Head-Related Transfer Functions and their Role in the Localization of Sound Sources in Human Listeners

Habilitationsschrift

Piotr Majdak

Universität für Musik und darstellende Kunst

Graz

November 2014

habilitation 6G
Modified: 10. Nov. 2014

Table of Contents

1. Preface

1.1. Motivation

Head-related transfer functions (HRTFs) describe the acoustic filtering of a sound source by listener's body. A binaural set of HRTFs describes all acoustic cues required for localization of sound sources, an essential task in the perception of speech and music.

It was 2001, as I have learned the actual meaning of “HRTFs” in its full depth. At that time, I was working (together with my colleague Markus Noisternig) on a student project at the Institute of Electronic Music and Acoustic (IEM). Markus' background was sound-engineering, my background was electronics and signal-processing, and thus, at the end of the project, we came up with a digital signal processor capable of spatializing sound sources in real time via headphones. The novelty of the system was a smooth spatial interpolation achieved by combining the Ambisonics with the acoustically measured HRTFs as filters for the spatialization.

Already in 2001, the project outcome arouse my interest in spatial hearing. Straightaway after finishing that project, while investigating the impact of HRTF-related details on sound localization for my Master thesis, I learned more about spatial hearing by conducting psychoacoustic experiments in normal-hearing (NH) listeners. Later, during my work at the Acoustic Research Institute (ARI) of the Austrian Academy of Sciences (ÖAW), I extended my research field to cochlear implants (CIs) where I combined psychoacoustics with signal processing in order to investigate spatial hearing in CI listeners. Since 2002, the start of my work at the ARI, my area of interest has continuously expanded to the field of spatial hearing, acoustics, and audio engineering.

This thesis outlines my scientific contributions to the field of spatial hearing. Despite the focus on spatial hearing, studies on other acoustic-related aspects (not included in this thesis) underlie my broad interests in the large field of acoustics and audio research.

In the remaining of this chapter, my scientific achievements are described. Actually, none of them are mine alone: in my research area, it is nearly impossible to do research on one's own. Collaborations are essential and the achievements described here have been only possible because of many collaborations I was able to establish. Bernhard Laback (the head of the group I belong to) and Peter Balazs (formerly the head of the Mathematics group and currently the institute director) are my “senior” collaborators at the ARI. Robert Baumgartner, Katharina

Egger, and Harald Ziegelwanger are PhD students working with me under my supervision. Naturally, all these and other project-related collaborators are co-authors in the corresponding publications.

The achievements provided in the remaining part of this chapter are sorted by research topics and projects in order to stress their connection. Chapter 2 lists the peer-reviewed publications included in this thesis, sorted by the research topic. Chapter 3 provides an introduction to spatial hearing emphasizing the role of HRTFs in sound localization. It also relates the scientific background to the conclusions which can be drawn from the individual studies included in this thesis. Chapter 4 provides a short outlook for my future work.

1.2. Scientific achievements

Generally, in my work, I aim at a better understanding of the mechanisms underlying spatial hearing and at applying that knowledge to improve acoustic and audio systems. While spatial hearing is a very broad research area, my work considered so far:

- Spatial hearing in *normal-hearing* listeners by means of audio reproduction systems providing spatial cues to the auditory system. Here, many aspects of the spatial hearing *beyond the horizontal plane* are still uncovered, and thus I worked on *understanding and modeling of sound localization in sagittal planes* and other aspects of *three-dimensional (3D) sound reproduction*.
- Spatial hearing in *hearing-impaired* listeners by means of providing algorithms and improving hearing-assist devices restoring the spatial cues. Here, more basic spatial effects like *sound localization in the horizontal plane* and *speech understanding in noise* are relevant. Interaural time differences (ITDs) are important cues for performing well in these tasks.[1] Thus, I work on *understanding and improving the ITD-based spatial hearing in CI listeners*.

In order to be able to work towards those goals, in addition to the scientific work, I also developed psychoacoustic methods and software algorithms. The solutions considered, depending on the research question, bilaterally synchronized electric stimulation of CI listeners, multichannel loudspeaker-based or headphone-based sound reproduction, as well as software packages and toolboxes. In the following sections, details of my scientific achievements are described, sorted by research topics and projects.

1 Note that ITDs can be derived from the interaural difference of HRTF's group delay, thus, a binaural set of HRTFs can be seen as a general form of describing an ITD.

1.2.1. Acquisition, approximation, and applications of HRTFs

My first study considered the development of the multiple exponential sweep method (MESM), which allowed us *to reduce the duration of HRTF measurements by factor of four* (Majdak et al., 2007). For many studies on HRTFs, listener-specific HRTFs are essential, thus, a setup for acoustic measurement of HRTFs is required. For the hundreds of required spatial positions, the state-of-the art measurement methods were rather slow at that time; it would take an hour to measure all the positions, a time period, in which the listener would need to sit absolutely still. By exploiting sparsity of frequency sweeps in the time-frequency domain, together with my colleague Peter Balazs, I have developed the MESM, which allows us to measure 1550 spatial positions within 16 minutes in our semi-anechoic chamber. The MESM can be applied to efficiently measure multiple electro-acoustic systems and in the meantime, it has been implemented in several labs, e.g., the Institute of Technical Acoustics (ITA), Aachen, Germany. As for the scientific recognition, my article was cited 22 times in other journals.

The MESM is implemented in ExpSuite, *a software framework for psychoacoustic experiments and acoustic measurements*. Within ExpSuite, for each study and experiment, applications can be developed, which, being stand-alone applications, benefit from the general framework functionality of ExpSuite. I have started to implement ExpSuite in 2003; the first ExpSuite applications considered ITD sensitivity in CI listeners and acoustic measurement of HRTFs with the MESM. Since then, 22 applications were created and were bases for many publications. As part of *my research philosophy to promote reproducible research*, ExpSuite is available on SourceForge as a free and open-source project.

After the development of the MESM, we started to measure HRTFs for various studies, continuously expanding our HRTF database. At the time of writing, the database includes measurement sets of over 100 listeners. During the last decade, more and more labs were measuring HRTFs, and an active exchange of HRTFs takes place. Unfortunately, most of the labs stored the measured HRTFs in their own format, causing difficulties for a simple exchange of the data. A unification of the various data is not as trivial as it may appear because the measurements are done using different methods and considering different aspects of the measurements. In 2012, I decided to create a file format which would fit all the needs. I developed the *spatially oriented file format for acoustics* (SOFA), which can handle not only HRTFs but also a variety of data like binaural room impulse responses (BRIRs), directional room impulse responses (DRIRs), headphone transfer functions (HpIR), or simulated acoustic data like numerically calculated HRTFs, etc. The main strength of SOFA is its *expandability*, which aims at being able to handle future needs *without breaking the backwards compatibility*. A file format can only be successful when it is widely recognized. Thus, for the first draft of SOFA (Majdak et al., 2013a), I worked with nine collaborators from all over the world. SOFA, being a set of open specifications, is currently under revision for an AES standard. A liaison with MPEG has been established and our website, sofaconventions.org, hosts over 1000 SOFA files.

SOFA files can be easily accessed from both Matlab and C++, showing its wide range of applicability. As for Matlab, I have created a toolbox (SOFA API), which provides general functions dealing with data stored in SOFA files. In 2013, the SOFA API received the *Reproducibility Research Prize* from the Queen Mary University of London.

Even though still not an official standard, SOFA has been worldwide accepted. SOFA is used to represent HRTFs by e.g., Tohoku University[2], and to represent BRIRs by e.g., the British broadcasting corporation (BBC)[3]; demonstrating its use for both research and applications in audio engineering.

As for the application of SOFA, I initiated a development of the plug-in *SOFAlizer* for the VLC media player (VLC)[4]. In SOFAlizer, arbitrary SOFA files can be used in order to create virtual loudspeakers simulated via headphones. Currently, a proof-of-concept version of SOFAlizer is implemented and published at SourceForge[5]. The source code is under review at the VideoLAN community aiming at a stable plug-in for the standard release of VLC-Player.

Given the long-winded acoustic HRTF measurements, I was seeking for other means of HRTF acquisition. Together with Wolfgang Kreuzer and Bernhard Laback, I started to work on alternative ways to capture HRTFs where we aimed at visually scanning listeners' heads and calculating HRTFs using numerical algorithms (fast-multipole coupled boundary-element method). Our results showed, for the first time, that the *numerical method is in general capable of calculating HRTFs even for high frequencies*. We also showed that the accuracy of the geometry acquisition and the perceptual interpretation of the HRTF simulations are essential issues when simulating HRTF (Kreuzer et al., 2009).

2 see http://www.riec.tohoku.ac.jp/pub/hrtf/index.html; last viewed on 7.7.2014.

3 http://www.bbc.co.uk/rd/publications/sbsbrir; last viewed on 7.7.2014.

4 http://www.videolan.org/vlc; last viewed on 4.8.2014.

5 http://sourceforge.net/p/sofacoustics/code/HEAD/tree/trunk/SOFAlizer; last viewed on 4.8.2014.

While these results showed that numerical methods are in general capable of calculating HRTFs for high frequencies, a method for an efficient geometry acquisition has not been settled yet. Thus, following the discussions with my colleague Wolfgang Kreuzer, we started to develop methods for efficient geometry acquisition and fast numerical calculation of HRTF. This work happens currently within the FWF-granted project[6] “Virtual Acoustics: localization model & numeric simulations (LocaPhoto)”.

Within LocaPhoto, first, in a collaboration with the Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH (VRVis), we showed that by using *photogrammetric reconstruction applied on a large set of photos, a reconstruction of pinna's geometry is, in principle, achievable* (Reichinger et al., 2013). These results were obtained by evaluating seven methods for three-dimensional scanning the human pinna. The detailed requirements for the photogrammetric reconstruction are still subject of investigations.

Second, *we determined details required for the calculation of perceptually-valid HRTFs*, i.e., HRTFs with spectral features evoking similar sound localization performance as compared to those from acoustically measured HRTFs (Ziegelwanger et al., 2013b). In particular, we showed that 1) *an arbitrary small receiver element of the mesh yields most probably perceptually valid HRTFs as long as it is placed within the simulated ear-canal area* and 2) *geometry accuracy in the range of 1 mm is required for the pinna in order to obtain perceptually valid HRTFs*. These conclusions were drawn from the systematical investigation of the role of the virtual microphone on simulated sound-localization performance, and from studying the effect of the mesh quality by systematically varying the average edge length in the mesh. The journal version has been recently submitted to the Journal of the Acoustical Society of America (JASA).

Third, we *improved the HRTF simulation efficiency by a factor of ten* (Ziegelwanger et al., 2014). While the geometry sampling requires at least 6 elements per wavelength, we showed evidence that violation of that rule for larger distances from the receiver element, i.e., at the contralateral side of the head, does not perceptually-significantly contribute to the calculation error. However, it reduces the number of elements in the representation of listener's geometry by factor of ten. Thus, we proposed *a non-uniform sampling scheme for the listener's geometry* in which the average edge length depends on the distance to the receiver element. By a further *combination of the non-uniform sampling with near-field simulation and range extrapo-*

6 My role: principle investigator (PI) and project leader.

lation in the spherical-harmonic domain, we were able *to reduce the calculation of perceptually valid HRTFs from hours to a few minutes*. The journal version is currently under preparation for a submission to Applied Acoustics.

In an other study, tempted by the beauty of time-frequency representation, I learned how to approximate HRTFs in subbands (Marelli et al., 2008). This very first step led to collaboration within the FWF-Lise-Meitner project[7], where we more extensively have worked on the subband representation of HRTFs. In 2012, I supervised Robert Baumgartner on his master thesis, which revealed that *a joint binaural optimization considering the HRTF phase is essential*. Since then, we have clarified the details of the approximation, performed sound-localization experiments evaluating the approximations (Marelli et al., 2014), and a journal article showing *how to approximate HRTFs in subbands under consideration of perceptual aspects* has been submitted to the IEEE Transactions on Audio, Speech, and Language Processing (TASLP).

1.2.2. Sound localization beyond the horizontal plane

I made several contributions to a better understanding of mechanisms underlying 3D sound localization.

First, in a sound localization experiment, I showed that 1) *the localization performance of naive listeners improves when a congruent visual environment is provided*; 2) *head and manual hand pointers provide similar localization response performance*; and 3) *initial training is essential in order to obtain stable and reproducible localization responses* (Majdak et al., 2010). These conclusions resulted from a study, in which we investigated the effect of the visual environment, the pointing method, and the training on sound localization in NH listeners. To this end, I have developed a setup for testing sound localization in a virtual auditory space (VAS). This development was required as a basis for future studies and direct prerequisite in the FWF-granted project on the role of spectral cues in CI listeners (CI-HRTF)[8]. The setup includes a virtual visual environment and auditory stimulation via headphones, allowing to test and train sound localization with arbitrary HRTFs. This setup is a basis for all sound localization studies included in this thesis.

7 My role: Co-PI, responsible for the psychoacoustic part of the project.

8 My role: Co-PI, responsible for the part related to HRTF measurement and sound-localization experiments. PI: Bernhard Laback.

As the next step, in a sound localization experiment with CI listeners, I showed that *CI listeners' ability of localizing sounds in sagittal planes is mainly based on the broadband level* (Majdak et al., 2011b). Although some of our CI listeners were able to localize sounds in the sagittal planes better than chance, the judgments of the sagittal position of a sound depended on the amount of overall level roving, leaving little evidence for processing directional spectral cues as provided by HRTFs in CI listeners.

Further, we evaluated parameters potentially limiting the poor sound localization ability in CI listeners. First, we investigated the effect of the number of spectral channels on sound localization in the median plane by testing NH listeners listening to a CI simulation (Goupell et al., 2010b). We found that *CIs may provide a sufficient number of spectral channels for some sagittal-plane sound localization capabilities* (although worse than NH listeners), without loss of speech understanding. Second, we performed spectral profile analysis in CI listeners in order to determine how spectral profiles and current-level discrimination may limit the processing of directional spectral cues (Goupell et al., 2008). As soon as the overall level was roved, there *was very little evidence for a spectral profile analysis, indicating a strong limitation for the processing of directional cues by CI listeners*.

In the last study of the project CI-HRTF, we investigated listener's ability to re-calibrate the auditory system for using drastically modified HRTFs in sound localization (Majdak et al., 2013c). Motivated by the parameter space for CI listeners, we spectrally warped listener-specific HRTFs and trained NH listeners three weeks on sound localization with such modified HRTFs. By including a control group listening to listener-specific but band-limited HRTFs, and performing sophisticated double-blinded experiments, *we were able to distinguish between procedural learning* (ability to perform in a localization experiment as well to use the equipment) *and perceptual learning* (actual re-calibration of the auditory system). Our results showed that the training effect for the warped condition exceeded that for procedural learning, suggesting *a stable auditory recalibration due to the training of only two hours per day*. After the training, performance with band-limited sounds was better than that with warped ones. Our results further show that *training can improve sound localization in cases where spectral cues have been reduced by band-limiting or remapped by warping*, suggesting that listeners having limited access to high frequencies, might also improve their localization ability when provided with spectrally warped or band-limited sounds and adequately trained on sound localization.

In an other project (a collaboration[9] with Bruno Masiero and Janina Fels, ITA Aachen), we have investigated the sound localization when listening via crosstalk cancellation (CTC) systems, which aim at providing binaural signals via loudspeakers to the listener. The research question was the *quality of binaural and spectral localization cues provided by CTC systems with listener-specific (i.e., individualized) and non-individualized HRTFs*. The novelty of our approach was the distinction between the same HRTF measurement of a listener (i.e., individualized matched) and two different HRTFs measurements of the same listener (i.e., individualized mismatched)[10]. We showed that the *individualized matched CTC systems are able to provide performance similar to that from the binaural listening*. We also showed that mismatch and lack of individualization yielded a substantially degraded performance for targets placed outside of the loudspeaker span and behind the listeners, indicating that individualized CTC systems are important for the reproduction of targets *behind* the listener. In contrast, *for targets in the same hemisphere as the loudspeakers, the individualization was of little advantage only* (Majdak et al., 2013b).

Our previous results on HRTF simulations (Kreuzer et al., 2009), showed that the perceptual evaluation of the simulations as an effect of the systematic parameter variation is a challenging task because of the large number of experimental conditions and the time-consuming psychoacoustic tests. Based on these results, it became clear that a sound-localization model predicting the localization performance is required to reduce the number of conditions for further behavioral evaluations in localization experiments. Thus, within the project LocaPhoto[6], we developed sound-localization models capable to perceptually evaluate HRTFs.

First, we *developed a model of the direction-continuous time-of-arrival in HRTFs* (Ziegelwanger and Majdak, 2014). This model robustly estimates a continuous function of the broadband delay from a set of HRTFs. This allows to evaluate HRTFs with respect to their directional temporal cues and, when applied on a binaural HRTF set, with respect to their ITDs. We showed that, when applied to acoustically measured HRTFs of over 140 listeners, the consideration of the position of the listeners during the HRTF measurement is essential in order to robustly model the broadband delay resulting from that measurement. When applied to simu-

9 An internally funded short project under my supervision.

10 Note that both HRTF sets (original and second measurement) were listener-specific, i.e., they yielded similar localization performance. However, they differed on the numerical level and on the level of spectral features.

lated HRTFs, a comparison of the model parameters (radius and ear positions) allows to validate the simulated HRTFs with respect to the broadband delays and ITDs.

Second, we *developed a model predicting the listener's sound-localization performance in sagittal planes*, i.e., based on the spectral cues in HRTFs (Baumgartner et al., 2014). This model was evaluated for various modifications of HRTFs, showing good correlation with the actual localization responses (correlation coefficients in the range of 0.8). The model allows *to evaluate HRTFs with respect to their directional spectral amplitude cues*.

While both models were applied for evaluation of simulated HRTFs, a simple version of that sagittal-plane localization model was applied to investigate sound localization in various spatial-audio applications (Baumgartner et al., 2013). Among other applications, we showed, *how to apply the model to evaluate and improve loudspeaker positions in audio-reproduction systems with respect the sagittal-plane sound localization*. Further, we applied the sagittal-plane localization model to investigate the contribution of acoustic and non-acoustic factors to listener-specific performance of sagittal-plane sound localization (Majdak et al., 2014a). In that study, we simulated a situation of listening with other's HRTFs and being completely re-calibrated to those HRTFs. Note that without our model, such a re-calibration would require an extensive training with modified HRTFs yielding in a demanding experimental effort including weeks of exposure to the modified cues. Our findings suggest that *the across-listener variability in sagittal-plane localization ability is only marginally determined by the acoustic factor*, i.e., the quality of directional cues found in typical human HRTFs. Rather, *the non-acoustic factors, supposed to represent the listeners' efficiency in processing directional cues, appear to be important*.

The sagittal-plane localization model and the time-of-arrival model have been implemented in the Auditory Modeling Toolbox (Søndergaard and Majdak, 2013). The AMT is a Matlab/Octave toolbox for developing and applying auditory perceptual models with a particular focus on binaural models. In addition to the model implementations, AMT provides published human data and model demonstrations. Thus, *model implementations can be evaluated by running so-called experiments* aimed at reproducing results from the corresponding publications – following the principle of *reproducible research*.

The AMT was created in 2009 by Peter Søndergaard. In 2013, because of my substantial contributions to the AMT, my programming skills, and my experience in handling open-source projects, *I was honored to become the main developer of the AMT (Majdak et al.,* *2014b)*. Since then, three releases have been published and four further models have been incorporated. Currently, the AMT consists of 22 various models, including models of the auditory periphery, models of monaural and binaural signal masking, models of speech recognition, and models of spatial percepts like sound-source localization.

The AMT is the umbrella project of the AABBA-group. The group "aural assessment by means of binaural algorithms" is a group of scientists with the goal of collaborating in the application of models of human binaural listening and understanding led by Jens Blauert. The last output of the AABBA group was a book published by Springer, where I contributed two chapters (Baumgartner et al., 2013; Søndergaard and Majdak, 2013). The AMT is, however, not limited to the contributions from the AABBA group – it is free and open for model developers and model users. The AMT, being an open-source project, is hosted on Sourceforge[11] and has been downloaded over 3000 times in the last three years.

1.2.3. Sound localization in electric hearing

CIs are hearing-devices which aim at providing hearing by means of direct electrical stimulation of the auditory system. CI systems consist of an implant driving an electrode array placed in the cochlea of a listener, and of a speech processor, providing the implant with the required acoustic information. In 2002, the state-of-the-art CI systems were not capable of providing bilateral CI listeners with sufficient ITD information and the reasons were not well understood. Thus, some of my studies[12] addressed various aspects of ITD-based sensitivity in CI listeners.[13]

We started at an investigation of the particular contribution of ITD in the fine-structure (analog to the carrier in modulated band-limited acoustic signals) and the ongoing envelope of a signal. We showed that *CI listeners are sensitive to ITD in the fine timing*, given its rate does not exceed a certain limit. Further, the results showed a weak contribution of ongoing envelope ITD (Majdak et al., 2006). In the meantime, this study was cited 35 times and had a huge impact on further development of stimulation strategies by CI manufacturers like MED-EL.

11 http://amtoolbox.sourceforge.net; last viewed on 7.7.2014.

12 Note that four of the described publications (Laback et al., 2007; Laback and Majdak, 2008a; Majdak et al., 2006; Majdak and Laback, 2009) have already been included in my PhD thesis. These publications are described here in order to underpin the consistency in the line of my research. They are not included in this habilitation thesis.

13 This work was done within an internally funded long-term project of the ÖAW. My role: Co-PI, with Bernhard Laback as PI.

Then, we more deeply investigated the contribution of ongoing envelope modulation, onset, and offset pulses. We showed that *the contribution of the onset pulses depends on the pulse rate and the offset pulses have a rather weak contribution at all* (Laback et al., 2007).

An important finding from these two studies was that *the fine-structure ITD sensitivity is limited to low pulse rates*. However, high pulse rates seem to be important to provide sufficiently loud auditory sensation and to appropriately encode speech information. In order to overcome the weak contribution of the fine-structure ITD cues in high-rate stimulation, we investigated the ITD-based sensitivity when stimulating with electric pulse trains with binaurally-synchronized timing jitter. In those stimuli, the ITD was maintained, but the timing of binaurally-synchronized pulse pairs was jittered. We showed evidence for large improvements in ITD sensitivity with those stimuli, and published the results in the prestigious journal PNAS (Laback and Majdak, 2008a). This publication received considerable world-wide attention in the scientific communities (cited 28 times so far) and triggered both psychophysical and physiological follow-up research in different labs, also with our contributions (Goupell et al., 2010a). Finally, this study yielded a European patent (Laback and Majdak, 2008b) and a U.S. patent (Laback and Majdak, 2008c).

We have contributed the effect of the binaurally-synchronized jitter to a kind of recovery from binaural adaptation, an effect well-known in NH listeners. In NH listeners, however, rapid changes in the timing also induce changes in the place of excitation at the basilar membrane. In order to investigate the contribution of place and timing, we investigated the effect of binaurally-synchronized jitter on ITD sensitivity in NH listeners. First, *we showed that also in NH listeners, jitter improved ITD sensitivity in high-rate stimulation*. Further, by using a physiologically-based model of auditory nerve and brainstem (in particular, medial superior olive neurons) we found that *jitter increased firing synchrony in the auditory periphery, suggesting that jitter is possibly related to changes in the temporal firing pattern, not spectral changes* (Goupell et al., 2009).

Sometimes it is advantageous to simulate CI listeners' performance in NH listeners. When simulating rate effects, the tonotopic place of stimulation in NH listeners, i.e., center frequency of acoustic stimuli, might play a role. Thus, we investigated the effects of center frequency and pulse rate on the sensitivity to ongoing envelope ITD in NH listeners using band-pass-filtered pulse trains (i.e., a simulation of the electric stimulation in CI listeners). For the tested center frequencies, sensitivity decreased with increasing pulse rate and the lack of an interaction between pulse rate and center frequency indicated that 1) *auditory filtering was not the rate limiting factor in ITD perception* and 2) *ITD sensitivity between CI and NH listeners stimulated with high-frequency bandpass-filtered pulse trains is not confounded by the choice of center frequency* (Majdak and Laback, 2009).

As for the rather weak sensitivity to the envelope-based ITD in CI listeners, we investigated the contribution of envelope shape properties to the ITD sensitivity. Our results showed that in *the sensitivity improved with increasing silent period in each modulation cycle and with increasing peak level* (Laback et al., 2011). These results provide important rules for the encoding of speech in electric stimulation strategies when aiming for improving the CI listener's sensitivity to envelope ITD cues.

While many aspects on ITD sensitivity are still unclear, also more complex binaural phenomena are interesting. An example is the binaural interference, a reduction of the sensitivity to binaural cues in a narrow-band target caused by the presence of a simultaneous narrow-band interferer in a remote spectral region. The basis for binaural interference is the binaural integration of information across spectral components, in which a single auditory object is formed when monaural signal properties indicate that these components correspond to the same acoustic object. In 2007, a "release" from binaural interference was shown by means of embedding temporally the target in a stream of identical stimuli, showing the relevance of temporal and spectral cues for auditory grouping (Best et al., 2007). It was, however, not clear if the same auditory grouping mechanisms apply also in the rather different electric hearing. This question is especially interesting because the tonotopic separation between the target and interferer is more weak in electric hearing and no evidence for a release in electric hearing would suggest a peripheral interference in acoustic hearing. Thus, *we investigated the contribution of ITDs and ILDs to binaural interference in CI listeners and showed that, under particular conditions, the binaural interference can appear*. We further showed *a robust release from binaural interference in electric hearing*, indicating that also in acoustic hearing, the interference was at least partly centrally mediated (Best et al., 2011).

Working so many years on ITD sensitivity in CI listeners allowed to accumulate much knowledge on that research topic. Thus, recently, we were invited to submit a review article about ITD sensitivity in CI listeners to Hearing Research. The online version of the article appeared on October 19th this year (Laback et al., 2014).

Currently, I work on two ITD-related projects. First, a project about the ITD sensitivity in a *multielectrode stimulation* was granted by the CI-company MED-EL.[14] Second, in collaboration with Bertrand Delgutte (Massachusetts Eye and Ear Infirmary, MIT, Boston, USA), an NIH-project was granted.[15] In that project, we will devise and test new approaches for delivering binaural information effectively, and test whether *chronic CI stimulation* can help the brain to develop the circuitry for binaural processing in animals with early-onset deafness. While the physiological aspects will be tested by the MIT team in animals, the corresponding psychoacoustic aspects will be tested by our team in human listeners.

1.2.4. Acoustics and audio-engineering beyond spatial hearing

Acoustics and audio engineering is a very wide area. It touches both fundamental scientific research and implementations of applications. In my work, I aim at linking these two complementary fields. As for the fundamental research, I am interested in very basic questions about our auditory system, e.g., my study about effects of HRTFs and non-acoustic factors on sound-localization (Majdak et al., 2014a, see previous sections). As for the applications, I am interested to implement the state-of-the-art findings in research tools and applications, e.g., SOFA API, AMT, and SOFAlizer.

Acoustics and audio engineering also touches fields like neurophysiology, psychoacoustics, signal processing, and electro-engineering. I aim at combining them, being aware of the impossibility of becoming an expert in all of those fields. Collaborations with neurophysiologists like Bertrand Delgutte (MIT, USA) on the one hand, and audio-engineers like Markus Noisternig (IRCAM, Paris) on the other hand, help me to be up-to-date with the big picture of acoustics and audio engineering even beyond the spatial hearing. At the ARI, with its interdisciplinary research focus within acoustics, I had several opportunities to collaborate outside the field of spatial hearing.

For example, I have developed *a time-frequency method for increasing the signal-to-noise ratio (SNR) in system identification with exponential sweeps* (Majdak et al., 2011a). This work was done within a WWTF-funded project MULAC.[16] With that method, the SNR in noisy measured impulse responses such as room impulse responses (RIRs) can be *post-hoc* im-

14 My role: PI, together with Bernhard Laback. Within that project, I am the PhD supervisor of Katharina Egger.

15 My role: Co-PI (PI: Bernhard Laback).

16 My role: key researcher, responsible for the applications in acoustics. PI: Peter Balazs.

proved. A journal version of the manuscript is under preparation for a submission to the IEEE TASLP. As an other example, I was investigating the effect of train type on annoyance and acoustic features of the rolling noise (Kasess et al., 2013). In that study[17], the mental state "annoyance" was investigated as an effect of acoustic presentation with train passbys recorded with various breaking systems. Statistical modeling showed that *the A-weighted energy equivalent sound level contributed most to the annoyance*. In conditions in which that level was not available as a cue, the *spectral spread significantly contributed*. These two studies are examples of my contribution to acoustics and audio-engineering in general.

1.3. Acknowledgments

I thank all my co-authors and cooperation partners, who provided me with a lot of interesting ideas, productive comments, and nice research projects. Thank you!

I thank Werner A. Deutsch, for taking me in the ARI family and providing me with perfect conditions for working in a productive, open, and multidisciplinary environment. I warmly thank Bernhard Laback for introducing me to the wonderful life of science and his ongoing support since the start of my work at the ARI. I thank Katharina Egger, Robert Baumgartner, Harald Ziegelwanger, and Michael Mihocic for working hard on translating my ideas and visions into action. I also thank all the other people at the ARI, especially Peter Balazs, for their discussions and help on almost every aspect of science. Thank you!

I am also grateful to many colleagues from the IEM, where while finishing my studies more than a decade ago, my interest in the research topic has started. I especially thank Robert Höldrich and Alois Sontacchi, for being vital at various points of my scientific career as well as for providing me protected freedom to work and teach. Thank you!

Many thanks also go to my family. I thank my wife Ali, for being a wonderful companion supporting my life. I thank my kids Chiara and Sarah for putting a smile on my face even when tired. And I am grateful to my parents Ela and Zbyszek, for patiently teaching me that education is an important part of life. Thank you!

17 My contribution: experimental design, psychoacoustic support, and writing. PI: Christian Kasess.

2. List of included publications

In total, 16 journal papers, five peer-reviewed proceedings, and one peer-reviewed book chapter are included in this thesis. For the list, the following applies:

- Only publications included in this thesis are listed; the full list is provided in the CV.
- The list is sorted by research topics, regardless the underlying projects.
- The manuscript Ziegelwanger, Majdak, and Kreuzer (#9 in the list) has been recently submitted to JASA. The average turn-around of my articles is 9 months, thus, I expect its publication in 2015.
- The manuscript Marelli, Baumgartner, and Majdak (#10) has been recently submitted to IEEE Transactions on Audio, Speech and Language Processing. I expect its publication in 2015.
- The article Laback, Egger, and Majdak (#22) has been online published on October 28th as an uncorrected proof. While our corrections have been already submitted to the editor, only the uncorrected proof is included in this thesis.
- The actual status of my publications can be tracked at http://tiny.cc/MajdakPubs.

2.1. HRTFs: Acquisition, approximation, and applications

1. Majdak, P., Balazs, P., Laback, B. **(2007)**. "Multiple Exponential Sweep Method for Fast Measurement of Head Related Transfer Functions," J. Audio. Eng. Soc. **55**, 623-637.
2. Marelli, D., Fu, M., Balazs, P., Majdak, P. **(2008)**. "An Iterative Method for Approximating LTI Systems using Subbands," in the proceedings of the *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas,* 3533-3536.
3. Kreuzer, W., Majdak, P., Chen, Z. **(2009)**. "Fast multipole boundary element method to calculate head-related transfer functions for the whole audible frequency range", J. Acoust. Soc. Am. **126**, 1280-1290.
4. Reichinger, A., Majdak, P., Sablatnig, R., Maierhofer, S. **(2013)**. "Evaluation of Methods for Optical 3-D Scanning of Human Pinnas," in Proceedings of the *3D Vision Conference 2013, Third Joint 3DIM/3DPVT Conference*, Seattle, WA, 390–397.
5. Majdak, P., Masiero, B., Fels, J. **(2013)**. "Sound localization in individualized and non-individualized crosstalk cancellation systems," J. Acoust. Soc. Am. **133**, 2055-2068.
6. Ziegelwanger, H., and Majdak, P. **(2014)**. "Modeling the direction-continuous time-of-arrival in head-related transfer functions", J. Acoust. Soc. Am. **135**, 1278-1293.
7. Marelli, D., Baumgartner, R., Majdak, P. **(2014)**. "Efficient Representation of Head-Related Transfer Functions in Subbands," in the proceedings of the *22nd European Signal Processing Conference (EUSIPCO)*, Lisabon.

8. Ziegelwanger, H., Majdak, P., Kreuzer, W. **(2014)**. “Non-uniform sampling of geometry for the numeric simulation of head-related transfer functions,” in the proceedings of the 21st *International Congress on Sound and Vibration (ICSV)*, Beijing, China, Proceedings Number 209.

9. Ziegelwanger, H., Majdak, P., and Kreuzer, W. **(submitted)**. “Numerical calculation of perceptually valid head-related transfer functions from listener’s geometry,” submitted to J. Acoust. Soc. Am.

10. Marelli, D., Baumgartner, R., Majdak, P. **(submitted)**. “Efficient Approximation of Head-Related Transfer Functions in Subbands for Accurate Sound Localization,” submitted to IEEE Trans. Audio Speech Lang. Proc.

2.2. Sound localization beyond the horizontal plane

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

12. Goupell, M. J., Majdak, P., 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.

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

14. Baumgartner, R., Majdak, P., Laback, B. **(2013)**. "Assessment of sagittal-plane sound localization performance in spatial-audio applications," in *The Technology of Binaural Listening* (Chapter 4, p. 93-119), edited by Blauert, J. (Springer, Berlin, Heidelberg).

15. Majdak, P., Walder, T., Laback, B. **(2013)**. "Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions," J. Acoust. Soc. Am. **134**, 2148-2159.

16. Baumgartner, R., Majdak, P., Laback, B. **(2014)**. “Modeling sound-source localization in sagittal planes for human listeners”, J. Acoust. Soc. Am., **136**, 791-802.

17. Majdak, P., Baumgartner, R., Laback, B. **(2014)**. “Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization,” Front. Psychol., **5**, 319(1–10).

2.3. Sound localization in electric hearing

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

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

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

21. Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W., Pok, S. **(2011)**. "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing," J. Acoust. Soc. Am. **130**, 1515-1529.

22. Laback, B., Egger, K., Majdak, P. **(2014)**. “Sensitivity to the interaural time differences in cochlear-implant listeners”, Hear. Res., DOI: 10.1016/j.heares.2014.10.004.

3. Introduction and summary

The publications included in this thesis address various research questions within the field of spatial hearing. In this section, that research field is framed within the habilitation discipline *acoustics and audio engineering*. Further, a short introduction is provided to those areas of spatial hearing that are addressed in the included publications.

3.1. Acoustics and audio engineering

Acoustics is the branch of science dealing with the study of all mechanical waves in gases, liquids, and solids including vibration and sound. Acoustics is an interdisciplinary science, usually being closely related to hearing: the word "acoustic" is related to the Greek words ἀκουστικός (*akoustikos*, meaning "of or for hearing, ready to hear") and ἀκουστός (*akoustos*, "heard, audible"). Nevertheless, the central subject of acoustics is wave propagation. Thus, acoustics includes also fields like ultrasound, infrasound, and vibration.

Audio engineering is the branch of engineering dedicated to recording, manipulation, and reproduction of sound. It links the creative and practical aspects of sounds (including speech and music) with the development of technologies and the advance in scientific understanding of hearing. Audio engineering can be thought as a part of *acoustic* engineering, which deals with sound and vibration. Both engineering branches rely on acoustics as their underlying scientific basis.

Hence, the topic "acoustics and audio engineering" spans a line from basic science to application, including hearing in humans, of which spatial hearing is an important aspect.

3.2. Spatial hearing

Spatial hearing refers to the ability of perceiving sound sources in the three-dimensional space we live in. It is a fundamental aspect of auditory perception. Being previously the basic requirement for survival (acoustically detecting the attack of enemies and tracking the location of a prey), nowadays, it is the primary ability for detecting the direction of sound events in critical situations like road traffic, for orienting in reverberant spaces, and for segregating sound sources in a multisource environment in order to improve communication between individuals. Spatial hearing has been subject of many investigations (for extensive reviews, see

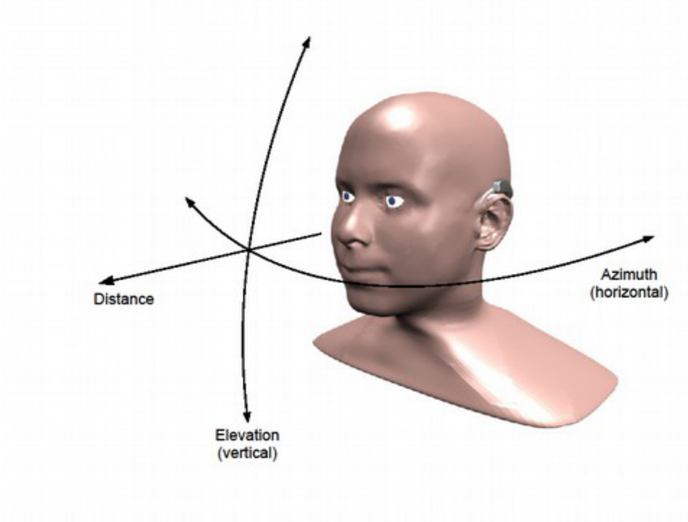

Figure 1: Three dimensions in spatial hearing.

Blauert, 1974; Gilkey and Anderson, 1997; Stecker and Gallun, 2012; Xie, 2013; Yost and Gourevitch, 1987).

From the psychoacoustic point of view, spatial hearing manifests in various perceptual effects like externalization, apparent source width, listener envelopment, spatial unmasking, and, of course, sound localization, i.e., the ability to estimate the position of sound sources. As for the sound-source position, one has to separate between the distance and the directional angles (see Fig. 1). Even from the geometrical point of view, the distance (as the radius) and the angles have distinct roles in a spherical coordinate system. This also applies to the ability and mechanisms for estimation of the distance and the direction of a sound source in humans. In this thesis, I focus on the sound localization as a function of direction and describe the role of HRTFs in the estimation of the sound direction. For extensive reviews on distance perception in humans, see e.g., Kopčo et al. (2012); Zahorik et al. (2005).

Usually, the spherical coordinate system is used to describe the position of a sound: the azimuth angle describes the horizontal direction and the elevation (or zenith) angle describes the vertical direction (see Fig. 2, right panel). However, the vertically placed poles in the spherical coordinate system do not correspond with the horizontally aligned ears of a listener, and as a consequence, the different types of localization cues (see Sec. 3.4) cannot be easily related to the angles. The interaural-polar coordinate system (see Fig. 2, left panel), sometimes also called double-pole or horizontal-polar system, better represents the correspondence between the angles and the localization cues (Morimoto and Aokata, 1984). In the interaural-polar coordinate system, the lateral angle represents the lateral direction of the sound and defines a

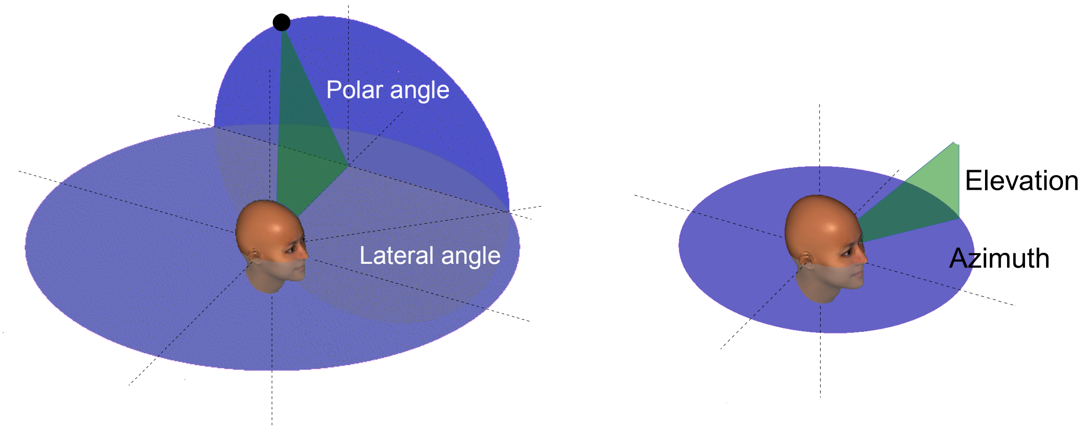

Figure 2: Left panel: Interaural-polar coordinate system for sound localization in humans. Right panel: geodetic spherical coordinate system usually used in acoustics and mathematics.

sagittal plane (SP, a plane parallel to the median plane). Within a sagittal plane, the polar angle defines the particular direction of the sound source. Since the interaural-polar coordinate system is more appropriate when discussing effects related to spatial hearing, it is also used in this thesis.

3.3. Head-related transfer functions

The sound arriving at the ear drum is filtered by the ear, head, and torso of the listener. For a spatially compact sound source, the effect of that acoustic filtering can be described in terms of a linear time-invariant system by the head-related transfer functions (HRTFs). HRTFs vary with the position of the sound source (see Fig. 3) and are given for the two ears. Thus, a binaural set of HRTFs for each position is required to describe the interaural and monaural spatial features available for sound localization.

HRTFs depend on both the source direction and the distance between listener and source. For distances larger than 1 m, HRTFs do not change much with the distance, thus they can be considered as “far-field HRTFs”. For distances below 1 m, HRTFs change with the distance, thus they are sometimes called “near-field HRTFs” (Brungart and Rabinowitz, 1999). In the near-field HRTFs, particularly for lateral directions, the low-frequency ILDs increase with decreasing distance, and they can be a cue for the estimation of the sound source distance even in free field (Brungart et al., 1999). Usually, when only directional effects are investigated, far-field HRTFs are implicitly assumed.

HRTFs also depend on the individual geometry of the listener and thus listener-specific HRTFs are required to achieve accurate localization performance particularly along the sagit-

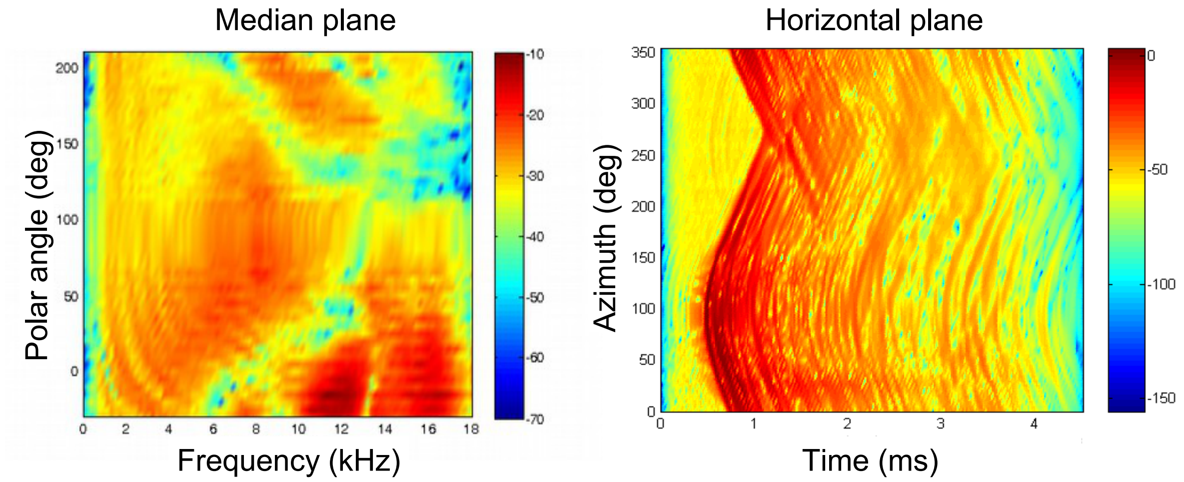

Figure 3: Head-related transfer functions of the left ear of NH5. Left: Amplitude spectra of the directional transfer functions in the median plane. Right: Energy-time curves of the head-related impulse responses in the horizontal plane. The color encodes the amplitude in dB re max.

tal planes (Bronkhorst, 1995; Middlebrooks, 1999a). HRTFs may contain both direction-dependent and direction-independent features and thus can be thought of as a series of two acoustic filters. The direction-independent filter, represented by the common transfer function (CTF), can be calculated from an HRTF set comprising many directions by averaging the log-amplitude spectra of all available HRTFs of a listener's ear (Majdak et al., 2010; Middlebrooks, 1999b). The directional features are represented by the directional transfer functions (DTFs) which can be calculated by filtering the corresponding HRTF with the inverse CTF. The CTF usually exhibits a low-pass filter characteristic because higher frequencies are attenuated as an effect of the head and pinna shadow. Thus, compared to HRTFs, DTFs pronounce spectral features above 4 kHz and are commonly used to investigate the nature of spectral cues in localization experiments in sagittal planes (Goupell et al., 2010b; Majdak et al., 2010; Middlebrooks, 1999b).

3.3.1. HRTF acquisition

HRTFs can be measured acoustically. The acoustic measurements are done by inserting a microphone into the ear canal of a listener and performing electro-acoustic system identification. HRTFs can be considered as linear time-invariant systems, thus, for the acquisition of HRTFs, tools from linear signal processing can be applied. Many system identification methods suitable for HRTF measurements exist (for a recent review, see Enzner et al., 2013)

The electro-acoustic chain of HRTF measurements usually consists of some weakly non-linear parts like loudspeakers and power amplifiers, requiring a special handling of the linear and non-linear components in the measurements. System identification with frequency sweeps allows to partially separate the linear and non-linear components in the measurements (Müller and Massarani, 2001). Thus, sweeps have been widely used in acoustics, not only in the measurement of HRTFs, but also in the measurement of room impulses and in-situ measurement of material parameters like absorption coefficients (Mommertz, 1995).

In an HRTF measurement, hundreds of directions are usually considered, rendering the measurement a longsome procedure lasting for tens of minutes. For human listeners, it is hard to keep still for such a long period of time without reduction of the spatial accuracy of the measurement. For a mannequin, such a fixation is easily feasible. Thus, several attempts have been proposed to speed-up the measurements for human listeners. They are based on the exploration of the orthogonality in time and frequency. For example, normalized least square methods have been used to retrieve HRTFs from uncorrelated noises presented simultaneously from different directions (Enzner et al., 2013). This method relies on the strict linearity of the electro-acoustic chain. Another approach is the multiple exponential sweep method (MESM) which uses sweeps overlapping in time such that 1) the linear responses from different directions can be separated, and 2) the nonlinearities can be separated and discarded (Majdak et al., 2007). With the MESM, in a semi-anechoic chamber, with 22 loudspeakers playing 1.7-s long sweeps, HRTF measurement for 1550 directions takes 16 minutes. Recently, the MESM has been further optimized for HRTF measurement in an anechoic chamber (Dietrich et al., 2013) allowing to measure HRTFs with a continuous rotation of 40 loudspeakers within 2 minutes (Dietrich et al., 2012).

HRTFs can also be obtained from simulation. The simulations can be based on empirical approaches where parameters of a structural model are determined (Brown and Duda, 1998; Spagnol et al., 2013a). The simulations can also be based on numerical calculations of the sound pressure based on the numeric representation of a listener's geometry (Katz, 2001). For the latter approach, boundary-element methods (BEM; Ciskowski and Brebbia, 1991) and finite-difference time-domain methods (Takemoto et al., 2012) have been successfully applied. In both approaches, an accurate representation of the listener's geometry is required, thus, the calculations are usually very time and resources consuming. For a long time, the simulated HRTFs were available only for either limited frequency range (Katz, 2001) or for parts of human morphology (Kahana and Nelson, 2007). Recently, by improving the numerical methods, namely, by coupling the BEM with the fast-mulitpole method (Greengard and Rokhlin, 1987), HRTF calculation for the whole audible frequency range has been established (Gumerov et al., 2010; Kreuzer et al., 2009). While the HRTF calculation for human mannequins seems to be achievable (Gumerov et al., 2010), for actual human bodies some work is still required. Recently, numerically calculated HRTFs have been validated in terms of providing sound localization performance equivalent to that usually found when listening with acoustically measured HRTFs (Ziegelwanger et al., 2013b). Further work on enabling HRTF simulations available for commercial applications includes the improvements in geometry capturing (Reichinger et al., 2013; Spagnol et al., 2013b) and speeding up the numerical calculations (Ziegelwanger et al., 2014).

Finally, listener-specific HRTFs can be obtained by customizing others' HRTFs. This can be done on the basis of anthropometric data (Middlebrooks, 1999a; Hwang and Park, 2008; So et al., 2010), acoustic features (Iida et al., 2014), psychoacoustic tests (Katz and Parseihian, 2012), or combinations of these approaches (Middlebrooks et al., 2000). All of these approaches showed improvements as compared to listening with non-individualized HRTFs. Nevertheless, the problem of acquiring listener-specific HRTFs for accurate sound localization seems to be not fully solved yet.

3.3.2. HRTF applications: binaural synthesis

The HRTFs filter sounds presented in free field via loudspeakers (Fig. 4, top-left). When the HRTFs are known, they can be used to simulate that filtering for the reproduction via headphones (Fig. 4, bottom-left). Such a technique is usually called binaural synthesis, and the whole apparatus can be called virtual auditory display (Xie, 2013).

The presentation of binaural signals can also be targeted in loudspeaker reproduction systems. To this end, crosstalk propagation paths between the loudspeakers and ears must be considered (Kirkeby et al., 1998). This can be done by using so-called crosstalk cancellation (CTC) systems (see Fig. 4, right). The concept of presenting virtual sound sources via CTC systems includes HRTFs used for the spatialization and HRTFs used to cancel the crosstalk propagation paths (“Setup HRTFs” in Fig. 4). In simulations, it has been shown that the sweet spot depends on the signal frequency, and the concept of “optimal source distribution” has been proposed in order to create robust and large sweet spots in CTC systems (Takeuchi and Nelson, 2002). In a simulation of the binaural processing of the auditory system, it has been shown that accurate binaural cues can only be provided when the setup HRTFs match the ac-

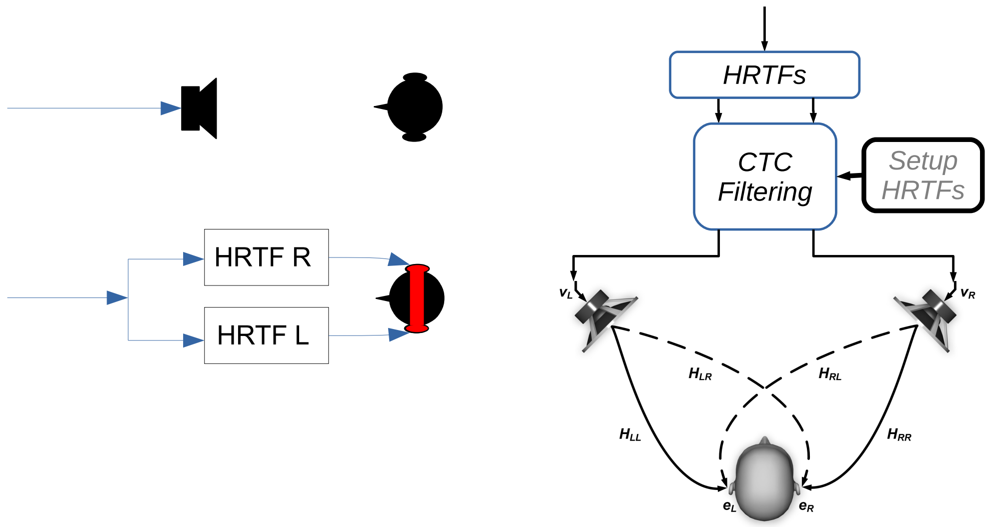

Figure 4: Top left: Free-field sound reproduction via loudspeakers. Bottom left: Binaural synthesis via headphones. Right: Binaural synthesis via loudspeakers.

tual propagation paths (Akeroyd et al., 2007). As shown in actual sound localization experiments, for localization along the lateral dimension, the exact knowledge of the propagation paths is not required, while it is required for the localization in the sagittal planes (Majdak et al., 2013b), especially for the simulation of sound sources placed behind the listener.

Despite the different ways of creating virtual sound sources, an exchange of HRTFs between the various audio systems is required. As the acquisition of HRTFs became more and more popular, each laboratory used its own format of storing the data. The first attempt to unify the different HRTF databases was limited to the Matlab file format (Andreopoulou and Roginska, 2011). A more general format is the spatially oriented format for acoustics (SOFA), which allows not only to store HRTFs, but also general multidimensional data recorded in various spatial arrangements, e.g., directional room impulse responses recorded with a multichannel microphone array (Majdak et al., 2013a). In SOFA, various conventions, i.e., rules defining the description of particular type of data, are used. Currently, conventions for free-field HRTFs, for binaural room impulse responses, as well as for headphone impulse responses are available.[18] SOFA is not restricted to Matlab; actually, netCDF (Rew and Davis, 1990), the numerical container used in SOFA, is widely used for storing multidimensional data in the field of climatology, meteorology, oceanography, and geographic information systems, offering interfaces for many programming languages. SOFA is currently undergoing a standardization by the Audio Engineering Society (Noisternig and Majdak, 2014).

18 Available at http://sofaconventions.org, last viewed on 23.10.2014

3.3.3. HRTF approximation

In virtual auditory displays, audio signals are filtered with HRTFs yielding virtual sound sources. In order to obtain a computationally efficient filtering process with a low latency, various approaches have been proposed. On the one hand, as an improvement of the overlap-add and overlap-save methods (Proakis and Manolakis, 1996), low-latency convolution approaches have been developed (Gardner, 1995; Wefers and Vorländer, 2012). These methods, here called segmented fast Fourier transform (SFFT) methods, are an attractive option in virtual auditory and auralization systems (Vorländer, 2007). They offer a trade-off between latency demands and computational complexity.

On the other hand, the computational efficiency can be improved by approximating HRTFs, i.e., allowing an error in a specific representation of the HRTFs. Various methods have been proposed, with the shortening of the impulse responses as the simple one (Senova et al., 2002). Further shortening of the impulse responses can be achieved by implementing them as a minimum-phase system combined with broadband delay (Mehrgardt and Mellert, 1977). This method relies on the assumption that the interaural differences of the broadband delays, i.e., the ITDs, represent the main cue for the sound localization in the horizontal plane (Kulkarni et al., 1999). Thus, it is not surprising that much effort has been put towards ITD approximations (Kuhn, 1977; Savioja et al., 1999; Wang et al., 2009). These ITD-based models rely on a robust estimation of the broadband delays for both ears, for which many methods exist and which is not a trivial task (Defrance et al., 2008). Recently, various ITD estimation methods have been compared and their advantages and disadvantages have been discussed (Katz and Noisternig, 2014). For a given direction, the estimation of the ITD relies on the binaural estimation of the time-of-arrival (TOA), of which the interaural difference is the ITD. Given a listener-specific HRTF set, the three-dimensional TOAs can be estimated for each ear. Recently, a direction-continuous model of the TOA has been proposed (Ziegelwanger and Majdak, 2014), the parameters of which are fit to a monaural HRTF set. The fitted model can then be used to direction-continuously estimate the TOAs, and thus approximate HRTFs in a more compact way (Ziegelwanger et al., 2013a).

Further approximation concepts consider the representation of HRTFs in various domains. For example, HRTFs have been represented by using orthogonal eigentransfer functions (Chen et al., 1995), spherical harmonics (Evans et al., 1998), subbands (Marelli et al., 2008, 2014), or a sparse number of directions combined with directional interpolation (Freeland et

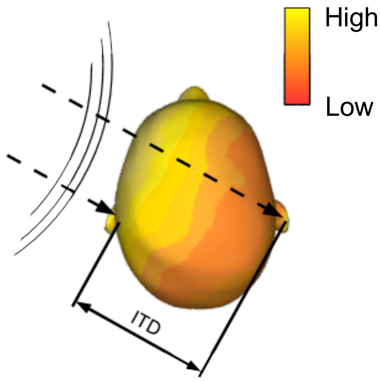

Figure 5: Binaural cues for sound sources located away from the median plane.

al., 2004; Matsumoto et al., 2004; Minnaar et al., 2005). Especially the representation of HRTFs in subbands (Marelli and Fu, 2007) seems to be promising because it offers a combination of the advantages of HRTF approximation considering perceptually relevant properties (ERB-scale, logarithmic amplitude scale) and the trade-off between the latency and computational complexity offered by the SFFT methods (Marelli et al., 2008).

Further, HRTFs have been approximated by pole-zero transfer functions (Blommer and Wakefield, 1997; Haneda et al., 1999; Kulkarni and Colburn, 2004). While a high efficiency can be achieved (only six poles and six zeros required), this representation suffers from the problem of coefficient commutation in real-time applications.

3.4. Spatial cues used for sound localization

3.4.1. Sound localization in the horizontal plane

Both ITD and ILD (see Fig. 5) are used by the binaural auditory system in order to estimate the position of sounds along the lateral dimension. For example, in an experiment conducted via headphones, pure tones can evoke a shift in the perceived lateral direction of the sound when an ITD or ILD is applied, an effect called lateralization. While ITDs occur across the entire audible frequency range (Kuhn, 1977), ILDs are more pronounced at high frequencies where the wavelength is in the range of physical dimensions of the head (Blauert, 1974).

The relevance of the interaural differences for sound localization has already been explored more than hundred years ago (Strutt, 1907) and the theory of the localization of sounds in the lateral dimension has come to be known as the duplex theory. However, already in 1876, John

Strutt (also known as Lord Rayleigh) realized that binaural hearing is not required for the localization of sounds in the sagittal planes:

"*The possibility of distinguishing a voice in front from a voice behind thus appear to depend on the compound character of the sound in a way that is not easy to understand, and for which the second ear would be of no advantage.*" (Strutt, 1876)

Macpherson and Middlebrooks (2002) revisited the duplex theory and investigated the contribution of ITDs, ILDs, monaural, and interaural spectral cues to the perception of the sound position along the lateral dimension. They found that the interaural cues are sufficient to achieve a sufficient lateralization accuracy, which is in agreement with the duplex theory. Furthermore, they could show that the spectral cues, both monaural and interaural, have only little impact on the perceived lateral angle. Thus, for lateralization, on the one hand, the binaural pair of HRTFs is required, on the other hand, their broadband average seems to be sufficient. Additionally, a comparison between ILD and ITD showed a higher weighting for ITD than for ILD, demonstrating the significance of the broadband TOA as an important cue in the lateralization of sounds.

While it is usually assumed that ITDs at lower frequencies (below 1.5 kHz) mainly contribute to sound lateralization, listeners are also sensitive to ITDs in high-frequency complex sounds (Klumpp and Eady, 1956). In particular, human listeners are sensitive to ITDs in the slowly-varying envelope of broadband signals (Bernstein and Trahiotis, 1994; Henning, 1974; Nuetzel and Hafter, 1976). This is also supported by physiological studies at the neural level (Shackleton et al., 2003; Skottun et al., 2001; Yin et al., 1984), showing strong evidence that our auditory system is able to extract timing information from the envelopes of high-frequency sound components and detect the ITD in the envelope. Thus, two types of ITDs can be defined (see Fig. 6). First, the envelope ITD (ENV ITD), which is the interaural delay in the signal envelope, i.e., the slowly-varying fluctuations of a signal. Second, the fine-structure ITD (FS ITD), which is the interaural delay in the fast-varying fine structure of a signal.

The ITD sensitivity to amplitude-modulated high-frequency sounds seems to be generally lower than that to low-frequency sounds (Henning, 1974). When transposed tones are used as stimuli (van de Par and Kohlrausch, 1997), aiming at mimicking the neural response of the auditory filters to low-frequency sounds, a better sensitivity can be achieved, especially for modulation rates below 300 Hz (Bernstein and Trahiotis, 2002). However, by increasing the modulation frequency, the performance for transposed tones decreases rapidly while the perfor-

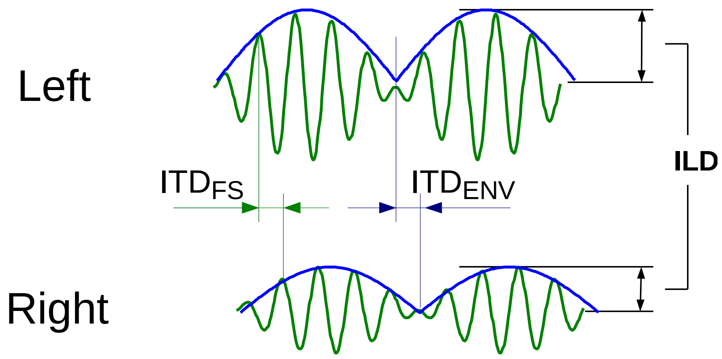

Figure 6: ILD, fine-structure ITD (ITD_{FS}) and envelope ITD (ITD_{ENV}) in a modulated stimulus.

mance for pure tones with corresponding frequencies remained approximately constant. This has been discussed in terms of a possible effect of peripheral processing, namely, auditory filtering, which smears the envelope fluctuations, resulting in a lower modulation depth of the internal representation of the acoustic signal, which in turn deteriorated ITD sensitivity. The rate limit in ITD sensitivity is still a matter of discussion (Bernstein and Trahiotis, 2002, 2014; Majdak and Laback, 2009).

3.4.2. Sound localization beyond the horizontal plane

While interaural differences in time and intensity are important for sound localization in the lateral dimension, monaural spectral cues are assumed to be the most salient cues for sound localization in the sagittal planes (Macpherson and Middlebrooks, 2002; Wightman and Kistler, 1997). In particular, monaural spectral cues are essential for the perception of the source elevation within a hemifield (Asano et al., 1990; Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002) and for front-back discrimination of the perceived auditory event (So et al., 2010; Zhang and Hartmann, 2010). The different contribution of the interaural and monaural spectral cues can also be observed by using the interaural-polar coordinate system (see Fig. 2). In that coordinate system, the contribution of the interaural cues can be well described in terms of the lateral angle, and the contribution of the spectral cues can be well described in terms of the polar angle (Best et al., 2005; Carlile and Pralong, 1994; Middlebrooks, 1999b).

Although spectral cues are processed monaurally, information from both ears affects the perceived location in most cases (Morimoto, 2001). The relative contribution of the ipsilateral ear (the one closer to the source) increases with increasing lateral angle of the sound source (Hofman and Van Opstal, 2003), becoming dominant over the contralateral ear for lateral angles beyond approximately 60°. Thus, even for SP sound localization, the lateral source position, mostly depending on the broadband binaural cues (Macpherson and Middlebrooks, 2002), is important to determine the binaural weighting of the monaural cues.

When it comes to understand the relevance of spectral features for sound localization, in general, the pinna plays a larger role for higher frequencies (Middlebrooks and Green, 1991) and the torso for lower frequencies (Algazi et al., 2001). Further details on the nature of the spectral features are still subject of investigations. On the one hand, the role of the macroscopic spectral features has been investigated (Asano et al., 1990; Goupell et al., 2010b; Hwang and Park, 2008; Kulkarni and Colburn, 1998, 2004; Langendijk and Bronkhorst, 2002; Macpherson and Middlebrooks, 2003; Senova et al., 2002). On the other hand, the role of only a few local spectral features in sound localization has been investigated (Iida et al., 2007, 2014; Middlebrooks and Green, 1992; Zhang and Hartmann, 2010). Generally, spectral-profile analysis (Green, 1988; Moore, 2012), which is the ability to discriminate different spectral shapes, is required to evaluate spectral features in a sound localization task. Recent findings show that the ability to discriminate spectral shapes correlates well with the ability to localize sounds in sagittal planes (Andéol et al., 2013). This finding seems to be in agreement with recent evidence that the listener-specific sensitivity (used to distinguish between good and poor localizers) is an important factor in explaining the listener-specific localization performance in sagittal planes (Majdak et al., 2014a).

Various models for predicting sound localization in sagittal planes have been proposed: first as a concept based on the comparison of amplitude spectra (Zakarauskas and Cynader, 1993), then as functional models comparing the actual sound and template spectra (Hofman and Van Opstal, 1998; Middlebrooks, 1992), and finally as models mapping those comparisons to statistical response probabilities (Baumgartner et al., 2014; Langendijk and Bronkhorst, 2002). The most recent model provides also psychoacoustic performance parameters used in localization experiments and it was evaluated under various conditions of HRTF modifications. Thus it can be considered as a tool for the evaluation of HRTFs in sound reproductions systems (Baumgartner et al., 2013) with respect to spectral cues of HRTFs.

In order to interpret the spectral features as directional filtering, the listener's auditory system is calibrated to its HRTFs. When the HRTFs are modified, the auditory system is able to recalibrate within a time course of several weeks only (Hofman et al., 1998; Majdak et al., 2013c). Note that this recalibration process seems to rely on a different mechanism than the adaptation processes shown to happen within minutes (Parseihian and Katz, 2012; Zahorik et al., 2006).

Besides the binaural and spectral features, sound localization performance also depends on other factors. For example, vision influences the perceived location of sounds, e.g., the ventriloquism effect (McGurk and MacDonald, 1976). Thus, the link between a visual and auditory stimulation has been investigated under various conditions (Lewald and Ehrenstein, 1998; Majdak et al., 2010; Shelton and Searle, 1980). Further, the amount of training on the task, or the response method can have an effect (Haber et al., 1993; Majdak et al., 2010; Seeber, 2002). Consideration of all those factors, and consideration of other modalities like tactile (Altinsoy, 2012) and proprioceptive (Grant and Magee, 1998) perception can potentially improve the quality of future auditory virtual environments.

3.5. Sound localization in electric hearing

A cochlear-implant (CI) system is a partially-implanted electronic device that provides a sense of sound to a severely hard of hearing or profoundly deaf person. Unlike conventional hearing aids, a CI system does not only amplify sound, but it captures the incoming sound and converts it to electric current pulses which directly stimulate parts of the cochlea. The CI system consists of an external speech processor (including a microphone, a signal-processing device, and a power supply) connected to a transmitter coil, which transmits power and the processed information, and the actual implant, which converts the received signal to electric currents and drives the electrodes placed in the cochlea (see Fig. 7).

Nowadays, common CI systems use multichannel, pulsatile stimulation (Wilson et al., 1991). The reason for using pulses instead of continuous stimulation is the broad excitation spread of electric stimulation in the cochlea. Simultaneous stimulation on more than one electrode results in an uncontrolled interaction through vector summation of the electric fields from each of the electrodes (de Balthasar et al., 2003; Pelizzone et al., 1999) which is avoided when stimulating only one electrode at a time. Most of the clinically available strategies are envelope-based, i.e., within each frequency channel, the signal envelope is extracted and used to modulate a train of electric pulses.

Bilateral CI systems become more and more popular and spatial hearing with respect to sound localization and speech understanding in noise became an important topic in the electric-hearing research (for recent reviews see van Hoesel, 2012) and Litovsky et al., 2012).

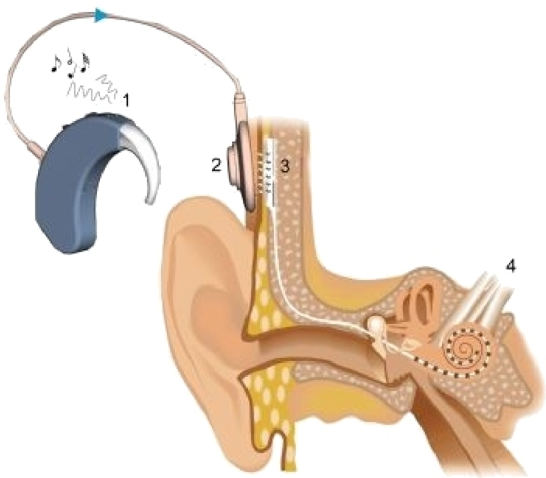

Figure 7: Cochlear implant system. (1) speech processor, (2) transmitter coil, (3) implant, (4) electrode array. From http://recorlsa.club.fr/implantcochleaire/descripgenerale.html (29.11.2007).

When used in clinical bilateral CI systems, the envelope-based strategies do not consider any special binaural coordination and the electric stimulation is controlled by two independently running speech processors. Nevertheless, for sounds positioned along the lateral dimension, the head shadow causes ILDs, to which access is easily available, assuming that speech processors have the same configuration (Noel and Eddington, 2007). Thus, it is not surprising that bilateral CI listeners using envelope-based stimulation strategies are able to estimate the lateral direction of single sound sources (Grantham et al., 2008; Laback et al., 2004; Majdak et al., 2011b; Seeber and Fastl, 2008; Senn et al., 2005).

As for ITDs, the wave propagation between the two ears acts as an acoustic synchronization, potentially offering access to ITD cues. The situation is, however, more complex (for a recent review, see (Laback et al., 2014). Generally, in the sense of bilateral CI systems using pulsatile stimulation, the ITDs from acoustic hearing translate to envelope and fine-structure ITDs in electric hearing (see Fig. 8). However, with envelope-based CI systems, the ITD can be transmitted in the envelopes only and the sensitivity has been found to be rather poor (thresholds around 250 µs in best conditions tested, (Laback et al., 2004). Interestingly, CI listeners seem to be much more sensitive to ITDs in the fine structure, at least for pulse rates up to 400 pps (van Hoesel, 2007; van Hoesel and Tyler, 2003; Majdak et al., 2006). That sensitivity remained even when tested with only four electric pulses (Laback et al., 2007).

Based on those findings, and utilizing the acoustic synchronization of the processors, clinical fine-structure-based stimulation strategies like fine-structure processing (FSP; Hochmair et

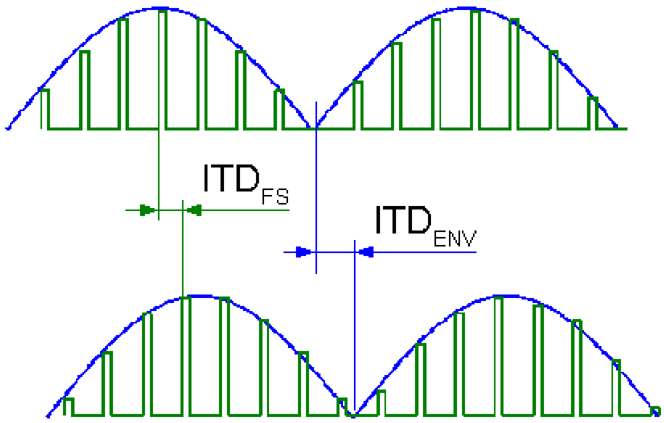

Figure 8: Fine-structure ITD (ITD_{FS}) and envelope ITD (ITD_{ENV}) in a modulated pulsatile electric stimulus. Only the positive phase of each pulse shown.

al., 2006) and fundamental asynchronous stimulus timing (FAST; Smith, 2010, 2013) have been implemented. These strategies use acoustically binaurally synchronized pulses driven at low rates, thus, transmitting fine-structure ITD to the auditory system. Note that there is also the peak-derived timing strategy, which, being a pioneer work with respect to encoding timing cues in electric stimulation (van Hoesel and Tyler, 2003), used too high rates and thus, has not come to clinical application.

While low stimulation rates are beneficial for ITD sensitivity, clinical systems usually use pulse rates above 800 pps in order to transmit the speech information (Arora et al., 2009; Loizou et al., 2000). One method to overcome that rate limit is the binaurally-synchronized jitter in the stimulation timing of the pulses (Laback and Majdak, 2008a). In that stimulation scheme, the timing between subsequent pulses is jittered while the ITD remains preserved. Jittered pulse trains have been shown to yield large improvements in ITD sensitivity as compared to regular pulse trains, even for average pulse rates up to 1515 pps. Thus, the application of binaurally-synchronized jitter in a stimulation strategy might be advantageous, particularly when using higher pulse rates. Similar effects of jitter have been shown in cats, recorded at the level of inferior colliculus (Goupell et al., 2010a; Hancock et al., 2012) where a recovery of neural firing has been shown in high-rate stimulation in the condition with binaurally-coherent jitter. Recent findings show that using two pulses within a short interval may be as effective as binaurally-synchronized jitter (Hancock et al., 2012) which might simplify its implementation in a clinical stimulation strategy.

Since the implementation of fine-structure ITD cues in electric hearing is not trivial, another way to improve spatial hearing abilities is to enhance the sensitivity to envelope ITDs.

While various parts of the envelope contribute differently to the ITD sensitivity, the off time (the sub-threshold part of the envelope) and the slope seem to substantially contribute (Laback et al., 2011). Similar results have been found in NH listeners (Klein-Hennig et al., 2011; Laback et al., 2011). Recently, an enhancement of the envelope, and thus, increase of the slope and off time, has been shown to be advantageous also in the bimodal (electric and acoustic) stimulation (Francart et al., 2014).

Sound localization can be interfered by other cues, an effect called binaural interference (for a review, see Best et al., 2007). For example, the ITD thresholds of a target can be elevated when another stimulus is simultaneously presented with a different ITD than the target (McFadden and Pasanen, 1976). In NH listeners, the binaural interference can be reduced by introducing sequential grouping, i.e., capturing the interferer in a sequence of identical bursts (Best et al., 2007). From that study, however, it was not clear, whether the binaural interference is a more peripheral or a more central process. In a follow-up study, binaural interference and its recovery have been investigated in CI listeners (Best et al., 2011), providing evidence that binaural interference is at least partially centrally mediated. Interestingly, sequential grouping yielded a recovery from binaural interference also in CI listeners.

Much of the described research was done by stimulating a single binaural electrode pair. While such a setup allows to best control of the experimental conditions, it does not correspond to the real-life situation, where multiple binaural electrode pairs are stimulated at the same time. Studies currently underway investigate the ITD sensitivity for multiple-electrode stimulation (Egger et al., 2014; Jones et al., 2013; Kan et al., 2013), indicating the direction of future research on ITD sensitivity in electric hearing.

4. Outlook

In NH listeners, spatial perception seems to dominate current acoustic and audio research. Further insights on sound localization, externalization, and more general aspects of sound quality will be required in order to better understand spatial hearing in NH listeners. The inclusion of top-down processes like attention and context-driven active listening will be as important as detailed modeling of peripheral auditory processes.

In CI listeners, where speech perception in noise and sound lateralization is currently the main field of investigations, a good ITD sensitivity for a variety of conditions will be the future research challenge. To that end, a better understanding of processes involved in the peripheral auditory system as well as the development of better methods to provide binaural acoustic information via electric interfaces are required.

Addressing research questions in studies with both NH and CI listeners has the advantage of the comparison between the acoustic and electric hearing modes. This allows to obtain new insights and to better understand the auditory system as a whole. Naturally, more knowledge about spatial hearing will yield better applications in the field of acoustics and audio engineering in the future.

5. References

Akeroyd, M. A., Chambers, J., Bullock, D., Palmer, A. R., Summerfield, A. Q., Nelson, P. A., and Gatehouse, S. (**2007**). "The binaural performance of a cross-talk cancellation system with matched or mismatched setup and playback acoustics," *J Acoust Soc Am*, **121**, 1056–1069.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001**). "Elevation localization and head-related transfer function analysis at low frequencies," *J Acoust Soc Am*, **109**, 1110–1122.

Altinsoy, M. E. (**2012**). "The Quality of Auditory-Tactile Virtual Environments," *J Audio Eng Soc*, **60**, 38–46.

Andéol, G., Macpherson, E. A., and Sabin, A. T. (**2013**). "Sound localization in noise and sensitivity to spectral shape," *Hear Res*, **304**, 20–27.

Andreopoulou, A. and Roginska, A. (**2011**). "Towards the Creation of a Standardized HRTF Repository," in *131th Convention of the Audio Engineering Society (AES)*, Convention Paper 8571.

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.

Asano, F., Suzuki, Y., and Sone, T. (**1990**). "Role of spectral cues in median plane localization," *J Acoust Soc Am*, **88**, 159–168.

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

Baumgartner, R., Majdak, P., and Bernhard, L. (**2013**). "Assessment of sagittal-plane sound localization performance in spatial-audio applications," in *The Technology of Binaural Listening*, edited by J. Blauert, (Springer, Berlin, Heidelberg), pp. 93–119.

Baumgartner, R., Majdak, P., and Laback, B. (**2014**). "Modeling sound-source localization in sagittal planes for human listeners," *J Acoust Soc Am*, 791–802.

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. (**2014**). "Sensitivity to envelope-based interaural delays at high frequencies: Center frequency affects the envelope rate-limitation," *J Acoust Soc Am*, **135**, 808–816.

Best, V., Carlile, S., Jin, C., and van Schaik, A. (**2005**). "The role of high frequencies in speech localization," *J Acoust Soc Am*, **118**, 353–363.

Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (**2007**). "Binaural interference and auditory grouping," *J Acoust Soc Am*, **121**, 1070–1076.

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

Blauert, J. (**1974**). *Räumliches Hören (Spatial hearing)*, (S. Hirzel, Stuttgart), pp. 256.

Blommer, M. and Wakefield, G. (**1997**). "Pole-zero approximations for head-related transfer functions using a logarithmic error criterion," *IEEE Trans Speech Audio Proc*, **5**, 278–287.

Bronkhorst, A. W. (**1995**). "Localization of real and virtual sound sources," *J Acoust Soc Am*, **98**, 2542–2553.

Brown, C. P. and Duda, R. O. (**1998**). "A structual model for binaural sound synthesis," *IEEE Trans Speech Audio Proc*, **6**, 476–488.

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

Brungart, D. S. and Rabinowitz, W. M. (**1999**). "Auditory localization of nearby sources Head-related transfer functions," *J Acoust Soc Am*, **106**, 1465–1479.

Carlile, S. and Pralong, D. (**1994**). "The location-dependent nature of perceptually salient features of the human head-related transfer functions," *J Acoust Soc Am*, **95**, 3445–3459.

Ciskowski, R. D. and Brebbia, C. A. (**1991**). *Boundary Element Methods in Acoustics*, (Springer, New York).

Defrance, G., Daudet, L., and Polack, J.-D. (**2008**). “Finding the onset of a room impulse response: straightforward?,” *J Acoust Soc Am*, **124**, EL248–EL254.

Dietrich, P., Masiero, B., Pollow, M., Krechel, B., and Vorländer, M. (**2012**). “Time efficient measurement method for individual HRTFs,” *Fortschritte Akust Tagungsband 38 DAGA*, 333–334.

Dietrich, P., Masiero, B., and Vorländer, M. (**2013**). “On the Optimization of the Multiple Exponential Sweep Method,” *J Audio Eng Soc*, **61**, 113–124.

Egger, K., Laback, B., and Majdak, P. (**2014**). “Across-Electrode Integration of Interaural Time Difference in Bilateral Cochlear Implant Listeners,” presented at the 37th MidWinter Meeting of the Association for Research in Otolaryngology (ARO),.

Enzner, G., Antweiler, C., and Spors, S. (**2013**). “Trends in Acquisition of Individual Head-Related Transfer Functions,” in *The Technology of Binaural Listening*, edited by J. Blauert, (Springer, Berlin, Heidelberg), pp. 57–92.

Evans, M. J., Angus, J. A. S., and Tew, A. I. (**1998**). “Analyzing head-related transfer function measurements using surface spherical harmonics,” *J Acoust Soc Am*, **104**, 2400–2411.

Francart, T., Lenssen, A., and Wouters, J. (**2014**). “Modulation enhancement in the electrical signal improves perception of interaural time differences with bimodal stimulation,” *J Assoc Res Otolaryngol JARO*, **15**, 633–647.

Freeland, F. P., Biscainho, L. W. P., and Diniz, P. S. R. (**2004**). “Interpositional Transfer Function for 3D-Sound Generation,” *J Audio Eng Soc*, **52**, 915–930.

Gardner, W. G. (**1995**). “Efficient convolution without input-output delay,” *J Audio Eng Soc*, **43**, 127–136.

Gilkey, R., H. and Anderson, T., R. (**1997**). *Binaural and Spatial Hearing in Real and Virtual Environments*, (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 840.

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

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.-D. (**2008**). “Current-level discrimination and spectral profile analysis in multi-channel electrical stimulation,” *J Acoust Soc Am*, **124**, 3142–3157.

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

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

Grant, S. C. and Magee, L. E. (**1998**). “Contributions of proprioception to navigation in virtual environments,” *Hum Factors*, **40**, 489–497.

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

Greengard, L. and Rokhlin, V. (**1987**). “A fast algorithm for particle simulations,” *J Comput Phys*, **73**, 325–348.

Gumerov, N. A., O’Donovan, A. E., Duraiswami, R., and Zotkin, D. N. (**2010**). “Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation,” *J Acoust Soc Am*, **127**, 370–386.

Haber, L., Haber, R. N., Penningroth, S., Novak, K., and Radgowski, H. (**1993**). “Comparison of nine methods of indicating the direction to objects: data from blind adults,” *Perception*, **22**, 35–47.

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.

Haneda, Y., Makino, S., Kaneda, Y., and Kitawaki, N. (**1999**). “Common-acoustical-pole and zero modeling of head-related transfer functions,” *IEEE Trans Speech Audio Proc*, **7**, 188–196.

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

Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schösser, H., Garnham, C., and Anderson, I. (**2006**). “MED-EL Cochlear implants: state of the art and a glimpse into the future,” *Trends Amplif*, **10**, 201–219.

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. (**2012**). “Bilateral Cochlear Implants,” in *Auditory Prostheses*, *Springer Handbook of Auditory Research*, edited by F.-G. Zeng, A. N. Popper, and R. R. Fay, (Springer New York), pp. 13–57.

Hofman, M. and Van Opstal, J. (**2003**). “Binaural weighting of pinna cues in human sound localization,” *Exp Brain Res*, **148**, 458–470.

Hofman, P. M. and Van Opstal, A. J. (**1998**). “Spectro-temporal factors in two-dimensional human sound localization,” *J Acoust Soc Am*, **103**, 2634–2648.

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

Hwang, S. and Park, Y. (**2008**). “Interpretations on pricipal components analysis of head-related impulse responses in the median plane,” *J Acoust Soc Am*, **123**, EL65–EL71.

Iida, K., Ishii, Y., and Nishioka, S. (**2014**). “Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener’s pinnae,” *J Acoust Soc Am*, **136**, 317–333.

Iida, K., Motokuni, I., Itagaki, A., and Morimoto, M. (**2007**). “Median plane localization using a parametric model of the head-related transfer function based on spectral cues,” *Appl Acoust*, **68**, 835–850.

Jones, H. G., Kan, A., and Litovsky, R. Y. (**2013**). “Binaural sensitivity of bilateral cochlear implanted patients to amplitude modulated stimulation presented on multiple electrodes,” presented at the Conference on Implantable Auditory Prostheses (CIAP),.

Kahana, Y. and Nelson, P. A. (**2007**). “Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models,” *J Sound Vib*, **300**, 552–579.

Kan, A., Jones, H., and Litovsky, R. Y. (**2013**). “Issues in binaural hearing in bilateral cochlear implant users,” in *Proceedings of Meetings on Acoustics*, **19**, 050049.

Kasess, C. H., Noll, A., Majdak, P., and Waubke, H. (**2013**). “Effect of train type on annoyance and acoustic features of the rolling noise,” *J Acoust Soc Am*, **134**, 1071–1081.

Katz, B. F. (**2001**). “Boundary element method calculation of individual head-related transfer function I Rigid model calculation,” *J Acoust Soc Am*, **110**, 2440–2448.

Katz, B. F. G. and Noisternig, M. (**2014**). “A comparative study of Interaural Time Delay estimation methods,” *J Acoust Soc Am*, **135**, 3530–3540.

Katz, B. F. G. and Parseihian, G. (**2012**). “Perceptually based head-related transfer function database optimization,” *J Acoust Soc Am*, **131**, EL99–EL105.

Kirkeby, O., Nelson, P. A., and Hamada, H. (**1998**). “Local sound field reproduction using two closely spaced loudspeakers,” *J Acoust Soc Am*, **104**, 1973–1981.

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.

Klumpp, R. G. and Eady, H. R. (**1956**). “Some measurements of interaural time difference thresholds,” *J Acoust Soc Am*, **28**, 859–860.

Kopčo, N., Huang, S., Belliveau, J. W., Raij, T., Tengshe, C., and Ahveninen, J. (**2012**). “Neuronal representations of distance in human auditory cortex,” *Proc Natl Acad Sci U S A*, **109**, 11019–11024.

Kreuzer, W., Majdak, P., and Chen, Z. (**2009**). “Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range,” *J Acoust Soc Am*, **126**, 1280–1290.

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.

Kulkarni, A. and Colburn, H. S. (**2004**). “Infinite-impulse-response models of the head-related transfer function,” *J Acoust Soc Am*, **115**, 1714–1728.

Kulkarni, A., Isabelle, S. K., and Colburn, H. S. (**1999**). “Sensitivity of human subjects to head-related transfer-function phase spectra,” *J Acoust Soc Am*, **105**, 2821–2840.

Laback, B., Egger, K., and Majdak, P. (**2014**). “Perception and Coding of Interaural Time Differences with Bilateral Cochlear Implants,” *Hear Res*, 10.1016/j.heares.2014.10.004.

Laback, B. and Majdak, P. (**2008a**). “Binaural jitter improves interaural-time difference sensitivity of cochlear implantees at high pulse rates,” *Proc Natl Acad Sci USA*, **105**, 814–817.

Laback, B. and Majdak, P. (**2008b**). “Binaural stimulation in neural auditory prostheses or hearing aids,” *EU Pat*, **Code # 08 759 294.5**, given on 17.9.2009.

Laback, B. and Majdak, P. (**2008c**). “Binaural stimulation in neural auditory prostheses or hearing aids,” *US Pat*, **Code # 20080319509**, Kind Code A1, granted on 25.12.2008.

Laback, B., Majdak, P., and Baumgartner, W.-D. (**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., Pok, S.-M., Baumgartner, W.-D., 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., Majdak, P., Baumgartner, W.-D., and Pok, S.-M. (**2011**). “Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing,” *J Acoust Soc Am*, **130**, 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.

Lewald, J. and Ehrenstein, W. H. (**1998**). “Auditory-visual spatial integration: A new psychophysical approach using laser pointing to acoustic targets,” *J Acoust Soc Am*, **104**, 1586–1597.

Litovsky, R. Y., Goupell, M. J., Godar, S., Grieco-Calub, T., Jones, G. L., Garadat, S. N., Agrawal, S., et al. (**2012**). “Studies on Bilateral Cochlear Implants at the University of Wisconsin’s Binaural Hearing and Speech Laboratory,” *J Am Acad Audiol*, **23**, 476–494.

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.

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.

Macpherson, E. A. and Middlebrooks, J. C. (**2003**). “Vertical-plane sound localization probed with ripple-spectrum noise,” *J Acoust Soc Am*, **114**, 430–445.

Majdak, P., Balazs, P., Kreuzer, W., and Dörfler, M. (**2011a**). “A time-frequency method for increasing the signal-to-noise ratio in system identification with exponential sweeps,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3812–3815.

Majdak, P., Balazs, P., and Laback, B. (**2007**). “Multiple exponential sweep method for fast measurement of head-related transfer functions,” *J Audio Eng Soc*, **55**, 623–637.

Majdak, P., Baumgartner, R., and Laback, B. (**2014a**). “Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization,” *Front Psychol*, **5**, 319(1–10).

Majdak, P., Carpentier, T., Nicol, R., Roginska, A., Suzuki, Y., Watanabe, K., Wierstorf, H., et al. (**2013a**). “Spatially Oriented Format for Acoustics: A Data Exchange Format Representing Head-Related Transfer Functions,” in *Proceedings of the 134th Convention of the Audio Engineering Society (AES)*, Convention Paper 8880.

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. (**2011b**). “Two-dimensional localization of virtual sound sources in cochlear-implant listeners,” *Ear Hear*, **32**, 198–208.

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., Laback, B., and Baumgartner, W.-D. (**2006**). “Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing,” *J Acoust Soc Am*, **120**, 2190–2201.

Majdak, P., Masiero, B., and Fels, J. (**2013b**). “Sound localization in individualized and non-individualized crosstalk cancellation systems,” *J Acoust Soc Am*, **133**, 2055–2068.

Majdak, P., Søndergaard, P., and Blauert, J. (**2014b**). “Die Auditory-Modeling Toolbox der AABBA-Initiative,” in *Fortschritte der Akustik – DAGA*, 584–585.

Majdak, P., Walder, T., and Laback, B. (**2013c**). “Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions,” *J Acoust Soc Am*, **134**, 2148–2159.

Marelli, D., Baumgartner, R., and Majdak, P. (**2014**). “Efficient Representation of Head-Related Transfer Functions in Subbands,” in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*, 1–5.

Marelli, D. and Fu, M. (**2007**). “A functional analysis approach to subband system approximation and identification,” *IEEE Trans Signal Process*, **55**, 493–506.

Marelli, D., Fu, M., Balazs, P., and Majdak, P. (**2008**). “An iterative method for approximating LTI systems using subbands,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3533–3536.

Matsumoto, M., Yamanaka, S., Tohyama, M., and Nomura, H. (**2004**). “Effect of arrival time correction on the accuracy of binaural impulse response interpolation: Interpolation methods of binaural response,” *J Audio Eng Soc*, **52**, 56–61.

McFadden, D. and Pasanen, E. G. (**1976**). “Lateralization of high frequencies based on interaural time differences,” *J Acoust Soc Am*, **59**, 634–639.

McGurk, H. and MacDonald, J. (**1976**). “Hearing lips and seeing voices,” *Nature*, **264**, 746–748.

Mehrgardt, S. and Mellert, V. (**1977**). “Transformation characteristics of the external human ear,” *J Acoust Soc Am*, **61**, 1567–1576.

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

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

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

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

Middlebrooks, J. C. and Green, D. M. (**1992**). “Observations on a principal components analysis of head-related transfer functions,” *J Acoust Soc Am*, **92**, 597–599.

Middlebrooks, J. C., Macpherson, E. A., and Onsan, Z. A. (**2000**). “Psychophysical customization of directional transfer functions for virtual sound localization,” *J Acoust Soc Am*, **108**, 3088–3091.

Minnaar, P., Plogsties, J., and Christensen, F. (**2005**). “Directional resolution of head-related transfer functions required in binaural synthesis,” *J Audio Eng Soc*, **53**, 919–929.

Mommertz, E. (**1995**). “Angle-dependent in-situ measurements of reflection coefficients using a subtraction technique,” *Appl Acoust*, **46**, 251–263.

Moore, B. C. J. (**2012**). *An introduction to the psychology of hearing*, (Brill Academic Publishers, Leiden, Netherlands), pp. 458.

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.

Müller, S. and Massarani, P. (**2001**). “Transfer-function measurement with sweeps,” *J Audio Eng Soc*, **49**, 443–471.

Noel, V. A. and Eddington, D. K. (**2007**). “Binaural localization ability can be predicted from basic measure of ILD and is robust to effects of AGC,” in *9th Conference on Implantable Auditory Prostheses*.

Noisternig, M. and Majdak, P. (**2014**). *Spatial acoustic data file format* (AES-X212), *Standard of the Audio Engineering Society.*

Nuetzel, J. M. and Hafter, E. R. (**1976**). “Lateralization of complex waveforms: effects of fine structure, amplitude, and duration,” *J Acoust Soc Am*, **60**, 1339–1346.

Parseihian, G. and Katz, B. F. G. (**2012**). “Rapid head-related transfer function adaptation using a virtual auditory environment,” *J Acoust Soc Am*, **131**, 2948–2957.

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.

Pelizzone, M., Cosendai, G., and Tinembart, J. (**1999**). “Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for ineraid implanted subjects,” *Ear Hear*, **20**, 228–237.

Proakis, J. G. and Manolakis, D. G. (**1996**). *Digital signal processing: principles, algorithms, and applications*, (Prentice Hall International, Englewood Cliffs, NJ).

Reichinger, A., Majdak, P., Sablatnig, R., and Maierhofer, S. (**2013**). “Evaluation of Methods for Optical 3-D Scanning of Human Pinnas,” in *Proceedings of the 3D Vision Conference*, 390–397.

Rew, R. K. and Davis, G. P. (**1990**). “NetCDF: An Interface for Scientific Data Access,” *IEEE Comput Graph Appl*, **10**, 76–82.

Savioja, L., Houpaniemi, J., Lokki, T., and Väänäen, R. (**1999**). “Creating interactive virtual acoustic environments,” *J Audio Eng Soc*, **47**, 675–705.

Seeber, B. (**2002**). “A new method for localization studies,” *Acta Acust United Acust*, **88**, 446–450.

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

Senn, P., Kompis, M., Vischer, M., and Haeusler, R. (**2005**). “Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors,” *Audiol Neurootol*, **10**, 342–352.

Senova, M., McAnally, K., and Marin, R. (**2002**). “Localization of virtual sound as a function of head-related impulse response duration,” *J Audio Eng Soc*, **50**, 57–66.

Shackleton, T. M., Skottun, B. C., Arnott, R. H., and Palmer, A. R. (**2003**). “Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of Guinea pigs,” *J Neurosci*, **23**, 716–724.

Shelton, B. R. and Searle, C. L. (**1980**). “The influence of vision on the absolute identification of sound-source position,” *Percept Psychophys*, **28**, 589–596.

Skottun, B. C., Shackleton, T. M., Arnott, R. H., and Palmer, A. R. (**2001**). “The ability of inferior colliculus neurons to signal differences in interaural delay,” *Proc Natl Acad Sci U A*, **98**, 14050–14054.

Smith, Z. (**2010**). “Improved sensitivity to interaural time differences with the FAST coding strategy,” presented at the 11th International Conference on Cochlear Implants and Other Implantable Auditory Technologies,.

Smith, Z. (**2013**). “Efficient Coding for Auditory Prostheses,” presented at the 36th Midwinter Meeting of the Association of the Research in Otolaryngology (ARO),.

Søndergaard, P. and Majdak, P. (**2013**). “The Auditory Modeling Toolbox,” in *The Technology of Binaural Listening*, edited by J. Blauert, (Springer, Berlin, Heidelberg), pp. 33–56.

So, R. H. Y., Ngan, B., Horner, A., Braasch, J., Blauert, J., and Leung, K. L. (**2010**). “Toward orthogonal non-individualised head-related transfer functions for Forward and backward directional sound: cluster analysis and an experimental study,” *Ergonomics*, **53**, 767–781.

Spagnol, S., Geronazzo, M., and Avanzini, F. (**2013a**). “On the relation between Pinna Reflection Patterns and Head-Related Transfer Function Features,” *IEEE Trans Audio Speech Lang Process*, **21**, 508–519.

Spagnol, S., Geronazzo, M., Rocchesso, D., and Avanzini, F. (**2013b**). “Extraction of Pinna Features for Customized Binaural Audio Delivery on Mobile Devices,” in *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, 514.

Stecker, G. C. and Gallun, F. J. (**2012**). “Binaural Hearing, Sound Localization, and Spatial Hearing,” in *Translational Perspectives in Auditory Neuroscience: Normal Aspects of Hearing*, edited by K. L. Tremblay and R. F. Burkard, (Plural Publishing, Inc., San Diego), pp. 383–433.

Strutt, J. W. (**1876**). “Our perception of the direction of a source of sound,” *Nature*, **14**, 32–33.

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

Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K. (**2012**). “Mechanism for generating peaks and notches of head-related transfer functions in the median plane,” *J Acoust Soc Am*, **132**, 3832–3841.

Takeuchi, T. and Nelson, P. A. (**2002**). “Optimal source distribution for binaural synthesis over loudspeakers,” *J Acoust Soc Am*, **112**, 2786–2797.

Vorländer, M. (**2007**). *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and A oustic Virtual Reality*, (Springer, RWTH Edition), pp. 335.

Wang, L., Yin, F., and Chen, Z. (**2009**). “A hybrid compression method or head-related transfer functions,” *Appl Acoust*, **70**, 1212–1218.

Wefers, F. and Vorländer, M. (**2012**). “Optimal Filter Partitions for Non-Uniformly Partitioned Convolution,” in *Audio Engineering Society Conference: 45th International Conference: Applications of Time-Frequency Processing in Audio*.

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

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.

Xie, B. (**2013**). *Head-related transfer function and virtual auditory display*, (J. Ross Publishing, Plantatation, FL), pp. 482.

Yin, T. C., Kuwada, S., and Sujaku, Y. (**1984**). “Interaural time sensitivity of high-frequency neurons in the inferior colliculus,” *J Acoust Soc Am*, **76**, 1401–1410.

Yost, W. A. and Gourevitch, G. (**1987**). *Directional Hearing*, (Springer-Verlag, New York).

Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., and Tam, C. (**2006**). "Perceptual recalibration in human sound localization: learning to remediate front-back reversals," *J Acoust Soc Am*, **120**, 343–359.

Zahorik, P., Brungart, D. S., and Bronkhorst, A. W. (**2005**). "Auditory Distance Perception in Humans: A Summary of Past and Present Research," *Acta Acust United Acust*, **91**, 409–420.

Zakarauskas, P. and Cynader, M. S. (**1993**). "A computational theory of spectral cue localization," *J Acoust Soc Am*, **94**, 1323–1331.

Zhang, P. X. and Hartmann, W. M. (**2010**). "On the ability of human listeners to distinguish between front and back," *Hear Res*, **260**, 30–46.

Ziegelwanger, H. and Majdak, P. (**2014**). "Modeling the direction-continuous time-of-arrival in head-related transfer functions," *J Acoust Soc Am*, **135**, 1278–1293.

Ziegelwanger, H., Majdak, P., and Kreuzer, W. (**2014**). "Non-uniform sampling of geometry for the numeric simulation of head-related transfer functions," in *Proceedings of the 21st International Congress on Sound and Vibration*, Proceedings Number 209.

Ziegelwanger, H., Majdak, P., and Laback, B. (**2013a**). "Modeling the broadband time-of-arrival of the head-related transfer functions for binaural audio," in *Proceedings of the 134th Convention of the Audio Engineering Society (AES)*, Convention Paper 8884.

Ziegelwanger, H., Reichinger, A., and Majdak, P. (**2013b**). "Calculation of listener-specific head-related transfer functions: Effect of mesh quality," in *Proceedings of Meetings on Acoustics*, **19**, 050017.