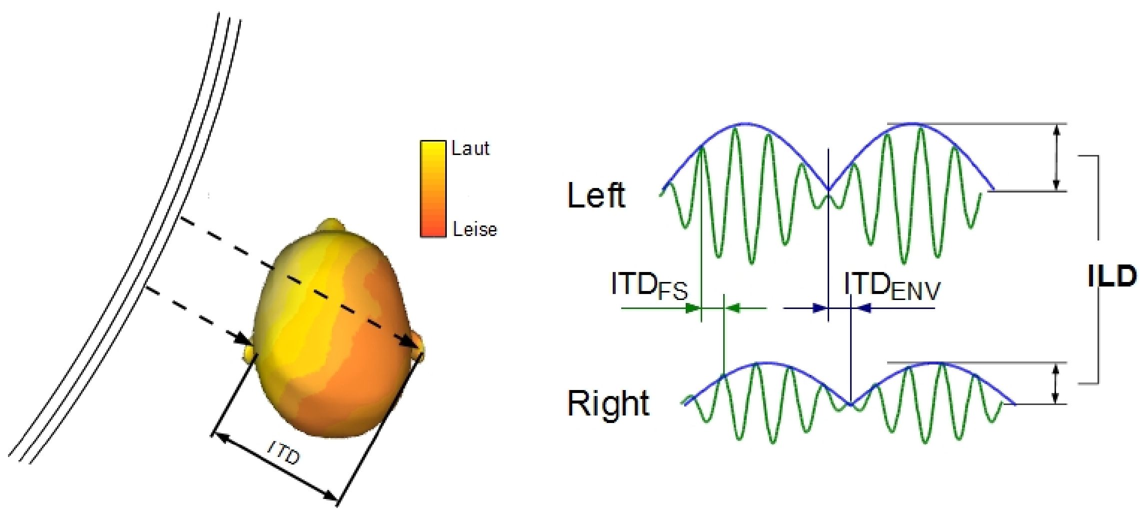

Fig. 2: Occurrence of ITD and ILD for a sound source that arrives from a location off the median plane. The right panel shows the input signals to the two ears. ITD_{FS} denotes fine structure ITD and ITD_{ENV} envelope ITD.

A sound wave arriving from a location off the median plane causes binaural disparities in the form of interaural level differences (ILDs) and interaural time differences (ITDs), as illustrated in Fig. 2. Because the stimulus shown in Fig. 2 is amplitude modulated, the ITD arises both in its fine structure and envelope, referred to as fine-structure ITD and envelope ITD. While ITDs occur across the entire audible frequency range and depend only slightly on frequency (Kuhn, 1977), ILDs are much more pronounced at high frequencies where the wavelength is small relative to the physical dimensions of the head and pinnae (Blauert, 1997). The left panel of Fig. 2 shows the ITD (in ms) as a function of the lateral angle for sinusoids of three sinusoidal frequencies and clicks. The right panels of Fig. 2 show the ILD (in dB) as a function of the lateral angle for three different frequencies (upper right panel) and as a function of the frequency for a fixed lateral angle of 90° (lower right panel).

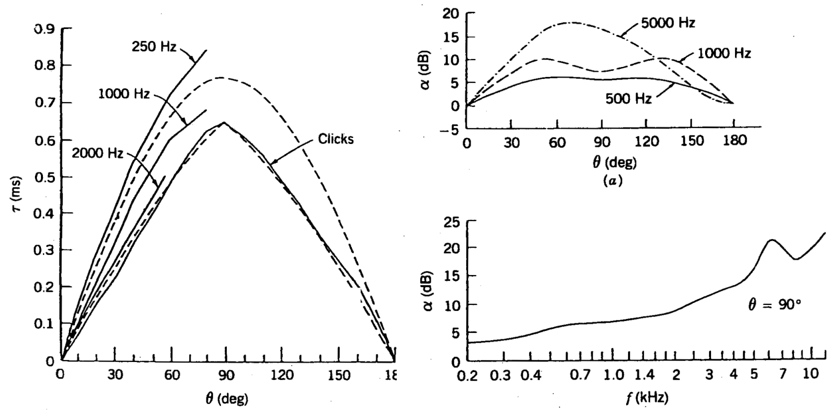

Fig. 3: Left panel: ITD (in ms) as a function of the lateral angle for sinusoids of three frequencies and clicks. Upper right panel: ILD as a function of the lateral angle for three different frequencies. Lower right panel: ILD as a function of the frequency for a fixed lateral angle of 90°. From Hafter and Trahiotis (1997).

The sensitivity to binaural disparities can be best measured by presenting stimuli via headphones, where ITDs and ILD at the ears can be accurately controlled and varied independently of each other. In order to measure the JND for one type of binaural disparity, the other type of disparity is usually set to zero. In a commonly used task, the perceived lateral position of a target sound with a non-zero binaural disparity is compared against the position of a *diotic* reference sound (i.e., with no binaural disparity).

Using such a paradigm to measure ITD-JNDs as a function of the frequency of an unmodulated sinusoid reveals increasing sensitivity up to about 1000 Hz and a decline in sensitivity thereafter (Fig. 4). The decline in sensitivity at high frequencies is a result of two processes. First, with increasing frequency of a sinusoid above about 1000 Hz the neural response is less and less correlated with the phase (or fine-structure) of the signal (Dreyer and Delgutte, 2006), resulting in impoverished ITD coding. This is referred to as the fine-structure rate limitation in ITD coding. Second, with increasing frequency, when the half-period of the sinusoid approaches the size of the ITD, the ITD cue in the phase of the ongoing signal becomes ambiguous. From the observer's point of view it is not possible to decipher on which side the ongoing signal is leading in time. For amplitude-modulated sounds at high frequencies, however, listeners can extract ITD information from the ongoing envelope (Henning, 1974; Yost, 1976; Bernstein and Trahiotis, 1994), given that the neural responses are phase-locked to the envelope. This envelope-ITD sensitivity depends on the shape of the ongoing envelope. Henning (1974) used sinusoidally amplitude-modulated (SAM) tones, and the JNDs were considerably higher than those for low-frequency pure tones. Yost (1976) and Hafter and Dye (1983) used high-frequency filtered pulse trains and the JNDs were on the order of those for low-frequency pure tones. Direct manipulation of the envelope shape has supported the idea that envelope-ITD sensitivity improves with increasing peakedness of the envelope (Bernstein and Trahiotis, 2002), as will be outlined below.

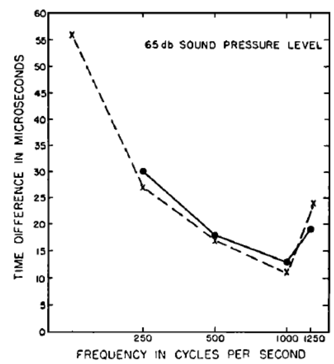

Fig. 4: ITD-JNDs as a function of the frequency of a sinusoid. From Zwislocki and Feldman (1956).

In case of ILD, the situation is simpler, as the ILD-JND is approximately constant across the audible frequency range, being in the order of 0.5-2 dB (Hershkowitz and Durlach, 1969; Grantham, 1984; Stellmack, Viemeister and Byrne, 2004; Hartmann and Constan, 2002).

Several studies measured the magnitude of lateralization evoked by various amounts of binaural disparity. For ILD, Fig. 5 shows the perceived sidedness of the auditory image as a function of the ILD, which has been measured using the method of magnitude estimation (Yost and Hafter, 1987). For ITD, Fig. 6 shows the ILDs of an acoustic pointer adjusted by the subjects to match the intracranial positions evoked by various ITD sizes (Bernstein, 2001). The circles and triangles show data for a low-frequency Gaussian noise (100-Hz-wide noise centered at 500 Hz) and a SAM tone with a center frequency of 4 kHz and a modulation frequency of 128 Hz, respectively. It is obvious that the SAM tone evokes much less lateralization than the low-frequency noise. It has been shown, however, that so-called transposed stimuli, which have been designed with the aim of providing high-frequency channels with temporal information similar to that for low-frequency sinusoids (van de Par

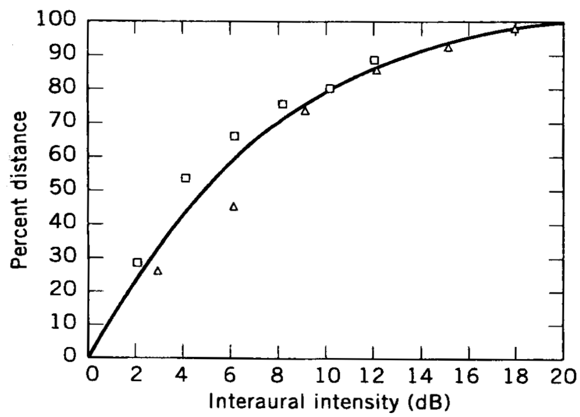

Fig. 5: Magnitude of lateralization evoked by ILD, determined by means of specifying the perceived sidedness of the auditory image as a function of the ILD. From Yost and Hafter (1987).

and Kohlrausch, 1997), elicit much higher ITD sensitivity, comparable to that for corresponding low-frequency sinusoids (Bernstein, 2001; Bernstein and Trahiotis, 2002; Bernstein and Trahiotis, 2003). This is demonstrated in Fig. 6 by the square symbols, showing that the magnitude of lateralization for a 128-Hz tone transposed to 4 kHz approaches that for the low-frequency noise.

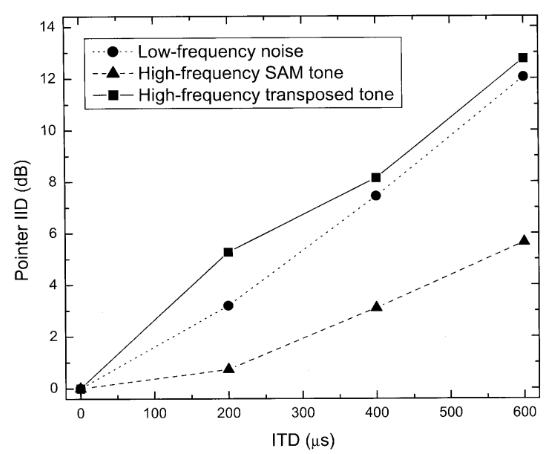

Fig. 6: Magnitude of lateralization evoked by ITD for different types of stimuli, determined by means of the ILD of an acoustic pointer that matched the intracranial position evoked the stimuli as a function of the ITD. From Bernstein (2001).

Hafter and colleagues provided evidence that the sensitivity to ITD in the ongoing signal, i.e., after the onset, decreases for high-frequency stimuli with amplitude modulation, like bandpass-filtered pulse trains, if the modulation frequency (or rate) is too high, an effect they referred to as binaural adaptation (Hafter and Dye, 1983; Hafter, Dye and Wenzel, 1983; Buell and Hafter, 1988; Hafter and Buell, 1990; Hafter, 1997). This effect is reflected in the data of Bernstein and Trahiotis (2002) that are presented in Fig. 7. For both SAM and transposed tones the ITD-JNDs fall off for modulation rates between 256 and 512 Hz while the ITD-JNDs for sinusoids with corresponding frequencies even improve within that range. Interestingly, it has been shown that, under certain conditions, introducing trigger signals during the ongoing signal causes a recovery from binaural adaptation, reflected by an improvement in ITD sensitivity (Hafter and Buell, 1990).

It is generally assumed that the binaural system extracts binaural information within the individual auditory frequency channels, referred to as auditory filters. This is supported by the finding that the binaural sensitivity declines with increasing mismatch of the carrier frequencies presented at the two ears. Fig. 8 shows the decreasing ability to discriminate between binaural in-phase vs. out-of-phase envelope modulation, a measure of ITD-sensitivity, with increasing amount of spectral mismatch between the sinusoidal high-frequency carriers (Blanks *et al.*, 2007). Similarly in case of ILD, Fig. 9 shows increasing ILD-JNDs (decreasing performance) with increasing spectral mismatch between 1/3 octave-band noises presented at the two ears (Francart and Wouters, 2007).

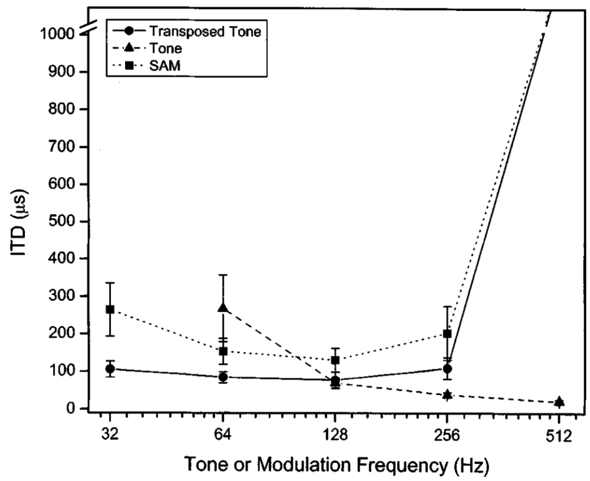

Fig. 7: ITD-JNDs as a function of pure tone or modulation frequency for different types of stimuli (see legend). From Bernstein and Trahiotis (2002).

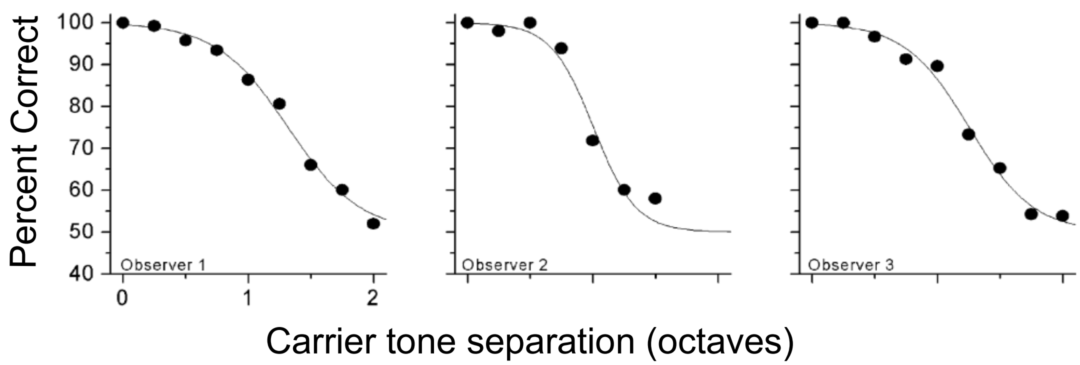

Fig. 8: Percent correct scores of discrimination between binaurally in-phase vs. out-of-phase envelope modulation, a measure of ITD-sensitivity, as a function of the spectral mismatch between the sinusoidal high-frequency carriers. The starting phases of the modulator were random from trial to trial. Results for three listeners are shown. From Blanks et al. (2007).

In contrast to the lateralization tasks mentioned above, testing the localization of sound sources requires the sound source to be played back by a loudspeaker placed at different positions around the listener or by means of binaural virtual acoustics (see section II.2.B.). In a typical horizontal-plane localization task an array of loudspeakers is distributed along a circle or semicircle centered at the listener's head. The listener indicates the perceived spatial position either by some type of pointing device or by some more abstract direction specification (e.g., the loudspeaker number). The accuracy of horizontal sound localization is generally highest for frontal sound sources and decreases with increasing lateral angle α (Blauert, 1997).

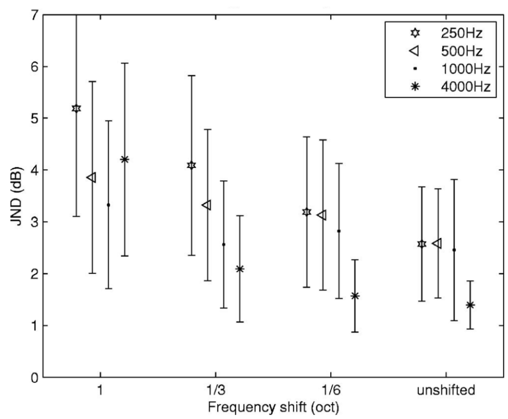

Fig. 9: ILD-JNDs as a function of the spectral mismatch between 1/3 octave-band noises presented at the two ears. From Francart and Wouters (2007).

Given that natural broadband sounds contain both ITD and ILD cues, an important question is how these cues contribute to sound localization in the left/right dimension. According to the duplex theory of sound localization (Strutt, 1907), the auditory system uses primarily ITD at low frequencies and ILD at high frequencies to determine the spatial position of a sound source. The development of binaural virtual acoustics techniques made it possible for the first time to test this theory. In binaural virtual acoustics, 2-D spatial hearing is simulated by presenting stimuli that are filtered with individually measured head-related transfer functions (HRTFs) via headphones (for more details about this technique see section II.2.B.). Using this technique allows to modify the phase and amplitude properties of HRTFs and thus to study the independent contribution of ITD and ILD cues to sound localization (Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002). These studies showed that ITD dominates left/right localization for stimuli containing low-frequency components. Only for high-pass filtered stimuli, ILD dominates for most listeners. When introducing amplitude modulation in high-pass filtered stimuli, the contribution of envelope ITD was shown to increase (Macpherson and Middlebrooks, 2002).

B. Vertical-Plane Localization

To the extent that the head and ears are symmetrical, sound waves arriving from different positions along a given vertical plane, e.g., the median plane, all result in constant binaural disparities. Fig. 10 shows contours of constant ITD (in µs) and constant ILD (in dB) from acoustic measurements on actual listeners. The observation that the contour for a given binaural disparity always fall on a single vertical plane suggests that binaural information does not help the listener to determine the vertical position of the sound source. Rather, listeners use the direction-specific filtering effect of the external ear, particularly the pinna, and of the head and torso to determine the sound's vertical position. The left panel of Fig. 11 illustrates that, depending on the vertical angle of incidence of a sound source, the sound wave reflects from different tiny structures of the pinna. The addition of these reflections with the respective direct sounds results in comb-filter effects causing direction-dependent peaks and notches in the spectrum of the sound entering the ear canal (see right panel of Fig. 11). The direction-dependent filter effect of the pinnae, head and torso can be described by the HRTFs. Because it has been shown that listeners use the spectral features of HRTFs in order to determine the vertical position of a sound source (e.g., Blauert, 1997; Wightman and Kistler, 1989b; Kistler and Wightman, 1992), these features have been termed spectral localization cues or simply spectral cues. Another common term is pinna cues.

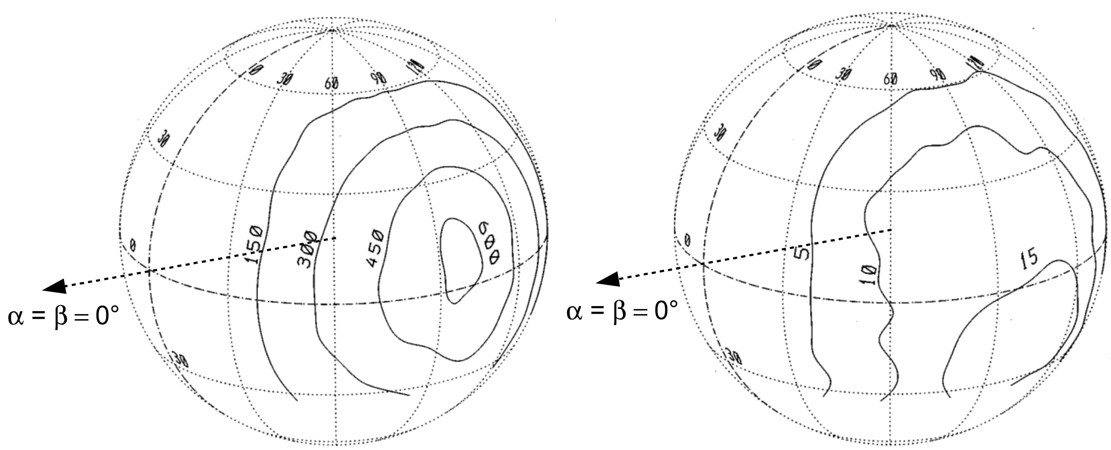

Fig. 10: Contours of constant ITD in µs (left panel) and constant ILD in dB (right panel) from acoustic measurements on actual listeners. The contours indicate that binaural cues do not help to determine the vertical position of a sound source. From Wightman and Kistler (1999).

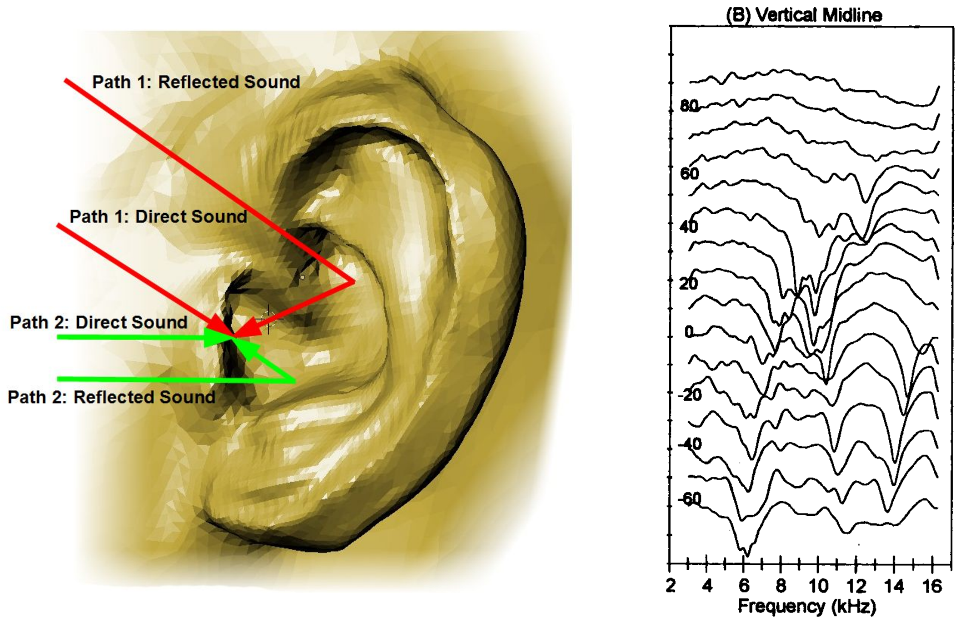

Fig. 11: Illustration of the direction-dependent reflections of sound waves from pinna structures (left panel), causing direction-dependent peaks and notches in the spectrum of the resulting sound at the ear canals (right panel). Figure comprising right panel from Middlebrooks (1992).

It has been shown that sounds which are presented in the free field and thus filtered by the natural HRTFs of the respective listener are perceived as "externalized", i.e. perceived outside of the head. In contrast, modifying the listener-specific HRTFs causes sounds to be perceived inside the head (Blauert, 1997). The technique of binaural virtual acoustics has made it possible to perceive and localize sounds presented via headphones such as if they were presented in the free-field. This technique requires a sound source to be processed, i.e. filtered, with the HRTF corresponding for the desired direction. In order to obtain listener-specific HRTFs, measurement microphones are placed at the ear canals of a listener. Presentation of excitation signals from various directions around the listener and subsequent system identification allows to derive a set of HRTFs (see e.g., Wightman and Kistler, 1989a; Middlebrooks, 1999a; Blauert, 1997). HRTFs differ considerably between individual persons, like fingerprints. Interestingly, experiments have shown that, in order to achieve natural externalization and accurate vertical-plane localization performance, listener-specific HRTFs are required. It is noteworthy that the front-back discrimination of sound sources does not exclusively rely on spectral-shape cues but can also be achieved by means of head movements combined with tracking of the resulting dynamic binaural cues (Perrett and Noble, 1997). However, is has been suggested that head movements have relatively little practical relevance in discriminating front from back and it is evident that they do not help at all for short sounds (Middlebrooks and Green, 1991).

An important requirement for evaluating spectral localization cues is the ability to discriminate different spectral shapes, an ability called spectral profile analysis (Green, 1988). Profile analysis is often thought to be based on comparison of excitation levels across auditory filters, thus, it should not depend on the absolute stimulus level. In order to rule out absolute level as a discrimination cue, experiments on profile analysis have often incorporated roving of the broadband stimulus level across the intervals of a trial. Although level roving involves some degree of uncertainty, listeners have been shown to be able to detect small spectral changes even with a level rove range of up to 40 dB (Mason *et al.*, 1984). There are indications that profile analysis may at least partly be based on the evaluation of changes in the temporal pattern at the output of the auditory filter centered at the frequency of the spectral change (Green *et al.*, 1992; Young and Sachs, 1979).

C. <u>Auditory Grouping Mechanisms in Spatial Hearing</u>

In order to structure visual and auditory information and, in particular, to decompose the visual or acoustic representation of mixtures of physical objects into separate perceptual objects, our brain relies on so-called Gestalt principles (Köhler, 1940; Wertheimer, 1945; Goldstein, 2009). Examples are the grouping of similar or proximate elements into an object. In the research field of auditory scene analysis (Bregman, 1990), several signal features have been identified on which the Gestalt principles are applied and which are thus important for the segregation of acoustic sources, including the temporal structure, spectral separation, intensity, pitch, timbre, synchrony, harmonicity, and location in space. Interestingly, the location cues alone appear to be weak segregation cues (e.g., Culling and Summerfield, 1995; Drennan, Gatehouse and Lever, 2003). However, in combination with monaural segregation cues, spatial cues have a strong effect, largely facilitating the intelligibility of a target speech source in spatially separated competing speech sources (Kidd *et al.*, 1998; Marrone, Mason and Kidd, 2008; Schwartz, McDermott and Shinn-Cunningham, 2012).

3. Cochlear Implants and Spatial Hearing

Cochlear implants are auditory prostheses that attempt to restore auditory perception in the deaf or profoundly hearing impaired. For recent reviews on cochlear implants, see Zeng, Popper and Fay (2004), Clark (2003), Wilson and Dorman (2008), or Zeng *et al.* (2008). The basic concept is to electrically stimulate the auditory nerve by means of an electrode array that is inserted into the cochlea (see Fig. 12), thus, bypassing the peripheral processing of acoustic signals up to the level of the auditory nerve. This peripheral processing in normal hearing mainly consists of directional filtering of the pinna, resonance of the ear canal, impedance-adjustment in the middle ear, frequency-to-tonotopic-place transformation in the cochlea, active amplification of mechanical vibrations in the cochlea by the outer hair cells, and mechano-electric transduction of the inner hair cells. In order to account for the frequency-to-place transformation in acoustic hearing, modern CIs separate the acoustic signal captured by the microphone into several frequency channels. Within each channel, the temporal envelope is extracted, compressed, and then used to modulate an electric pulse train. The output of each channel is allocated to one of the tonotopically distributed electrode contacts in the cochlea. Such a processing scheme is referred to as envelope-based stimulation strategy. Recently, researchers and manufacturers have started to encode not only the envelope information but also the temporal fine-structure information in each channel by locking the timing of electrical pulses to the temporal fine structure in each channel. Such strategies are referred to as fine-structure-based stimulation strategies.

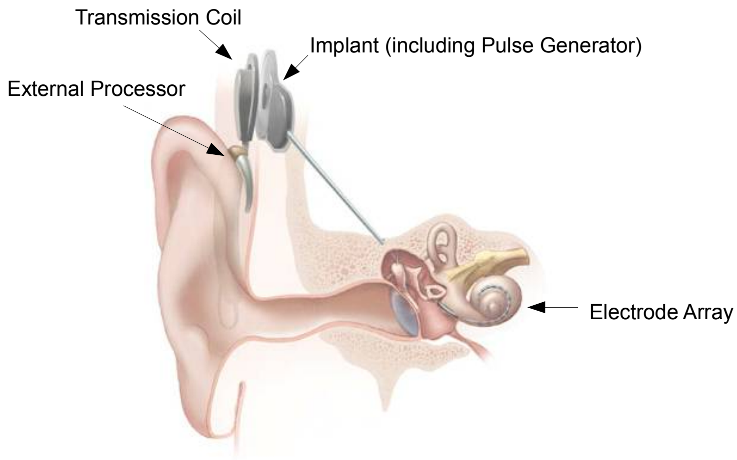

Fig. 12: Schematics of a cochlear implant system.

As of 2010, about 220.000 people worldwide have received CIs, approximately 50% of which are children. Postlingually deafened implantees or children implanted within the first few years often achieve high levels of speech understanding in quiet (e.g., Wilson *et al.*, 1991).

In order to provide CI listeners with the advantages of spatial hearing in the left/right dimension, CIs have been increasingly implanted at both ears (bilaterally). Although current CIs have not been designed for binaural hearing, several studies have shown perceptual advantages of bilateral supply. The accuracy of horizontal sound localization is higher compared to unilateral supply (e.g., Nopp, Schleich and D'Haese, 2004; Grantham *et al.*, 2007; van Hoesel and Tyler, 2003). With respect to speech understanding in noise, bilateral CI listeners benefit from head shadow effects in case of spatial separation of target and noise sources (Schleich, Nopp and D'Haese, 2004; van Hoesel and Tyler, 2003; van Hoesel *et al.*, 2008). However, the sound localization performance of bilateral CI listeners is still largely degraded compared to NH listeners (Litovsky, Parkinson and Arcaroli, 2009; Grantham *et al.*, 2007; Seeber, Baumann and Fastl, 2004; Majdak, Goupell and Laback, 2011) and the binaural unmasking effect, i.e. the improvement of signal (e.g., speech) audibility in spatially separated noise due to auditory processing of binaural cues, is very small or not observed at all (van Hoesel and Tyler, 2003; van Hoesel *et al.*, 2008). It is noteworthy that bilateral implantation has the additional advantage of providing a backup system in case of system failure and the potential advantage of an increased number of independent frequency channels.

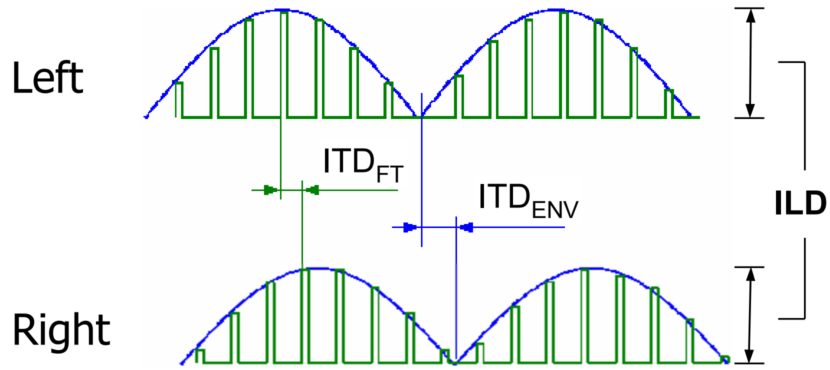

Fig. 13: Different types of binaural cues for amplitude-modulated electric pulse trains. Note that only the positive phases of the pulses are shown. ITD$_{FT}$ denotes fine timing ITD and ITDENV envelope ITD.

There are several potential reasons why binaural hearing may be degraded in bilateral CI listeners. First, current bilateral CI systems may not provide appropriate binaural information. Fig. 13 shows the different types of binaural cues that can be encoded by amplitude-modulated electric pulse trains. Note that ITD in the electric carrier pulse train is referred to as fine-timing ITD, instead of the term fine-structure ITD as used for acoustic stimuli. Second, the basic sensitivity to binaural cues (ITDs and ILDs) might be reduced in CI listeners, even under optimal representation of these cues by means of direct and interaurally coordinated presentation of electrical stimuli via a CI research system. Such a deficit could be due to a potential deprivation of the binaural auditory system as a consequence of lack of appropriate binaural input. It could also be caused by the specific properties of electric stimulation compared to acoustic stimulation. This including, for example, the failure to stimulate the most apical tonotopic region assumed to be important for ITD coding (Macpherson and Middlebrooks, 2002), the high degree of phase locking and across-fiber synchrony in the neural response to electric stimulation (Dynes and Delgutte, 1992), or the altered temporal dispersion across nerve fibers due to the lack of cochlear traveling wave (Colburn *et al.*, 2009).

In the studies described in section II.4., we attempted to address the following questions with respect to spatial hearing in the left/right dimension with CIs:

a) How sensitive are bilateral CI listeners to binaural cues presented via a research system?

b) To what extent do current bilateral CI systems encode binaural information?

c) How could the sensitivity to binaural information be improved?

In contrast to spatial hearing in the left/right dimension, spatial hearing in the vertical dimension with CIs has been a completely new field when we started the project CI-HRTF. Vertical-plane localization with CIs is complicated by a number of factors. First, the spectral resolution is coarse compared to normal (acoustic) hearing. Although state-of-the-art CIs have up to 22 electrodes, channel interactions limit the number of perceptually independent channels to maximally about 10-12 in most cases (e.g., Cohen *et al.*, 1996; Friesen *et al.*, 2001). Thus, it is not clear if CI listeners are able to resolve spectral localization cues. Second, it is commonly assumed that vertical-plane localization in normal hearing relies on frequencies from about 4 to 16 kHz (e.g., Hebrank and Wright, 1974; Wightman and Kistler, 1997). However, the characteristic frequency of the most basal neurons that can be stimulated with CIs is usually not higher than about 8.5 kHz (Vermeire *et al.*, 2008; Carlyon *et al.*, 2010). While the high-frequency localization cues can be mapped to the stimulation range available in electric hearing, it is unclear if the auditory system is able to adapt to such a manipulation. Third, it is unclear if CI listeners are able to discriminate between different spectral shapes, i.e. to perform spectral profile analysis. The studies published so far (Drennan and Pfingst, 2006) did not test conditions where absolute level could be ruled out as a discrimination cue. This is important because in real-life situations the listener does not *a priori* know the level of the sound to be localized. In the studies described in II.5. we attempted to address all these questions.

4. Specific Studies Towards Better Understanding of Left/Right Localization

A. <u>ILD and Envelope-ITD Perception with Clinical CI Systems and in Normal Hearing</u>

In Laback *et al.* (2004) we addressed the question to what extent state-of-the-art CI systems provide bilateral CI listeners with ILD and ITD cues for sound lateralization. NH listeners were used as a control group. Using different types of broadband stimuli, we found that ILD sensitivity was generally high, irrespective of the spectral or temporal stimulus properties, while ITD sensitivity was generally low and JNDs were only measurable for stimuli with a pronounced and slow temporal modulation. The second part of the study focused on narrowband stimuli. Because it is known that in normal hearing the binaural system extracts ILD and ITD cues within auditory filters, thus requiring frequency matching of the carrier signals at the two ears (Perrott, Briggs and Perrott, 1970; Scharf, Florentine and Meiselman, 1976; Blanks *et al.*, 2007; Francart and Wouters, 2007), extensive pre-experiments had to be performed with the CI listeners in order to determine interaural electrode pairs best matching in pitch. For unmodulated sinusoids passed through the CI processors and mapped to pitch-matched electrode pairs, the CI listeners showed no ITD sensitivity at all. In accordance with an analysis of the electric output of the processors, this demonstrated that fine-structure ITD information in the acoustic input stimuli is completely discarded, resulting in no usable lateralization cues. ILD sensitivity for sinusoids with pitch-matched electrodes was found to be comparable to broadband stimuli and to decrease systematically with increasing pitch-mismatch of the stimulating electrodes.

B. <u>Sensitivity to Fine-structure, Envelope, and Gating-ITD</u>

In two subsequent studies (Majdak, Laback and Baumgartner, 2006; Laback, Majdak and Baumgartner, 2007) we studied the sensitivity of CI listeners to ITD presented in different signal parts, namely in the fine timing, the ongoing envelope modulation, the onset, and the offset. By using an interaurally synchronized research interface, the stimulus timing of the pulses could be controlled with an accuracy better than 2.5 μs. NH listeners were tested with an acoustic simulation of pulsatile electric stimulation, namely bandpass-filtered acoustic pulse trains. One main finding from these studies was that CI listeners are sensitive to fine-timing ITD if the pulse rate does not exceed a few hundreds pulses per second (pps). Furthermore, low-rate ongoing envelope modulation (typical for speech) was shown to contribute relatively little to lateralization when fine-timing ITD was also present. Finally, CI listeners showed some sensitivity to onset ITD if the pulse rate was high and the stimulus duration was short, but they showed no sensitivity at all to offset ITD. An important conclusion from these studies was that in order to provide CI listeners with salient ITD cues, the ITD information has to be encoded in the pulse timing, which is not the case for commonly used envelope-based strategies.

C. Improvement of ITD Sensitivity with Binaurally-Synchronized Jitter

The preceding studies showed the potential for providing CI listeners with salient spatial information about sound sources by encoding fine-timing ITD. However, the observed limitation of fine-timing ITD sensitivity to low pulse rates (< 300 pps) poses a severe practical limitation, because higher pulse rates are required in order to properly encode speech information. While addressing this issue, we found an interesting analogy between the pulse rate limitation in electric hearing and a limitation in ongoing envelope-ITD sensitivity at higher modulation rates, the so-called binaural adaptation phenomenon reported in the NH literature and described in section II.2.A.. Even more interesting was the finding of follow-up studies showing that introducing trigger signals during the ongoing signal caused a release from binaural adaptation. Inspired by these findings and assuming that the electric pulse-rate limitation and the binaural adaptation phenomenon have related origins, we tested the hypothesis that introducing binaurally synchronized jitter in binaural electric pulse sequences (thus preserving the ongoing ITD cue) improves ITD sensitivity. By testing CI listeners, we demonstrated that for pulse rates exceeding 400 pps, jitter indeed causes large improvements in ITD sensitivity compared to periodic pulse trains (Laback and Majdak, 2008c; Laback and Majdak, 2008a). In a follow-up study with NH listeners (Goupell, Laback and Majdak, 2009) we showed comparable effects of jitter for acoustic pulse trains. That study also revealed that the jitter effect in normal hearing is most likely based on temporal changes, in contrast to the interpretation of literature studies that attributed the release from binaural adaptation to the concomitant spectral changes induced by the trigger signals. Furthermore, using a physiologically-based model of the auditory nerve and brainstem we found that the random timing of the jittered pulses increases firing synchrony in the auditory periphery, causing improved rate-ITD tuning. In a subsequent cooperation study with the Massachusetts Eye and Ear Infirmary (Harvard Medical School), the jitter effect had been replicated in physiological measurements in the inferior colliculus of bilaterally implanted cats and compared to the effects in CI and NH listeners (Goupell *et al.*, 2010; Laback, 2012).

The jitter effect has been replicated subsequently in another acoustic-hearing study (Brown and Stecker, 2011) and further physiological studies with bilaterally implanted cats and rabbits (Hancock, Chung and Delgutte, 2012; Hancock *et al.*, 2011). Particularly the physiological studies supported the interpretation that at least part of the jitter effect can be attributed to peripheral neural refractory effects. On the other hand, the detailed mechanism underlying the jitter effect, particularly the potential contribution of more central mechanisms, and the relation to the binaural adaptation effect are not fully understood yet. With respect to application for CIs, we proposed a scheme to incorporate binaurally-synchronized jitter in future bilateral CI systems or hearing aids (Laback and Majdak, 2008b; Laback and Majdak, 2009) which appears to be promising to improve the CI listeners' access to ITD cues and thus to improve their performance in sound localization and speech understanding in noise.

D. Does the Place of Stimulation Affect the Rate Limitation in Acoustic ITD Perception?

Acoustical simulations of pulsatile electric stimulation in NH listeners often use pulse trains that are bandpass-filtered in a frequency region high enough so that the individual partials are unresolved by the auditory system. An important question arising with such simulations is if the temporal smearing effect resulting from auditory filtering and thus the choice of the center frequency of the bandpass filter is responsible for the described pulse-rate limitation in ITD perception (which also occurs in rate-pitch perception). Because auditory filtering is bypassed in electric hearing, such an influence could potentially confound the comparison between the NH and CI listener's performance.

We investigated the potential effect of auditory filtering in acoustic hearing by systematically studying the effects of center frequency and pulse rate on ITD sensitivity in high-frequency filtered pulse trains in NH listeners (Majdak and Laback, 2009). For all center frequencies tested (up to 9.2 kHz), ITD sensitivity decreased with increasing rate, yielding a rate limit of approximately 500 pps. On average, the sensitivity slightly decreased with increasing center frequency. The lack of an interaction between pulse rate and center frequency indicated that auditory filtering is *not* responsible for the pulse rate limitation. Thus, we concluded that the comparison between NH and CI listeners in previous studies was not confounded by the choice of the center frequency of stimulation in NH listeners. Furthermore, the results revealed surprisingly high ITD sensitivity at very high center frequencies where previous studies found much poorer sensitivity. This can be attributed to the stimulus bandwidth being a constant proportion of the center frequency in our study, in contrast to the constant absolute bandwidth used in previous studies that resulted in relatively less pronounced envelope modulation at high center frequencies.

E. Effects of Envelope Shape on Envelope-ITD Perception

Several normal-hearing studies are available which showed that the sensitivity to envelope-ITD depends on the peakedness of the temporal envelope (e.g., Bernstein and Trahiotis, 2002; Dreyer and Delgutte, 2006; Griffin *et al.*, 2005). These studies were, however, not designed to reveal which aspects of the envelope shape are important for ITD perception. In experiments with NH and CI listeners (Laback *et al.*, 2011b; Laback, Zimmermann and Majdak, 2010), we systematically and independently varied different features of the ongoing envelope shape, including the off time (the silent period in each modulation cycle), the slope steepness, and the peak level. For the CI listeners, ITD sensitivity improved with increasing off time and peak level (Laback *et al.*, 2011b; Laback, Zimmermann and Majdak, 2010), while the slope steepness had no systematic effect. Based on those results we suggested that envelope ITD perception in electric hearing could be improved by signal processing methods that sharpen the envelope representation.

For the NH listeners, also the slope steepness had a significant effect and there was no interaction between the effects of the off time and slope. We showed that the improvement in ITD sensitivity for transposed tones compared to SAM tones as reported in the literature can be explained by the combined effects of the off time and the slope. The overall effects of the parameters off time and slope were consistent with an acoustic-hearing study performed at the same time at the University of Oldenburg (Klein-Hennig *et al.*, 2011), in which a different type of envelope was used. However, steepness of ITD JNDs as a function of the off time was found to differ considerable between the two studies. In a follow-up co-operation study with the University of Oldenburg (Dietz *et al.*, 2013) we compared the two types of envelopes using matched parameters and showed that the differences mostly vanish when employing a perception-based metric for the off time that takes the modulation depth sensitivity into account.

F. Multiple-channel Stimulation, Binaural Interference, and Auditory Grouping

The studies comparing binaural sensitivity (ITD and ILD) in direct electric stimulation and acoustic stimulation described so far involved the stimulation of a single interaural electrode pair and a narrowband stimulus, respectively. To our knowledge, no published studies investigated binaural sensitivity with simultaneous[2] stimulation of multiple electrodes. It is well known from monaural CI studies that adjacent and often even more distant electrodes are subject to neural-population interactions (see e.g., Shannon, 1993).

As a first step towards more real-life conditions, i.e. broadband sounds and/or multiple sound sources, we studied binaural sensitivity with multiple-electrode stimulation. In particular, we studied the so-called binaural interference effect and the influence of monaural auditory grouping mechanisms (Best, Laback and Majdak, 2011). In normal hearing, binaural interference refers to the reduced sensitivity to binaural cues of a target sound by the presence of a simultaneous interferer sound in a remote frequency region. This effect appears to be the consequence of a monaural simultaneous grouping mechanism that groups spectral components which are likely to belong to the same sound source into a single auditory object (based on factors responsible for grouping described in section 2.C.) . We tested the occurrence of this effect in electric hearing by asking six CI listeners to lateralize a target pulse train presented at a single electrode pair, as a function of stimulus ITD or ILD, either in isolation or in presence of a quasi-diotic interferer pulse train that was tonotopically maximally remote from the target pair. Five out of six CI listeners showed binaural interference, i.e. a reduced magnitude of lateralization of the target in presence of the interferer. In an additional test condition, the interferer was presented as part of an ongoing stream of identical interferer pulse trains. For that condition, all five listeners who showed binaural interference also showed a robust recovery from interference, consistent with the hypothesis that as a result of sequential grouping the listeners now perceived the interferer as part of an ongoing stream of auditory events and no longer fused the interferer with the target. Indeed, for that condition the listeners reported to hear two dis-

[2] When using the term "simultaneous" we refer also to conditions were pulses are presented in an interleaved fashion across different electrodes in order to avoid electric-field interactions, thus, strictly-speaking representing non-simultaneous stimulation.

tinct sound sources. The recovery from binaural interference indicated that binaural interference arises at least partly at a central auditory processing level. Overall, the results represent an objective demonstration of simultaneous and sequential grouping mechanisms in electric hearing.

5. Specific Studies Towards Better Understanding of Vertical-Plane Localization

The work on these studies was funded by the FWF, project P18401-B15, granted to myself in 2005.

A. Development of a Method for Learning and Testing 2-D Sound Localization

Providing CI listeners access to spectral localization cues will inevitably require some form of adaptation or remapping of their spatial auditory system to the new spatial information. Therefore, the first step of our series of studies was to develop a method that allows to both learn and test 2-D sound localization. An intuitive way of learning new sound localization cues is by means of audio-visual calibration. Therefore, we developed an experimental system where the listener is presented with sound sources in virtual 2-D-space while being immersed in a spherical virtual visual environment. In Majdak, Goupell and Laback (2010) we studied the effect of a) visual feedback condition (visual environment or darkness) and b) the type of pointing method (head or manual pointing). The visual environment was found to result in better localization performance compared to darkness, but the pointing method had no significant effect. The results led to a reliable localization learning/testing method with a virtual audio-visual environment and a hand-held pointing device. The manual pointing device was chosen because of its ergonomic advantages compared to the head-pointing method.

B. Vertical-Plane Localization of CI Listeners Using Clinical Devices

We attempted to determine if or to what extent bilateral CI listeners are able to localize in the vertical planes using their clinical speech processors. Using the 2-D localization method developed in the previous step, virtual acoustic stimuli, based on subject-specific HRTFs, were presented directly into the audio inputs of the processors. The CI listeners showed generally much worse performance compared to NH listeners both in the horizontal and vertical dimensions. Interestingly, with a small amount of level roving, the front/back-discrimination performance was above chance level. However, the performance decreased to chance level when the amount of level roving was increased. Correlation analyses between the target angle, response angle, and signal level supported the conclusion that front-back discrimination heavily relied on the signal level rather than on spectral localization cues. Given that in real-life situations the signal level is not *a priori* known, these results demonstrated the need to explicitly encode spectral localization cues in CIs.

C. Spectral Sensitivity and Profile Analysis in Electric Hearing

Vertical-plane localization relies mostly on the perception and discrimination of spectral features of the sounds entering the ears. Therefore, crucial questions are to what extent CI listeners are able 1) to resolve spectral features such as peaks and notches and 2) to discriminate different spectral shapes. With respect to the resolution of spectral features, our study on the recognition and auditory representation of vowellike stimuli in electric hearing (Laback, Deutsch and Baumgartner, 2004) provided valuable insight. In that study, confusion matrices for eight steady-state vowels were determined and then masking patterns, i.e., internal spectral representations, for these stimuli were measured. The analysis of results revealed that both the peak frequencies and the between-peak spectral structure were reasonably well represented in the masking patterns and those features were actually used by the listeners to discriminate between the vowels. Overall, assuming a comparable spectral complexity of vowels and spectral localization cues, these results suggested that the spectral resolution of CIs may be sufficient for vertical-plane localization.

With respect to the discrimination between different spectral shapes in electric hearing, no studies were available that excluded absolute level as a discrimination cue. Thus, we performed a systematic spectral-shape discrimination study with six CI listeners (Goupell *et al.*, 2008a). In experiment 1, listeners detected spectral peaks or notches imposed on a constant-loudness background. The listeners' performance was good without level roving, but dropped dramatically with level roving, especially for notches. In experiment 2, listeners discriminated changes in the peak heights and notch depths. The outcome was roughly similar as in experiment 1. In experiment 3, listeners detected a change in the electrode position of peaks or notches. The JNDs amounted to about 1 electrode. In experiment 4, all previous experiments were repeated with larger level roving. We found no evidence that CI listeners performed an across-channel comparison in any of these tasks. In addition, the effect of the parameters peak/notch width and tonotopic peak/notch position were tested. Overall, the results suggested that CI listeners may not be able to discriminate different spectral shapes, as required for vertical-plane localization if the overall stimulus level is not *a priori* known.

Two follow-up studies found some evidence of across-electrode comparison in spectral-shape discrimination of CI listeners (Won *et al.*, 2011; Azadpour and McKay, 2012), although they did not apply such a rigorous method to rule out within-channel and overall level cues. In any case, it should be considered that our stimuli (and those of Azadpour and McKay (2012)) provided exclusively spectral cues. Given that temporal cues may play some role in profile analysis in normal hearing (Green *et al.*, 1992), it is possible that explicitly encoding temporal cues in electric hearing would allow for true profile analysis in CI listeners. Studying the potential contribution of temporal cues will be the focus of a follow-up project.

D. Frequency Channels and Electrodes Required for Speech Intelligibility: Room for Spatial Hearing

Encoding spectral localization cues with CIs could potentially compromise speech intelligibility by altering important spectral speech cues. It is therefore important to determine which frequency channels and corresponding electrodes are required for understanding speech in quiet and noise which in turn allows to identify those electrodes that are "free" for encoding spectral localization cues. To that end, we performed speech intelligibility experiments with six CI listeners implanted with an 12-electrode array that covers a wide tonotopic range (Goupell *et al.*, 2008b). Assuming a much larger potential for "free" electrodes in the basal (high-frequency) region, we independently varied the upper boundaries of the processing frequency range and the electric stimulation range while keeping the respective lower boundaries fixed. For the matched conditions, where the estimated tonotopic stimulation places corresponded to the channel frequencies, we found that only 8-10 electrodes are required to obtain high speech intelligibility in quiet and noise. Thus, 2 to 4 electrodes appear to be available for exclusive coding of spectral localization cues. For the unmatched conditions, involving either spectral expansion or compression, changes in frequency-place mapping of less than about 0.8 octaves caused no degradation in performance. The results provide a basis for the optimal allocation of spectral channels to electrodes with respect to speech understanding in noise while leaving maximal room for encoding spectral localization cues.

E. Number of Channels Required for Vertical-Plane Localization

The quite limited number of frequency channels available with current CI systems may pose a fundamental restriction for vertical-plane localization. Studies with NH listeners indicated that broadly tuned spectral cues are important rather than fine spectral details (e.g., Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002), which may suggest that the small number of channels available with CIs is sufficient for vertical-plane localization. However, the particular signal manipulations used to reduce the spectral resolution in those normal-hearing studies make it difficult to conclude on the critical number of channels in a vocoder system such as a CI. To that end, we performed median-plane localization experiments with eight NH listeners using a vocoder, varying the number of spectral channels from 3 to 24 (Goupell, Majdak and Laback, 2010). The channels were logarithmically spaced within the range from 0.3 to 16 kHz. Without feedback training, the performance did not significantly change with decreasing number of channels from 18 to 9, although it was slightly worse than for unprocessed broadband click trains or noise stimuli. A separate experiment with feedback training was performed using two different 12-channel mappings, one with logarithmic channel spacing and one allocating more channels to the speech range. With the latter mapping, called speech-localization (SL) mapping, we attempted to allow for both good speech understanding and vertical-plane localization. For both conditions the performance was slightly worse than for the unprocessed conditions. Furthermore, speech perception in quiet and noise for the SL condition was as good as for a standard configuration used in clinical CI processors. These experiments suggest that the number of spectral channels available with current CIs could be sufficient for vertical-plane localization.

F. Learning Vertical-Plane Localization with Altered Spectral Cues in Normal Hearing

Providing CI listeners with high-frequency spectral localization cues that range from about 4 to 16 kHz requires to map these cues to the available electric stimulation range which usually has an upper limit of about 8.5 kHz. Because this involves a shift (or more precisely a warping) of spectral information, it can be assumed that listeners cannot use this information immediately, but it may be possible that listeners can adapt to the manipulation by means of training. Experiments on monaural speech intelligibility with spectrally shifted speech (e.g., Fu, Nogaki and Galvin, 2005) or on binaural vertical-plane localization with alterations of spectral localization cues within a given frequency range (Hofman, Van Riswick and Van Opstal, 1998) showed strong adaptation effects.

In a long-time localization-training experiment (Majdak, Walder and Laback, conditionally accepted; Majdak, Walder and Laback, 2011), we studied the effect of warping frequency channels from 2.8 to 16 kHz to a stimulation range from 2.8 to 8.5 kHz, leaving the channels from 0.3 to 2.8 kHz unmodified. Using the virtual audio-visual environment described in section II.5.A., thirteen NH listeners completed two-hour daily training over a period of three weeks. One group of listeners (target group) listened to frequency-warped stimuli, the other group (control group) to stimuli low-pass filtered at 8.5 kHz. The study was implemented using a double-blind design. The control group showed a large localization error rate at the beginning, while at the end of the training they essentially reached their baseline performance (for stimuli with unmodified spectral cues). The target group showed a very large error rate at the beginning and a slow but continuous improvement that does not appear to have saturated at the end of the training. These results are promising with respect to the idea to present high-frequency spectral localization cues to the stimulation range available with CIs. The strong adaptation effect seen for the control group is encouraging for people suffering from high-frequency hearing loss. The results also have implications for the general understanding of spatial hearing mechanisms. While previous studies have already shown the plasticity of the spatial hearing mechanisms with respect to alterations of spectral cues within a given (high) frequency region, our results show also strong plasticity when mapping high-frequency spectral cues towards lower frequencies or when removing them at all.

6. Potential Implications for Speech Perception in Noise

It is well known that having access to spatial information about our acoustic environment facilitates understanding speech in competing noise (e.g., Bronkhorst, 2000). It can therefore be assumed that improving the access of CI listeners (and possibly also users of other hearing devices) to spatial cues not only enhances their sound localization performance but also improves their ability to understand speech in noise. In NH listeners, advantages of spatial hearing for speech understanding arise when the target speech source, *T*, is spatially separated from the interfering source(s), *I*. One advantage results from the acoustic head shadow, i.e. the ability of the listener to focus on the ear with the better signal-to-noise ratio (SNR). Another advantage results from auditory processing of spatial cues, i.e., ITDs, ILDs, and monaural spectral localization cues. In the following we focus on the advantages resulting from the processing of specific types of spatial cues.

A. Spatial-Hearing Advantage for Lateral Separation of Target and Interferer

The spatial-hearing advantages for configurations involving lateral (left/right) separation of *T* and *I* have been excessively studied with NH listeners. In a free-field configuration with *T* presented from the front and *I* from the side, listening with both ears improves speech intelligibility relative to listening with the ear remote from *I* only (thus, having the higher SNR), an effect referred to as binaural unmasking. In headphone listening, a classic demonstration of binaural unmasking is the much better audibility of a diotic tone presented in noise when the noise is interaurally phase-inverted (S_0N_π) than when it is in phase (S_0N_0). Binaural unmasking can be understood as a release from energetic masking due to binaural auditory processing (Durlach, 1963). Recent studies indicate that ITDs are particularly important for binaural unmasking (van der Heijden and Joris, 2010; Culling, Hawley and Litovsky, 2004). Under laboratory conditions, CI listeners have shown some binaural unmasking (Long *et al.*, 2006; Lu, Litovsky and Zeng, 2010; Van Deun *et al.*, 2009), however, the effect is much smaller than in NH listeners.

When spatially coincident interfering sources consist also of speech, and particularly if their properties, such as gender, are similar to the target speech, the intelligibility of the target speech can be further reduced in addition to the energetic masking, an effect referred to as informational masking (Brungart, 2001; Kidd *et al.*, 1998). In such situations, spatial separation of *T* and *I* helps to reduce informational masking by means of improving the perceptual segregation between the individual sound sources (Kidd *et al.*, 1998; Marrone, Mason and Kidd, 2008; Hawley, Litovsky and Culling, 2004; Best *et al.*, 2006). It is assumed that both ITD and ILD are important for spatial release from informational masking.

It can be expected that improving the access of CI listeners to binaural cues results in enhanced speech intelligibility in noise, by means of reducing both energetic and informational masking. Improvements in binaural cue sensitivity can be expected from the following three strategies that emerged from the studies included in this treatise: a) better encoding of fine-timing ITD, b) introducing binaurally-synchronized jitter to improve the sensitivity to fine-timing ITD at high pulse rates, and c) enhancing envelope-ITD cues by manipulating (e.g., sharpening) the temporal envelope.

B. Spatial-Hearing Advantage for Vertical Separation of Target and Interferer

Spatial-hearing advantages for speech-in-speech listening based on vertical separation of T and I have been studied much less extensively. A recent study by Martin *et al.* (2012) found a significant spatial release from masking due to median-plane separation of T and I. A study by Worley and Darwin (2002) found that median-plane separation of T and I drastically improved the listeners' ability to attend to the target source, thus, also causing a spatial release from masking. The improvement was much larger for free-field sound presentation than when using binaural virtual acoustics with non-individual HRTFs, suggesting that subject-specific spectral cues are important for the effect to occur. Aaronson, Rakerd and Hartmann (2009) measured spatial release from masking for T and I being separated along the front-back dimension. They found 2 to 4 dB masking release, which they attributed to the ability of the listeners to localize the target and distractors based on spectral localization cues.

With respect to CIs, it seems possible that improving the access to spectral cues for vertical-plane localization, as described in the studies included in this treatise, also results in spatial release from masking in listening situations involving vertical (e.g., front/back) source separation and, thus, contributes to improved speech intelligibility in noise.

7. Implications for the Design of Future Bilateral Hearing Devices

Although the studies included in this habilitation treatise were primarily intended to advance our understanding of auditory processes in electric and acoustic hearing, several results have relevance for the design of future bilateral hearing devices such as CIs or hearing aids.

With respect to the lateral dimension and stimulation at a single electrode pair, our studies (Laback, Majdak and Baumgartner, 2007; Majdak, Laback and Baumgartner, 2006; Laback and Majdak, 2008c) and also those of other groups (van Hoesel, 2007; van Hoesel, Jones and Litovsky, 2009) clearly showed that ITD in the fine timing of pulses at low rates (< 300 pps) produces a robust lateralization cue. Thus, encoding fine-timing ITD with bilateral CI systems has the potential to improve spatial hearing in the lateral dimension. Two stimulation strategies have been proposed (PDT: van Hoesel *et al.*, 2008; FSP: Riss *et al.*, 2008) that encode fine-timing ITD by locking the stimulation pulses to either the peaks or zero-crossings of the temporal waveform in a given channel. However, a direct comparison between the PDT strategy and conventional envelope-based strategies showed no improvement in lateralization or speech intelligibility in noise (van Hoesel *et al.*, 2008). It appears very likely (and is mentioned in van Hoesel *et al.*, 2008) that the lack of improvements is due to the fact that in PDT the stimulation rates at most stimulation channels (except for the apical-most channel) exceeded the maximum rate eliciting fine-timing based ITD cues. An apparently straight-forward solution for this problem, to simply lower the pulse rate, would have the disadvantage that it worsens the representation of the speech envelopes and thus likely would have negative effects on speech perception (Loizou, Poroy and Dorman, 2000; Arora *et al.*, 2009). Furthermore, achieving sufficient loudness with low pulse rates would require either higher stimulation amplitudes, which is restricted by the voltage supply, or longer stimulation pulses, which reduces the flexibility in pulse timing. A more promising solution appears to be to introduce binaurally-synchronized jitter in the stimulation timing of the pulses. For a single interaural electrode pair, we have shown that this approach results in considerable improvements in ITD sensitivity at pulses rates up to about 1500 pps (Laback and Majdak, 2008c). While such a high rate may not be required or even be disadvantageous for robust speech understanding, intermediate rates of about 500-800 pps could be optimal, i.e., rates for which binaurally-synchronized jitter has the potential to improve the access to ITD cues. Thus, jitter may allow to provide both salient ITD cues and precise monaural temporal information.

Interestingly, fine-timing ITD sensitivity for low-rate pulse trains has been shown to be at least as good at basal electrodes as at apical electrodes (Best, Laback and Majdak, 2011; Laback, Best and Majdak, 2011; van Hoesel, Jones and Litovsky, 2009), thus, jitter could be beneficial for encoding ITD cues even at basal electrodes. We described a possible way to implement binaurally-synchronized jitter (Laback and Majdak, 2008b; Laback and Majdak, 2009) or a deterministic variant (Hancock, Chung and Delgutte, 2012; Laback *et al.*, conditionally accepted) in bilateral CI or hearing aid systems. The deterministic variant may be advantageous for preserving monaural speech cues, although this still has to be investigated.

Our results on envelope-ITD perception (Laback *et al.*, 2011b) suggest that the access of CI listeners to ITD cues with bilateral CI systems could be improved by sharpening the temporal envelope of the channel signals, i.e., by increasing the slope steepness and thereby prolonging the lower-level segments, i.e., off times, and by enhancing the peak levels. Because these modifications have opposed effects on loudness, they could be combined while keeping the loudness approximately constant. Such an "enhancement" of envelope-ITD cues could be achieved by modifying the acoustic-to-electric amplitude-mapping function or explicitly by means of an algorithm which modifies the temporal envelope. A similar method could be used for conventional (acoustic) hearing aids or for other types of hearing systems that encode binaural information (such as auditory displays).

The situation with multiple-electrode stimulation in CIs is more complex and likely involves peripheral interactions between adjacent electrodes (Jones, Litovsky and van Hoesel, 2009). Our study on binaural interference (Best, Laback and Majdak, 2011), intentionally exploring the situation with a large tonotopic separation between the target and interferer electrodes and revealing no indication for peripheral interactions, was a first step in that direction. There are no published studies available yet on the constraints in encoding ITD cues in multiple-electrode stimulation with closer electrode spacings, leaving large room for extensive future research. Assuming strong interference effects for narrowly-spaced electrodes, a promising general approach appears to be sparse stimulation, which restricts the stimulation events to those which are sufficiently distant in time-place space.

With respect to a possible strategy to encode spectral cues for vertical-plane localization with CIs, our studies yield mixed results. On the one hand, we showed that the spectral resolution and the number of frequency channels provided by state-of-the-art CI electrodes could be sufficient to encode both spectral localization and speech cues (Laback, Deutsch and Baumgartner, 2004; Goupell *et al.*, 2008b; Goupell, Majdak and Laback, 2010). On the other hand, we found no indications that CI listeners are able to discriminate between different spectral shapes when level cues were excluded (Goupell *et al.*, 2008a), a basic requirement for vertical-plane localization in practical situations where the sound level is not *a priori* known. However, it is possible that profile analysis in electric hearing requires temporal cues which we did not provide in our experiments. Thus, future studies will have to determine to what extent temporal cues contribute to profile analysis in electric hearing and in which way salient temporal cues can be incorporated in CI strategies.

Irrespective of this open issue, requirements for encoding spectral localization cues with future CI systems are to place the microphones at the entrance of the ear canals and to extend the processing frequency range in order to capture the high-frequency pinna cues. According to our results, the mapping of frequency channels should provide at least the 8 most apical electrodes with speech cues (from about 0.3 to 3 kHz) and the remaining electrodes with vertical-plane-localization cues (from about 3 to 16 kHz). These results were obtained for a 12-electrode array with a 2.8-mm spacing between the electrodes. For other array configurations, the values have to be adapted accordingly. An possible alternative strategy, inspired by our long-term localization training experiment, could be to omit high-frequency cues (> 8-10 kHz) at all and train listeners to fully exploit the spatial information contained in the remaining, lower-frequency region.

It is evident that the proposed strategies for improving spatial hearing with future CIs and hearing aids represent basic approaches. Several fundamental research issues have to be solved in order to design true binaural CI or hearing aid systems.

8. Overall Approach of Comparing Acoustic and Electric Hearing

The studies described in this treatise were intended to advance our understanding of psychophysical mechanisms involved in spatial hearing in both acoustic and electric stimulation. The comparison of results for the two hearing modes allowed to obtain new insights into the role of mechanisms in normal hearing. For example, the results showed that the beneficial effect of timing jitter on ITD sensitivity can be attributed to the temporal changes, a conclusion which could not have been drawn from acoustic experiments alone because temporal changes always cause concomitant spectral changes. Furthermore, the comparison between the two hearing modes allowed to identify specific perceptual deficits in electric hearing, e.g., the lack of profile analysis ability of CI listeners when the level is roved. In conclusion, the overall approach of performing parallel psychophysical studies in acoustic and electric hearing appears to have a large potential and should be pursued in the future.

III. LIST OF INCLUDED ARTICLES

1. Binaural Hearing in CI and NH Listeners

A. ILD- and Envelope-ITD Perception with Clinical CI Systems and in Normal Hearing

Laback, B., Pok, S. M., Baumgartner, W. D., Deutsch, W. A., Schmid, K. (2004). Sensitivity to Interaural Level and Envelope Time Differences of Two Bilateral Cochlear Implant Listeners using Clinical Sound Processors, Ear & Hearing 25, 5, 488-500.

B. Fine-structure, Envelope, and Gating-ITD Perception

Laback, B., Majdak, P., Baumgartner, W. D. (2007). Lateralization discimination of interaural time delays in four-pulse sequences in electric and acoustic hearing, J. Acoust. Soc. Am. 121, 2182-2191.

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

C. Does the Place of Stimulation Affect the Rate Limitation in Acoustic ITD Perception?

Majdak, P. and Laback, B. (2009). Effect of center frequency and rate on the sensitivity to interaural delay in high-frequency click trains, J. Acoust. Soc. Am. 125, 3903-3913.

D. Binaurally-Synchronized Jitter in ITD Perception

Laback, B. and Majdak, P. (2008). Binaural jitter improves interaural time difference sensitivity of cochlear implantees at high pulse rates, Proc. Natl. Acad. Sci. U S A (PNAS) 105, 2, 814-817.

Laback, B., and Majdak, P. (2008). Reply to van Hoesel: Binaural jitter with cochlear implants, improved interaural time-delay sensitivity, and normal hearing, Letter to Proc Natl Acad Sci USA (PNAS) 12, 105, 32.

Goupell, M., Laback, B., and Majdak, P. (2009). Enhancing sensitivity to interaural time differences at high modulation frequencies by introducing temporal jitter, J. Acoust. Soc. Am. 2511-2521.

Goupell, M. J., Hancock, K., Majdak, P., Laback, B., and Delgutte, B. (2010) "Binaurally-coherent jitter improves neural and perceptual ITD sensitivity in normal and electric hearing," in Advances in Auditory Physiology, Psychophysics and Models, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), 303-313.

Laback, B. (2012). "Neural Basis Of Improved ITD Sensitivity with Jitter. Focus on "Neural ITD coding with bilateral cochlear implants: Effect of binaurally-coherent jitter," J Neurophysiol 108, 712-713.

Laback, B., and Majdak, P. (2008). "Binaural stimulation in neural auditory prostheses or hearing aids,"

- U. S. Patent, Code #20080319509, Kind Code A1, granted at 25.12.2008.
- EU patent #08 759 294.5, granted at 17.9.2009.

E. <u>Effects of Envelope Shape on Envelope-ITD Perception</u>

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W. D., and Pok, M. (2011). Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing, J. Acoust. Soc. Am. 139, 1515-1529.

Dietz, M., Wendt, T., Ewert, S., Laback, B., and Hohmann, V. (2013). Comparing the effect of pause duration on threshold interaural time differences between exponential and squared-sine envelopes, J. Acoust. Soc. Am. 133, 1-4 (L).

F. <u>Binaural Interference and Auditory Grouping in Electric Hearing</u>

Best, V., Laback, B., and Majdak, P. (2011). Binaural interference in bilateral cochlear-implant listeners, J. Acoust. Soc. Am. 130, 2939-2950.

2. Spectral Cues for Auditory Localization in CI and NH Listeners

A. Three-Dimensional Sound Localization

Majdak, P., Goupell, M., and Laback, B. (2010). 3-D localization of virtual sound sources: effects of visual environment,pointing method, and training, Attention, Perception, & Psychophysics 72, 454-469.

Majdak, P., Goupell, M., and Laback, B. (2011). Two-dimensional Localization of Virtual Sound Sources in Cochlear-Implant Listeners, Ear & Hearing 32, 198-208.

B. Spectral Cue Sensitivity and Profile Analysis in CI Listeners

Laback, B., Deutsch, W. A., Baumgartner, W. D. (2004). Coding of vowel-like signals in cochlear implant listeners, J. Acoust. Soc. Am. 116, 1208-1223.

Goupell, M., Laback, B., Majdak, P., and Baumgartner, W. D. (2008). Current-level discrimination and spectral profile analysis in multi-channel electrical stimulation, J. Acoust. Soc. Am. 124, 3142-3157.

Goupell, M., Laback, B., Majdak, P., and Baumgartner, W. D. (2008). Effects of frequency-place mapping and number of channels on speech intelligibility in electrical stimulation, J. Acoust. Soc. Am. 123, 2295-2309.

C. Number of Channels Required for Vertical-Plane Sound Localization

Goupell, M., Majdak, P., and Laback, B. (2010). Median-plane sound localization as a function of the number of spectral channels using a channel vocoder, J. Acoust. Soc. Am. 127, 990-1001.

D. Plasticity in Vertical-Plane Localization with Altered Spectral Cues

Majdak, P., Walder, T., and Laback, B. (conditionally accepted). "Effect of spectral warping and long-term training on sound localization performance," J. Acoust. Soc. Am.

IV. REFERENCES

Aaronson, N. L., Rakerd, B., and Hartmann, W. M. (**2009**). "Release from speech-on-speech masking in a front-and-back geometry," J Acoust Soc Am **125**, 1636-1648.

Arora, K., Dawson, P., Dowell, R., and Vandali, A. (**2009**). "Electrical stimulation rate effects on speech perception in cochlear implants," Int J Audiol **48**, 561-567.

Azadpour, M., and McKay, C. M. (**2012**). "A psychophysical method for measuring spatial resolution in cochlear implants," J Assoc Res Otolaryngol **13**, 145-157.

Balazs, P., Laback, B., Eckel, G., and Deutsch, W. A. (**2010**). "Time-Frequency Sparsity by Removing Perceptually Irrelevant Components Using a Simple Model of Simultaneous Masking," IEEE Transactions on Audio, Speech and Language Processing **18,1**, 34-49.

Bernstein, L. R. (**2001**). "Auditory processing of interaural timing information: new insights," J Neurosci Res **66**, 1035-1046.

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**, 3561-3567.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J Acoust Soc Am **112**, 1026-1036.

Bernstein, L. R., and Trahiotis, C. (**2003**). "Enhancing interaural-delay-based extents of laterality at high frequencies by using "transposed stimuli"," J Acoust Soc Am **113**, 3335-3347.

Best, V., Gallun, F. J., Ihlefeld, A., and Shinn-Cunningham, B. G. (**2006**). "The influence of spatial separation on divided listening," J Acoust Soc Am **120**, 1506-1516.

Best, V., Laback, B., and Majdak, P. (**2011**). "Binaural interference in bilateral cochlear-implant listeners," J Acoust Soc Am **130**, 2939-2950.

Blanks, D. A., Roberts, J. M., Buss, E., Hall, J. W., and Fitzpatrick, D. C. (**2007**). "Neural and behavioral sensitivity to interaural time differences using amplitude modulated tones with mismatched carrier frequencies," J Assoc Res Otolaryngol **8**, 393-408.

Blauert, J. (**1997**). *Spatial hearing, 2nd edition* (MIT-Press, Cambridge, MA).

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge).

Bronkhorst, A. (**2000**). "The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions," Acustica - Acta Acustica **86**, 117-128.

Brown, A. D., and Stecker, G. C. (**2011**). "Temporal weighting functions for interaural time and level differences. II. The effect of binaurally synchronous temporal jitter," J Acoust Soc Am **129**, 293-300.

Bruce, V., Green, P. R. and Georgeson, M. A. (**1996**). *Visual perception: Physiology, Psychology, and Ecology (3rd ed.)* (Psychology Press, Hove, East Sussex).

Brungart, D. S. (**2001**). "Informational and energetic masking effects in the perception of two simultaneous talkers," J Acoust Soc Am **109**, 1101-1109.

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**, 2063-2066.

Carlyon, R. P., Macherey, O., Frijns, J. H. M., Axon, P. R., Kalkman, R. K., Boyle, P., Baguley, D. M., Briggs, J., Deeks, J. M., Briaire, J. J., Barreau, X., and Dauman, R. (**2010**). "Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear," J Assoc Res Otolaryngol **11**, 625-640.

Carterette, E. C. and Friedman, M. P. (**1974**). *Handbook of Perception, Volume II: Psychophysical Judgement and Measurement* (Academic Press, New York).

Clark, G. (**2003**). *Cochlear implants: Fundamentals and applications* (Springer, New York).

Cohen, L. T., Busby, P. A., Whitford, L. A., and Clark, G. M. (**1996**). "Cochlear implant place psychophysics 1. Pitch estimation with deeply inserted electrodes," Audiol Neurootol **1**, 265-277.

Colburn, H. S., Chung, Y., Zhou, Y., and Brughera, A. (**2009**). "Models of brainstem responses to bilateral electrical stimulation," J Assoc Res Otolaryngol **10**, 91-110.

Culling, J. F., and Summerfield, Q. (**1995**). "Perceptual separation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay," J Acoust Soc Am **98**, 785-797.

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**, 1057-1065.

Dietz, M., Wendt, T., Ewert, S., Laback, B., and Hohmann, V. (**2013**). "Comparing the effect of pause duration on threshold interaural time differences between exponential and squared-sine envelopes," J. Acoust. Soc. Am. **133**, 1-4.

Drennan, W. R., and Pfingst, B. E. (**2006**). "Current-level discrimination in the context of interleaved, multichannel stimulation in cochlear implants: effects of number of stimulated electrodes, pulse rate, and electrode separation," J Assoc Res Otolaryngol **7**, 308-316.

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**, 2178-2189.

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**, 2327-2341.

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

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.

Ellermeier, W., and Faulhammer, G. (**2000**). "Empirical evaluation of axioms fundamental to Stevens's ratio-scaling approach: I. Loudness production," Percept Psychophys **62**, 1505-1511.

Fechner, G. T. (**1860**). *Elemente der Psychophysik* (Breitkopf und Härtel, Leipzig [Reissued 1964 by Bonset, Amsterdam]).

Francart, T., and Wouters, J. (**2007**). "Perception of across-frequency interaural level differences," J Acoust Soc Am **122**, 2826-2831.

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**, 1150-1163.

Fu, Q., Nogaki, G., and Galvin, J. J. 3. (**2005**). "Auditory training with spectrally shifted speech: implications for cochlear implant patient auditory rehabilitation," J Assoc Res Otolaryngol **6**, 180-189.

Gescheider, G. (**1997**). *Psychophysics: the fundamentals (3rd ed.)* (Lawrence Erlbaum Associates, New Jersey).

Gilkey, R. H. and Anderson, T. R. (**1997**). *Binaural and Spatial Hearing in Real and Virtual Environments* (Lawrence Erlbaum Associates, NJ).

Goldstein, E. B. (**2009**). *Sensation and Perception (8th ed.)* (Wadsworth, Belmont).

Goupell, M. J., Laback, B., and Majdak, P. (**2009**). "Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter," J Acoust Soc Am **126**, 2511-2521.

Goupell, M. J., Laback, B., Majdak, P., and Baumgartner, W. (**2008a**). "Current-level discrimination and spectral profile analysis in multi-channel electrical stimulation," J Acoust Soc Am **124**, 3142-3157.

Goupell, M. J., Laback, B., Majdak, P., and Baumgartner, W. (**2008b**). "Effects of upper-frequency boundary and spectral warping on speech intelligibility in electrical stimulation," J Acoust Soc Am **123**, 2295-2309.

Goupell, M. J., Majdak, P., and Laback, B. (**2010**). "Median-plane sound localization as a function of the number of spectral channels using a channel vocoder," J Acoust Soc Am **127**, 990-1001.

Goupell, M., Hancock, K., Majdak, P., Laback, B. and Delgutte, B. (**2010**). "Binaurally-coherent jitter improves neural and perceptual ITD sensitivity in normal and electric hearing," in *The Neurophysiological Bases of Aditory Perception* (pp. 303-313), edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York).

Grantham, D. W. (**1984**). "Interaural intensity discrimination: insensitivity at 1000 Hz," J Acoust Soc Am **75**, 1191-1194.

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., and Haynes, D. S. (**2007**). "Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants," Ear Hear **28**, 524-541.

Green, D. M. (**1988**). *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, New York).

Green, D. M. and Swets, J. A. (**1966**). *Signal Detection theory and psychophysics* (Wiley, New York).

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., and Nguyen, Q. (**1992**). "Spectral shape discrimination of narrow-band sounds," J Acoust Soc Am **92**, 2586-2597.

Gridi-Papp, M. and Narins, P. M. (**2008**). "Sensory Ecology of Hearing," in *The Senses: A Comprehensive Reference* (pp. 1089-1102), edited by P Dallos and D Oertel (Elsevier, San Diego).

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**, 3463-3478.

Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiol Rev **90**, 983-1012.

Hafter, E. and Trahiotis, C. (**1997**). "Functions of the binaural system," in *Encyclopedia of Acoustics* (pp. 1461-1479), edited by M. J. Crocker (Wiley & Sons, New York).

Hafter, E. R. (**1997**). "Binaural adaptation and the effectiveness of a stimulus beyond its onset," in *Binaural and Spatial Hearing in Real and Virtual Environments* (pp. 211-232), edited by R. H. Gilkey, and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ).

Hafter, E. R., and Buell, T. N. (**1990**). "Restarting the adapted binaural system," J Acoust Soc Am **88**, 806-812.

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**, 644-651.

Hafter, E. R., Dye, R. H. J., and Wenzel, E. (**1983**). "Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number," J Acoust Soc Am **73**, 1708-1713.

Hancock, K. E., Chung, Y., and Delgutte, B. (**2012**). "Neural ITD coding with bilateral cochlear implants: effect of binaurally coherent jitter," J Neurophysiol **108**, 714-728.

Hancock, K. E., Chung, Y., Blasbalg, S., Colburn, S., and Delgutte, B. (**2011**). "Effect of binaurally-coherent jitter on neural ITD coding with bilateral cochlear implants," presented at the *presented at the 34th Meeting of the Association for Research in Otolaryngology*, Baltimore.

Hartmann, W. M., and Constan, Z. A. (**2002**). "Interaural level differences and the level-meter model," J Acoust Soc Am **112**, 1037-1045.

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

Hebrank, J., and Wright, D. (**1974**). "Spectral cues used in the localization of sound sources on the median plane," J Acoust Soc Am **56**, 1829-1834.

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

Hershkowitz, R. M., and Durlach, N. I. (**1969**). "Interaural time and amplitude jnds for a 500-Hz tone," J Acoust Soc Am **46**, 1464-1467.

Hofman, P. M., Van Riswick, J. G., and Van Opstal, A. J. (**1998**). "Relearning sound localization with new ears," Nat Neurosci **1**, 417-421.

Jones, F. N. (**1974**). "History of psychophysics and judgment," in *Handbook of Perception, Volume II: Psychophysical Judgement and Measurement* (pp. 1-22), edited by E. C. Carterette, and M. P. Friedman (Academic Press, New York).

Jones, G. L., Litovsky, R., and van Hoesel, R. (**2009**). "Relationship of monaural and binaural channel interaction effects in bilateral cochlear implant users," Poster presented at the *Conference on Implantable Auditory Prostheses (CIAP)*, Tahoe City, CA.

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

Kistler, D. J., and Wightman, F. L. (**1992**). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J Acoust Soc Am **91**, 1637-1647.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," J Acoust Soc Am 129 , 3856-3872.

Köhler, W. (**1940**). *Dynamics in psychology* (Liveright, New York).

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," J Acoust Soc Am **62**, 157-167.

Kulkarni, A., and Colburn, H. S. (**1998**). "Role of spectral detail in sound-source localization," Nature **396**, 747-749.

Laback, B. (**2012**). "Neural Basis Of Improved ITD Sensitivity with Jitter. Focus on "Neural ITD coding with bilateral cochlear implants: Effect of binaurally-coherent jitter," J Neurophysiol **108**, 712-713.

Laback, B., and Deutsch, W. A. (**2001**). "Perceptual grouping of tone sequences in cochlear implant stimulation," in proceedings of the *International Congress on Acoustics*, Rome, CD-Rom.

Laback, B., and Majdak, P. (**2008a**). "Reply to van Hoesel: Binaural jitter with cochlear implants, improved interaural time-delay sensitivity and normal hearing," Proceedings of the National Academy of Sciences of the USA (PNAS) **105**, E52.

Laback, B., and Majdak, P. (**2008b**). "Binaural stimulation in neural auditory prostheses or hearing aids," U.S. Patent **Code # 20080319509**, Kind Code A1, granted on 25.12.2008.

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

Laback, B., and Majdak, P. (**2009**). "Binaural stimulation in neural auditory prostheses or hearing aids," EU patent **Code # 08 759 294.5**, granted on 17.9.2009.

Laback, B., Balazs, P., Necciari, T., Savel, S., Ystad, S., Meunier, S., and Kronland-Martinet, R. (**2011a**). "Additivity of nonsimultaneous masking for short Gaussian-shaped sinusoids," J Acoust Soc Am **129**, 888-897.

Laback, B., Best, V., and Majdak, P. (**2011**). "Tonotopic symmetry of ITD-based lateralization and channel interference," presented at the *Conference on Implantable Auditory Prostheses (CIAP)*, Asilomar, Pacific Grove, CA, USA.

Laback, B., Deutsch, W. A., and Baumgartner, W. (**2004**). "Coding of vowellike signals in cochlear implant listeners," J Acoust Soc Am **116**, 1208-1223.

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**, 2182-2191.

Laback, B., Necciari, T., Balazs, P., Savel, S., and Ystad, S. (**conditionally accepted**). "Simultaneous masking additivity for short Gaussian-shaped tones: Spectral effects," J. Acoust Soc Am.

Laback, B., Pok, S., Baumgartner, W., Deutsch, W. A., and Schmid, K. (**2004**). "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors," Ear Hear **25**, 488-500.

Laback, B., Zimmermann, I. and Majdak, P. (**2010**). "Binaural Processing and Spatial Hearing," in *Perception of Interaural Time Differences in Electric and Acoustic Hearing* (pp. 403-416), edited by J. M. Buchholz, T. Dau, J. C. Dalsgaard, and T. Poulsen (Danavox Jubilee Foundation 2010, Ballerup).

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W. D., and Pok, S. M. (**2011b**). "Effects of envelope shape on interaural delay sensitivity in acoustic and electric hearing," J Acoust Soc Am **139**, 1515-1529.

Langendijk, E. H. A., and Bronkhorst, A. W. (**2002**). "Contribution of spectral cues to human sound localization," J Acoust Soc Am **112**, 1583-1596.

Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (**2009**). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear **30**, 419-431.

Loizou, P. C., Poroy, O., and Dorman, M. (**2000**). "The effect of parametric variations of cochlear implant processors on speech understanding," J Acoust Soc Am **108**, 790-802.

Long, C. J., Carlyon, R. P., Litovsky, R. Y., and Downs, D. H. (**2006**). "Binaural unmasking with bilateral cochlear implants," J Assoc Res Otolaryngol **7**, 352-360.

Lu, T., Litovsky, R., and Zeng, F. (**2010**). "Binaural masking level differences in actual and simulated bilateral cochlear implant listeners," J Acoust Soc Am **127**, 1479-1490.

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**, 2219-2236.

Majdak, P., and Laback, B. (**2009**). "Effects of center frequency and rate on the sensitivity to interaural delay in high-frequency click trains," J Acoust Soc Am **125**, 3903-3913.

Majdak, P., Goupell, M. J., and Laback, B. (**2010**). "3-D localization of virtual sound sources: effects of visual environment, pointing method, and training," Atten Percept Psychophys **72**, 454-469.

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

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**, 2190-2201.

Majdak, P., Walder, T., and Laback, B. (**2011**). "Learning to localize band-limited sounds in vertical planes," presented at the *34th Mid-Winter Meeting of the Association for Research in Otolaryngology*, Baltimore, MD, USA.

Majdak, P., Walder, T., and Laback, B. (**conditionally accepted**). "Effect of spectral warping and long-term training on sound localization performance," J. Acoust. Soc. Am.

Marrone, N., Mason, C. R., and Kidd, G. J. (**2008**). "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," J Acoust Soc Am **124**, 3064-3075.

Martin, R. L., McAnally, K. I., Bolia, R. S., Eberle, G., and Brungart, D. S. (**2012**). "Spatial release from speech-on-speech masking in the median sagittal plane," J Acoust Soc Am **131**, 378-385.

Mason, C. R., Kidd, G. J., Hanna, T. E., and Green, D. M. (**1984**). "Profile analysis and level variation," Hear Res **13**, 269-275.

Middlebrooks, J. C. (**1992**). "Narrow-band sound localization related to external ear acoustics," J Acoust Soc Am **92**, 2607-2624.

Middlebrooks, J. C. (**1999a**). "Individual differences in external-ear transfer functions reduced by scaling in frequency," J Acoust Soc Am **106**, 1480-1492.

Middlebrooks, J. C. (**1999b**). "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," J Acoust Soc Am **106**, 1493-1510.

Middlebrooks, J. C., and Green, D. M. (**1991**). "Sound localization by human listeners," Annu Rev Psychol **42**, 135-159.

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

Morimoto, M. (**2001**). "The contribution of two ears to the perception of vertical angle in sagittal planes," J Acoust Soc Am **109**, 1596-1603.

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

Necciari, T., Savel, S., Laback, B., Menieur, S., Balazs, P., Kronland-Martinet, R., and Ystad, S. (**submitted, 2011**). "Time-frequency spread of auditory masking for spectrally and temporally maximally-compact stimuli," J. Acoust. Soc. Am , .

Nopp, P., Schleich, P., and D'Haese, P. (**2004**). "Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants," Ear Hear **25**, 205-214.

Perrett, S., and Noble, W. (**1997**). "The effect of head rotations on vertical plane sound localization," J Acoust Soc Am **102**, 2325-2332.

Perrott, D. R., Briggs, R., and Perrott, S. (**1970**). "Binaural fusion: its limits as defined by signal duration and signal onset," J Acoust Soc Am **47**, 565-568.

Riss, D., Arnoldner, C., Baumgartner, W., Kaider, A., and Hamzavi, J. (**2008**). "A new fine structure speech coding strategy: speech perception at a reduced number of channels," Otol Neurotol **29**, 784-788.

Scharf, B., Florentine, M., and Meiselman, C. H. (**1976**). "Critical band in auditory lateralization," Sens Processes **1**, 109-126.

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

Schwartz, A., McDermott, J. H., and Shinn-Cunningham, B. (**2012**). "Spatial cues alone produce inaccurate sound segregation: the effect of interaural time differences," J Acoust Soc Am **132**, 357-368.

Seeber, B. U., Baumann, U., and Fastl, H. (**2004**). "Localization ability with bimodal hearing aids and bilateral cochlear implants," J Acoust Soc Am **116**, 1698-1709.

Shannon, B. (**1993**). "Psychophysics," in *Cochlear Implants: Audiological Foundations* (pp. 357-388), edited by R. S. Tyler (Singular, San Diego).

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2004**). "Monaural and interaural intensity discrimination: level effects and the "binaural advantage"," J Acoust Soc Am **116**, 1149-1159.

Stevens, S. S. (**1957**). "On the psychophysical law," Psychol Rev **64**, 153-181.

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

Thurstone, L. L. (**1927**). "Psychophysical analysis," Am. J. of Psychology **38**, 369-389.

Titchener, E. B. (**1905**). *Experimental Psychology, Vol II. Quantitative Experiments. Part I. Students Manual* (Macmillan, New York, reprinted 1927).

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**, 1671-1680.

van der Heijden, M., and Joris, P. X. (**2010**). "Interaural correlation fails to account for detection in a classic binaural task: dynamic ITDs dominate N0Spi detection," J Assoc Res Otolaryngol **11**, 113-131.

Van Deun, L., van Wieringen, A., Francart, T., Scherf, F., Dhooge, I. J., Deggouj, N., Desloovere, C., Van de Heyning, P. H., Offeciers, F. E., De Raeve, L., and Wouters, J. (**2009**). "Bilateral cochlear implants in children: binaural unmasking," Audiol Neurootol **14**, 240-247.

van Hoesel, R. J. M. (**2007**). "Sensitivity to binaural timing in bilateral cochlear implant users," J Acoust Soc Am **121**, 2192-2206.

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

van Hoesel, R. J. M., Jones, G. L., and Litovsky, R. Y. (**2009**). "Interaural time-delay sensitivity in bilateral cochlear implant users: effects of pulse rate, modulation rate, and place of stimulation," J Assoc Res Otolaryngol **10**, 557-567.

van Hoesel, R., Böhm, M., Pesch, J., Vandali, A., Battmer, R. D., and Lenarz, T. (**2008**). "Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies," J Acoust Soc Am **123**, 2249-2263.

Vermeire, K., Nobbe, A., Schleich, P., Nopp, P., Voormolen, M. H., and Van de Heyning, P. H. (**2008**). "Neural tonotopy in cochlear implants: an evaluation in unilateral cochlear implant patients with unilateral deafness and tinnitus," Hear Res **245**, 98-106.

Weber, E. H. (**1834**). *De pulsu, resorptione, auditu et tactu. Anatationes anatomicae et physiologicae* (Köhler, Leipzig).

Wertheimer, M. (**1945**). *Productive thinking* (Harper, New York).

Wightman, F. L., and Kistler, D. J. (**1989a**). "Headphone simulation of free-field listening. I: Stimulus synthesis," J Acoust Soc Am **85**, 858-867.

Wightman, F. L., and Kistler, D. J. (**1989b**). "Headphone simulation of free-field listening. II: Psychophysical validation," J Acoust Soc Am **85**, 868-878.

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**, 1648-1661.

Wightman, F. L., and Kistler, D. J. (**1997**). "Monaural sound localization revisited," J Acoust Soc Am **101**, 1050-1063.

Wightman, F. L., and Kistler, D. J. (**1999**). "Resolution of front-back ambiguity in spatial hearing by listener and source movement," J Acoust Soc Am **105**, 2841-2853.

Wilson, B. S., and Dorman, M. F. (**2008**). "Cochlear implants: current designs and future possibilities," J Rehabil Res Dev **45**, 695-730.

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

Won, J. H., Jones, G. L., Drennan, W. R., Jameyson, E. M., and Rubinstein, J. T. (**2011**). "Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners," J Acoust Soc Am **130**, 2088-2097.

Worley, J. W., and Darwin, C. J. (**2002**). "Auditory attention based on differences in median vertical plane position," in proceedings of the *Proceedings of the 2002 international conference on auditory display*, Kyoto, .

Yost, W. A. (**1976**). "Lateralization of repeated filtered transients," J Acoust Soc Am **60**, 178-181.

Yost, W. A. and Hafter, E. R. (**1987**). "Lateralization," in *Directional Hearing* (pp. 49-84), edited by W. Yost, and G. Gourevitch (Springer-Verlag, New York).

Young, E. D., and Sachs, M. B. (**1979**). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers," J Acoust Soc Am **66**, 1381-1403.

Zeng, F. G., Popper, A. N. and Fay, R. (**2004**). *Cochlear implants: Auditory prostheses and electric hearing* (Springer, New York).

Zeng, F., Rebscher, S., Harrison, W., Sun, X., and Feng, H. (**2008**). "Cochlear implants: system design, integration, and evaluation," IEEE Rev Biomed Eng **1**, 115-142.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J Acoust Soc Am **28**, 860-864.