

**Temporal and spatial aspects
of hearing as revealed by
auditory evoked potentials**

Von der Fakultät für Mathematik und Naturwissenschaften
der Carl von Ossietzky Universität Oldenburg
zur Erlangung des Grades und Titels eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation

Dirk Junius

geb. am 18. Juli 1973

in Bremen

Gutachter: Prof. Dr. Dr. Birger Kollmeier

Zweitgutachter: Prof. Dr. Volker Mellert

Tag der Disputation: 14.07.2005

Contents

| | | |
|----------|--|----------|
| 1 | General introduction | 1 |
| 2 | Influence of cochlear processing on ABR | 5 |
| 2.1 | Introduction | 7 |
| 2.2 | Method | 10 |
| 2.2.1 | Subjects | 10 |
| 2.2.2 | Apparatus | 10 |
| 2.2.3 | Stimuli | 11 |
| 2.2.4 | Procedure | 13 |
| 2.2.5 | Statistical Analysis | 13 |
| 2.3 | Results | 14 |
| 2.3.1 | Experiment 1: The chirp embedded in two tones | 14 |
| 2.3.2 | Experiment 2: Responses to rising-falling chirp trains | 17 |
| 2.3.3 | Experiment 3: Effect of within-train rate on wave-V amplitude and latency | 21 |
| 2.4 | Discussion | 27 |
| 2.4.1 | Influence of the cochlear traveling wave on brainstem responses | 27 |
| 2.4.2 | Effect of within-train repetition rate on wave V | 28 |
| 2.4.3 | Summary and conclusions | 31 |

| | | |
|----------|--|-----------|
| 3 | An externalization detector at early stages of the auditory system? | 33 |
| 3.1 | Introduction | 35 |
| 3.2 | Methods | 38 |
| 3.2.1 | Subjects | 38 |
| 3.2.2 | Stimuli | 38 |
| 3.2.3 | EEG-Recording and procedure | 40 |
| 3.2.4 | Data post-processing | 42 |
| 3.2.5 | Analysis of amplitudes and latencies | 43 |
| 3.2.6 | Dipole source analysis | 43 |
| 3.3 | Results | 45 |
| 3.3.1 | Waveform morphology | 45 |
| 3.3.2 | Analysis of amplitudes and latencies | 46 |
| 3.3.3 | Dipole source analysis | 50 |
| 3.4 | Discussion | 56 |
| 4 | Influence of spatial position on chirp-evoked LAEPs and MMN | 61 |
| 4.1 | Introduction | 63 |
| 4.2 | Methods | 65 |
| 4.2.1 | Subjects | 65 |
| 4.2.2 | Stimuli | 66 |
| 4.2.3 | EEG-Recordings and procedure | 68 |
| 4.2.4 | Data post-processing | 68 |
| 4.2.5 | Analysis of amplitudes and latencies | 69 |
| 4.2.6 | Dipole source analysis | 70 |
| 4.3 | Results | 72 |
| 4.3.1 | Waveform morphology | 72 |

| | | |
|----------|--|-----------|
| 4.3.2 | Analysis of amplitudes and latencies | 73 |
| 4.3.3 | Dipole source analysis | 76 |
| 4.4 | Discussion | 78 |
| 5 | Summary and perspectives | 85 |
| | Bibliography | 89 |

Zusammenfassung

Die vorliegende Dissertation untersuchte die neuronalen Mechanismen bei akustischer Darbietung realistischer räumlicher Stimuli. Dazu wurden an Versuchspersonen über Elektroden auf der Kopfoberfläche nicht-invasiv die akustisch evozierten Potenziale (AEP) abgeleitet. Im Vergleich zu anderen Methoden liefern AEP eine hohe zeitliche Auflösung und sind gleichzeitig in der Lage, auch tiefliegende neuronale Aktivität (z.B. des Hirnstamms) zu registrieren. Zur Verbesserung des Signal/Rausch-Verhältnisses wurde ein optimierter Chirp-Stimulus zur Auslösung der AEP verwendet. Die Hypothese eines neuronalen Externalisations- bzw. Lokalisationsdetektors, welcher erhöhte Antworten auf Stimuli mit realistischen Kombinationen von interauralen und spektralen Merkmalen liefert, wurde überprüft.

Zunächst wurden in Kapitel 2 einige grundlegende monaurale Eigenschaften des Chirps untersucht. Es zeigte sich, dass Antworten auf langdauernde Stimuli, die den Chirp enthalten, für niedrige und mittlere Pegel äquivalent zu den Antworten auf Chirps bei Einzeldarbietung sind. Bei diesen Pegeln hängt die frequenzübergreifende Integration der neuronalen Aktivität nicht vom akustischen Kontext ab und verhält sich linear und zeitinvariant. Erst bei Pegeln oberhalb von etwa 50 dB HL zeigten sich Unterschiede zwischen den beiden Stimulusparadigmen, welche wahrscheinlich durch neuronale Sättigung hervorgerufen wurden. Zur Untersuchung der zeitlichen Verarbeitung wurden die Amplituden und Latenzen auf Chirp- und Klickfolgen in Abhängigkeit von der Wiederholfrequenz analysiert. Während die Latenz der Welle V mit der zeitlichen Position im Stimulus für alle Wiederholfrequenzen anstieg, wurde die Amplitude der Welle V nur von der höchsten Wiederholrate (250 Hz) beeinflusst. Damit stellt die Latenz der Welle V wahrscheinlich einen sensitiveren Indikator für neuronale Synchronizität dar als ihre Amplitude. Die Ergebnisse von Kapitel 2 verdeutlichen die Bedeutung der kochleären Verarbeitung für die Generierung der Hirnstammpotenziale.

In Kapitel 3 und 4 wurde der Einfluss der räumlichen Position der Schallquelle auf chirp-evozierte Mehrkanal-AEP untersucht. Die Schalldarbietung erfolgte mittels virtueller Akustik, basierend auf individuellen Außenohrimpulsantworten. Diese Art der Schalldarbietung unterschied sich von den meisten AEP-Studien zu Schalllokalisation in der Literatur, bei denen nur interaurale Zeit- und Pegeldifferenzen untersucht wurden, die lediglich zu einer Im-Kopf-Lokalisation (Lateralisation) führen. Im Gegensatz dazu erzeugte die hier verwendete virtuelle Akustik eine realistische, externalisierte Wahrnehmung der akustischen Stimuli sowohl bei Variation des Azimuts als auch der Elevation. Da das auditorische System die interauralen und spektralen Merkmale irgendwo im neuronalen System zusammenfassen muss, um ein lokalisiertes Schallobjekt zu bilden, wurde der Frage nachgegangen, ob sich verschiedene Schallelevationen in den Potenzialantworten unterscheiden lassen bzw. ob externalisierte Stimuli generell höhere Antworten auslösen als nicht-externalisierte. In Kapitel 3 wurden frühe und mittellatente AEP analysiert, deren Quellen im Hirnstamm bzw. im primären auditorischen Kortex liegen. Die gemessenen Antworten auf externalisierte Stimuli waren nicht höher als bei diotischer Stimulation. Bei den frühen AEP zeigte sich eine deutliche Abhängigkeit der binauralen Potenziale und binauralen Differenzpotenziale von der seitlichen Auslenkung des Stimulus. Dieser Effekt trat auch in der Zeitentwicklung eines modellierten rotierenden Hirnstamm-dipols auf. Die Elevation der Schallquelle hatte keinen Einfluss auf die frühen AEP. Die Abhängigkeit vom Azimut der Schallquelle war bei den mittellatenten Antworten nur schwach in den Latenzen der Differenzpotenziale und stärker in den Momenten zweier fixierter Dipole sichtbar.

Mit Kapitel 4 wurden die Messungen des vorigen Kapitels hin zu höheren Stationen der Hörbahn erweitert. Dazu wurden späte AEP mit externalisierten räumlichen Stimuli ausgelöst. Im Gegensatz zum vorherigen Experiment wurden die Stimulusbedingungen hier entsprechend dem sogenannten „oddball“ Paradigma angeboten, d.h. ein häufiger Standardstimulus (von vorne) wurde von seltenen, zufällig gesetzten Devianten (übrige Schalleinfallrichtungen) ersetzt. Die Potenzialantworten zeigten auch hier eine Abhängigkeit vom Grad der seitlichen Auslenkung der Stimuli. Änderungen der Elevation beeinflussten die Antworten nicht. Die sogenannte „Mismatch Negativity“ (MMN), berechnet aus der Differenz zwischen den Antworten auf die Devianten und der Antwort auf den Standard, zeigte ebenfalls eine Abhängigkeit vom Azimut der Schallquelle, aber nicht von der Elevation.

Die Ergebnisse aus Kapitel 3 und 4 deuten darauf hin, dass der Schallazimut zumindest bis hin zu den Generatoren der Welle P2 und der MMN in der Hörbahn kodiert wird. Da die Schallelevation nur einen sehr schwachen und nicht-systematischen Einfluss auf die Potenziale hat und die Antworten auf externalisierte Stimuli nicht höher ausfallen als die auf einen diotischen (nicht-externalisierten) Stimulus, muss die Hypothese eines spezifischen Externalisations- bzw. Lokalisationsdetektors verworfen werden.

Die Lokalisationsleistung entlang der Midsagittalebene könnte in Kapitel 3 und 4 aufgrund der kurzen Chirpdauer nicht optimal gewesen sein. Längere Stimuli (z.B. Breitbandrauschen oder Chirp-Sequenzen) könnten in zukünftigen Studien die Lokalisationsleistung für die Elevation verbessern und dadurch möglicherweise zu einem systematischen Einfluss der Schallelevation auf die späten AEP führen.

In dieser Arbeit wurden die berechneten Dipolorte relativ zum Kugelschalenmodell angegeben. Individuelle anatomische Kopfbilder, wie sie z.B. die Kernspintomographie liefert, würden zu realistischeren Kopfmodellen, absoluten Positionen der Dipolquellen und einer höherer Genauigkeit der Dipolquellenanalyse führen. Da AEP im Vergleich zur Kernspintomographie zwar eine höhere Zeitauflösung, aber eine wesentlich schlechtere Ortsauflösung ermöglichen, könnte ein vergleichbares Experiment mit der Kernspintomographie die schallrichtungsabhängige Aktivierung des Auditorischen Kortex näher beleuchten.

Zusätzliche Experimente könnten die Rolle der Aufmerksamkeit auf die späten AEP untersuchen. Bei der Aufnahme später AEP schauen die Versuchspersonen üblicherweise Stummfilme oder lesen Bücher während sie „passiv“ eine Sequenz von akustischen Stimuli hören, d.h. die Aufmerksamkeit ist nicht gerichtet. Ein psychoakustisches „alternative forced-choice“ Verfahren, bei dem die Versuchsperson aktiv eine Lokalisationsaufgabe während der AEP-Messung absolvieren muss, würde die Aufmerksamkeit auf die Stimuli richten und interessante Vergleiche zu psychoakustischen Studien und konventionellen AEP-Studien ermöglichen. Die notwendigen Erweiterungen der Stimulus-Präsentationssoftware werden momentan implementiert.

Zusammenfassend liefern diese Doktorarbeit einen wichtigen Beitrag zum Verständnis der Mechanismen, die der Generierung monauraler und binauraler AEP unterliegen und erlauben wertvolle Folgerungen für die zukünftige Forschung. Die Tatsache, dass kein Externalisations- bzw. Lokalisationsdetektor nachgewiesen werden konnte, deutet darauf hin, dass die Integration interauraler und spektraler Merkmale auf Stationen oberhalb der Generatoren der Welle P2 und der MMN stattfindet.

Chapter 1

General introduction

One of the most elaborated features of the auditory system is its ability to accurately localize sound sources in three-dimensional space. The most reliable localization cues depend on our ability to compare signals between both ears, such as differences in time (interaural time differences, ITD) and level (interaural level differences, ILD). This comparison allows us to localize sounds in the horizontal plane with exceptional high resolution: For sinusoidal stimuli, the smallest detectable change in angular position can be as low as 1° , corresponding to an ITD of about $10 \mu\text{s}$ (Mills, 1958, 1972). As an additional cue for localization, the direction-dependent spectral filtering of the sound is analyzed. This spectral shaping is a consequence of reflections at the head, torso, and pinnae before the sound enters the ear canals and allows discriminating between sound directions with quasi-identical interaural time and level differences, i.e., along the midsagittal plane. Furthermore, the spectral cues are responsible for externalization, i.e., the perception of auditory objects outside the listener's head.

Numerous psychophysical and physiological studies have investigated the mechanisms underlying sound localization in the auditory system. Yet the exact neural processes responsible for spatial perception are still not fully understood. For non-invasive studies in humans, electroencephalography (EEG) is the only technique capable of registering neuroelectric activity stemming from deep sources of the brain with sufficient time resolution. For investigations of the auditory modality, acoustic stimuli which specifically activate neurons along the auditory pathway are presented. The acquired auditory evoked potential (AEP) is about one order of magnitude smaller than the spontaneous EEG activity and can be extracted by

an averaging and filtering process. The current dissertation uses this technique to investigate the basic mechanisms underlying spatial perception in humans.

AEPs play a key role when investigating the neural processes underlying the various aspects of hearing in humans. usually grouped by their latency, i.e., the time of occurrence after the onset of the stimulus. This grouping corresponds roughly to the site of generation (e.g., [Picton et al., 1974](#); [Scherg, 1991](#)). The auditory brain-stem response (ABR) generally consists of seven peaks within the first 10 ms after stimulus onset, evoked by the simultaneous discharge of a large number of nerve cells in the ABR (e.g., [Jewett et al., 1970](#); [Jewett and Williston, 1971](#)). The potentials up to a latency of about 50 ms are classified as the middle latency response (MLR). The most prominent waves N19 (a negative peak roughly 19 ms after stimulus onset) and P30 (a positive peak about 30 ms after stimulus onset) are generated bilaterally in the primary auditory cortices situated in the temporal plane([Scherg, 1991](#)). Longer-latency potentials are referred to as late auditory evoked potentials (LAEPs), reflecting higher order processes of auditory function. The earlier LAEPs (N1 and P2, at latencies of about 100 ms and 175 ms, respectively) are generated close to the MLR sources in the auditory cortex. As latency increases, numerous higher brain centers are also involved in the generation of the LAEP. The mismatch negativity (MMN), with latencies between 100 and 200 ms, is a pre-attentive component derived by subtracting the LAEP to a frequent ‘standard’ stimulus from that to a rarely presented ‘deviant’ stimulus. This type of stimulation is commonly known as *oddball paradigm*. The MMN is regarded as a powerful tool for studying pre-attentive auditory sensory memory (e.g., [King et al., 1995](#); [Näätänen and Alho, 1995](#); [Ritter et al., 1995](#)). In summary, the various groups of AEPs can help to focus on specific neural areas along the pathway of auditory information processing.

Most studies investigating spatial hearing by means of AEPs used only stylized spatial sounds with ITDs and ILDs. These stimuli produce only unrealistic spatial percepts inside the listener’s head. The aim of the present thesis is to investigate the neural representation of realistic spatial sounds in humans. Therefore, the direction dependent natural combinations of ITD, ILD, and spectral cues are incorporated in the stimuli by an individual virtual acoustics. Binaural AEPs and binaural difference potentials (BD) are analyzed in dependence on sound direction. The BD is derived by subtracting the summed monaural responses from the binaural response and is therefore regarded as the specific binaural representation of the stimulus. However,

the signal-to-noise ratio (SNR) of the BD is substantially smaller than that of the binaural potentials. Therefore an optimized chirp stimulus is used, which, in contrast to the commonly used click stimulus, accounts for cochlear processing and leads to an enhancement of the AEPs.

To gain fundamental knowledge about the temporal properties of chirp-evoked potentials, chapter 2 examines the effect of the chirp onset on the ABR. Differences between transient and steady-state stimulation are studied by comparing the responses to single chirps with responses to chirps embedded in longer-duration stimuli. A chirp is embedded in two sinusoids to study the frequency following response (FFR) to the low-frequency tone in relation to the low-frequency components of the chirp. Furthermore, various trains of chirps are used to investigate the effect of neural synchronization without interference with stimulus onset. Finally, the effect of within-train rate on click and chirp-evoked responses is examined. By increasing the within-train rate, the role of adaptation is studied by investigating the transition from the unadapted ABR (at the beginning of the train) to the adapted ABR (at the end of the train). The results of chapter 2 elucidate the role of cochlear processing for the formation of ABR.

The neural mechanisms responsible for spatial hearing are investigated in chapters 3 and 4. In contrast to most existing studies that only present lateralized sounds without spectral localization cues, a virtual acoustics paradigm is incorporated to obtain realistic, externalized sound objects. The dependence of evoked potentials and modeled dipolar sources on the virtual stimulus direction (both on the horizontal and midsagittal plane) is examined. Chapter 3 focuses on neural processes occurring at the level of the brainstem and primary auditory cortices. The responses to spatial stimuli are compared to a diotic reference condition to test if response amplitudes benefit from stimuli with ‘natural’ combinations of ITD, ILD, and spectral cues.

Chapter 4 extends the work of the preceding chapter to more central stages of the auditory system. Properties of the LAEP and MMN components to spatial stimuli are systematically analyzed. Furthermore, this chapter tests for the existence of hemispheric dominance, i.e., a general predominant activation of one of the auditory cortices.

Finally, chapter 5 summarizes the results of chapters 2 to 4 and gives an outlook to possible future research on basis of the results of this thesis.

Chapter 2

Influence of cochlear traveling wave and neural adaptation on auditory brainstem responses¹

Abstract

The present study investigates the relationship between evoked responses to transient broadband chirps and responses to the same chirps when embedded in longer-duration stimuli. It examines to what extent the responses to the composite stimuli can be explained by a linear superposition of the responses to the single components, as a function of stimulus level. In the first experiment, a single rising chirp was temporally and spectrally embedded in two steady-state tones. In the second experiment, the stimulus consisted of a continuous alternating train of chirps: each rising chirp was followed by the temporally reversed (falling) chirp. In both experiments, the transitions between stimulus components were continuous. For stimulation levels up to approximately 70 dB SPL, the responses to the embedded chirp corresponded to the responses to the single chirp. At high stimulus levels (80-100 dB SPL), disparities occurred between the responses, reflecting a nonlinearity in the processing when neural activity is integrated across frequency. In the third experiment, the effect of within-train rate on wave-V response was investigated. The response to the chirp presented at a within-train rate of 95 Hz exhibited the same

¹A modified version of this chapter was accepted for publication in *Hearing Research* ([Junius and Dau, 2005](#))

amplitude as that to the chirp presented in the traditional single-stimulus paradigm at a rate of 13 Hz. For a corresponding experiment with bandlimited chirps of 4 ms duration, where the within-train rate was 250 Hz, a clear reduction of the response amplitude was observed. This nonlinearity in terms of temporal processing most likely reflects effects of short-term adaptation. Overall, the results of the present study further demonstrate the importance of cochlear processing for the formation of brainstem potentials. The data may provide constraints on future models of peripheral processing in the human auditory system. The findings might also be useful for the development of effective stimulation paradigms in clinical applications.

2.1 Introduction

Evoked responses represent the summation of responses from many neurons, recorded from electrodes placed at the surface of the head (e.g., [Jewett et al., 1970](#)), remote from individual neurons. Auditory evoked potentials can be recorded from all levels of the auditory system. They are usually grouped by the time of occurrence after the onset of the stimulus, and this grouping corresponds roughly to the site of generation. The auditory brainstem response (ABR) is considered a synchronized potential, evoked by the simultaneous discharge of a large number of nerve cells in the auditory brainstem. It is generally assumed that the ABR is an electrophysiological event evoked by either the onset or the offset of an acoustic stimulus ([Hecox et al., 1976](#); [Kodera et al., 1977](#); [Debruyne and Forrez, 1982](#); [Gorga and Thornton, 1989](#); [Campen et al., 1997](#)). Because of its abrupt onset, an acoustic click is generally considered to be an ideal stimulus for eliciting an effective ABR. The click-evoked ABR waveform generally consists of seven peaks, all occurring within the first 10 ms after signal onset. Of the seven peaks, waves I, III, and V are sufficiently robust to be used clinically. The most robust peak, wave V, can be elicited at near-threshold levels.

However, in the cochlea, the response to a click is not entirely synchronous; the peak of the response occurs several milliseconds later in low-frequency channels than in high-frequency channels (e.g., [Békésy, 1960](#); [Kiang et al., 1965](#); [Kiang, 1975](#)). This staggered activation reflects the change of stiffness along the cochlear partition. As a consequence, the phase velocity of the traveling wave depends upon its spatial position along the basilar membrane. It takes more time for the low-frequency region to reach maximal displacement at the apical end of the cochlea. Along the cochlear partition, high and low frequencies are not excited simultaneously, resulting in dispersion of activity associated with the click.

Recent studies have shown that a chirp rising in frequency, which is tailored to activate the entire cochlea concurrently, evokes a larger wave-V amplitude than a traditional click presented at the same sensation level ([Dau et al., 2000](#); [Wegner and Dau, 2002](#); [Fobel and Dau, 2004](#)). The rising chirp theoretically produces simultaneous displacement maxima by cancelling traveling-time differences along the cochlear partition. The equations determining the temporal course of the chirp were derived on the basis of a cochlear model ([de Boer, 1980](#)), and were calculated to be

the inverse of the delay-line characteristic of the human cochlear partition. The use of a broadband rising chirp was shown to reflect activity also from low-frequency regions, whereas neural synchrony across the cochlear partition is decreased for the click in accordance with the reduction in traveling wave velocity in the apical region of the cochlea (Dau et al., 2000; Wegner and Dau, 2002; Rupp et al., 2002). The use of a temporally reversed (falling) chirp leads to a further decrease in synchrony as reflected in ABR responses that are smaller than those from a click, at low and moderate stimulus levels. The studies therefore demonstrated that an appropriate temporal organization, determined in part by basilar-membrane (BM) traveling-wave properties, can significantly increase the synchrony of neural discharges. The findings were compatible with earlier results from recordings of compound action potentials (Shore and Nuttall, 1985) reflecting activity at the level of the auditory nerve. Since the ABR components considered here reflect neural response from the brainstem, the effect of an optimized synchronization at the peripheral level can also be observed at the brainstem level, at least at the level of processing where wave V is generated.

However, at high stimulus levels, onset effects may have contributed to the recorded potential pattern, due to spreading excitation along the basilar membrane. Thus, the response to the low-frequency portion in the rising chirp may have been superimposed with activity from cochlear regions tuned to higher frequencies (Dau et al., 2000). In addition to the onset effects, at high sound pressure levels, the response to the lower frequency components of the chirp (about 200-500 Hz) may also contain activity associated with the frequency following response (FFR), which is likely to be generated by activity stemming from mid and high frequencies (Dau, 2003). The advantage of a higher wave-V response amplitude for the chirp stimulus relative to the click was observed only at low and moderate stimulation levels, where it can be assumed that *all* frequencies contribute to the evoked response. However, at high stimulation levels, with respect to wave-V amplitude, no advantage was found for the chirp over the click (Dau et al., 2000).

The present study investigated the effect of the stimulus onset on chirp responses. Specifically, the differences between transient and steady-state stimulation were studied by comparing the responses to chirps presented in the traditional “single-stimulus” paradigm, with responses to the same chirps when embedded in longer-duration stimuli. In the composite-stimulus sequences, the transitions between the

stimulus components were always continuous. In the first experiment, the stimulus was a low-frequency tone (320 Hz), followed by a chirp (320 to 8000 Hz), and then a high-frequency tone (8000 Hz). This stimulus configuration allowed the investigation of the relation between the FFR to the low-frequency tone and the low-frequency components in the chirp, as well as the “traditional” ABR to the chirp. In the second experiment, the stimulus was a continuous alternating train of broadband chirps. Each rising chirp was followed by its temporally reversed (i.e., falling) counterpart. In this way, the effect of neural synchronization for the rising versus the falling chirp could be studied without interference with stimulus onset responses, as is the case for transient stimulation. In the third experiment, the influence of within-train rate on click and chirp-evoked responses was examined. The within-train rate of the chirp was varied in order to record the transition from the unadapted ABR (in the beginning) to the adapted ABR (at the end of the train). Adaptation is the variation in response which occurs during a constant stimulus. In peripheral auditory signal processing, the change in neural activity is typically maximum at onset and then decays or adapts to a smaller sustained change in response (e.g., [Smith, 1977](#); [Westerman and Smith, 1984](#)). Adaptation of firing rates in auditory-nerve fibers was found to consist of at least two decaying components. Rapid adaptation occupies the first few milliseconds and is superimposed upon short-term adaptation, which has a time constant of about 40-60 ms (e.g., [Smith and Brachman, 1982](#); [Westerman and Smith, 1984, 1985](#)). Adaptation at higher levels of brainstem neural processing, like the cochlear nucleus and inferior colliculus, was found to be more variable and complex (e.g., [Abbas and Gorga, 1981](#); [Kaltenbach et al., 1993](#); [Walton et al., 1995](#)).

If the responses associated with individual chirps were similar to those of the embedded chirps (with respect to wave V), the system could be considered as behaving linearly with respect to the integration of neural activity across frequency and time. A disparity between single-chirp and embedded-chirp response magnitude would imply nonlinearities such as saturation and adaptation. Consequently, the results of the experiments should provide a better understanding of the role of cochlear processing in the formation of ABR and FFR.

2.2 Method

2.2.1 Subjects

Five normal-hearing subjects (one female and four males) with audiometric thresholds of 10 dB HL or better for frequencies between 500 Hz and 4000 Hz, and 20 dB HL or better at 8000 Hz participated in this study. No history of hearing problems was reported by any subject. All subjects were between 27 and 37 years of age, and either volunteered or were paid for the experiments. A subset of three subjects (all male) participated in the first two experiments. All five subjects participated in experiment 3.

2.2.2 Apparatus

The experiments were carried out with two PC-based computer systems that controlled stimulus presentation and recording of evoked potentials. A DSP-card (Ariel DSP32C) converted the digitally generated stimulus (44.1 kHz, 16 bit) to an analog waveform. The output of the DSP-card was connected to a digitally controlled audiometric amplifier, which presented the stimulus through an insert earphone (Etymotic Research ER-2) monaurally to the subject. To avoid electrical leakage, the

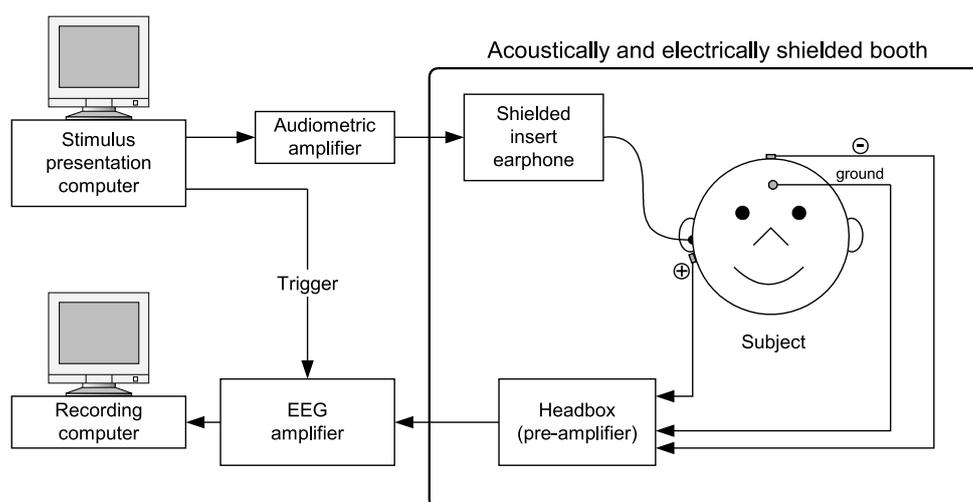


Figure 2.1: Scheme of the EEG recording setup used for the present study. The stimuli were presented in an electrically and acoustically shielded booth using shielded insert earphones. The electroencephalographic activity was recorded from surface electrodes.

insert earphone was encased in a copper box.

Electroencephalographic activity was recorded differentially between vertex (negative) and ipsilateral mastoid (positive), using silver/silver chloride electrodes. The ground electrode was placed on the forehead. Interelectrode impedance was maintained below 5 k Ω . Electrode signals were passed to an EEG amplifier (Synamps 5803) controlled by a PC. The EEG amplifier provided 74 dB of gain prior to anti-alias filtering (analog 2nd-order low-pass filter with 2 kHz cut-off frequency). A/D-conversion was made with a sampling rate of 10 kHz. The digitized data were stored on disk for subsequent analysis. After the recordings, the EEG-data were segmented into stimulus-related epochs and bandpass filtered using a finite-impulse-response (FIR) filter (Granzow et al., 2001) with 500 taps. Pass-band frequencies were set to 70 Hz and 1600 Hz. Epochs were averaged using an iterative weighted-averaging algorithm (Riedel et al., 2001). The standard error of the mean (SEM) was calculated for all waveforms. The EEG recording setup is shown schematically in Fig. 2.1.

2.2.3 Stimuli

The chirp stimuli used in the present study were generated on the basis of the equations described in Dau et al. (2000). The chirp started at low frequencies and swept towards higher frequencies with an instantaneous frequency increasing nonlinearly with time. Fig. 2.2 illustrates the stimuli schematically. The topmost row shows the stimulus used in the first experiment, referred to in the following as the “combination stimulus”. It consists of a 30-ms 320-Hz tone, followed by a 5.2-ms rising chirp (320-8000 Hz) and a 20-ms 8000-Hz tone. Thus, the chirp was temporally and spectrally embedded in the two steady-state tones. The chirp duration was determined by the choice of the two edge frequencies. The duration of the entire sequence was 55.2 ms. Four-ms long Hanning ramps were applied to the beginning and the end of the stimulus. The sequence was repeated at an average inter-stimulus rate of 10.3 Hz, and the period of silence between the sequences was randomized between 55 ms and 80 ms by a temporal jitter.

The second row shows the train of rising and falling broadband chirps used in the second experiment. In this case, the nominal edge frequencies of the chirp were 100 Hz and 10 kHz, the same as used in Dau et al. (2000). The corresponding duration of this chirp was 10.5 ms. Each train contained four consecutive rising-falling

chirps. Thus, the duration of the train was 84 ms. The silence interval between trains was randomized between 63 ms and 104 ms. The average inter-stimulus rate was 7.5 Hz.

The third row represents the broadband rising-chirp train used in the third experiment, with a within-train rate of 95.2 Hz (i.e., without any pause between the chirps in the train). The mean inter-stimulus rate was 6.7 Hz in this condition. The bottom row shows the train of bandlimited (0.5-10 kHz) chirps, also investigated in experiment 3, and presented at a mean inter-stimulus rate of 9.1 Hz. Here, the corresponding duration of the chirp was only 4.0 ms because of the absence of the frequencies below 0.5 kHz. For direct comparison, click recordings were obtained at the same two within-train rates (95.2 and 250 Hz, not indicated in the figure).

In experiments 1 and 2, the evoked responses were obtained for stimulation levels in

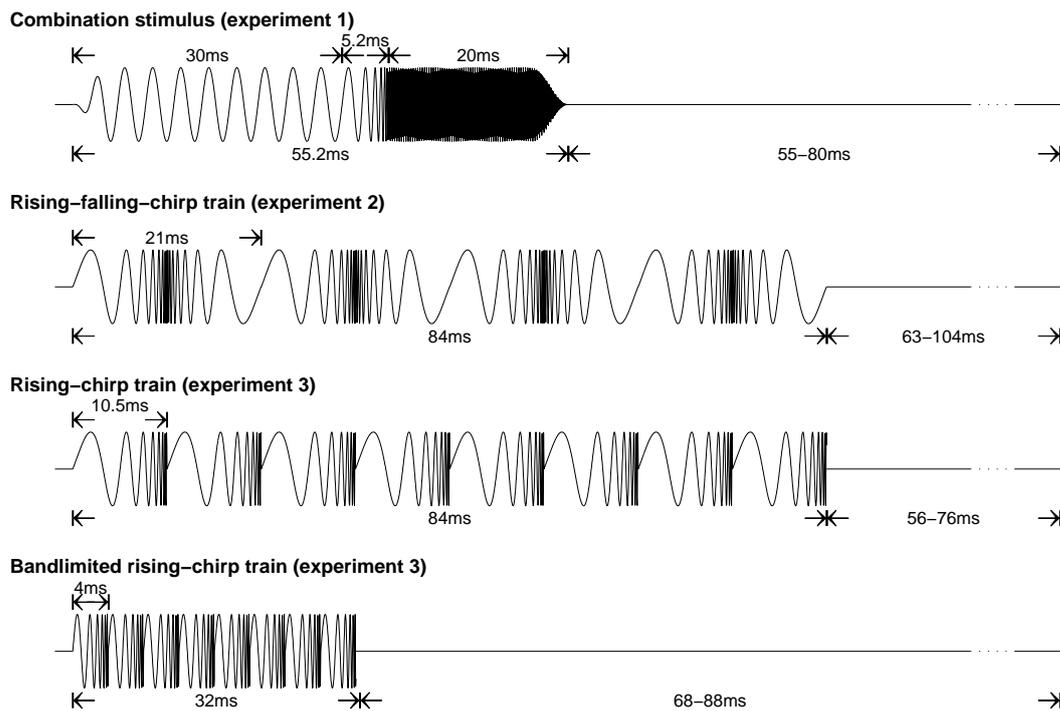


Figure 2.2: Stimuli used in the present study. Upper row: rising chirp (320-8000 Hz), embedded between steady-state sinusoids (experiment 1). Second row: Train of four broadband rising (100-10,000 Hz) - falling (10,000-100 Hz) chirps (experiment 2). Third and fourth row: train of broadband (100-10,000 Hz) and bandlimited (500-10,000 Hz) rising chirps, presented at within-train rates of 95.2 Hz and 250 Hz, respectively (experiment 3). The duration of the stimulus components, as well as the mean interval between sequences, are also indicated in the figure (see text for details).

the range between 50 and 100 dB sound pressure level (SPL)², in steps of 10 dB. In experiment 3, the responses were recorded only at 40 dB sensation level (SL) since the focus here was on effects of within-train rate. Constant SL was used instead of SPL, in order to permit a direct comparison between chirp and click data.

2.2.4 Procedure

For the EEG recordings, each subject lay on a couch in the recording chamber. The subjects were instructed to keep movement at a minimum, and to sleep if possible. The lights were turned off at the beginning of the session. Electrode impedances were checked at least every 20 minutes to ensure good quality of the recordings. Each trial consisted of 4000 stimulus presentations. The total recording time for the three experiments was about 6 h and was completed in two to three sessions on separate days. For the additional two subjects that only participated in experiment 3, the recording could be completed within a single session of about 1 h duration. The ear of stimulation was chosen randomly, i.e., for each subject one ear was chosen and then maintained.

In the first two experiments, the individual chirps and the composite stimuli were presented in interleaved fashion. The corresponding responses were obtained in the same recording session, which allowed a direct comparison of the responses obtained using the two stimulation paradigms.

To determine the sensation level for the various stimuli used in experiment 3, detection thresholds for the chirp trains were measured using an adaptive three alternative forced-choice (3AFC) procedure. The average level obtained over three repetitions for each stimulus train was considered as representing 0 dB SL for an individual subject.

2.2.5 Statistical Analysis

Wave-V peak-to-peak amplitude was analyzed for all stimulus conditions. The amplitude was measured from the peak to the largest negativity directly following it. For the experiments involving chirp and click trains (experiments 2 and 3), the wave-V response associated with each chirp and click was evaluated separately.

²throughout this chapter, SPL values represent *peak-equivalent* sound pressure levels

In addition, the average amplitude pooled across subjects was calculated for each instance of wave V. A Wilcoxon matched-pairs signed-rank test ($\alpha = 0.05$) was performed in experiment 3 to determine whether response amplitude and latency of the initial wave-V peak differed significantly from following peaks.

Throughout this chapter, responses are shown for a single representative subject in each experimental condition. Mean data for wave-V amplitude and latency, averaged across subjects, are summarized in additional figures.

2.3 Results

2.3.1 Experiment 1: The chirp embedded in two tones

Figure 2.3 shows for a single subject (HR) the responses evoked by the combination stimulus, comprising the rising chirp embedded between two tones. The portion of the stimulus containing the chirp is highlighted in gray. First, it can be seen that, at all stimulus levels, the data show a clear wave-V peak in response to the rising chirp. Neural activity in the auditory nerve is assumed to be synchronously activated at chirp *offset*. Wave-V latency relative to chirp offset is about 5.5-7.5 ms, depending on stimulus level. This corresponds to typical wave-V latency values obtained with click stimulation (e.g., Pratt and Sohmer, 1976; Jacobson et al., 1980). Second, at levels up to 70 dB SPL, no significant responses were evoked by the low-frequency tone at the beginning of the stimulus sequence. The response patterns change considerably at higher levels. For levels of 80 dB SPL and higher, the low-frequency tone generates a periodic response that corresponds to the “classical” FFR. The chirp-evoked wave-V response amplitude at these high sound pressure levels is somewhat smaller than at lower SPLs. The response becomes less distinctive and multiple peaks are apparent. Third, as expected, the high-frequency (8-kHz) tone does not produce any periodic response at any stimulation level since, at 8 kHz, the ability of the system to phase-lock to the fine structure is absent.

The right panel of Fig. 2.3 replots the data for the time interval between 25 ms and 55 ms. The wave V peak is indicated by a “V”. In addition, the responses obtained with single-chirp stimulation for the same subject are indicated (with an appropriate time-shift) by the gray curves. The responses in the embedded-chirp and the single-chirp condition are very similar for levels up to 70 dB SPL. At higher

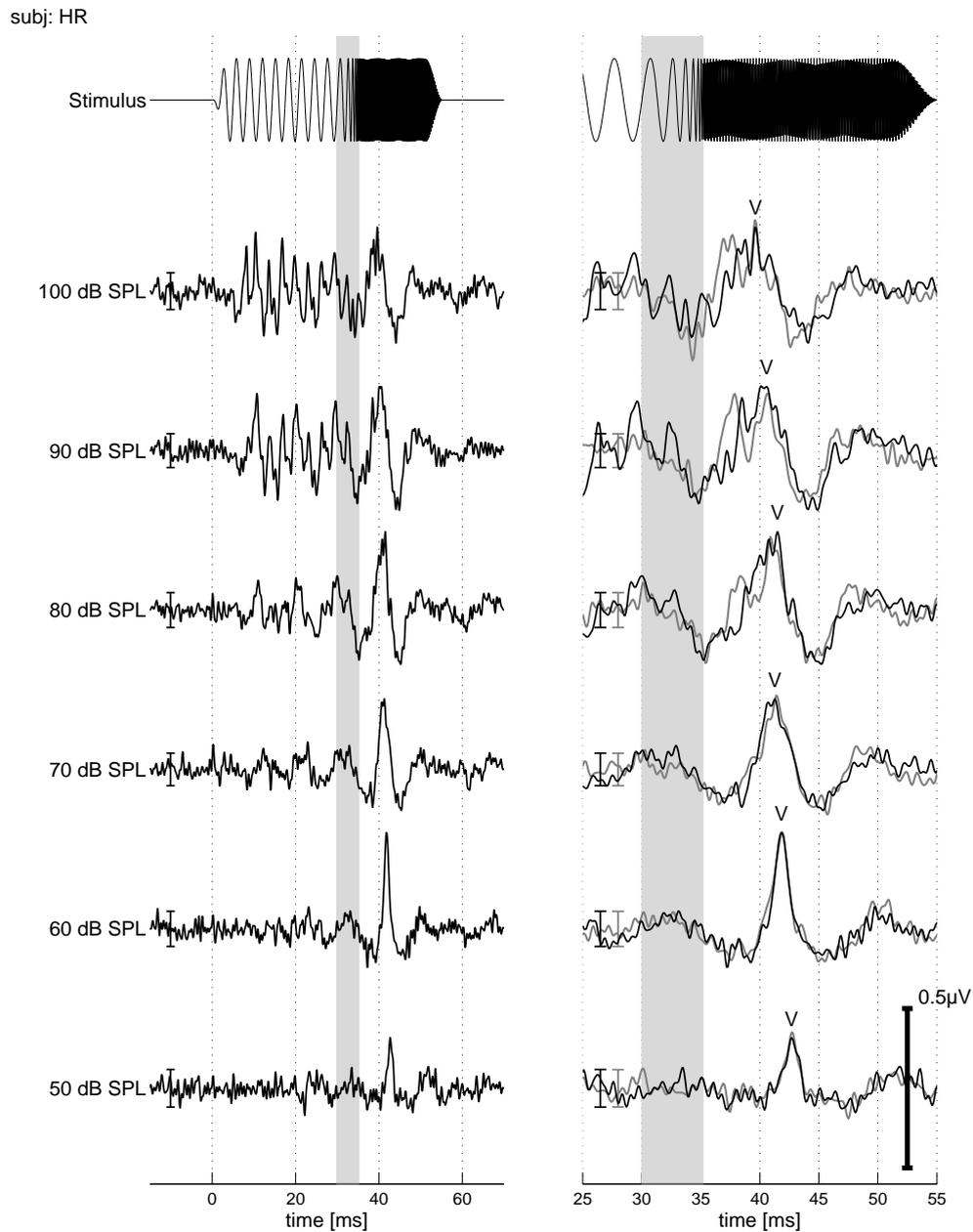


Figure 2.3: Evoked responses to the “combination stimulus” for subject HR. The upper row shows a single stimulus sequence. The time interval containing the chirp is highlighted in gray. The remaining rows represent the responses for stimulation levels from 50 to 100 dB SPL. The right panel shows a replot of the data in the left panel, using a different time scale for better visibility of some of the details. In addition, the responses to single-chirp stimulation are also shown in the right panel, indicated by the gray curves. The error bars in the figure indicate ± 3 SEM

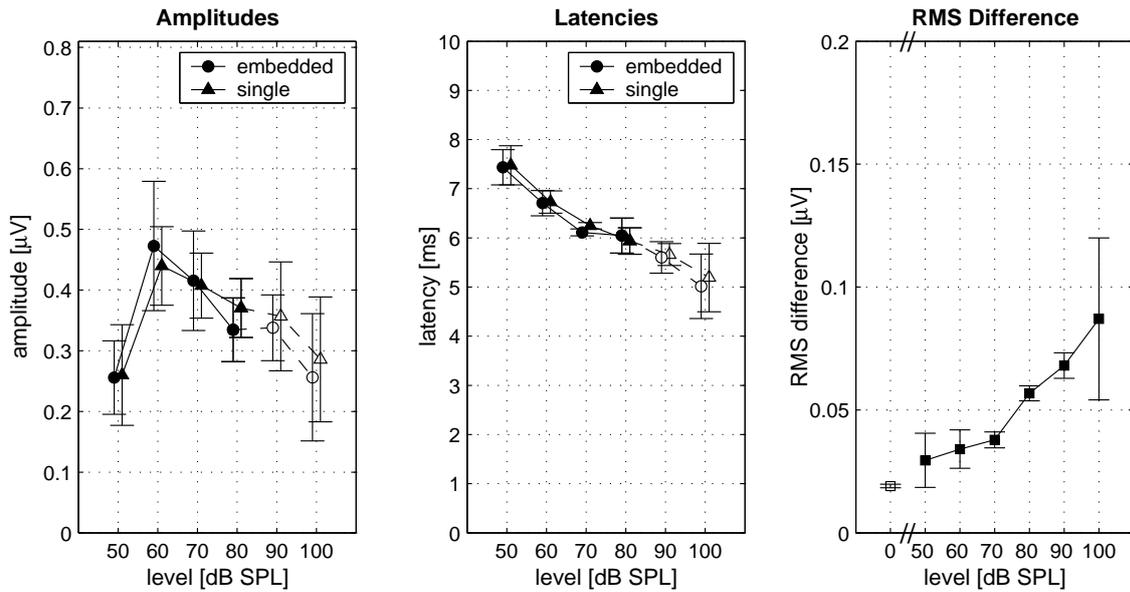


Figure 2.4: Average data obtained for the combination stimulus. Wave-V amplitude (left panel) and relative latency (middle panel) are shown as a function of stimulation level. The results for the embedded chirp and the single chirp are represented by circles and triangles, respectively. For better visibility, the results are plotted with an offset of 2 dB on the abscissa. The error bars denote the inter-individual standard deviations. For the two highest levels, the results are represented by open symbols and connected with dashed lines since the definition of wave V was less clear. The right panel shows the mean RMS of the difference waveform between embedded-chirp and single-chirp responses. For comparison, the RMS of the noise activity in the silence condition is represented by the open rectangle, at “0 dB”.

levels, differences between the responses occur. In particular, the responses to the single chirps show stronger earlier peaks (prior to wave V) than the corresponding responses to the embedded chirp. This earlier activity represents the response to the onset of the chirp at these high stimulus levels, where the early low-frequency energy in the chirp stimulates basal cochlear regions and produces a response (see also Fig. 2 in [Dau et al., 2000](#), for comparison).

Wave-V amplitudes and latencies were derived from the waveforms and averaged across subjects. The results are shown in Fig. 2.4. The latencies were defined relative to chirp offset. The amplitude (left panel) and latency (middle panel) values do not differ significantly between embedded-chirp (circles) and single-chirp (triangles) stimulation. Wave-V amplitude increases up to 60 dB SPL and decreases at higher SPLs. Wave-V latency decreases monotonically with level. The standard deviations for the amplitude data are relatively large (compared to the latency val-

ues), and reflect a high variability across subjects. However, the general shape of the amplitude-level function is very similar for all subjects. Because the response patterns are more complicated at high sound pressure levels (e.g., multiple peaks for most subjects), computation of wave-V amplitude and latency is less clear than at lower stimulus levels. The estimates at 90 and 100 dB SPL are therefore indicated as open symbols and connected by dashed lines.

The right panel of Fig. 2.4 shows the root mean square (RMS) of the waveform differences between embedded-chirp and single-chirp responses, averaged across subjects. The calculation was obtained for the time interval ± 10 ms relative to wave-V latency. For comparison, the RMS value of the recording noise differences (embedded-chirp versus single-chirp response) in the pre-stimulus interval (-10.0 ms to -0.1 ms), averaged across stimulus levels and subjects, is indicated by the open symbol. For stimulus levels of 50 dB SPL and 60 dB SPL, RMS differences are very small and do not differ significantly from the recording noise. At levels of 70 dB SPL and higher, RMS differences increase substantially with level and are significantly higher than the recording noise (i.e., differ by more than 2 standard deviations). The main reason for the differences at the high levels is the occurrence of the onset response to the single chirp that most likely results from effects of cochlear upward spread of excitation (see also discussion below).

2.3.2 Experiment 2: Responses to rising-falling chirp trains

Fig. 2.5 shows the rising-falling chirp train (upper trace) and the corresponding evoked responses (black curves) for subject DJ. For levels up to 80 dB SPL, the responses show four prominent peaks whereby the time interval between peaks corresponds to the period of the rising-falling chirp. As in the first experiment, the latency of the peaks (relative to chirp offset) corresponds to wave-V latency values typically found for traditional click stimulation. At each stimulus level, the amplitude of the peak remains roughly constant throughout the train. It increases with stimulation level up to about 70 dB SPL. At higher levels (80-100 dB SPL), the evoked pattern exhibits several peaks prior to wave V and may represent an FFR to the low-frequency portion of the rising chirp. The FFR-like pattern seems particularly pronounced for the first chirp in the train while the pattern in the responses to the following chirps show a somewhat reduced variation.

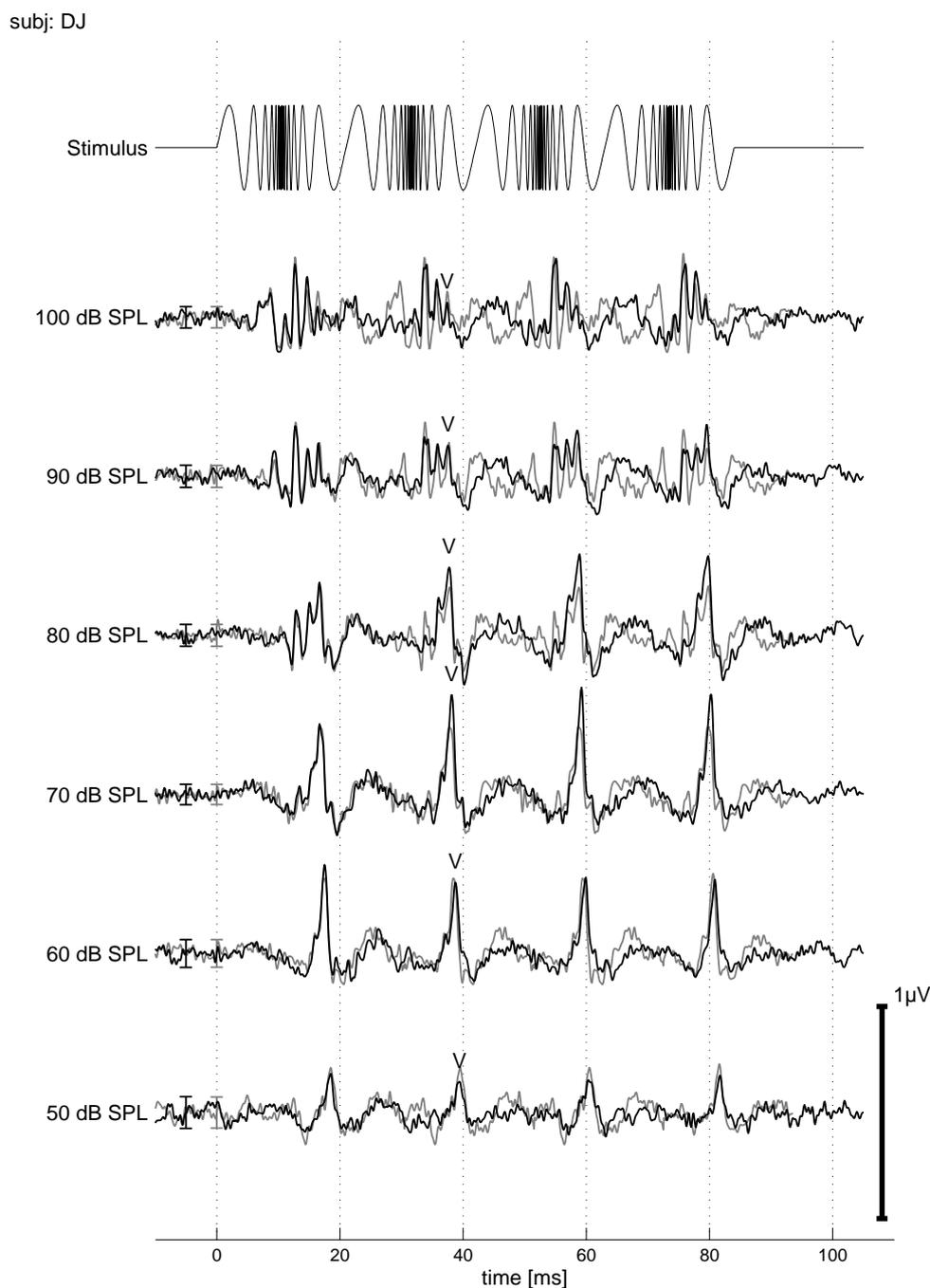


Figure 2.5: Evoked responses to the rising-falling chirp train for subject DJ. The upper row shows the stimulus. The remaining rows represent the responses for stimulation levels from 50 to 100 dB SPL. In addition, the responses to single chirps, convolved with a delta comb with the appropriate period, are indicated by the gray curves. The error bars in the figure indicate ± 3 SEM. The “V”-symbol indicates the peaks of the second wave V in the response.

The gray curves in Fig. 2.5 show the measured responses to the individual rising chirp, convolved with a delta comb whose period corresponds to that of the rising-falling chirp in the train (i.e., 21 ms). For levels up to 60 dB SPL, the single-chirp data closely match the data obtained with the train. At higher sound pressure levels (70-100 dB SPL), the single-chirp responses only closely correspond to the first chirp in the train. For the subsequent peaks, disparities between the responses to the single chirp and the chirp train become apparent: first, at sound pressure levels of 70 and 80 dB, the wave-V responses to the second to fourth chirp in the train are higher than the response to the first chirp, and thus higher than that to the single chirp. Second, at sound pressure levels of 80 dB and higher, the responses to the second to fourth chirp in the train show less pronounced “earlier” peaks compared to the response to the single chirp.³

Fig. 2.6 shows the average data across subjects. Wave-V amplitude and latency values were extracted from the response waveforms. The latencies were defined relative to the respective offset of the rising chirp in the train. The left panel represents the amplitude of the first wave-V peak as a function of the stimulus level. The results for the rising-falling chirp train (circles) and the single chirps (triangles) are very similar. Wave-V amplitude increases up to 60 dB SPL, and decreases at higher sound pressure levels. Compared to the results of experiment 1, the amplitude is slightly larger here due to the larger bandwidth (100-10,000 Hz) of the chirp than in experiment 1 (320-8000 Hz). Thus, more neurons may have contributed to the evoked chirp response. As in experiment 1, the results at high stimulus levels (80-100 dB SPL) are represented by open symbols and connected with dashed lines. This was done in order to indicate that wave V was generally less clearly defined than at lower stimulus levels.

The second panel of Fig. 2.6 represents wave-V amplitude, averaged across all four peaks. For levels up to 60 dB SPL, results for the rising-falling chirp train are close to those of the single chirps. However, at higher sound pressure levels, wave-V amplitude for the chirp train is larger than that for the single chirp. As indicated in

³Responses to single rising-falling chirps were also recorded. The results are not shown here since the recordings were not obtained in the same session as it was the case for the other comparisons made in the present study. However, the results obtained with the individual rising-falling chirp were essentially identical to those for the rising chirp (i.e., the falling chirp had only a marginal effect on wave-V amplitude). This was as expected, as it produces desynchronized neural activity, at least at the low and moderate stimulus levels (Dau et al., 2000).

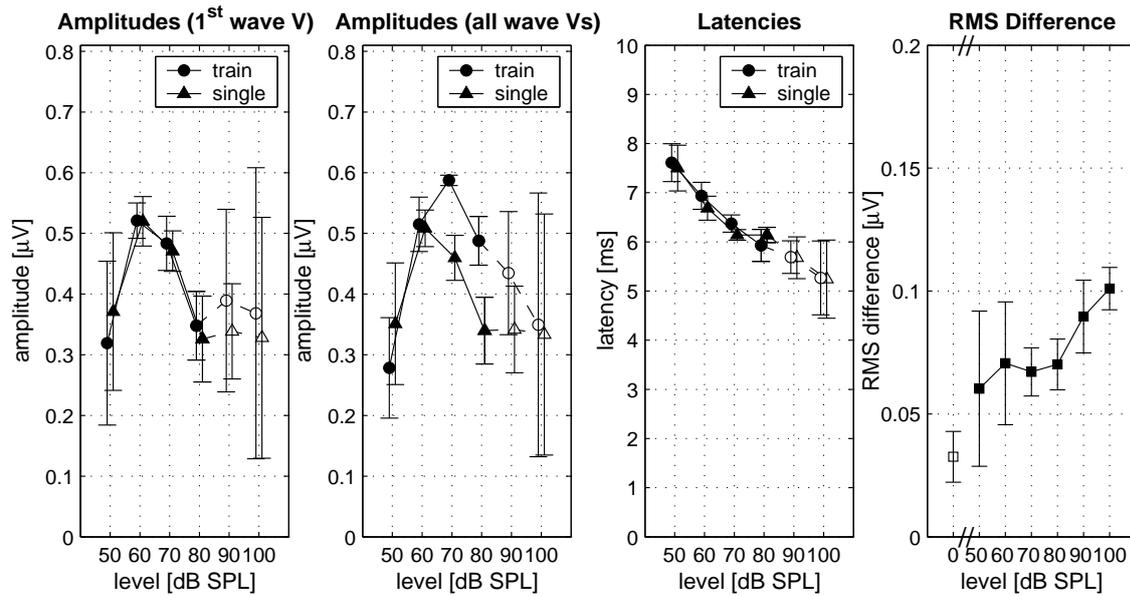


Figure 2.6: Average data obtained for the rising-falling chirp train. Wave-V amplitude (first and second panel) and relative latency (third panel) are shown as a function of stimulus level. The results for the rising chirp in the train are represented by the circles, while those for the single rising chirp are shown by the triangles. For better visibility, the results are plotted with an offset of 2 dB on the abscissa. In the first panel, only the initial wave V is considered, while the second panel represents the average across all four wave-V peaks. The error bars denote the inter-individual standard deviations. For the two highest stimulus levels, the results are represented by open symbols and connected with dashed lines since the definition of wave V was less clear than at lower levels. The right panel shows the mean RMS of the difference waveform between embedded-chirp and single-chirp responses. For comparison, the RMS of the noise activity in the silence condition is represented by the open rectangle, at “0 dB”.

Fig. 2.5 for the individual subject DJ, this effect is due to the increased amplitude of the second to fourth wave V relative to the initial wave V. The third panel of Fig. 2.6 shows wave-V latency, averaged over all four wave-V peaks. The latency decreases with increasing level. The function is essentially the same for the initial wave V alone which is not shown explicitly.

The RMS value of the difference between the response waveforms obtained in the chirp-train and the single-chirp paradigm was calculated in the interval between 0 ms and 90 ms, as a function of stimulus level. The results are shown in the right panel of Fig. 2.6. The RMS value of the recording noise differences (embedded-chirp versus single-chirp response) in the pre-stimulus interval (-10.0 ms to -0.1 ms), averaged

across levels and subjects, is indicated by the open symbol. As in experiment 1, for stimulation levels of 50 dB SPL and 60 dB SPL, the response differences do not differ significantly from the recording noise. At levels of 70 dB SPL and higher, the RMS differences increase substantially with level and are significantly higher than the recording noise.

In summary, the responses to the embedded chirp and the single chirp in experiments 1 and 2 were essentially the same at low sound pressure levels. At higher stimulus levels, some disparities between the responses to embedded and single chirps occurred. First, onset responses clearly became dominant in single-chirp stimulation for levels of 80-100 dB SPL. Second, for the chirp train, the wave-V amplitude of the peaks following the initial wave V were enhanced at stimulus levels of 70-80 dB SPL.

2.3.3 Experiment 3: Effect of within-train rate on wave-V amplitude and latency

The previous experiments showed that wave-V amplitude associated with the rising chirp remains unchanged throughout the rising-falling chirp train for levels up to 60 dB SPL. This was the case even for a within-train rate of 47.6 Hz. This rate lies within a range for which a strongly reduced wave-V amplitude was observed in other studies (Thornton and Coleman, 1975; Jiang et al., 1991). The effect of within-train rate on wave V for the chirp was investigated further in experiment 3. For comparison, click responses were collected at the same within-train rates as used for the chirps. In this experiment, the stimulation level was held constant at 40 dB SL, corresponding to an average sound pressure level of about 65 dB for the chirps (and 80 dB for the clicks). According to the results from the previous experiments with chirps, this level still lies in the region where embedded-chirp and single-chirp stimulation produce comparable results.

The upper left panel of Fig. 2.7 illustrates the response to the train of rising-falling chirps as in experiment 2, for subject KW. As described earlier, the train evokes four prominent peaks elicited by the rising chirps. The peak-amplitude is stable throughout the train. The upper right panel shows the results for a train consisting only of rising chirps. The time interval between the chirps was zero, i.e., the within-train rate was 95.2 Hz, corresponding to a chirp duration of 10.5 ms. The response pattern shows eight peaks, each with essentially the same amplitude as associated

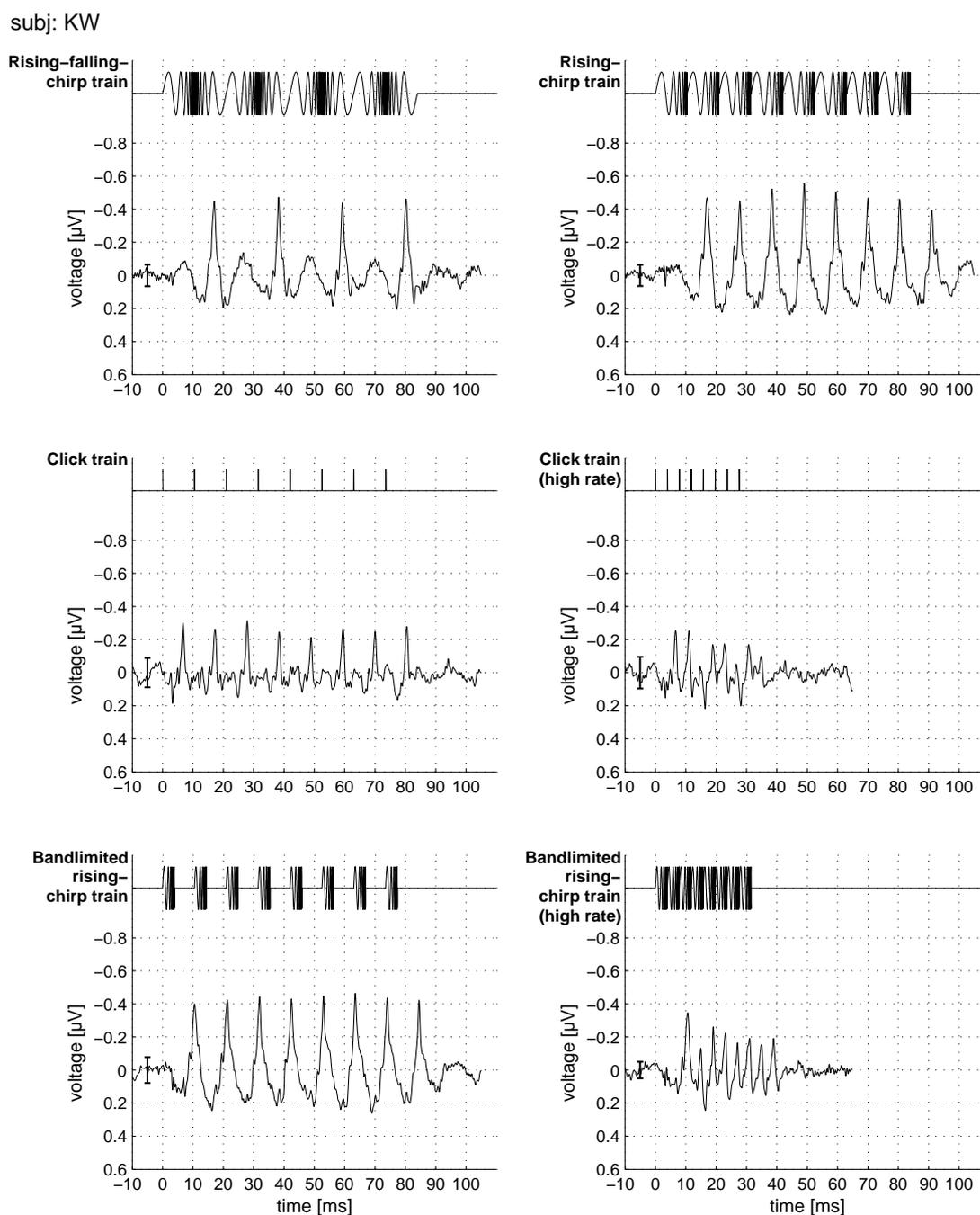


Figure 2.7: Evoked responses to the stimulus trains in experiment 3. The corresponding stimuli are shown above the responses. Only one stimulation level was used (40 dB SL). Subject: KW. The error bars indicate ± 3 SEM

with the previous condition.

For comparison, clicks at the same within-train rate (95.2 Hz) were presented. The corresponding responses are shown in the middle left panel of the figure. Again, eight single wave-V-peaks of about the same amplitude can be seen. For click stimuli,

wave-V amplitude is substantially smaller than for chirps. This is consistent with the results obtained in earlier studies (Dau et al., 2000; Wegner and Dau, 2002; Fobel and Dau, 2004). The advantage of the chirp over the click can also be observed in the lower left panel which shows the response to bandlimited chirps (0.5-10 kHz) presented at the same within-train rate of 95.2 Hz. This chirp has a duration of only 4.0 ms. Wave-V amplitude is significantly larger than for the click, and is almost as large as recorded in response to the broadband chirp.

The main reason for investigating the bandlimited chirp was to drive the auditory system to its limit with respect to temporal processing. The lower right panel of Fig. 2.7 shows the responses to the same bandlimited chirp but presented at the maximum within-train rate (without stimulus overlap) of 250 Hz. The stimulus train was presented at the same sound pressure level as that presented at 95.2 Hz. Each chirp in the train evokes a clear wave-V peak but the amplitude varies significantly throughout the response at this high within-train rate. The first peak is larger than the succeeding ones. This can also be seen in the average data shown in the corresponding panel of Fig. 2.8. For direct comparison, the click responses obtained at the same within-train rate are shown in the middle right panel of Fig. 2.7 (for the individual subject KW) and Fig. 2.8 (for the average results). As for the chirp stimulation at this high within-train rate, the click-evoked wave-V response amplitude is clearly reduced from the second chirp on.

Fig. 2.9 shows an analysis of wave-V latency. The results indicate the average across subjects. In each panel, the average latency of each peak in the train is indicated as open circles. The mean over the latency values across the train is represented by the closed circles in each panel. For all trains, the first component has the shortest wave-V latency. All subsequent peaks show a significantly increased latency relative to the first one. The latency difference between the first and the following peaks is more pronounced for the trains with the highest within-train rate of 250 Hz.

For the rate of 250 Hz the superposition of response components from separate stimuli in the train might have led to interference affecting individual components of the brainstem response. Fig. 2.10 shows the individual evoked responses to clicks at this rate for all five subjects that participated in the third experiment. For direct comparison, simulated superimposed responses for single-click stimulation are plotted and indicated as gray curves for the same subjects. The simulated waveforms were obtained by convolving the individual recorded single-click results

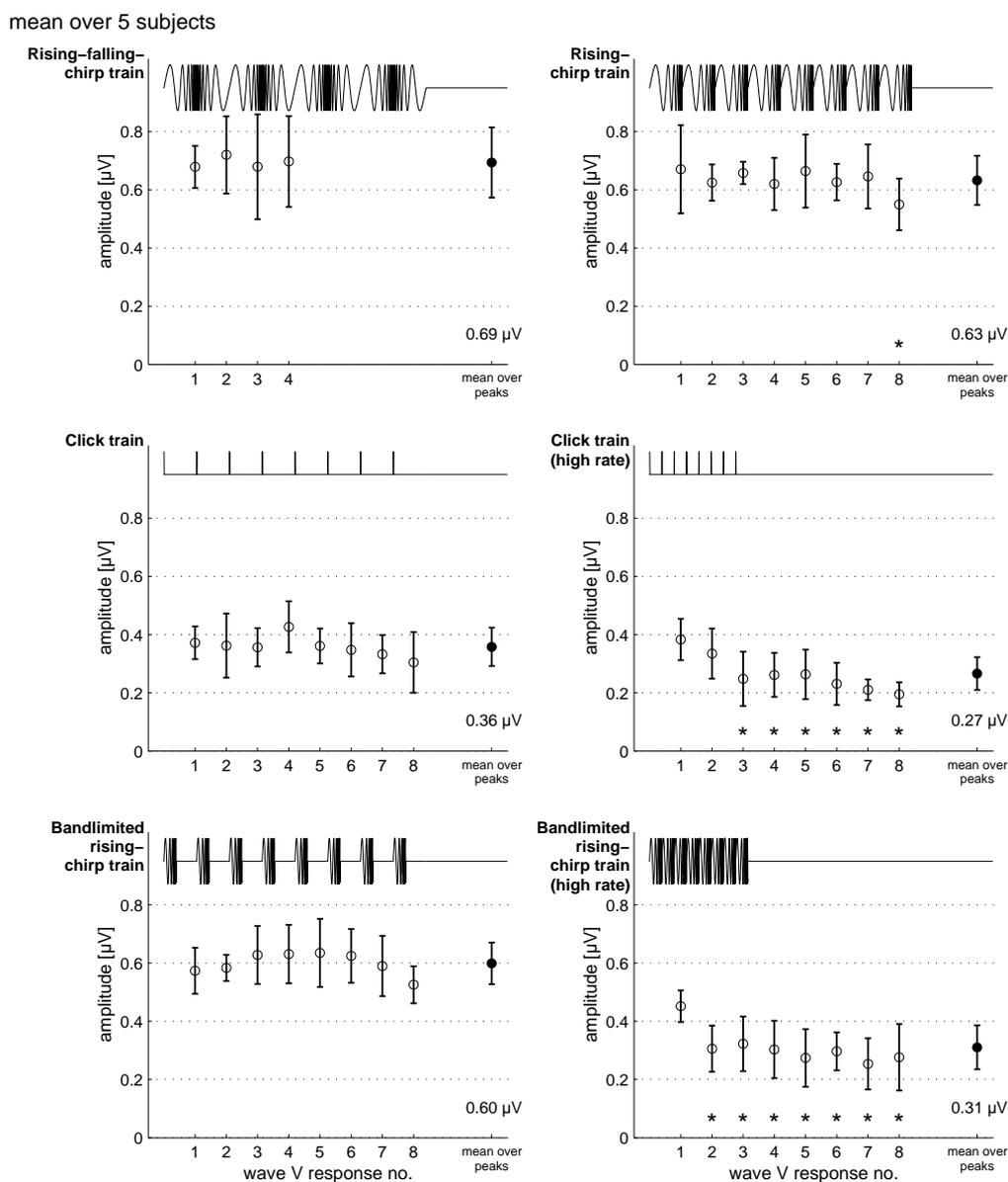


Figure 2.8: Wave-V amplitude values for the six stimulus trains investigated in experiment 3, averaged across subjects. Each single peak was averaged separately (open symbols). The mean over all peaks is plotted as closed circle in each panel. The error bars indicate inter-individual standard deviations. The asterisks indicate that the amplitude is significantly smaller than the response to the first chirp in the train.

with a delta comb whose period was 4 ms. The comparison indicates the latency change of the peaks in the click train towards larger values, as described above, which is not reflected in the simulation. With respect to wave-V amplitude, there is *no* systematic decrease across the simulated train response. Additional simulations (not shown) showed that the within-train rate (tested between 95 Hz and 250 Hz) had some effect on wave V due to the varying degree of constructive or destructive

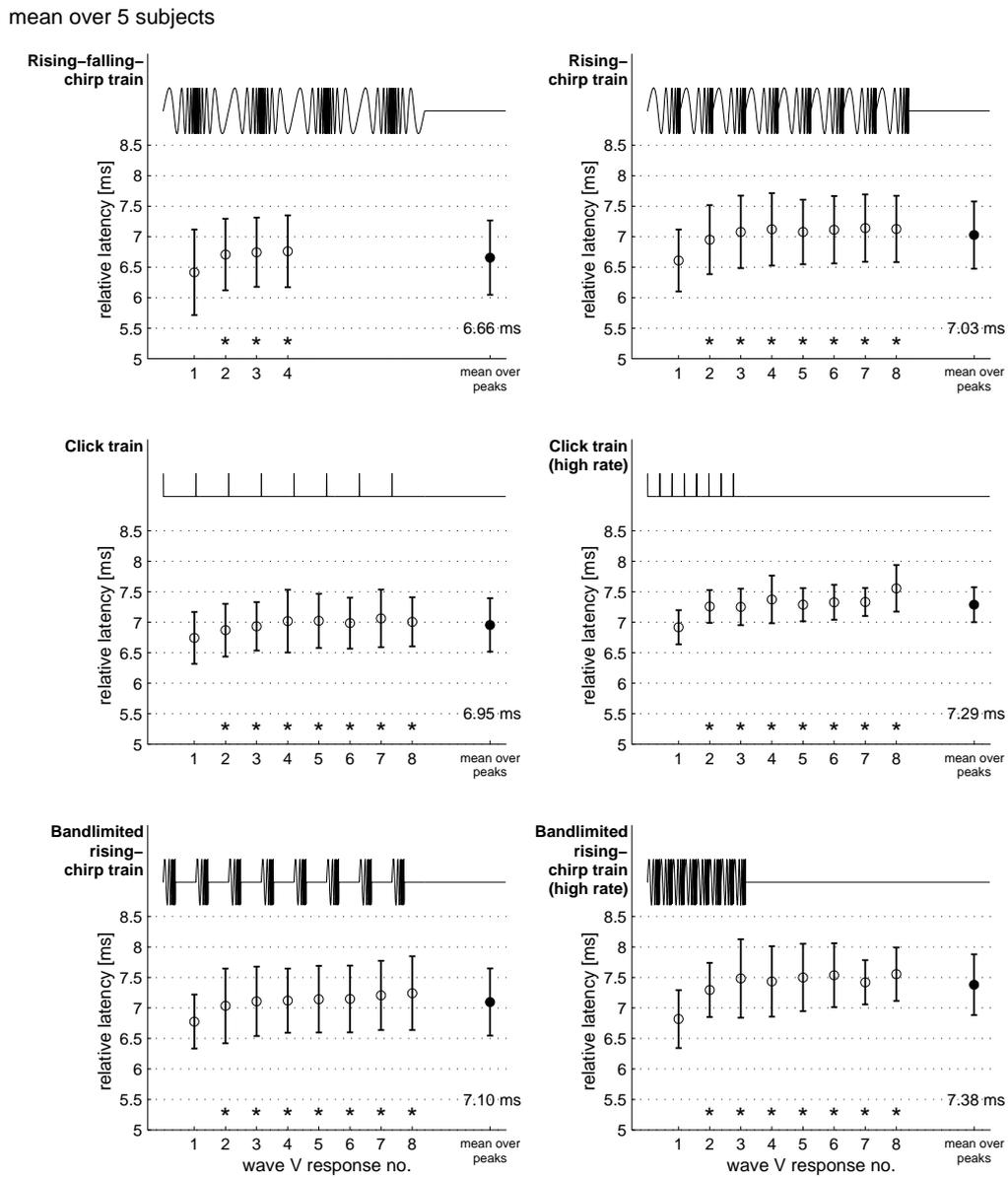


Figure 2.9: Wave-V latency values for the six stimulus trains investigated in experiment 3, averaged across subjects. Each single peak was averaged separately (open symbols). The mean over all peaks is plotted as closed circle in each panel. The error bars indicate inter-individual standard deviations. The asterisks indicate that the latency is significantly larger than that for the first chirp in the train. For the chirps, the latency values are given relative to chirp offset.

interference. However, a systematic decrease of wave-V amplitude within the train, as found in the experimental data, was not obtained. Thus, the observed decrease cannot be accounted for by an overlap of the separate response components.

In summary, experiment 3 showed that the within-train rate can be as high as approximately 100 Hz without any effect on wave-V amplitude. The mean amplitude

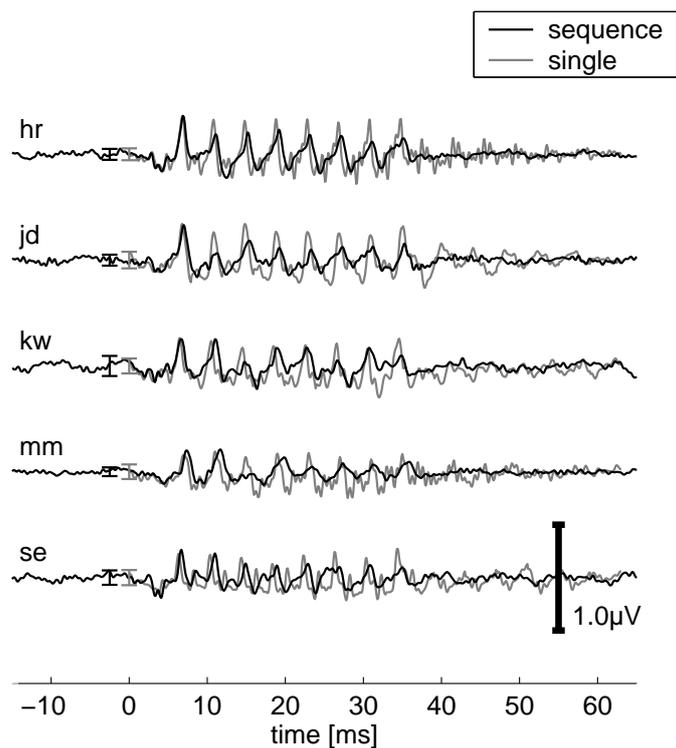


Figure 2.10: Individual responses to the click train with a within-train rate of 250 Hz (black curves). The gray curves show single-click responses convolved with a delta-comb with a period of 4 ms.

of the single-chirp response was the same as for the rising-falling chirp train and the rising-chirp train, and about twice as large as the corresponding click response. For the within-train rate of 250 Hz, wave-V amplitude was reduced from the second stimulus in the train on both for the chirp and the click. Wave-V latency increased as the number of chirps/clicks increased within the stimulus train. This was the case for all within-train rates tested in the experiment (47.6, 95.2, 250 Hz), whereby the effect was strongest for the highest rate.

2.4 Discussion

2.4.1 Influence of the cochlear traveling wave on brainstem responses

The results of the first two experiments demonstrated a close similarity between the responses to embedded and single chirps for stimulus levels up to 60-70 dB SPL. The chirp was equally effective in both stimulation paradigms, the acoustic context did not affect the responses. In this level region, wave V is the only significant component that contributes to the brainstem response while the earlier components, waves I and III, are typically not visible. With respect to wave-V amplitude, the auditory system behaves like a linear system in terms of the integration of neural activity across frequency in the conditions investigated. The evoked response, and particularly the analysis of specific components of it, only reflects certain aspects of neural processing, while other properties cannot be “seen” by this method.

At high stimulation levels (80 dB SPL and higher), nonlinear effects became apparent. The disparity between the responses evoked by the two stimulus paradigms increased at higher sound pressure levels. As discussed in a previous study (Dau et al., 2000), at these sound pressure levels, the chirp onset causes a significant basal spread of excitation of the early low-frequency signal components, and leads to a response peak with a latency of about 6 ms relative to the stimulus onset, which corresponds to wave V. This superimposes on the later activity from the mid and high frequencies in the chirp, finally resulting in a more complex response pattern than observed at lower stimulus levels. At the highest sound pressure levels, this more complex response pattern showed characteristics similar to that of an FFR. According to the results of a modeling study by Dau (2003), FFR to low-frequency tones or low-frequency tonal sweeps represent synchronized activity mainly stemming from mid-to-high frequency units, and not activity from units tuned to frequencies around the signal frequency. It was argued that, due to the high velocity of the traveling wave, basal channels are in relatively close phase relationships with each other and, thus, the synchronized activity in the basal region of the cochlear partition essentially initiates the FFR. Therefore, at high stimulus levels, the response to the chirp is most likely dominated by the activity from the mid and high frequencies, while at low and moderate levels, all frequency components in the chirp effectively con-

tribute to the evoked response. This is the case both for the single chirp as well as for the embedded chirp. However, in the embedded chirp, (i) the onset response is diminished or strongly reduced (see Fig. 2.5), and (ii) the variation of the FFR at the highest levels is typically reduced relative to the transient response. Most likely, effects of saturation in the case of the embedded stimulation cause the change of the evoked activity between single-chirp versus embedded-chirp stimulation.

The results from experiments 1 and 2 support the hypothesis that the critical factor for eliciting an ABR wave V is not the onset or offset of an acoustic stimulus (e.g., Hecox et al., 1976; Kodera et al., 1977; Debruyne and Forrez, 1982; Gorga and Thornton, 1989; Campen et al., 1997), but rather the degree of synchrony of activation along the tonotopic array originating in the cochlea. An appropriate temporal organization, determined by traveling wave properties, can enhance neural synchrony.

2.4.2 Effect of within-train repetition rate on wave V

A number of earlier studies have investigated the effect of click rate on processing in the brainstem. Most electrophysiological studies that employed amplitude and latency measures of ABR used the traditional single-stimulus paradigm (e.g., Jewett et al., 1970; Pratt and Sohmer, 1976; Burkard and Hecox, 1983; Lasky, 1991). In those studies investigating the effect of click rate, ABR latencies generally increased and ABR amplitude decreased as click rate increased. For example, Jiang et al. (1991) found a 33% decrease in wave-V amplitude when the click rate was raised from 10 Hz to 90 Hz at 40 dB hearing level (HL). Scott and Harkins (1978) found an amplitude decrease of 19% when increasing the rate from 10 Hz to 100 Hz. Harkins et al. (1979) found a latency increase from 5.7 ms to 6.2 ms when increasing the click rate from 10 Hz to 100 Hz at 70 dB HL.

However, the values of wave-V latency and amplitude depend on the stimulus paradigm. In several studies, including the present one, short trains of stimuli were presented in order to investigate the transition from the responses to the stimuli in the beginning of the train to those at the end of the train. For example, Don et al. (1977) presented trains of 10 clicks with a within-train rate of 100 Hz, at 40 dB SL. They reported that the ABR latencies in response to the fourth or fifth click (and subsequent clicks) were increased by about 0.6 ms relative to the response to the

first click, and were the same as to single clicks presented at a rate of 100 Hz. This latency increase, however, is about twice as large as the increase found in the present study for clicks presented at a within-train rate of 95 Hz. Unfortunately, [Don et al. \(1977\)](#) did not analyze wave-V amplitude in addition to wave-V latency. [Thornton and Coleman \(1975\)](#), using trains of four clicks with a within-train rate of 67 Hz, found an amplitude decrease of 48% when comparing the last wave-V peak of the train with the first. This is different from the results of the present study (where no amplitude change was found at within-train rates of 47.6 Hz and 95 Hz) while their observed latency increase of 4% (4th peak relative to 1st peak) corresponds to the results of the current study for the click train. In a recent study, [Polyakov and Pratt \(2003a\)](#) examined both amplitude and latency behavior when stimulating with trains of 10 clicks at a within-train rate of 91 Hz and an inter-stimulus rate of 5.13 Hz. [Polyakov and Pratt \(2003a\)](#) did not observe significant amplitude changes throughout the train for wave V, and only a moderate latency increase of 0.3 ms. Thus, their results are very close to the findings in the present study.

For the within-train rate of 250 Hz, a substantial reduction of wave-V amplitude (by about 35%) was observed in the present study for the peaks following the first. In addition, there was a larger increase in wave-V latency for these peaks in comparison to the results at the lower within-train rates (see Fig. 8). The simulation results presented in Fig. 2.10 indicated that linear superposition of overlapping response components in the train do not account for the above observations. Some possible mechanisms underlying the nonlinearities in terms of temporal processing with respect to wave-V behavior are discussed in the following.

The auditory system requires a finite period of time following a stimulus to fully recover its responsiveness. Major determinants of neural recover time could be refractory periods of neural elements and neural adaptation. The neural refractory period can not account for the latency shift seen with increasing within-train rates because the time course of this phenomenon is rapid (1-2 ms) compared to the time intervals between the stimuli considered here ([Don et al., 1977](#)). Adaptation refers to the variation of response which occurs during a sustained stimulus. The change in neural activity is typically maximum at onset and then decays or adapts to a smaller sustained change in response. The neural mechanisms underlying adaptation have been studied in peripheral and more central stages of the auditory nervous system (e.g., [Smith, 1977](#); [Kramer and Teas, 1982](#); [Westerman and Smith, 1987](#); [Arehole](#)

et al., 1989; Kaltenbach et al., 1993). At the level of the auditory nerve, adaptation can be described using two time constants: “rapid”, which occurs over 1-10 ms, and “short-term” in which the time constant is approx. 60 ms (Smith, 1977; Westerman and Smith, 1984). Moreover, the population response of the auditory nerve, the compound action potential (CAP), was also found to behave similarly (Gorga and Abbas, 1981). The CAP is related to the wave-I component of the ABR. Effects of adaptation (in a more variable and complex form than at auditory nerve level) were also observed at higher neural stages in the brainstem, in the cochlear nucleus (Abbas and Gorga, 1981) and the inferior colliculus (Arehole et al., 1987), and were also found to affect ABR wave V (Walton et al., 1995; Polyakov and Pratt, 2003a; Pratt et al., 2004). The data of the present study suggest that rapid adaptation is the main cause for the observed nonlinearities in wave-V behavior with respect to temporal processing.

Wave V represents a far-field summation of neural activity of many elements. The latency of its peak may represent either the modal value of the elements comprising the response, or the value of the strongest component, or a combination of the two (Don et al., 1977). In the present study it was found that, at within-train rates up to 95 Hz, the latency increased (across the stimulus train) while the amplitude remained constant, suggesting that the synchronicity was still large enough to generate a stable wave-V amplitude. Thus, in this condition, synchronization was possibly slightly reduced but overall neural activity remained unchanged. Wave-V latency might therefore be more sensitive to changes in synchronization than wave-V amplitude. In terms of wave-V amplitude, the system behaves essentially linearly at within-train rates up to about 100 Hz. For the earlier brainstem components, waves I and III, this is not the case anymore, as was shown in other studies (e.g., Thornton and Coleman, 1975; Pratt and Sohmer, 1976). For very high within-train rates, such as 250 Hz, synchronicity might be diminished further, as reflected in an increased wave-V latency shift. This could be the reason for the reduction of wave-V amplitude, because the temporal smearing of the neural activity might “exceed” the duration of the effective integration window at the stage of processing where wave V is generated.

2.4.3 Summary and conclusions

For stimulation levels up to 60-70 dB SPL (corresponding to roughly 40 dB HL), the evoked responses obtained with the embedded chirp closely corresponded to the responses elicited by single chirps. For these conditions, the auditory system behaves linearly with respect to wave-V amplitude, i.e., the integration of neural activity across frequency does not depend on the acoustic context. At high stimulation levels (typically 80 dB SPL and above), differences between the responses obtained in the two stimulation paradigms occurred, probably associated with effects of neural saturation.

With respect to temporal processing, the influence of within-train repetition rate on wave-V latency and amplitude was investigated for the click and the chirp. Wave-V latency increased with temporal position in the train for all within-train rates (47.6, 95.2, 250 Hz), whereby the effect was strongest for the highest rate. Wave-V amplitude, however, was only affected at the highest within-train rate tested (250 Hz), and behaved linearly at the lower rates. Wave-V latency therefore probably reflects a more sensitive indicator of neural synchronization than wave-V amplitude.

For a more quantitative understanding of the relation between the neural processing in the brainstem and the evoked far-field potential, modeling work is needed. Recent attempts to predict ABR, using a state-of-the-art model of auditory-nerve preprocessing have been successful in some aspects but failed in predicting other important aspects (Dau, 2003). For example, wave-V latency as a function of stimulation level could not be accounted for accurately. Some of the details of brainstem processing and their relation to the evoked far-field responses are still unknown. The experimental results of the present study may provide constraints on future models of peripheral and brainstem processing.

Overall, the results of the present study may further demonstrate the importance of cochlear processing for the formation of ABRs. The findings might have interesting implications for clinical applications. In studies investigating wave-V amplitude behavior, the stimulus-train paradigm allows higher mean stimulus rates than the traditional single-stimulus paradigm. The chirp stimulus might be useful as an objective indicator of hearing threshold since it enables the inclusion of contributions from the lower frequencies while click responses mainly reflect activities only from the mid and high frequencies.

Chapter 3

Is there an externalization detector at early stages of the auditory system? A chirp-evoked auditory evoked potential study

Abstract

The effect of externalization and spatial cues on the generation of auditory brainstem responses (ABRs) and middle latency responses (MLRs) was investigated in this study. Most previous potential studies used click stimuli with variations of interaural time (ITDs) and interaural level differences (ILDs) which merely led to a lateralization of sound inside the subject's head. In contrast, in the present study potentials were elicited by a virtual acoustics stimulus paradigm with 'natural' spatial cues and compared to responses to a diotic, non-externalized reference stimulus. Spatial sound directions were situated on the horizontal plane (corresponding to variations in ITD, ILD, and spectral cues) or the midsagittal plane (variation of spectral cues only). An optimized chirp was used, which had proven to be advantageous over the click since it compensates for basilar membrane dispersion. ABRs and MLRs were recorded from 32 scalp electrodes and both binaural potentials (B) and binaural difference potentials (BD, i.e., the difference between binaural and summed monaural responses) were investigated. The amplitudes of B and BD to spatial stimuli were not higher than those to the diotic reference. ABR amplitudes

decreased and latencies increased with increasing laterality of the sound source. A rotating dipole source exhibited characteristic patterns in dependence on the stimulus laterality. Changes in the elevation had no effect on the ABRs. For the MLR data, stimulus laterality was reflected in the latencies of the BDs. In addition, dipole source analysis revealed a systematic magnitude increase for the dipole contralateral to the azimuthal position of the sound source. For the variation of elevation, the right dipole source showed a stronger activation for stimuli away from the horizontal plane. The results indicate that at the level of the brainstem and primary auditory cortex binaural interaction is mostly affected by interaural cues (ITD, ILD). An enhancement of the interaction component due to natural combinations of ITD, ILD, and spectral cues could not be observed, thus providing no evidence for an externalization detector at these early stages of the human auditory system.

3.1 Introduction

The ability of the auditory system to localize sounds in three-dimensional space is determined by three acoustic cues: The interaural differences in sound arrival time (interaural time difference, ITD) and sound level (interaural level differences, ILD) are used to determine the sound direction on the horizontal plane (e.g., [Blauert, 1997](#)). For vertical localization, spectral cues, resulting from sound reflections on the head, torso, and pinnae must be utilized. This direction-dependent filtering also helps to resolve localization ambiguities for sound directions resulting in identical ITDs and ILDs (commonly referred to as ‘cone of confusion’, e.g., [Woodworth, 1938](#); [Mills, 1972](#)).

Presenting auditory stimuli over headphones with missing spectral cues (i.e., with ITD and ILD alone), results in a sound percept inside the listener’s head, typically along an axis from one ear to the other. This degraded type of localization is often referred to as lateralization. In order to invoke a true representation of three-dimensional space via headphones, the direction-dependent spectral filtering must be incorporated in the stimulus presentation. This is typically accomplished by the method of ‘virtual acoustics’: Head-related impulse responses (HRIRs) are measured with microphones in the subjects’ or artificial head’s ear canals for all relevant directions (and distances). These recorded HRIRs contain the ‘natural’ combination of all acoustic cues necessary to localize the position of a sound in space (ITD, ILD, and spectral cues, i.e., the general attenuation/gain as a function of frequency). When presenting a stimulus convolved with a pair of HRIRs to both ears over headphones, a percept similar to a free-field condition can be obtained (e.g., [Mehrgardt and Mellert, 1977](#); [Genuit, 1984](#); [Middlebrooks and Green, 1991](#); [Wightman and Kistler, 2005](#)).

Auditory evoked potentials (AEPs) play a key role when investigating the neural processes underlying sound perception in humans. They can be recorded from all levels of the auditory system and are usually grouped by their latency, i.e., the time of occurrence after the onset of the stimulus. This grouping corresponds roughly to the site of generation (e.g., [Picton et al., 1974](#); [Scherg, 1991](#)). AEPs represent the summation of responses from many neurons, recorded from the far-field, i.e., from electrodes placed on the surface of the head (e.g., [Jewett et al., 1970](#)). More specifically, the potentials recorded with binaural stimuli can help to understand the

mechanisms of sound localization. The binaural difference potential (BD) is commonly regarded as the neural correlate of binaural processing. It is calculated as the difference between binaural (B) and summed monaural (L+R) auditory evoked potentials, symbolically $BD=B-(L+R)$. The first negative deflection of this difference wave, named DN1 (after the convention introduced by [Ito et al., 1988](#)), occurs at a latency shortly after wave V of the auditory brainstem response (ABR), and is the most stable brainstem component of the ABR BD. Numerous studies investigated the dependence of DN1 (sometimes also referred to as β , e.g., [Levine, 1981](#)) on interaural time and level differences ([Wrege and Starr, 1981](#); [Gerull and Mrowinski, 1984](#); [Furst et al., 1985](#); [Jones and Van der Poel, 1990](#); [McPherson and Starr, 1995](#); [Brantberg et al., 1999](#); [Riedel and Kollmeier, 2002a, 2003](#)). All studies found increasing DN1 latencies for increasing ITDs. Some authors reported constant DN1 amplitudes for ITDs up to about 1 ms ([Furst et al., 1985](#); [Jones and Van der Poel, 1990](#); [Brantberg et al., 1999](#)). In contrast to that, [McPherson and Starr \(1995\)](#) found a monotonic decrease of DN1 amplitude with ITD. For increasing ILD, [Furst et al. \(1985\)](#) and [McPherson and Starr \(1995\)](#) report a monotonic increase of latencies and a monotonic decrease of amplitudes. [Riedel and Kollmeier \(2002a\)](#) investigated the binaural wave V and the DN1, varying ITD and ILD at the same time. Their results to trading stimuli indicate that ITD and ILD are not processed independently in the brainstem, but rather combined to roughly represent lateralization. In the middle-latency range, BD waves occur at latencies close to 19 ms, 30 ms, and 40 ms ([McPherson and Starr, 1993, 1995](#)). [McPherson and Starr \(1995\)](#) reported a monotonic amplitude decrease for increasing ITD and ILD for all three components. However, latencies were not affected by changes to ITD and ILD, except for the 30-ms component which showed an increase for increasing ITDs.

[Polyakov and Pratt \(2003b\)](#) used click-stimuli synthesized with non-individualized HRIRs from the KEMAR mannequin ([Gardner and Martin, 1995](#)). Analyzing an equivalent brainstem dipole source on basis of three bipolar, orthogonal electrode channels (three-channel-Lissajous' trajectory, e.g., [Pratt et al., 1983](#); [Jewett et al., 1987](#)), they found significant effects of stimulus position on latency and magnitude of the BD's equivalent dipole. However, the results for dipole magnitudes were mostly asymmetric for left and right lateralized stimuli, which contradicts the ABR data elicited by stimuli with ITD and ILD alone. In a follow-up study ([Polyakov and Pratt, 2003c](#)), the same authors investigated MLR recordings with the same

methods. Again, some significant effects on magnitude and latency of the BD's equivalent dipole were found for different azimuths and elevations. The results of these two studies support the notion that the processing of spectral cues in sound localization may have significant effects on ABR and MLR recordings, and these effects may differ from those caused by mere interaural cues.

In all these studies in the literature, only non-individualized HRIRs or even no HRIRs at all have been used to produce spatialized stimuli, but these sounds do not necessarily evoke a clearly externalized and distinctly localized spatial impression. It therefore remains unclear whether the observed binaural interactions in AEPs are negatively influenced by these unnatural cues and might benefit from using natural spatial sounds as provided by individual HRIRs. In real life, auditory objects are always externalized sounds. It is therefore reasonable to assume that the neural structures of the peripheral auditory system effectively operate as 'externalization detector' which not only analyzes ITD and ILD separately for each frequency, but is sensitive to natural combinations of ITD, ILD, and spectral cues across a wider range of frequencies. Such a detector would optimally be activated by natural or individualized spatial stimuli and would be considerably less activated by unnatural, non-externalized spatial cues across frequency. The current study tests this 'externalization detector hypothesis' at the level of the brainstem and primary auditory cortex in humans.

Instead of the commonly used click stimulus a rising chirp is used in this study. The chirp was designed to activate all cochlear locations simultaneously and has shown to evoke a larger ABR wave V amplitude than a click presented at the same sensation level (Dau et al., 2000; Wegner and Dau, 2002; Fobel and Dau, 2004). This advantage of the chirp over the click has been verified also for binaural potentials and BDs, at least for low and medium presentation levels (Riedel and Kollmeier, 2002b). Further studies reported that the chirp stimulus also increases the amplitudes of the middle latency response (MLR), which is generated at the level of the primary auditory pathway. In magnetoencephalographic studies, Rupp et al. (2002) found a significantly enhanced N19m-P30m complex for the rising chirp compared to the click. This demonstrates that phase delays between channels in the auditory pathway are preserved at least up to the primary auditory cortex.

The aim of this study is to test the 'externalization detector hypothesis' by investigating binaural potentials and binaural difference potentials, generated at the level

of the brainstem (ABR) and of the primary auditory cortex (MLR), incorporating realistic, externalized spatial stimuli with an individual virtual acoustics paradigm. Responses to spatial stimuli are compared to diotic chirp stimuli to test if response amplitudes benefit from externalized stimuli. A 32-channel electrode setup allows for modeling of dipolar sources. The dependence of the single channel waveform data and equivalent dipoles on both sound azimuth and elevation is examined.

3.2 Methods

3.2.1 Subjects

Six normal-hearing subjects (one female, five male), aged from 29 to 39, volunteered in the experiments. Subjects had no history of audiological or neurological problems and their audiometric thresholds were 10 dB HL or better for frequencies from 500 Hz to 4000 Hz, and 20 dB HL or better for 8000 Hz.

3.2.2 Stimuli

Stimuli were digitally generated at a rate of 50 kHz (16 bit resolution), D/A-converted by a DSP-card (Ariel DSP32C), and level adjusted by a digitally controlled audiometric amplifier. Stimuli were presented through insert earphones (Etymotic Research ER-2) encased in copper-boxes to avoid electrical artifacts in the AEP.

A rising chirp stimulus (Dau et al., 2000), designed to activate all cochlear locations simultaneously, was used instead of the conventional click. By attenuating the lower frequencies, the spectrum of the chirp was flat in the range between the nominal edge frequencies of 100 Hz and 18,000 Hz, resulting in -3 dB-points of about 130 Hz and 15,000 Hz for the stimulus. Fig. 3.1 shows the acoustic waveforms and spectra of the chirp. The duration of the chirp was 10.4 ms.

Virtual sound directions of the stimulus were generated by convolving the chirp digitally with individual head-related impulse responses (HRIRs). The HRIRs were recorded in an anechoic chamber (Otten, 2001, Chapter 3). For each stimulus condition, the absolute travel-time delay for the leading side was eliminated from the left and right HRIR.

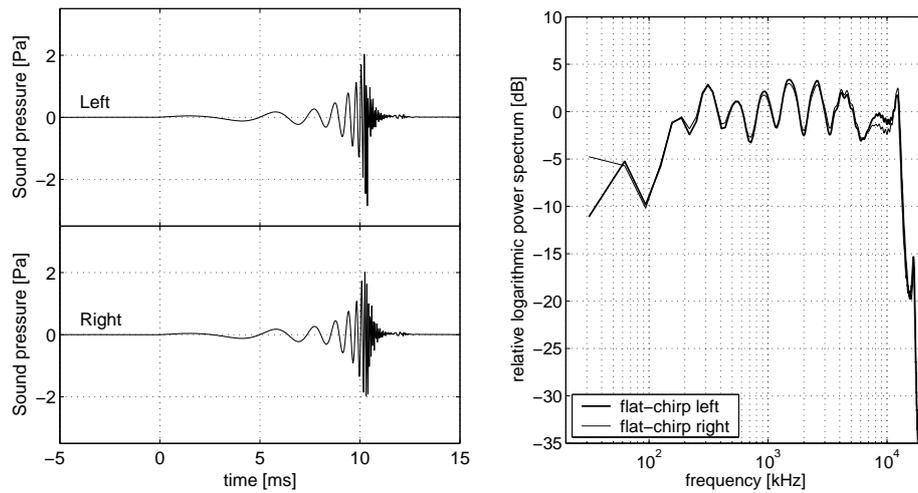


Figure 3.1: Optimized broadband chirp stimulus with flat spectrum. The chirp compensates for the basilar membrane dispersion providing synchronous discharges of auditory nerve fibers along the length of the entire cochlear partition (Dau et al., 2000). **Left panel:** Acoustic waveforms of the left and right flat-chirp in the time domain, measured at 102 dB peSPL, corresponding to 54 dB nHL. **Right panel:** Acoustic spectra of the left and right chirp.

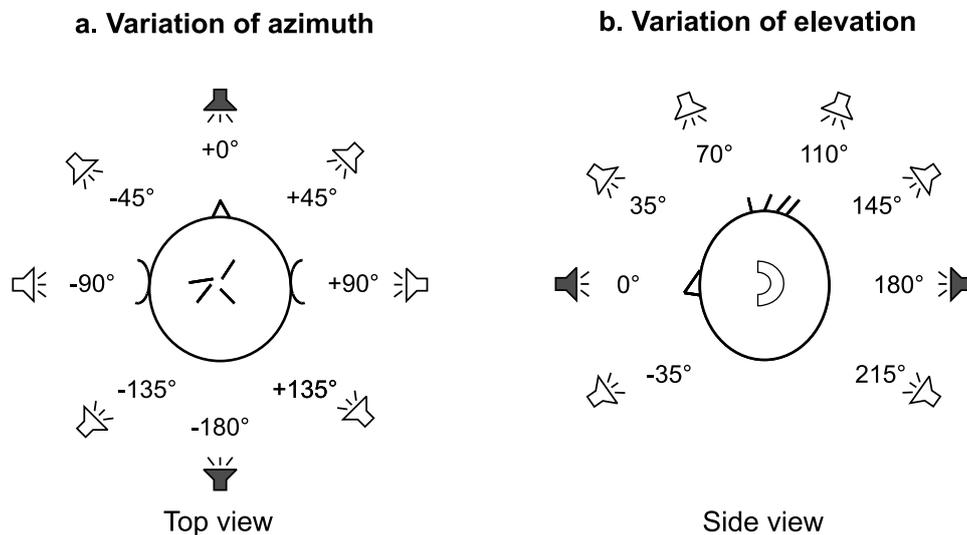


Figure 3.2: Stimulus conditions: The chirp was convolved with individual HRIRs corresponding to **a.** eight virtual directions in the horizontal plane and **b.** eight virtual directions in the midsagittal plane. In addition, the unconvolved chirp was presented as a diotic reference condition (not shown). The directions indicated by the dark speakers lie in both the horizontal and the midsagittal plane and were only measured once. For **a.**, numbers denote the azimuthal angle φ with the midsagittal plane, for **b.** the elevation angle ϑ with the horizontal plane.

A total of 15 different sound conditions was used in this study (Fig. 3.2): a. eight directions on the horizontal plane (variation of azimuth φ , the frontal stimulus has $\varphi = 0^\circ$); b. eight directions on the midsagittal plane (variation of elevation ϑ , the frontal stimulus has $\vartheta = 0^\circ$). Dark speakers indicate the directions that are identical for a. and b., i.e., they lie in both the horizontal and midsagittal plane. In addition, the unconvolved chirp was presented as a diotic reference condition.

Prior to the EEG-recordings, detection thresholds were obtained from all subjects to determine the sensation level for the diotic reference stimulus and the frontal stimulus direction. An adaptive three alternative forced-choice (3AFC; [Levitt, 1971](#)) procedure was employed, estimating the 70.7% point on the psychometric function. The average across three repetitions was considered as representing 0 dB SL for the respective stimulus and subject. For the EEG stimulus presentation, the level was set to 40 dB SL for the diotic reference stimulus. The virtual stimulus directions were jointly adjusted in level, such that the frontal stimulus condition had a level of 40 dB SL as well. The comparatively low stimulus levels were chosen to avoid the influences of acoustic crosstalk and the middle ear reflex ([Levine, 1981](#)).

3.2.3 EEG-Recording and procedure

Subjects wore a flexible cap (Easy Cap, Falk Minow Services) equipped with a set of Ag/AgCl-electrodes. Electroencephalographic activity was recorded differentially from 32 head positions according to the extended 10-20-system ([Jasper, 1957](#); [Sharbrough et al., 1991](#)). An electrode at the vertex (Cz) served as reference electrode, an electrode at the forehead (Fpz) as common ground. Interelectrode impedances were maintained below 5 k Ω . The electrode positions were measured before and after each recording session with an ultrasonic head-tracker (CMS30P, Zebris Medizintechnik). The multichannel EEG recording setup is shown schematically in Fig. 3.3.

During the recordings subjects lay comfortably on a couch in an acoustically and electrically shielded sound booth. They were asked to relax and encouraged to sleep during the recordings. Inside the booth, electrode signals were pre-amplified by a factor of 150 (Neuroscan Headbox) and passed to the main EEG amplifier (Neuroscan Synamps 5803) outside the booth for further amplification by a factor of 33.3, yielding an overall amplification of 74 dB. EEG potentials were then anti-alias filtered (analog 2nd order lowpass at a cut-off frequency of 2 kHz), A/D-converted at

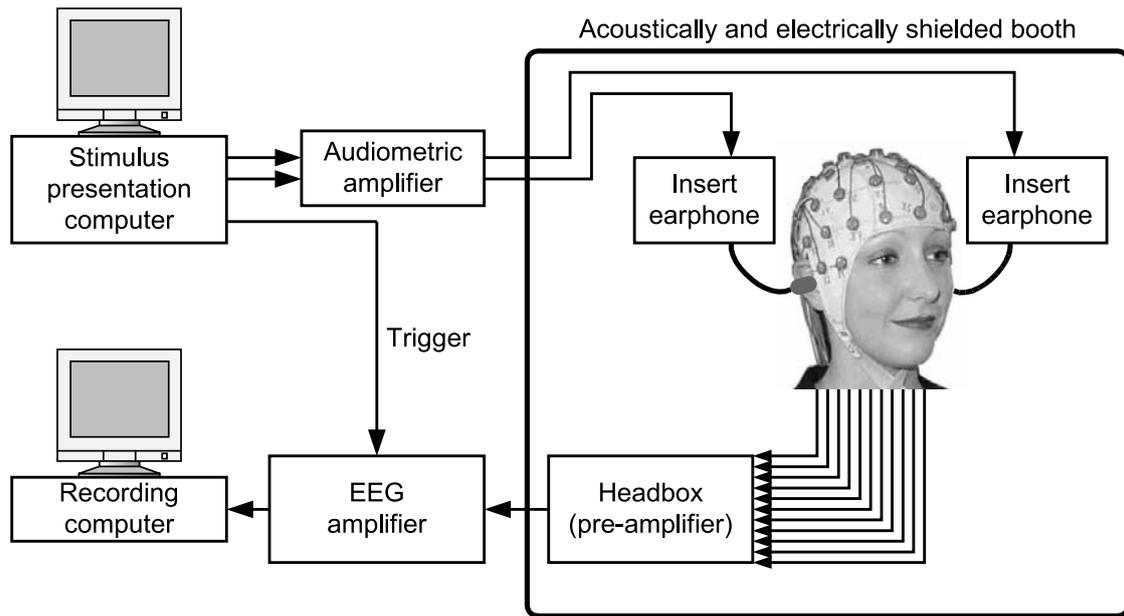


Figure 3.3: The EEG recording setup. Stimuli were presented in an electrically and acoustically shielded booth over insert earphones. Electroencephalographic activity was recorded from 32 surface electrodes (picture of woman's head taken from www.easycap.de by permission of Falk Minow Services).

a rate of 10 kHz, and continuously stored to the harddisk of the recording computer.

In a first recording session, only the responses to the variation of azimuth and the diotic reference condition were recorded (Fig. 3.2, a.). Each of the 9 stimulus directions was presented monaurally left, monaurally right, and binaurally with 10,000 repetitions per condition. This total of 270,000 stimuli was presented quasi-simultaneously in a random fashion. Consequently, any long-term variations in noise level had equal effect on all responses, yielding about the same signal-to-noise ratio for all averages (Ito et al., 1988). The interstimulus interval was uniformly distributed in the interval between 62 ms and 72 ms (temporal jitter), corresponding to a mean presentation rate of 14.9 Hz. The total recording time of five hours was divided into 24 runs, each lasting about 12.5 minutes. After each run, electrode impedances were checked to ensure a good quality of the recordings. Most of the subjects accomplished the 24 runs in two sessions of 12 runs on separate days. The duration of each of the two sessions, including the preparation of the electrode cap and measuring of the electrode positions prior and after each session, was typically 3.5-4.5 hours.

The additional responses to the variation of elevation (white speakers in Fig. 3.2, b.)

were acquired in a second set of recording sessions. These 180,000 stimuli (10,000 repetitions of 6 conditions, 2 monaural plus binaural each) were again presented in random order. The total recording time of almost 3.5 hours was divided into 16 runs, each lasting about 12.5 minutes. Most of the subjects accomplished the 16 runs in two sessions of 8 runs on separate days. The duration of each of the two sessions was typically 2.5-3 hours.

3.2.4 Data post-processing

All data post-processing and offline analysis (see sections 3.2.5 and 3.2.6) was performed in MATLAB (The Mathworks). EEG-data was segmented into stimulus-related epochs and bandpass filtered using a finite-impulse-response (FIR) filter with 200 taps (Granzow et al., 2001). The time-range of the epochs was from -5 to 45 ms relative to stimulus onset for ABR evaluation, and from -5 to 65 ms for MLR evaluation. The filter cut-off frequencies were set to 100-1500 Hz for ABR data and 20-300 Hz for MLR data. Epochs were averaged using an iterative weighted averaging algorithm (Riedel et al., 2001). The residual noise was calculated as the time-averaged standard error of the mean (SEM) σ for all waveforms.

The BD was calculated in the convention used by Ito et al. (1988), i.e., by subtracting the sum of the monaural responses (L+R) from the binaural response (B). For each of the 15 stimulus conditions, the BD was calculated from the corresponding three responses, e.g.,

$$\text{BD}_{\varphi=45^\circ} = \text{B}_{\varphi=45^\circ} - (\text{L}_{\varphi=45^\circ} + \text{R}_{\varphi=45^\circ}) \quad (3.1)$$

for the frontal right direction.

The SEMs for the monaural sum and BD were calculated from the SEMs of the measured waveforms according to

$$\sigma_{\text{L+R}} = \sqrt{\sigma_{\text{L}}^2 + \sigma_{\text{R}}^2} \quad (3.2)$$

$$\sigma_{\text{BD}} = \sqrt{\sigma_{\text{L}}^2 + \sigma_{\text{R}}^2 + \sigma_{\text{B}}^2} \quad (3.3)$$

For a quasi-simultaneous recording of B, L, and R, the SEMs of the three responses

are expected to be almost equal. Therefore the SEMs of the monaural sum and the BD increase by about a factor of $\sqrt{2}$ and $\sqrt{3}$, respectively.

3.2.5 Analysis of amplitudes and latencies

For the evaluation of amplitudes and latencies of ABR and MLR peaks, a subset of four channels showing the largest AEPs of the electrode configuration was chosen, namely positions A1, A2, PO9, and PO10. EEG data in these channels were up-sampled by a factor of 10 (yielding 100 kHz sampling rate) to increase the accuracy of the latency and amplitude of the peak measurements. The upsampling was done by zero-padding in the spectral domain, resulting in an almost perfect interpolation of the original data. A peak-finding algorithm was utilized, marking all peaks in the waveforms exceeding a peak-to-peak voltage greater than $\sqrt{2} \cdot 2\sigma$. ($\sqrt{2}$ since the variances of both peaks in the pair add up). Latency errors were derived from the curvature of the peaks and the amplitude errors (Hoth, 1986). The labeling of the identified peaks was done manually. The naming convention for the BD was adopted according to Ito et al. (1988). Except for ABR wave V, the amplitude of which was derived baseline-to-peak, all other amplitudes were evaluated as peak-to-peak values between the labeled waves (cf. Table 3.1). For the ABR, averages across the four channels were obtained for the peak amplitudes and latencies. For the MLR, amplitude and latency averages were derived separately for the two left (A1, PO9) and two right (A2, PO10) channels to account for possible hemispheric differences due to the larger spatial distance of the two primary auditory cortices in comparison to the brainstem structures. A Wilcoxon matched-pairs signed-rank test ($\alpha = 0.05$) was performed to ascertain whether response amplitudes and latencies differed significantly between stimulus directions.

3.2.6 Dipole source analysis

Multi-channel EEG-data was modeled by means of dipole source analysis. This is generally accomplished by minimizing a cost function that describes the difference between the measured and the modeled EEG. Equivalent current dipoles (ECDs) were used as a source model. These are focal sources representing the center of a small brain region in which many cells are synchronously activated (Scherg and Cramon, 1990; Scherg, 1991).

| Type of wave | | Latencies | Amplitudes |
|--------------|----|-----------|------------|
| ABR | B | V | V |
| | BD | DP1 | DP1–DN1 |
| | BD | DN1 | |
| MLR | B | N19 | P30–N19 |
| | B | P30 | |
| | BD | DN3 | DP4–DN3 |
| | BD | DP4 | |

Table 3.1: Overview of waves evaluated by latency and amplitude. *B* denotes a binaural potential, *BD* a binaural difference potential.

Electrode positions were averaged over measurements for each subject and fitted to a sphere by minimizing the least-square error. The radius of this sphere served as outer boundary r for the outer shell of a three-shell head model. This head model was employed with the following radius ranges s_1 , s_2 , s_3 and conductivities g_1 , g_2 , g_3 :

$$\begin{aligned}
 s_1 &= [0.0000 \quad 0.8977] \cdot r, & g_1 &= 0.33 \frac{1}{\Omega\text{m}} \\
 s_2 &= [0.8977 \quad 0.9659] \cdot r, & g_2 &= \frac{1}{80} \cdot g_1 \\
 s_3 &= [0.9659 \quad 1.0000] \cdot r, & g_3 &= g_1
 \end{aligned}$$

The first shell models the brain tissue, the second shell the bone, and the third shell the skull. For the subjects in this study, the radii r of the fitted sphere were in the range of 8.3 cm to 9.2 cm.

A rotating dipole was assumed to describe the binaural ABRs. In this model, the location is fitted but held constant for the investigated time interval. For each time sample, the three moment parameters (x -, y -, and z -components of the dipole vector) were optimized to best explain the data. The time interval for the fit started 1 ms prior to the latency of wave V and ended 2 ms after it. The latency of wave V was taken as the mean latency over channels A1, A2, PO9, and PO10.

For the binaural MLRs, two constrained fixed dipoles served as a source model. Location and orientation were fitted, but, in contrast to the rotating dipole model, remained fixed for the entire time interval investigated. Therefore, only the moment magnitude was fitted for each time sample. The pair of dipoles was constrained to account for hemispheric symmetry: The x -coordinates (left–right) of the locations

and the azimuth ϕ of the moments were constrained to be mirrored, i.e., had opposite signs. The y - and z -coordinates and the elevation θ were set to be the same for both dipoles. As additional constraint, the locations and orientations were required to be identical for all stimulus conditions. The time interval had a duration of 10 ms and was centered at the latency of wave P30. The latency of wave P30 was taken as the mean latency over channels A1, A2, PO9, and PO10.

3.3 Results

3.3.1 Waveform morphology

Fig. 3.4 illustrates the waveform morphology for the ABR (left panel) and MLR (right panel) in response to the diotic chirp. All curves represent data of channel PO10 for an exemplary subject (rh). The top three traces show the recorded monaural responses (L, R) and their derived sum (L+R). Drawn below are the recorded binaural response (B) and the calculated BD, i.e., the difference between B and L+R. The error bars indicate $\pm 3\sigma$ and are almost of equal size for all recorded responses conditions (L, R, B), as expected for a quasi-simultaneous recording. For derived potentials (L+R, BD), the error bars increase due to the summation/subtraction of the waveforms (as described in section 3.2.4). Triangles indicate relevant waveform peaks that were labeled for the evaluation of amplitudes and latencies. In comparison to click-stimulation, all latencies are shifted approximately by the duration of the chirp, i.e., 10 ms. This is due to the properties of the chirp, which synchronously activates the entire cochlea at chirp offset. A lowpass filtered version of wave V can be seen in the MLR data, but is not taken into account in the MLR evaluation. Contrary to that, the low-frequency MLR waves cannot be seen in the ABR-filtered data.

Fig. 3.5 shows the ABR data for all electrode channels distributed over the scalp. The monaural responses are not plotted here to enhance clarity. Note that the largest activity for all three traces can be observed at both mastoids (A1, A2) and the parieto-occipital channels (PO9, PO10).

3.3.2 Analysis of amplitudes and latencies

Results of the evaluation of amplitudes and latencies are shown in Figs. 3.6 to 3.9. All figures present the data in polar plots, i.e., the radius from the center codes the amplitude and latency values, respectively. The left panel displays data for the variation of azimuth φ , the right panel data for the variation of elevation ϑ . For a given peak, amplitude and latency data points of all virtual directions are connected with lines. The diotic reference condition is plotted without interconnections in both panels. The amplitudes and latencies for the relevant waves were averaged over the four channels A1, A2, PO9, and PO10. Finally, this channel average was

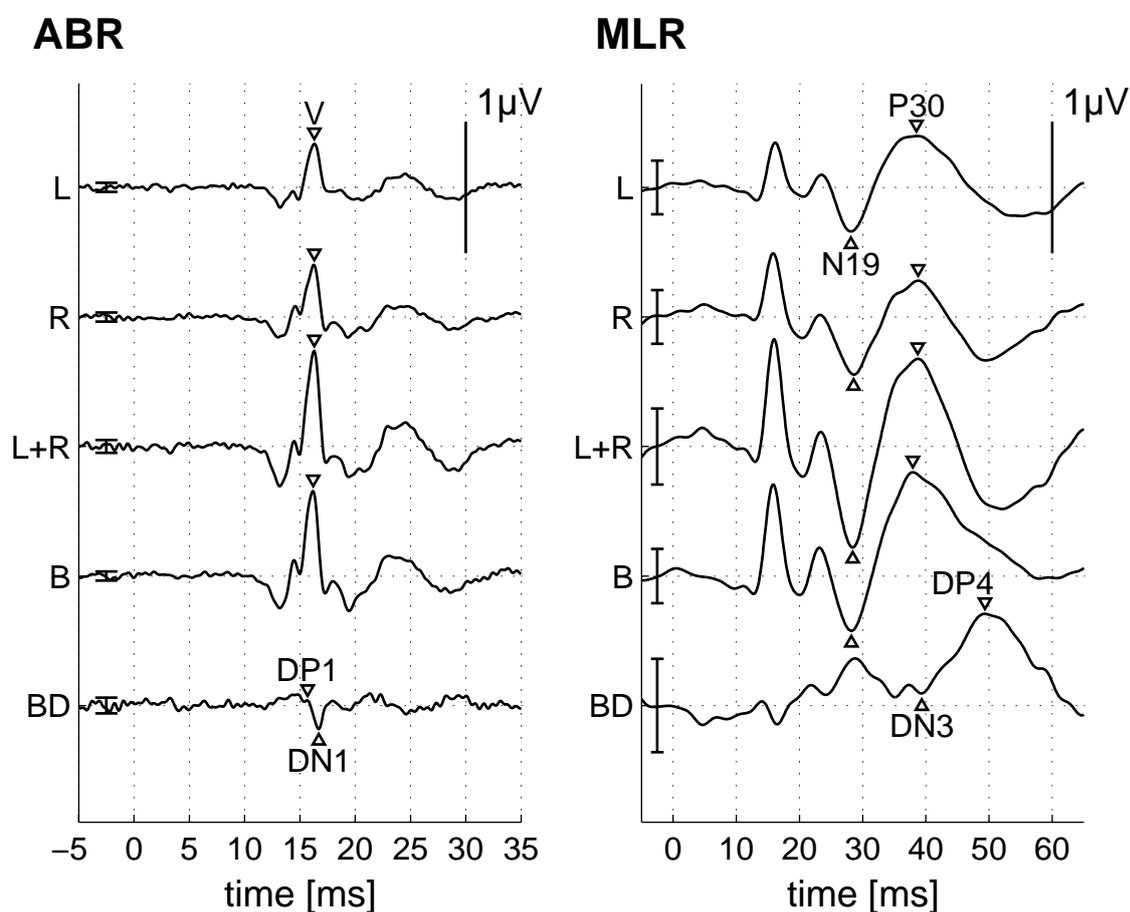


Figure 3.4: The binaural difference potential (BD) is calculated as the difference between binaural (B) and summed monaural (L+R) responses. Responses to the diotic (unconvolved) chirp, electrode position PO10 for subject rh. **Left panel:** ABR filter settings (100-1500 Hz, 200 taps FIR). Labels denote prominent ABR peaks. **Right panel:** MLR filter settings (20-300 Hz, 200 taps FIR). Labels denote prominent MLR peaks. Error bars indicate $\pm 3\sigma$ in all traces. Note that due to the properties of the chirp, all latencies are about 10 ms longer than for click stimulation.

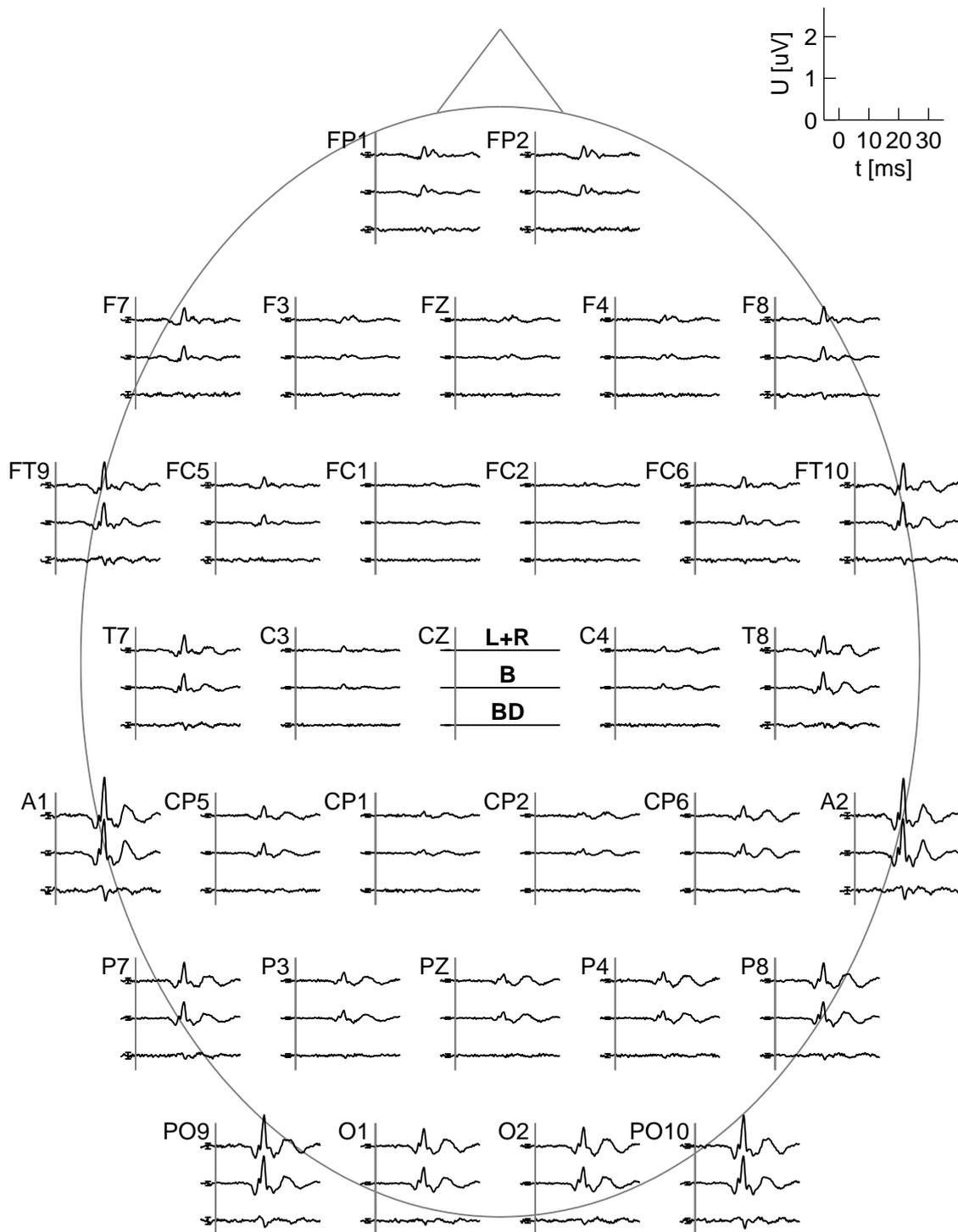


Figure 3.5: Scalp distribution of the ABR data to the diotic chirp for subject *rh*. **Top traces:** Summed monaural responses ($L+R$). **Middle traces:** Binaural potentials (B). **Bottom traces:** Binaural difference potentials (BD). Channel Cz represents the reference electrode, with zero voltage for all time instances. Error bars indicate $\pm 3\sigma$ in all traces.

taken to calculate the grand mean over subjects. Peaks that were not significant for any subject or condition were not included in the average. Error bars indicate the interindividual standard deviations and were only plotted to one side for better visibility.

Fig. 3.6 presents the ABR amplitudes for wave V of the binaural potentials and the amplitude difference DP1–DN1 for the BDs. For the variation of azimuth, the amplitudes of both wave V and DP1–DN1 decrease monotonically with increasing laterality (i.e., with increasing distance from the midline). While the laterality is clearly represented in the amplitude data, the front-back position of the stimulus is not: A comparison of the frontal stimuli with the corresponding back stimuli of the same laterality (e.g., $\varphi = 45^\circ$ versus $\varphi = 135^\circ$) yields no difference in amplitude. The amplitudes of the diotic condition are not significantly larger than the frontal and back condition, as revealed by the Wilcoxon test. Amplitude data do not exhibit a significant dependence on stimulus elevation.

The latency data for the ABR is plotted in Fig. 3.7 for waves V, DP1, and DN1. For the variation of the azimuth, only the latencies of the BDs increase monotonically with the laterality of the stimulus. The latency of wave V remains constant over all conditions. The diotic reference condition has significantly shorter latencies than the other conditions for all waves ($p < 0.01$). As the amplitudes, latency data do not exhibit a significant dependence on stimulus elevation.

Fig. 3.8 displays the MLR amplitudes of wave P30–N19 of the binaural potentials in the top row and the amplitudes of DP4–DN3 for the BDs in the bottom row. In contrast to the ABR data, the channel average was obtained separately for the two left (A1, PO9) and right channels (A2, PO10). For wave P30–N19 (but not DP4–DN3), the amplitudes of the right channels are higher than those for the left channels for all stimulus conditions, but this difference is not significant. In contrast to the ABR amplitudes, there is no pronounced dependence on the laterality of the stimulus direction, neither in the binaural potentials nor in the BDs.

The latency data for the MLR binaural potentials and BDs are shown in Fig. 3.9. For all waves, differences between left and right channels are not significant. Latencies of waves N19 (top row) and P30 (second row) do not vary with the direction of the stimulus. Waves DN3 (third row) and DP4 (bottom row) show prolonged latencies for lateralized stimuli. Varying the elevation of the stimulus does not affect the latencies systematically. Latencies between left and right channels differ slightly

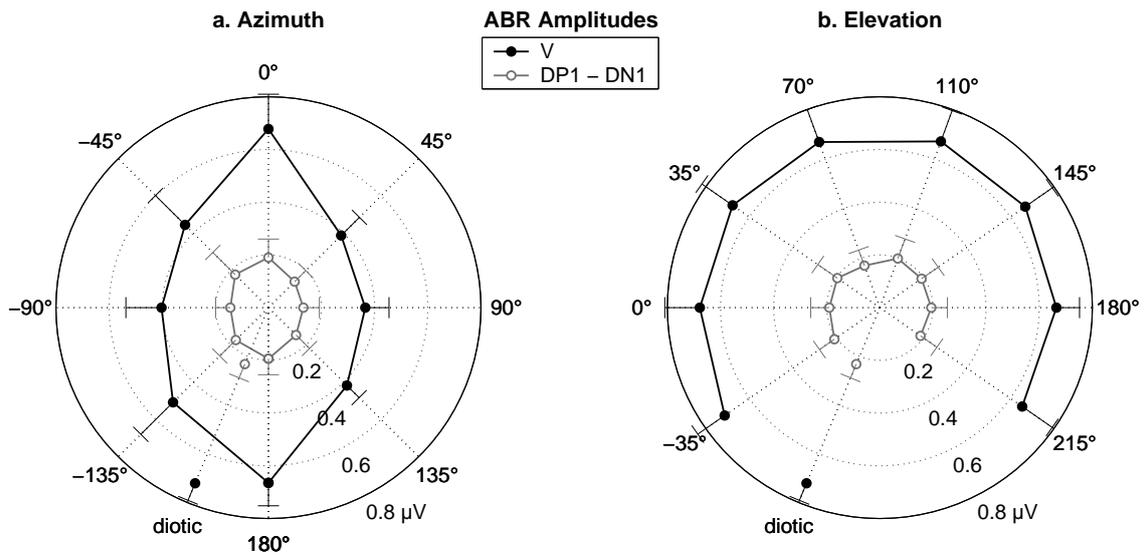


Figure 3.6: ABR amplitudes for waves V and $DP1-DN1$, averaged over four channels and six subjects. **Left panel:** Dependence on azimuth φ . **Right panel:** Dependence on elevation ϑ . The diotic reference condition is plotted in both panels. Error bars indicate the interindividual standard deviations. For both wave V and $DP1-DN1$, the amplitudes decrease with increasing laterality of the stimulus. Changing the elevation of the sound source does not influence the ABR amplitudes.

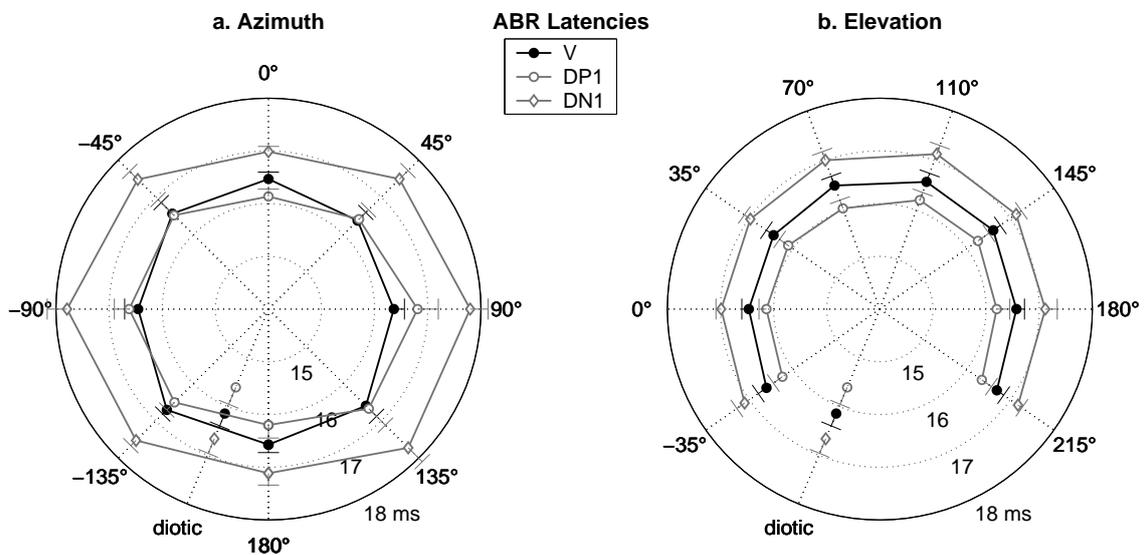


Figure 3.7: ABR latencies for waves V , $DP1$, and $DN1$, averaged over four channels and six subjects. **Left panel:** Dependence on azimuth φ . **Right panel:** Dependence on elevation ϑ . The diotic reference condition is plotted in both panels. Error bars indicate the interindividual standard errors. While wave- V latency remains about constant for all directions, the latency for $DP1$ and $DN1$ increases monotonically with increasing laterality. There is no dependence on elevation of the sound source.

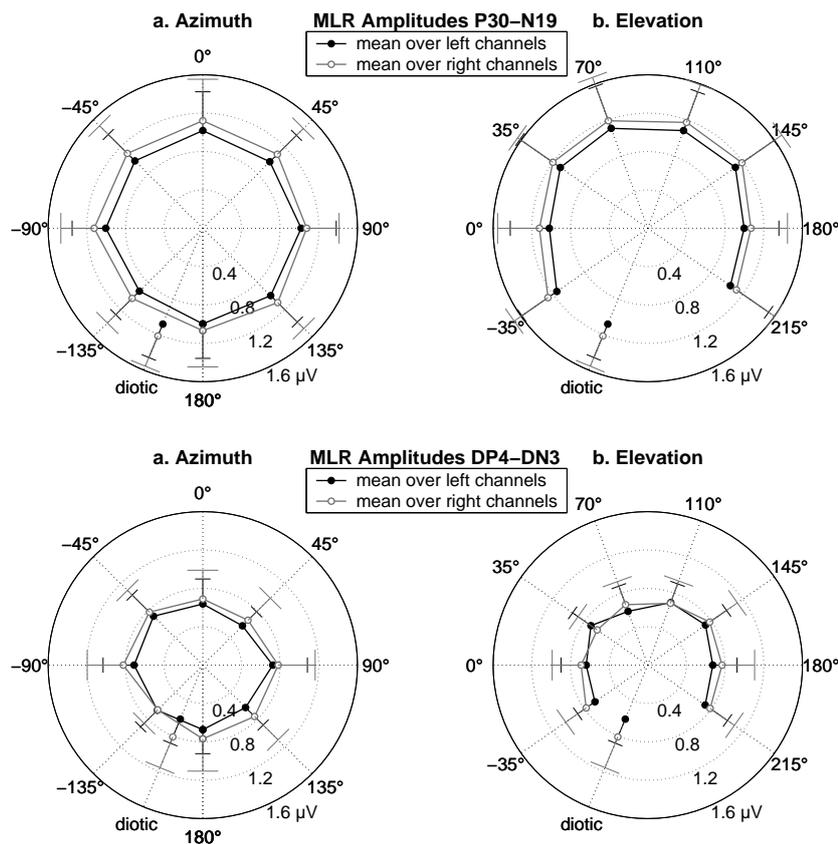


Figure 3.8: MLR amplitudes for waves P30–N19 (top row) and DP4–DN3 (bottom row), mean over six subjects. **Left panels:** Dependence on azimuth φ . **Right panels:** Dependence on elevation ϑ . Channel averages were obtained separately for two left (A1, PO9; black solid circles) and two right channels (A2, PO10; gray open circles). The diotic reference condition is plotted in both panels. Error bars indicate the interindividual standard deviations.

for waves DN3 and DP4, but these differences show no systematic dependence on stimulus direction.

3.3.3 Dipole source analysis

Figures 3.10 to 3.12 show the results for the dipole source analysis as the mean over all individual fits. The left panel represents data for the dependence on azimuth, the right panel for the dependence on elevation.

Locations of the fitted rotating dipoles for the binaural ABR are plotted in Fig. 3.10 for all three planes. x points to the right, y to the front and z to the top. The y - z -plane is plotted at the top left, the x - z -plane at the top right, and the x - y -plane at the bottom right for each panel. The error-ellipse indicates the 95% confidence

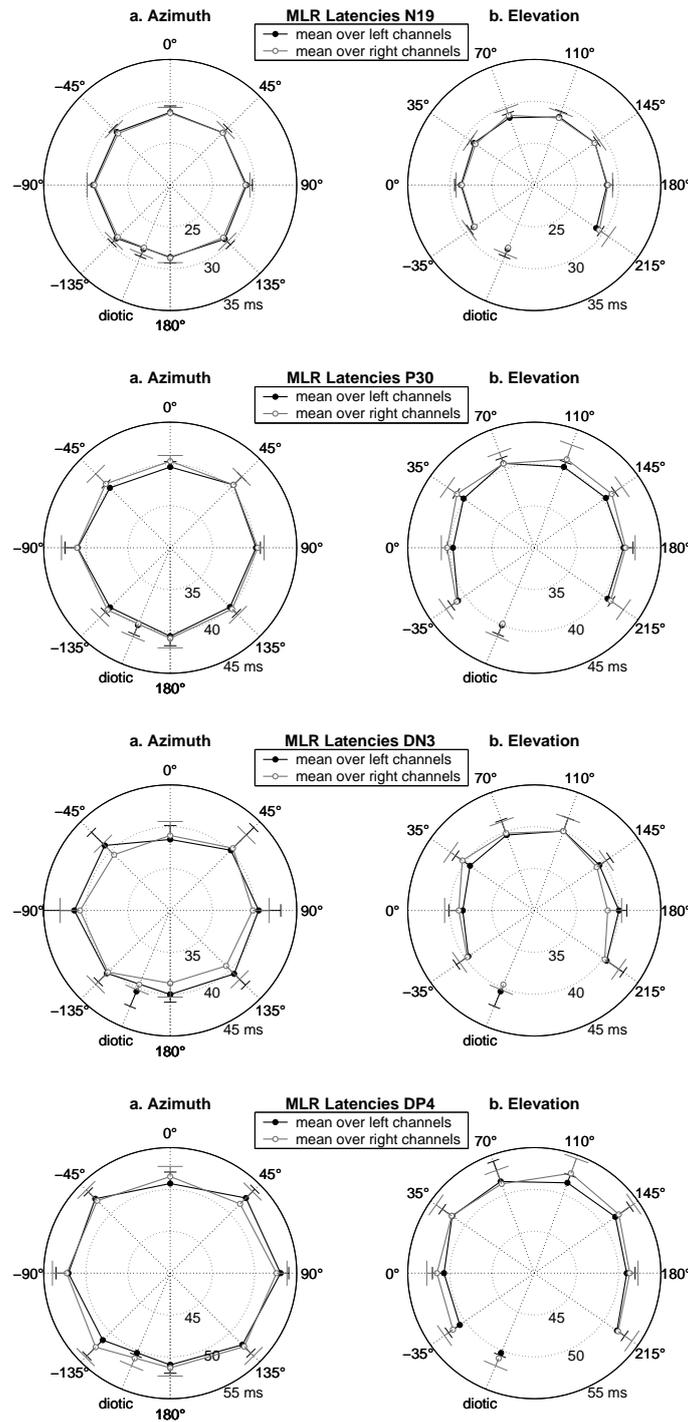


Figure 3.9: MLR latencies for waves N19 (top row), P30 (second row), DN3 (third row), and DP4 (bottom row), mean over six subjects. **Left panel:** Dependence on azimuth φ . **Right panel:** Dependence on elevation ϑ . Channel averages were obtained separately for two left (A1, PO9; black solid circles) and two right channels (A2, PO10; gray open circles). The diotic reference condition is plotted in both panels. Error bars indicate the interindividual standard deviation.

region of the dipole location fitted to the response to the diotic stimulus. The stimulus condition is coded by different markers in the plot (cf. the legend at the bottom left of each panel). All dipoles lie in a volume element of size $1.5 \times 3 \times 3 \text{ mm}^3$, which is smaller than the 95% confidence region of the fits. No systematic variation of location with stimulus direction can be observed. The brain region activated lies on average about 2.5 cm below the center of the head sphere.

Fig 3.11 illustrates the time course of the rotating dipole by means of moment trajectories for the binaural ABR. The top two graphs show the x - z -plane, the bottom two graphs the data in the y - z -plane. Each of the 3×3 subplots represents a stimulus condition, indicated by the text in the upper right corner. Note that the scaling of the x -axis is augmented for better conspicuity. For all stimulus directions dipole moments are largest in z -direction at the latency of wave V, indicating that most of the energy constituting wave V is due to an upward current. Furthermore, the dipole is tilted towards the front (y -direction). The time courses of the moments in x - and z -direction vary strongly with the laterality of the sound source: For

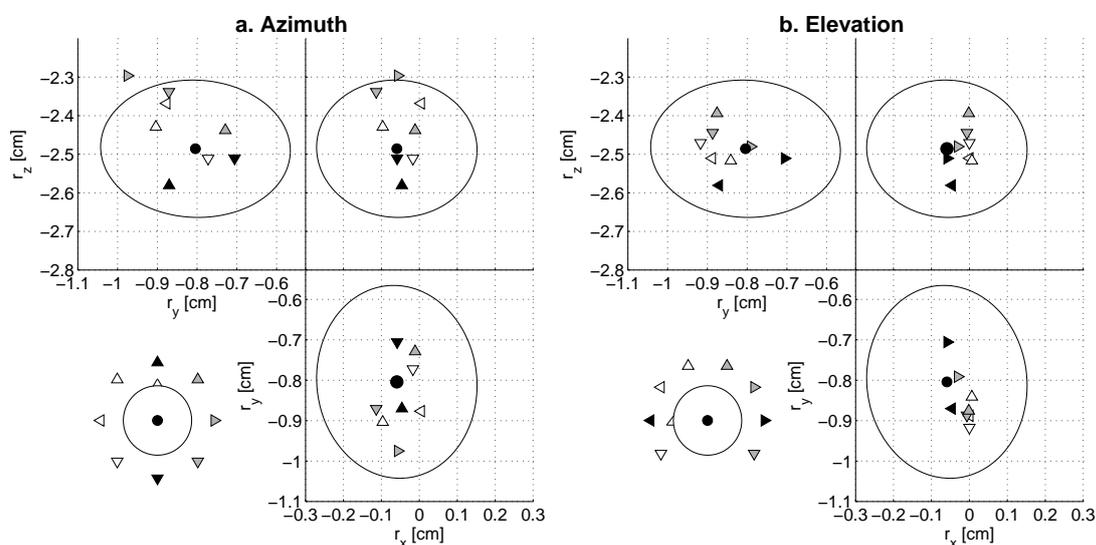


Figure 3.10: ABR dipole locations: Source analysis of binaural potentials, mean over subjects. For all stimulus conditions (**Left panel:** Dependence on azimuth, **right panel:** Dependence on elevation), the locations of a fitted rotating dipole in the three planes are shown. x points to the right, y to the front, and z to the top. The fit interval started 1 ms before to wave V and ended 2 ms after wave V. Error-ellipses indicate the 95% confidence region of the diotic stimulus. For all stimulus conditions the same brainstem region is activated. No systematic variation of location with stimulus direction can be observed.

directions from the right, the trajectory starts with negative horizontal moments and rotates clockwise over time to positive horizontal moments. For directions from the left, the trajectory behaves the opposite way, i.e., rotates counter-clockwise. The moment in z -direction is significantly lower for the left/right directions than for the directions in the midsagittal plane. There is no dependence on the elevation of the

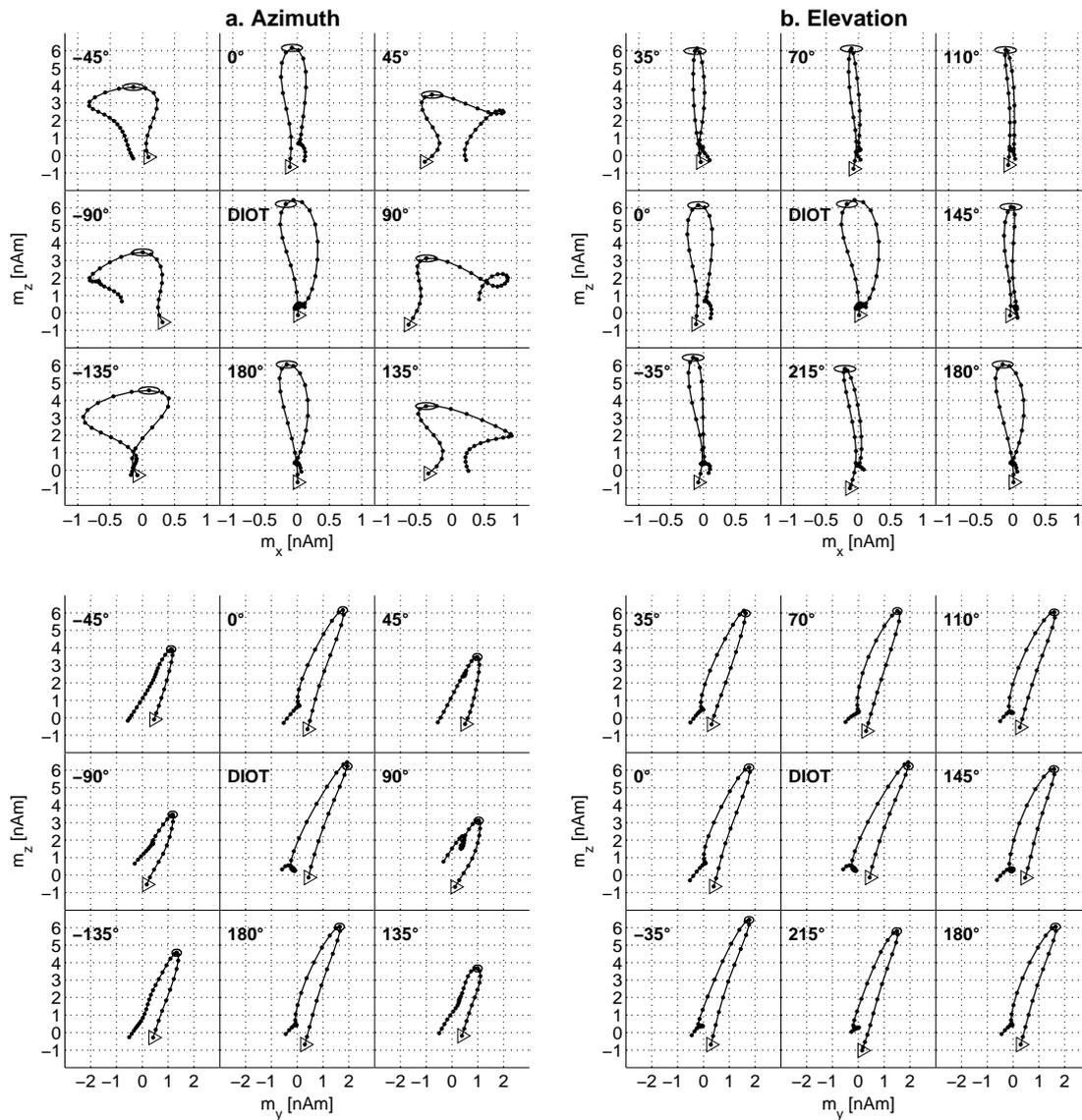


Figure 3.11: ABR dipole moments: Moment trajectories of a rotating dipole in the x - z -plane (top two panels) and y - z -plane (bottom two panels). x points to the right, y to the front, z to the top. Binaural potentials, mean over subjects. **Left panels:** Variation of azimuth. **Right panels:** Variation of elevation. Each subplot represents one stimulus direction, indicated by the text in the top right corner. Trajectories start 1 ms before wave V at the point marked with a triangle. At the latency of wave V, error ellipses indicate 95% confidence regions.

sound source.

For the binaural MLR data, the fitted locations and orientations for the two dipoles (constrained to be equal for all conditions, cf. section 3.2.6) were as follows:

| | x | y | z | ϕ | θ |
|---------------|---------|---------|---------|--------|----------|
| Left dipole: | -3.5 cm | +0.5 cm | +1.5 cm | -45° | +60° |
| Right dipole: | +3.5 cm | +0.5 cm | +1.5 cm | +45° | +60° |

x , y , and z denote the location in cartesian space, ϕ the azimuthal angle ($\phi = 0^\circ$ corresponds to the midline), θ the inclination from the horizontal plane. With the constraints used for the fitting procedure, the two dipoles could well be separated and lie close to where the primary auditory cortices are expected bilaterally in the temporal lobes.

Fig. 3.12 illustrates the time courses of the dipole moment magnitudes for the left (black lines) and right (gray lines) dipoles. Dipoles were averaged across all subjects. For the variation of azimuth (left panel), the time courses are rather similar for the central (front, back, and diotic) stimuli. The right stimuli show a stronger activation of the left dipole compared to the right dipole. Conversely, the left stimuli lead to a much stronger activation of the right dipole compared to the left dipole. For the variation of elevation, the left dipole shows about the same activation for all conditions except the diotic. The right dipole is more strongly activated than the left for the six stimulus conditions with elevational cues (i.e., all conditions but the front, back, and diotic).

For further analysis, the maxima of the individual MLR dipole magnitudes were extracted and averaged across subjects. The results are presented in Fig. 3.13 in a polar plot. For the variation of azimuth (left panel), the maximal dipole magnitudes increase with growing stimulus lateralization to the contralateral side. This increase is higher for the right dipole than for the left one. The maximal magnitudes remain constant for both dipoles when varying the elevation of the stimulus, with two exceptions: First, the left dipole exhibits a significantly higher magnitude in the diotic condition compared to the remaining conditions ($p < 0.05$). Secondly, for the right dipole, the frontal ($\vartheta = 0^\circ$) and rear ($\vartheta = 180^\circ$) stimuli result in lower maximal magnitudes than the other conditions ($p < 0.01$). The right dipole shows significantly higher magnitudes than the left dipole for stimulus conditions off the horizontal plane ($p < 0.07$). An analysis of individual data reveals that in three out

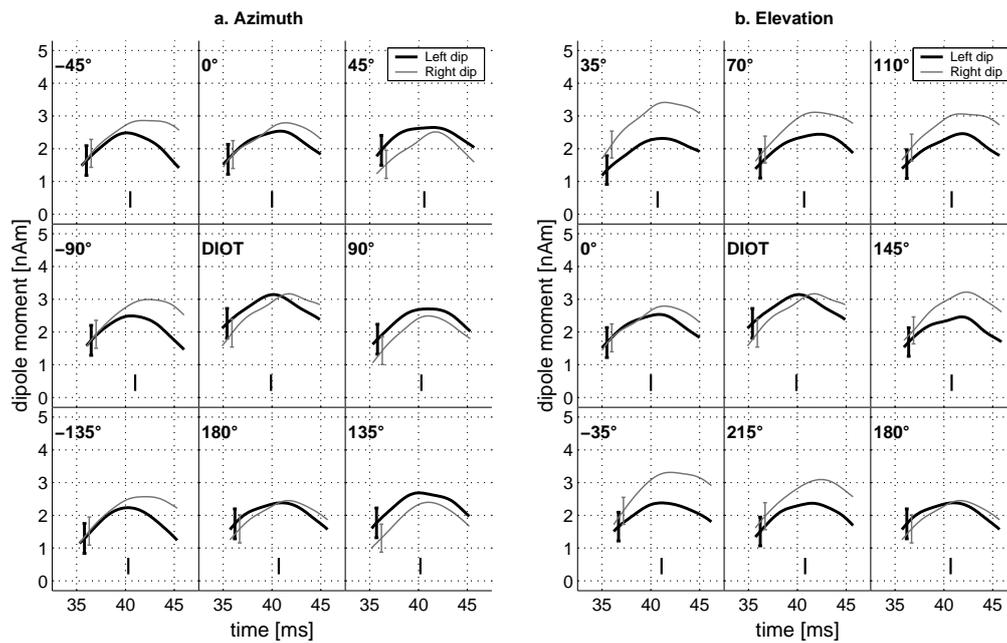


Figure 3.12: MLR dipole moment magnitudes of two constrained fixed dipoles, mean over subjects. **Left panel:** Variation of azimuth. **Right panel:** Variation of elevation. Each subplot represents one stimulus direction, indicated by the text in the top right corner. Black thick lines show the dipole in the left hemisphere, gray thin lines the one in the right hemisphere. Vertical bars denote the mean latency of wave P30, error bars the 95% confidence regions. The time interval starts 5 ms before wave P30 and ends 5 ms after it.

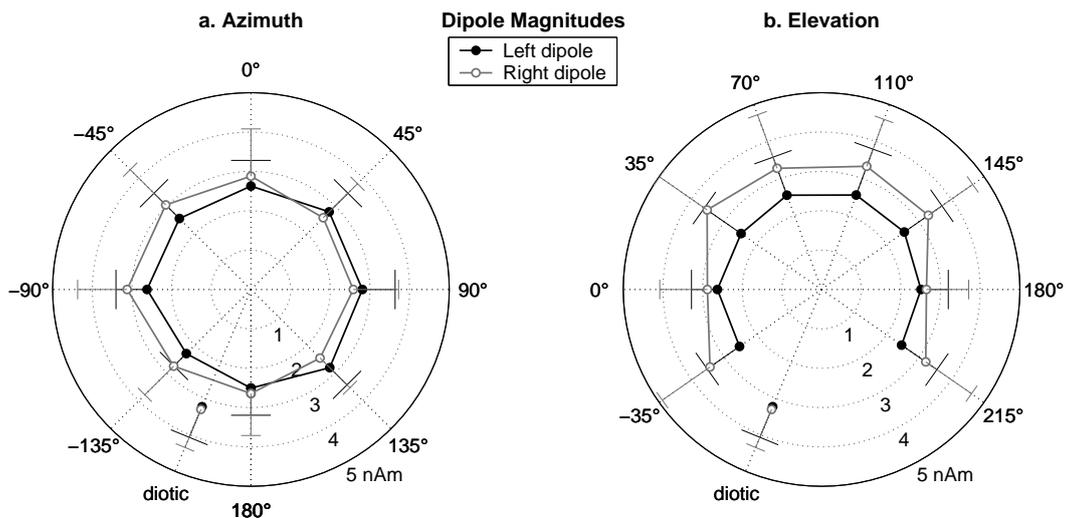


Figure 3.13: Maximal magnitudes of the dipole moments for the left (black solid circles) and right (gray open circles) dipole. Mean over subjects. **Left panel:** Variation of azimuth. **Right panel:** Variation of elevation. Error bars denote the interindividual standard deviations.

of six subjects the dipole in the right hemisphere shows higher magnitudes than the left dipole for the stimulus conditions off the midsagittal plane. The other three subjects showed no such hemispheric bias.

3.4 Discussion

Recordings of ABRs and MLRs were performed for 14 virtual sound directions and a diotic reference condition. The virtual acoustics paradigm resulted in externalized sound objects with azimuth and elevation cues. Single epochs were recorded and analyzed offline allowing for an improved filtering and averaging technique as well as for an estimation of the residual noise on a single-sweep basis (Granzow et al., 2001; Riedel et al., 2001). Artifact sources, such as acoustic crosstalk or middle ear reflex, were ruled out by use of randomized stimulation, moderate stimulation levels and insert earphones (Levine, 1981).

Amplitudes of both binaural potentials and BDs were not higher for the externalized stimuli than for the diotic reference and therefore did not benefit from incorporating realistic spatial cues. The hypothesis of a specific ‘externalization detector’, producing higher neural activity for realistic combinations of ITD, ILD, and spectral cues across frequency, has to be rejected. Instead, the mere interaural properties (ITD and ILD) were reflected by the responses. ABRs and MLRs showed a systematic dependence on the azimuthal position of the sound source. For the ABRs, this effect was observable in both single channel analysis and dipole source analysis: Increasing laterality led to decreasing amplitudes and increasing latencies in the binaural potentials and BDs. For the rotating dipole model, the moment trajectories revealed characteristic patterns depending on extent and direction of laterality.

The single channel analysis of the MLR data exhibited longer latencies for lateralized conditions in the BDs. The encoding of sound azimuth could more prominently be observed in the evaluation of the dipole moments, underlining the general information gain obtained by source analysis compared to the single channel analysis alone. Lateralized sounds resulted in a stronger activation of the contralateral dipole for wave P30. This contralateral dominance is also supported by neurophysiological data: In cat auditory cortex the majority of spatially selective neurons were found to be contralaterally-selective, with fewer ipsilaterally- and medially-selective cells (e.g., Reale and Brugge, 1990).

MLR dipole sources for wave P30 were situated bilaterally in the temporal lobes close to the expected location of primary auditory cortices (Leonard et al., 1998; Patterson et al., 2002). However, due to the lack of anatomical data, as derived by neuro-imaging methods such as magnetic resonance imaging (MRI), no absolute coordinates for the dipole locations could be depicted. A constrained dipole fit for wave N19 was also tested, but in contrast to the fit of wave P30 the results did not exhibit stable solutions for all subjects.

For stimulus conditions off the midsagittal plane, half of the subjects revealed a right hemispheric dominance. Most studies reporting on hemispheric dominance focus on more central stages of the auditory pathway by means of magnetoencephalography (MEG) (McEvoy et al., 1993; Sams et al., 1993; Palomäki et al., 2005). While some authors find a stronger activation in the right hemisphere as well (e.g., Palomäki et al., 2005), others report a lack of inter-hemispheric differences (McEvoy et al., 1993; Sams et al., 1993). More studies are necessary to further elucidate the issue of hemispheric dominance.

Further systematic influences, e.g., elevational cues or front/back-differences, could only be observed for the MLR data: The P30 dipole in the right hemisphere exhibited larger maximal magnitudes for the stimulus conditions off the horizontal plane, i.e., with elevational cues (cf. Fig. 3.13). However, data for these six conditions were acquired in separate recording sessions (cf. section 3.2.3). In contrast to ABRs, the MLRs are dependent on the subjects' vigilance. Therefore, the divergence in the data may have to be attributed to a different state of vigilance for these separate recording sessions with the six elevational conditions.

The relatively small number of six subjects was the 'price to be paid' for using individual HRIRs. HRIRs had originally been recorded for twelve subjects, six of which left the area of Oldenburg by the time this study was completed.

The results of this study were, at least qualitatively, in accordance to the numerous studies investigating the influence of ITD and ILD alone (Wrege and Starr, 1981; Gerull and Mrowinski, 1984; Furst et al., 1985; Jones and Van der Poel, 1990; McPherson and Starr, 1995; Brantberg et al., 1999; Riedel and Kollmeier, 2002a, 2003). A lateral acoustic object, as invoked by either ITD alone, ILD alone, or a combination of both with or without spectral cues, results in a substantial decrease of wave-V amplitude and increase of wave-V latency. In addition, DP1–DN1 amplitudes decrease and DN1 latencies increase with the laterality of the sound source.

Furthermore, this decrease in amplitudes and increase of latencies is monotonic with the degree of laterality. The rotating dipole model exhibits characteristic trajectories (i.e., development of dipole moments over time) depending on the hemifield and extent of lateralization. [Riedel and Kollmeier \(2003\)](#) found similar dipole trajectories for their synergistic conditions and with either ITD or ILD alone. As in this study, they found no systematic dependence of dipole location on the binaural conditions, which reflects the poor spatial resolution for brainstem sources obtained by EEG recordings.

[McPherson and Starr \(1995\)](#) investigated the relationship of MLR BDs on either ITD or ILD alone (i.e., not in conjunction). Opposed to the results in this work, the amplitudes of their waves DN3 and DP4 (termed D30 and D40 in their work) decreased monotonically with increasing ITD or ILD, while the amplitude difference DP4–DN3 remained constant in the results shown here. In addition, while in the results presented here both DN3 and DP4 showed increased latencies for lateralized stimuli, [McPherson and Starr \(1995\)](#) reported increasing latencies only for their DN3 as ITD was increased. The discrepancy between both studies may be due to different methodologies, especially concerning the validity of peak evaluation in the averaged AEP data. E.g., [McPherson and Starr \(1995\)](#) obtained averages consisting of only 2000 sweeps (rather than 10,000 sweeps here) and did not estimate the residual noise. In addition, they measured the amplitudes baseline-to-peak rather than peak-to-peak, which might also systematically affect the potentials, especially because they chose comparatively ‘open’ filter settings.

Two other studies utilize the spectral cues necessary to invoke a realistic, externalized sound perception during the recording of ABR ([Polyakov and Pratt, 2003b](#)) and MLR ([Polyakov and Pratt, 2003c](#)). Both studies use non-individualized HRIRs from an artificial head ([Gardner and Martin, 1995](#)) rather than individual HRIRs that were employed here. Moreover, no dipole source analysis based on multi-channel data and a head-model was performed, but equivalent dipole trajectories were estimated in voltage space (three-channel Lissajous’ trajectory, e.g., [Pratt et al., 1983](#); [Jewett et al., 1987](#)) only for the BD waveforms. Source analysis of BD data was also performed in this study, but the fit results exhibited large confidence intervals and showed no systematic effects with stimulus condition. Therefore, only the single channel analysis of BD amplitudes and latencies allows a comparison between studies. In their brainstem study ([Polyakov and Pratt, 2003b](#)), similar effects of stimulus

laterality on the latency of the BD component DN1 (termed Be_I in their study) were found. The diotic condition (without spatial cues) exhibited the shortest latencies when compared to all spatial conditions, which is also consistent with the results presented here. However, [Polyakov and Pratt \(2003b\)](#) also reported some latency effects when varying the elevation of the stimulus, which could not be found here in the single channel analysis. In addition, [Polyakov and Pratt \(2003b\)](#) found asymmetries in dipole magnitude for left and right positioned stimuli. This asymmetry is in opposition to the results found here and to other studies reporting on ITD/ILD effects on BD magnitude.

In their follow-up study, [Polyakov and Pratt \(2003c\)](#) investigated MLR recordings in dependence on the spatial position of the sound source. As in their preceding study, they analyzed the magnitudes and latencies of one equivalent dipole based on the three-channel Lissajous' trajectories, although a single dipole generator is clearly inadequate to model MLR sources. In contrast to the results reported in this study, [Polyakov and Pratt \(2003c\)](#) did not find significantly prolonged BD dipole latencies for their lateralized stimuli. Further significant effects in their study included a decrease of dipole magnitude for the frontal (only DP4 component) and the backward-positioned stimuli (DN3 and DP4 components) when compared to the condition with the stimulus positioned at the top. Although in the present work dipole analysis was only successful for binaural responses (and not BDs), here the dipole magnitudes for the front and back stimuli exhibited smaller magnitudes than *all* stimulus conditions with elevational cues, but only for the right dipole. Furthermore, the MLR latencies of the BDs showed a symmetric dependence on laterality, which is in opposition to the results from [Polyakov and Pratt \(2003c\)](#). Some methodological differences between studies might account for these diverging results: First, using click stimuli (with much shorter duration than the chirp used here) and non-individualized HRIRs might provide a weaker or less reliable externalization than in this study, especially on the midsagittal plane (see [Hofman and Van Opstal, 1998](#), for the effect of stimulus duration on the elevation localization performance). Secondly, [Polyakov and Pratt \(2003b,c\)](#) might not have eliminated the absolute travel-time delays for all HRIRs, which is crucial to derive exact *relative* latency effects between stimulus conditions. Third, [Polyakov and Pratt \(2003b,c\)](#) did not provide an estimate of the residual noise in their EEG data, which might have led to some misidentified peaks in the waveforms. Apart from that, the dif-

ferent modeling of dipolar sources might account for some discrepancies. Finally, although Polyakov and Pratt (2003b,c) used fifteen rather than six subjects, the interindividual standard deviations in the latency and amplitude analysis were not lower than in this study. Further research is necessary to reliably find out whether the differences between studies are a consequence of some of these differing measurement procedures or due to the statistical measurement error.

In summary, the results of this multi-channel AEP study indicate that the use of realistic, externalized stimuli has no advantageous influence on the responses when compared to a diotic reference condition. Mostly the degree of stimulus laterality, as produced by ITD and ILD, is encoded in the neural generators forming the ABR and MLR. The hypothesis for an ‘externalization detector’ at these early stages of the auditory pathway must be rejected.

Chapter 4

Influence of spatial position on chirp-evoked potentials: Test of the localization detector hypothesis at central auditory processing stages

Abstract

In this study, the role of realistic spatial auditory stimuli for the formation of late auditory evoked potentials (LAEPs) was examined. The hypothesis of a ‘localization detector’, providing distinct neural activation for natural combinations of ITD, ILD, and spectral cues over frequency, was tested. An optimal chirp stimulus, designed to activate the entire cochlear partition synchronously, was convolved with individual head-related impulse responses (HRIRs). Stimulus conditions were chosen to be six locations either on the horizontal plane (variation of azimuth) or midsagittal plane (variation of elevation) or both (front, back). The six stimulus directions were presented in an oddball paradigm in which the frontal stimulus served as a standard, i.e., was presented most of the time ($p = 85\%$). LAEPs and difference curves (DCs, with respect to the standard) were evaluated. For LAEPs and DCs, amplitudes and latencies were dependent on the laterality of the sound source in

the frontal horizontal plane. The DCs exhibited a mismatch negativity (MMN) component which was most prominent for the frontal right and right stimulus direction. Dipole source analysis for the LAEP component P2 and DC component M1 showed a clear dependence on laterality of the sound source. No hemispheric bias was found for conditions on the midsagittal plane and no differences between conditions along the midsagittal plane were found. Topographic voltage maps showed frontal MMN activity for the frontal right and right stimulus. Results indicate that across-frequency integration of ITD, ILD, and spectral cues, as assumed by the hypothesis of a frequency-integrating 'localization detector', does not occur at the level of the sources generating the P2 and MMN components in humans, but might be processed at more central stages of the auditory pathway.

4.1 Introduction

The auditory system, unlike the visual system, does not directly map the outer world onto the sensory receptors. Instead, three-dimensional auditory space is computed from three acoustic cues: The interaural differences in sound arrival time (interaural time differences, ITDs) and sound level (interaural level differences, ILDs) are used to code the left-right localization (e.g., [Blauert, 1997](#)). For vertical localization, spectral cues, resulting from sound reflections on the head, torso, and pinnae must be utilized. This direction-dependent filtering also helps to resolve localization ambiguities for sound directions resulting in identical ITDs and ILDs (commonly referred to as ‘cone of confusion’, e.g., [Woodworth, 1938](#); [Mills, 1972](#)).

Presenting auditory stimuli over headphones with missing spectral cues (i.e., with ITD and ILD alone), results in a sound percept inside the listener’s head, typically along an axis from one ear to the other. This perception is often referred to as ‘lateralization’, a somewhat degraded form of real localization. In order to invoke a true representation of three-dimensional space via headphones, the direction-dependent spectral filtering must be incorporated in the stimulus presentation. This is typically accomplished by the method of ‘virtual acoustics’: Head-related impulse responses (HRIRs) are measured with microphones in the subject’s or artificial head’s ear canals for all relevant directions (and distances). These recorded HRIRs contain the ‘natural’ combination of all acoustic cues necessary to localize the position of a sound in space (i.e., ITD, ILD and spectral filtering). When presenting a stimulus convolved with a pair of HRIRs to both ears over headphones, a percept similar to a free-field condition can be obtained (e.g., [Mehrgardt and Mellert, 1977](#); [Genuit, 1984](#); [Middlebrooks and Green, 1991](#); [Wightman and Kistler, 2005](#)). Moreover, in contrast to free-field acoustics, the virtual acoustics paradigm is easy to implement, independent of the listening environment and subjects’ movement, and highly reproducible.

Auditory evoked potentials (AEPs) play a key role when investigating the neural processes underlying sound perception in humans. They can be recorded from all levels of the auditory system and are usually grouped by their latency, i.e., the time of occurrence after the onset of the stimulus. This grouping corresponds roughly to the site of generation (e.g., [Picton et al., 1974](#); [Scherg, 1991](#)). AEPs represent the summation of responses from many neurons, recorded from the far-field, i.e., from

electrodes placed on the surface of the head (e.g., [Jewett et al., 1970](#)).

As pointed out in chapter 3, auditory brainstem responses (ABRs) and middle latency responses (MLRs) are highly affected by sound azimuth, but only little by sound elevation. As a consequence, utilizing virtual acoustics showed no advantage over traditional ITD/ILD stimulation, since differences in the spectral cues could not be resolved in the ABR and MLR. Therefore, these more peripheral processing stages might represent an ITD/ILD map, rather than a complete spatial map that incorporates spectral cues, hence integrating spectral and temporal localization cues across a broad frequency range. However, such an integrating processing stage, or ‘localization detector’, must reside somewhere along the auditory pathway. The objective of this study was to find evidence for such a ‘localization detector’ at higher levels of the auditory pathway, by looking for effects of realistic externalized stimuli on long-latency AEPs.

Late auditory evoked potentials (LAEPs) are generated in the latency range above about 50 ms. The most prominent peaks are wave N1 at a latency of about 100 ms and wave P2 at a latency of roughly 150–200 ms. The generation sites are the primary and secondary auditory cortices, situated bilaterally in the temporal lobes ([Scherg and Cramon, 1985](#); [Scherg et al., 1989](#)).

The mismatch negativity AEP (MMN) was suggested as a powerful tool for studying various aspects of central auditory information processing, such as pre-attentive auditory sensory memory (e.g., [King et al., 1995](#); [Näätänen and Alho, 1995](#); [Ritter et al., 1995](#)). The MMN is elicited by infrequent changes (‘deviants’) in a regular series of identical stimuli (‘standards’). Occurrence and amplitude of the MMN directly depend on the discriminability of the deviants in the series: The more salient the deviant, the higher the amplitude of the MMN. Since the MMN is even evoked when the acoustic stimuli are task-irrelevant (i.e., not attended), it seems to reflect an automatic, pre-attentive brain response to stimulus deviance. The MMN component is thought to be produced by a process comparing the neural trace of any incoming auditory stimulus with the established memory trace of the standard stimulus ([Näätänen, 1985](#)). If a deviant stimulus is detected, a neural mismatch process is initiated which generates the MMN.

LAEPs and MMNs (respectively their equivalents in magnetoencephalographic studies) have been studied quite extensively to elucidate the processing of auditory directional information ([Paavilainen et al., 1989](#); [McEvoy et al., 1991](#); [Sams et al.,](#)

1993; McEvoy et al., 1993; Schröger, 1996; Schröger and Eimer, 1996; Schröger and Wolff, 1996; Teder-Sälejärvi and Hillyard, 1998). Most studies utilized stimulus presentation via earphones with ITDs and ILDs or a combination of both.

Recent studies have shown that a rising chirp stimulus, designed to activate all cochlear locations simultaneously, evokes a larger ABR wave V amplitude than a click presented at the same sensation level (Dau et al., 2000; Wegner and Dau, 2002; Fobel and Dau, 2004). This advantage of the chirp over the click has been verified also for binaural ABRs and binaural difference potentials, at least for low and medium presentation levels (Riedel and Kollmeier, 2002b). Magnetoencephalographic studies have shown that the chirp stimulus can also increase the amplitudes of the middle latency response (MLR), which is generated at the level of the primary auditory cortex (Rupp et al., 2002). This demonstrates that phase delays between channels in the auditory pathway are preserved at least up to the primary auditory cortex. So far there is no proof that the chirp stimulus is also advantageous to the click for LAEPs and MMN. Yet, the chirp will be used in this study for sake of consistency to chapter 3, which utilizes the chirp for ABR and MLR responses.

The aim of this study is to investigate the dependence of binaural cortical potentials on the direction of the sound source. A virtual acoustics paradigm with individual HRIRs is incorporated to permit for varying both azimuth and elevation of the sound source. A 32-channel electrode setup allows for modeling of dipolar sources. The dependence of the single channel waveform data and equivalent dipoles on both sound azimuth and elevation is examined to test for the ‘localization detector hypothesis’.

4.2 Methods

4.2.1 Subjects

Five normal-hearing subjects (one female, four male), aged from 30 to 40, volunteered in the experiments. Subjects had no history of audiological or neurological problems and their audiometric thresholds were 10 dB HL or better for frequencies from 500 Hz to 4000 Hz, and 20 dB HL or better for 8000 Hz.

4.2.2 Stimuli

Stimuli were digitally generated at a rate of 50 kHz with 16 bit resolution, D/A-converted by a DSP-card (Ariel DSP32C), and level adjusted by a digitally controlled audiometric amplifier. Stimuli were presented through insert earphones (Etymotic Research ER-2) encased in copper-boxes to avoid artifacts in the AEP due to electrical leakage.

A rising chirp stimulus (Dau et al., 2000), designed to activate all cochlear locations simultaneously, was used instead of the conventional click. By attenuating the lower frequencies, the spectrum of the chirp was flat in the range between the nominal edge frequencies of 100 Hz and 18,000 Hz, resulting in -3 dB-points of about 130 Hz and 15,000 Hz for the stimulus. Fig. 3.1 shows the acoustic waveforms and spectra of the chirp. The duration of the chirp was 10.4 ms.

Virtual sound directions of the stimuli were generated by convolving the chirp digitally with individual head-related impulse responses (HRIRs). The HRIRs were recorded in an anechoic chamber (Otten, 2001, Chapter 3). For each stimulus condition, the absolute travel-time delay for the leading side was eliminated from the

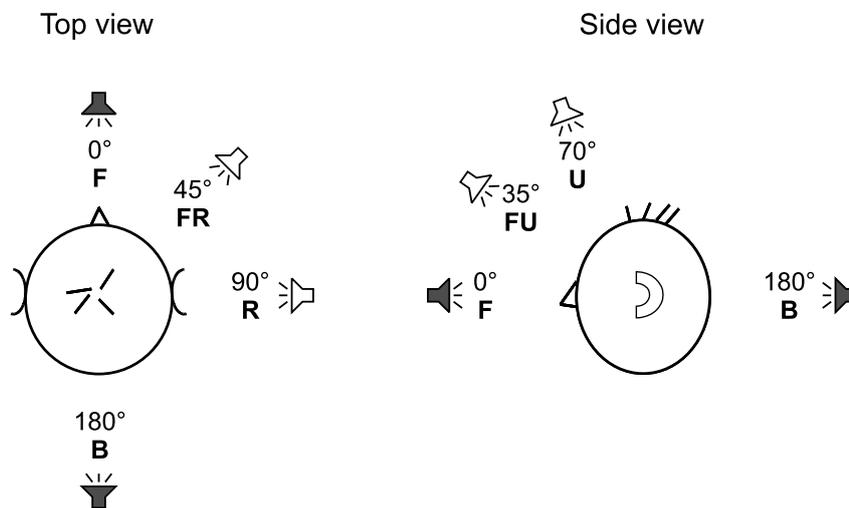


Figure 4.1: Stimulus conditions: The chirp was convolved with individual HRIRs corresponding to six virtual directions. **Left panel:** Top view, i.e., directions on the horizontal plane. Numbers denote the azimuthal angle φ with respect to the front direction. **Right panel:** Side view, i.e., directions on the midsagittal plane. Numbers denote the elevation angle ϑ with respect to the front direction. The directions indicated by the gray speakers are identical directions and were only measured once.

| Stimulus condition | Azimuth φ | Elevation ϑ | Repetitions | Probability |
|--------------------|-------------------|-----------------------|-----------------|------------------|
| Frontal (F) | 0° | 0° | 27,580 | 84.65% |
| Frontal Right (FR) | 45° | 0° | 1000 | 3.07% |
| Right (R) | 90° | 0° | 1000 | 3.07% |
| Frontal Up (FU) | 0° | 35° | 1000 | 3.07% |
| Up (U) | 0° | 70° | 1000 | 3.07% |
| Back (B) | 0° | 180° | 1000 | 3.07% |
| | | | Σ 32,580 | Σ 100.00% |

Table 4.1: Azimuth and elevation of the stimulus conditions used in the oddball paradigm, including the respective numbers of repetitions and probability.

left and right HRIR. Six different sound conditions were used in this study (Fig. 4.1 and Tab. 4.1): Frontal (F), frontal right (FR), right (R), frontal up (FU), up (U), and back (B).

The stimulus sequence was designed according to an oddball paradigm. The frontal stimulus (F) had a probability of about 84.65% and served as standard. Each of the remaining five stimuli had a probability of 3.07% resulting in a total probability of 15.35% for the deviants. The sequence was constrained to have at least three and maximally eight standards between each two deviants. All allowed numbers of intermediate standards (3-8) had the same occurrence. Each sequence contained 100 instances of each deviant and 2758 instances of the standard. The duration of the sequence was 21.7 minutes.

Prior to the EEG-recordings, detection thresholds were obtained from all subjects to determine the sensation level for the frontal stimulus direction. An adaptive three alternative forced-choice (3AFC) procedure was employed, estimating the 70.7% point on the psychometric function. The average level across three repetitions was considered as representing 0 dB SL for an individual subject. For the EEG presentation, all stimulus conditions were jointly adjusted in level, such that the frontal condition had a level of 40 dB SL.

4.2.3 EEG-Recordings and procedure

Subjects wore a flexible cap (Easy Cap, Falk Minow Services) equipped with a set of Ag/AgCl-electrodes. Electroencephalographic activity was recorded differentially from 29 head positions according to the extended 10-20-system (Jasper, 1957; Sharbrough et al., 1991). The vertex electrode (Cz) served as reference electrode, an electrode at the forehead (Fpz) as common ground. An additional electrode was placed on the tip of the nose (Nz) to allow for later re-referencing to nose-reference, as commonly done in MMN literature. Electroocular activity was recorded from two bipolar channels. The vertical EOG was recorded from electrodes placed above and below the right eye. The horizontal EOG was recorded from electrodes lateral to the outer canthi of both eyes. Interelectrode impedances were maintained below 5 k Ω . The electrode positions were measured before and after each recording session with an ultrasonic head-tracker (CMS30P, Zebris Medizintechnik). The multichannel EEG recording setup is shown schematically in Fig. 3.3.

During the recordings subjects sat comfortably in an armchair in an acoustically and electrically shielded sound booth. They were instructed to watch subtitled silent movies on an LC-display positioned about 1.5 m in front of them and to ignore the auditory stimulation. Inside the booth, electrode signals were pre-amplified by a factor of 150 (Neuroscan Headbox) and passed to the main EEG amplifier (Neuroscan Synamps 5803) outside the booth for further amplification by a factor of 33.3, yielding an overall amplification of 74 dB. EEG potentials were anti-alias filtered (analog 2nd order lowpass at a cut-off frequency of 200 Hz), A/D-converted at a sampling rate of 1 kHz, and continuously stored to a computer harddisk.

Each stimulus sequence (cf. section 4.2.2) was played 10 times, yielding 10 runs of EEG recordings with a total recording time of 3.6 hours per subject for the 32,580 stimuli altogether. After each run, electrode impedances were checked to ensure a good quality of the recordings. All subjects accomplished the 10 runs in two separate sessions of 5 runs each. These two sessions were performed on separate days.

4.2.4 Data post-processing

All data post-processing and offline analysis (see sections 4.2.5 and 4.2.6) was performed in MATLAB (The Mathworks). Data was segmented into stimulus-related

epochs (from -50 to 350 ms relative to stimulus onset) and bandpass filtered (1-20 Hz) using a zero-phase forward-backward Butterworth filter (24 dB/octave). Epochs were averaged using an iterative weighted averaging algorithm (Riedel et al., 2001). Epochs to the standard stimulus following directly an epoch to a deviant stimulus were not included in the average of the standards. The standard error of the mean (SEM) σ was calculated for all waveforms.

The evaluation of LAEP components was based directly on the vertex-referenced recording data. Additionally, for the evaluation of MMN components, the recorded data was re-referenced to the nose-electrode (Nz) by subtracting data in the nose channel from all other channels for each recording epoch prior to filtering and weighted averaging. Difference curves (DC) were derived only for these nose-referenced data by subtracting the mean response to the standard stimulus (i.e., the frontal stimulus F) from the mean response to each deviant stimulus separately. The SEM for the DC was calculated from the SEM of the nose-referenced data according to

$$\sigma_{DC_D} = \sqrt{\sigma_{Nz_S}^2 + \sigma_{Nz_D}^2} \quad (4.1)$$

with σ_{DC_D} denoting the SEM of the DC for any deviant and σ_{Nz_S} and σ_{Nz_D} signifying the SEM of the nose-referenced responses to the frontal standard and any deviant, respectively. Given a similar SEM for the responses of all directions, equation 4.1 yields an increase of the SEM by a factor of $\sqrt{2}$ for the DC.

Grand average waveforms over subjects were calculated for both vertex-referenced LAEP and nose-referenced DC for all stimulus conditions.

4.2.5 Analysis of amplitudes and latencies

Amplitudes and latencies of LAEP and DC peaks were evaluated from a subset of channels: For the LAEP data, the classical vertex-referenced derivations (A1, A2, and OZ) were analyzed. For the nose-referenced DC data, 11 channels covering the frontal and parieto-frontal areas (Fp1, Fp2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, and Fz) were chosen. Waveform data in these subsets of channels was upsampled to a sampling rate of 10 kHz (i.e., by a factor of 10) to increase the accuracy of the latency and amplitude of the peak measurements. The upsampling was done by

| Curve type | Latencies | Amplitudes |
|-------------------------------------|-----------|------------|
| LAEP _(vertex-referenced) | N1 P2 | N1–P2 |
| DC _(nose-referenced) | M1 M2 | M1–M2 |

Table 4.2: Overview of peaks evaluated by latency and amplitude.

zero-padding in the spectral domain, resulting in an almost perfect interpolation of the original data. A peak-finding algorithm was utilized, marking all peaks in the waveforms exceeding a peak-to-peak voltage greater than $\sqrt{2} \cdot 2\sigma$. ($\sqrt{2}$ since the variances of both peaks in the pair add up). Latency errors were derived from the curvature of the peaks and the amplitude errors (Hoth, 1986). The labeling of the identified peaks was done manually. Amplitudes were evaluated as peak-to-peak values between the labeled peaks (Tab. 4.2). A Wilcoxon matched-pairs signed-rank test ($\alpha = 0.05$) was performed to ascertain whether response amplitudes and latencies differed significantly between stimulus directions.

4.2.6 Dipole source analysis

Multi-channel LAEP and DC data was modeled by means of dipole source analysis. This is generally accomplished by minimizing a cost function that describes the difference between the measured and the modeled EEG. Equivalent current dipoles (ECDs) were used as a source model. These are focal sources representing the center of a small brain region in which many cells are synchronously activated (Scherg and Cramon, 1990; Scherg, 1991).

Electrode positions were averaged over measurements for each subject and fitted to a sphere by minimizing the least-square error. The radius of this sphere served as outer boundary r for the outer shell of a three-shell head model. The three-shell head model was employed with the radius ranges s_1 , s_2 , s_3 and conductivities g_1 , g_2 , g_3 :

$$\begin{aligned}
s_1 &= [0.0000 \quad 0.8977] \cdot r, & g_1 &= 0.33 \frac{1}{\Omega_m} \\
s_2 &= [0.8977 \quad 0.9659] \cdot r, & g_2 &= \frac{1}{80} \cdot g_1 \\
s_3 &= [0.9659 \quad 1.0000] \cdot r, & g_3 &= g_1
\end{aligned}$$

The first shell models the brain tissue, the second shell the bone, and the third shell the skull. For the subjects in this study, the radii r of the fitted sphere were in the range of 8.3 cm to 9.2 cm.

Since separated neural activity in both auditory cortices was expected, two fixed dipoles with hemispheric constraints were used as a source model. Location and orientation were fitted, but remained fixed for the time interval investigated. Therefore, only the moment magnitude was fitted for each time sample. The pair of dipoles was constrained to take account of hemispheric symmetry: The x -coordinates (left–right) of the location and the azimuth ϕ of the moment were constrained to be mirrored, i.e., had opposite signs. The y -coordinates (back–front) and z -coordinates (bottom–top) and the elevation θ were set to be identical for both dipoles. As additional constraint, the locations and orientations were required to be identical for all stimulus conditions. The time interval for the LAEP fit had a length of 90 ms and was centered around wave P2. For the DC fit, the length of the time window was 100 ms, centered around wave M1 (the peak of the MMN component). Latencies of waves P2 and M1 were determined by taking the average across the same respective subset of channels used for the analysis of amplitudes and latencies (cf. section 4.2.5).

Since topographic voltage maps (rather than true source modeling) are commonly used in the MMN literature, the electric scalp distributions for the DC data at the amplitude maximum of the MMN component were also derived and visualized with ASA (ANT Software). For that reason, DC data was re-referenced to average reference, i.e., the mean over all channels served as reference.

4.3 Results

4.3.1 Waveform morphology

Fig. 4.2 illustrates the vertex-referenced LAEP waveforms for all six stimulus conditions at the occipital channel Oz for the mean over subjects. For all stimulus conditions, peaks for wave N1 and P2 are easily identifiable. The peak-to-peak amplitudes for the N1–P2-complex are highest for conditions FR and R, i.e., for directions away from the midsagittal plane, and P2-latencies are markedly increased for these conditions. The SEM is much smaller for the frontal than for the other stimulus conditions, reflecting the higher number of averaged epochs.

The EEG traces for the nose-referenced data are presented in Fig. 4.3. Channel Fz (top row), a typical frontal MMN derivation site, exhibits higher amplitudes of wave N1 and higher latencies of waves N1 and P2 for deviants at stimulus conditions FR and R when compared to the standard curve. As a consequence, the appendant difference curves show a conspicuous biphasic component, with peak M1 representing the typical MMN component. At the left mastoid (channel A1, bottom row), the DC reveals polarity inversion. This effect is expected for electrodes positioned below the sylvian fissure.

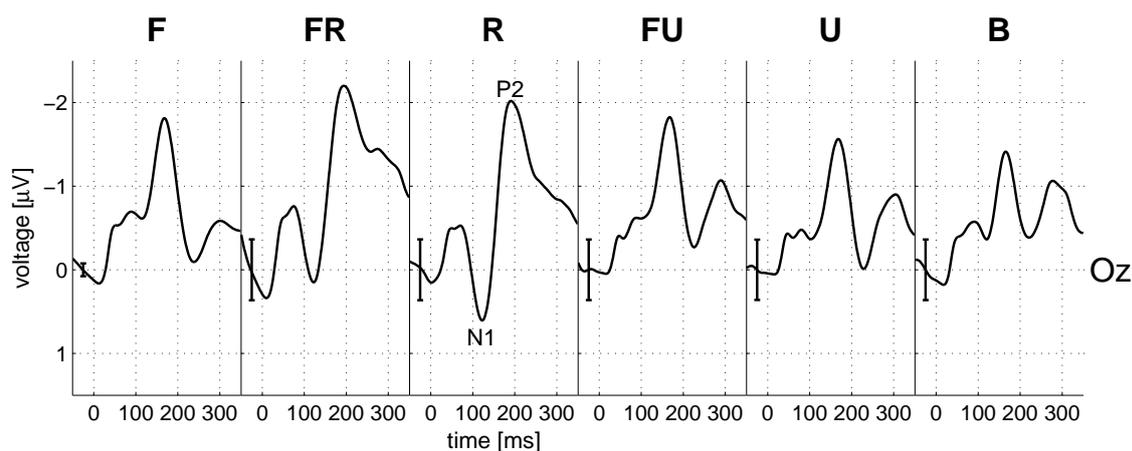


Figure 4.2: Waveform morphology for the (vertex-referenced) LAEP, mean over subjects, channel Oz. The six panels represent the responses for the six stimulus directions, denoted by the title labels (nomenclature according to Tab. 4.1). Exemplary peak labels mark the prominent waves N1 and P2 in the third panel. Error bars indicate $\pm 3\sigma$.

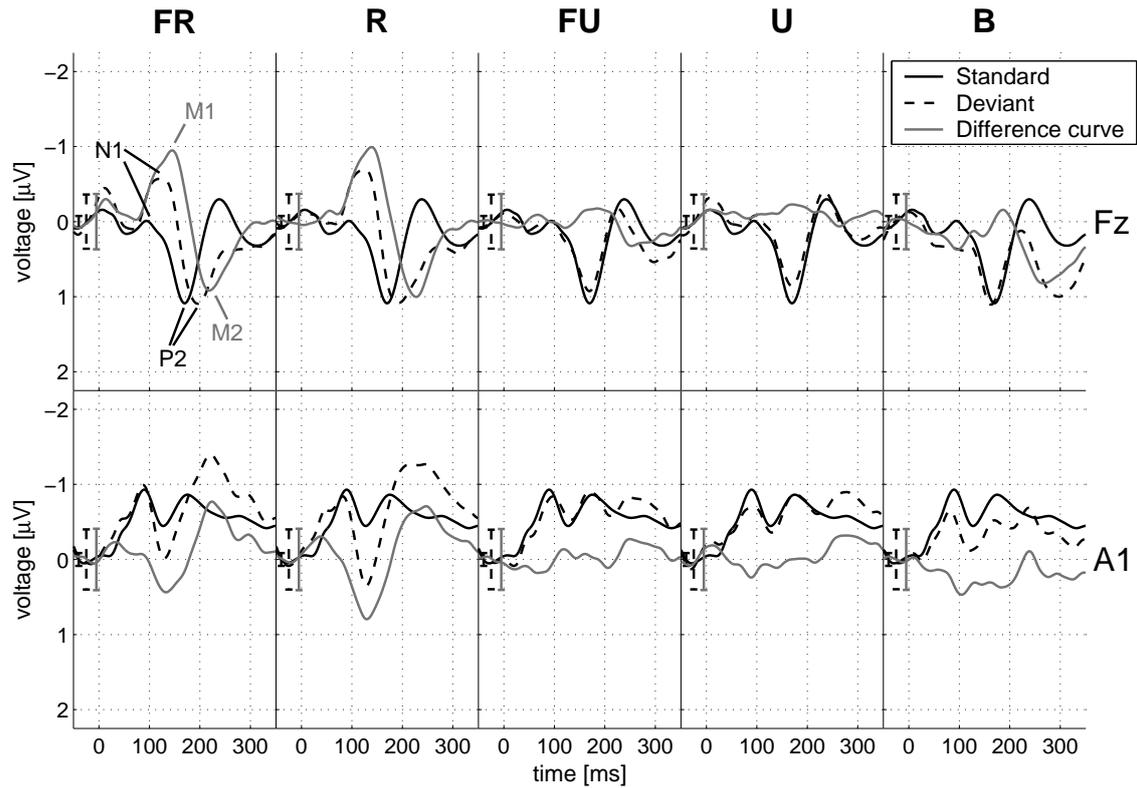


Figure 4.3: Waveform morphology for the nose-referenced data, mean over subjects: **Top row:** Channel Fz. **Bottom row:** Channel A1 (left mastoid). The five columns represent the five deviant stimulus directions indicated by the title labels (nomenclature according to Tab. 4.1). Each panel shows the traces for the response to the standard (frontal direction, black, identical within each row), the respective deviant response (dashed), and the difference curve (derived by subtracting the standard from the deviant response, gray). Exemplary peak labels in the top left panel denote the prominent waves N1/P2 and M1/M2 (difference curves). Error bars indicate $\pm 3\sigma$ separately for each trace.

4.3.2 Analysis of amplitudes and latencies

Results of the evaluation of amplitudes and latencies are shown in Fig. 4.4 and 4.5. In both figures, on the abscissa the different stimulus conditions are indicated, on the ordinate the amplitudes and latencies, respectively. In all panels, the wide black bars present the data averaged over subjects and channels. Error bars denote the interindividual standard deviations. Additionally, averages over subjects for each single channel were drawn in narrow white and gray bars to the left of each black bar.

Fig. 4.4 presents the amplitudes for the N1–P2-complex in the top panel and the latencies for wave N1 and P2 in the middle and bottom panels, respectively. The

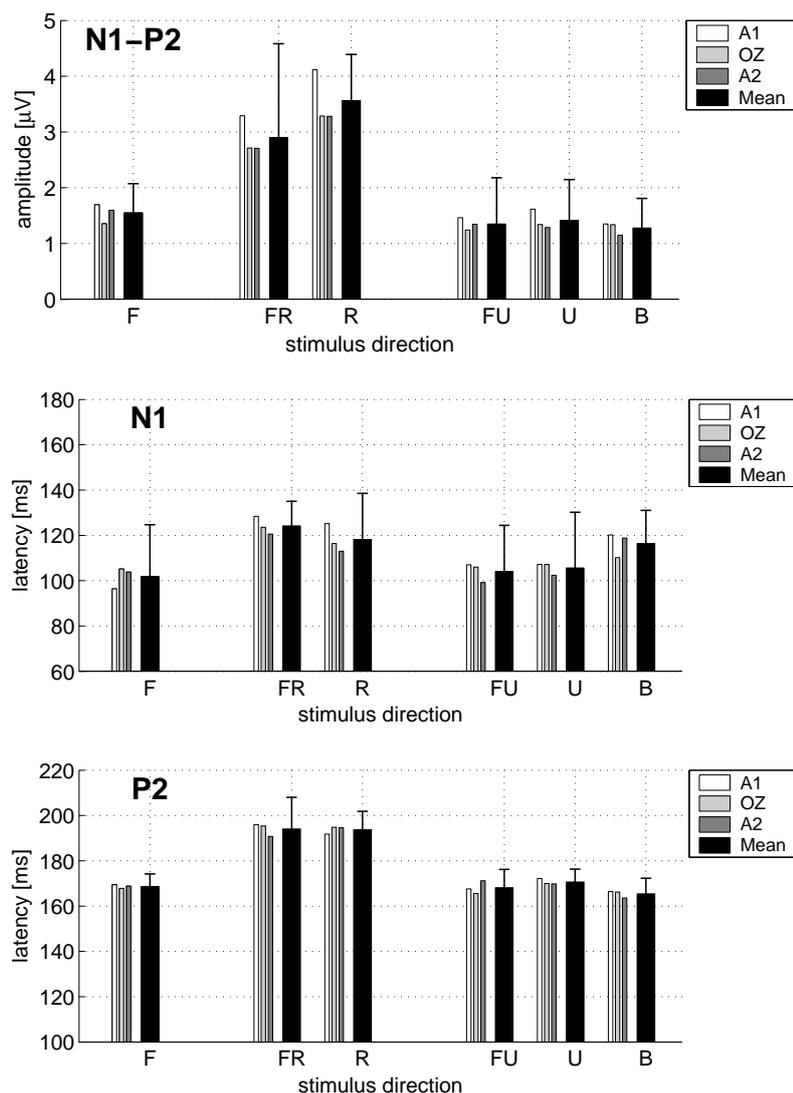


Figure 4.4: LAEP amplitudes and latencies for all stimulus conditions, mean over subjects. **Top panel:** N1–P2 amplitudes. **Middle panel:** N1 latencies. **Bottom panel:** P2 latencies. Wide black bars denote the mean over channels for the six stimulus directions. Narrow white and gray bars to the left of each black bar show the single channel average over subjects. Error bars represent standard deviations over subjects for the mean channel data.

stimulus directions on the midsagittal plane, i.e., F, FU, U, and B, show quite similar amplitudes and latencies. For the conditions FR and R, as revealed by the Wilcoxon test, amplitudes are significantly greater than for the other conditions, and the amplitude difference between FR and R is also significant. In addition, the latencies for wave P2 are significantly higher for FR and R compared to the midsagittal conditions. For N1, latencies of FR and R differ significantly only from conditions F and FU, but not from U and B. For the N1, the latency of the back

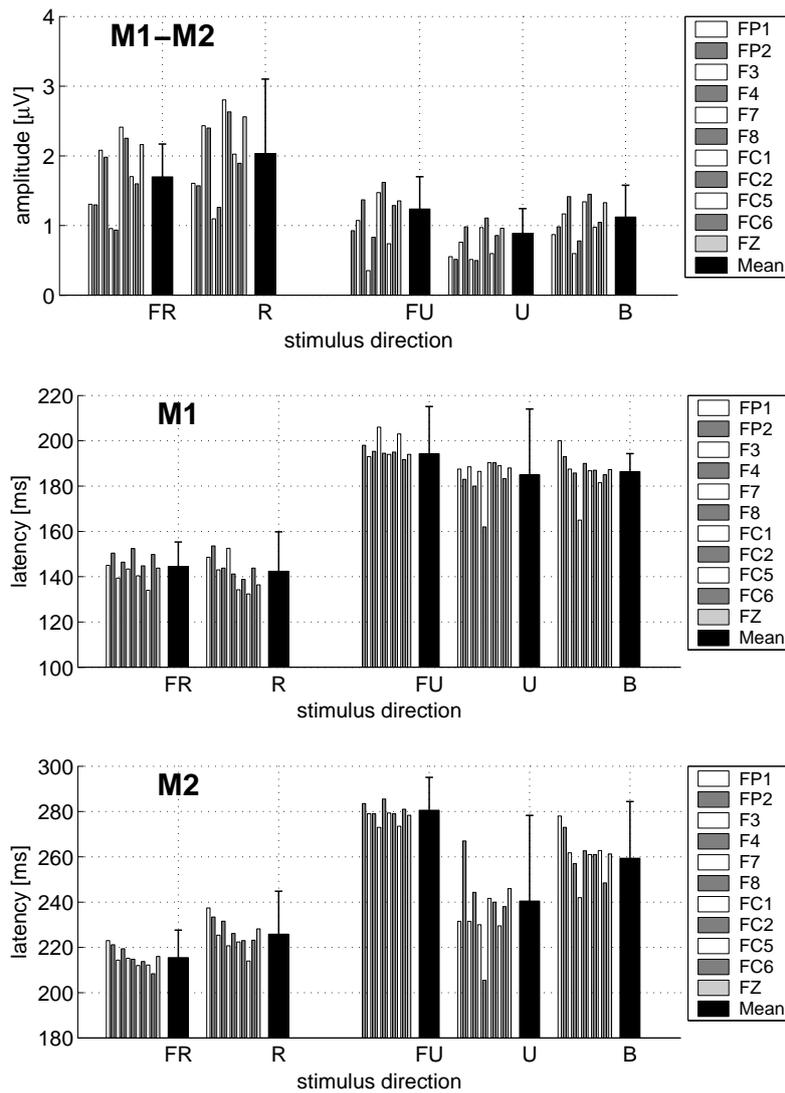


Figure 4.5: DC amplitudes and latencies for all stimulus conditions, mean over subjects. **Top panel:** M1–M2 amplitudes. **Middle panel:** M1 latencies. **Bottom panel:** M2 latencies. Wide black bars denote the mean over channels for the five deviant stimulus directions. Narrow white and gray bars to the left of each black bar show the single channel average over subjects (same order left-right as legend top-down). Error bars represent standard deviations over subjects for the mean channel data.

condition (B) is significantly greater than that of the frontal condition (F). For the P2, the latency of the back condition is slightly (but significantly) smaller compared to the frontal condition (F).

The top panel of Fig. 4.5 shows the amplitude differences for wave M1–M2, the middle and bottom panels present the latencies for wave M1 and M2, respectively. Complementary to the LAEP data (Fig. 4.4), amplitudes for conditions FR and R

are significantly greater than for the remaining conditions. In contrast to the LAEP data, the amplitude difference between conditions FR and R is not significant. For the latencies of wave M1, FR and R show a significant decrease compared to the other directions, but the difference between FR and R is not significant. For wave M2, latencies for FR and R are also shorter compared to the other conditions. However, the decrease is not significant compared to condition U.

4.3.3 Dipole source analysis

The dipole fit results for the LAEP, averaged across the individual fit results, are presented in Tab. 4.3 and Fig. 4.6. The fitted locations and orientations (constrained to be equal for all conditions, cf. section 4.2.6) are summarized in the top half of Tab. 4.3. The two dipoles could well be separated and lie 7 cm apart in the left and right hemisphere of the head. The dipole moments are orientated in almost frontal direction, with a tilt to the top. Fig. 4.6 illustrates the time courses of the dipole moment magnitudes for the left (thick black lines) and right (thin gray lines) dipoles. The moments of the dipole in the left hemisphere increase for stimulus conditions FR and R, i.e., depend on the laterality of the stimulus. This behavior reflects the stronger contralateral activation by stimuli away from the midline. Both dipoles show similar moments for conditions F, FU, U, and B: sounds originating from the midsagittal plane lead to an almost symmetrical activation in both hemispheres. There is virtually no influence on the dipole moments when changing the elevation of the stimulus.

The dipole fit results for the DCs, averaged across the individual fit results, are presented in the bottom half of Tab. 4.3 and in Fig. 4.7. Both dipole locations and orientations are very similar to the LAEP sources, with slightly more lateral and posterior locations. The magnitude time courses (Fig. 4.7) exhibit strongest activation of the left dipole for conditions FR and R. For the conditions on the midsagittal plane (bottom row), both dipoles show a rather weak activation.

Topographic voltage maps for the DC data are plotted in Fig. 4.8. For each stimulus condition, the M1 latency averaged across the frontal and parieto-frontal electrodes served as time instant for the maps. Conditions FR and R show similar activation patterns: A negative field frontally on top of the head and a transition to positive voltages to the lower back of the head indicate a typical MMN voltage distribution.

For condition R, the negative field is lateralized more to the right compared to condition FR. For all remaining conditions, the activation is rather weak. While the activation pattern for condition U is qualitatively similar to that of conditions FR and R, conditions FU and U show a negative field localized more backwards and lateralized to the right.

| P2 | x | y | z | ϕ | θ |
|---------------|---------|---------|---------|--------|----------|
| Left dipole: | -3.5 cm | -0.1 cm | +0.9 cm | +4° | +41° |
| Right dipole: | +3.5 cm | -0.1 cm | +0.9 cm | -4° | +41° |
| M1 | x | y | z | ϕ | θ |
| Left dipole: | -3.9 cm | -0.9 cm | +1.0 cm | -4° | +35° |
| Right dipole: | +3.9 cm | -0.9 cm | +1.0 cm | +4° | +35° |

Table 4.3: Results of the dipole fit for the LAEP wave P2 (top half) and DC wave M1 (bottom half), dipole positions and orientations, mean over subjects. x , y , and z denote the location in cartesian space (x points to the right, y to the front, and z to the top). ϕ is the azimuthal angle ($\phi = 0^\circ$ corresponds to the frontal direction), θ the inclination from the horizontal plane.

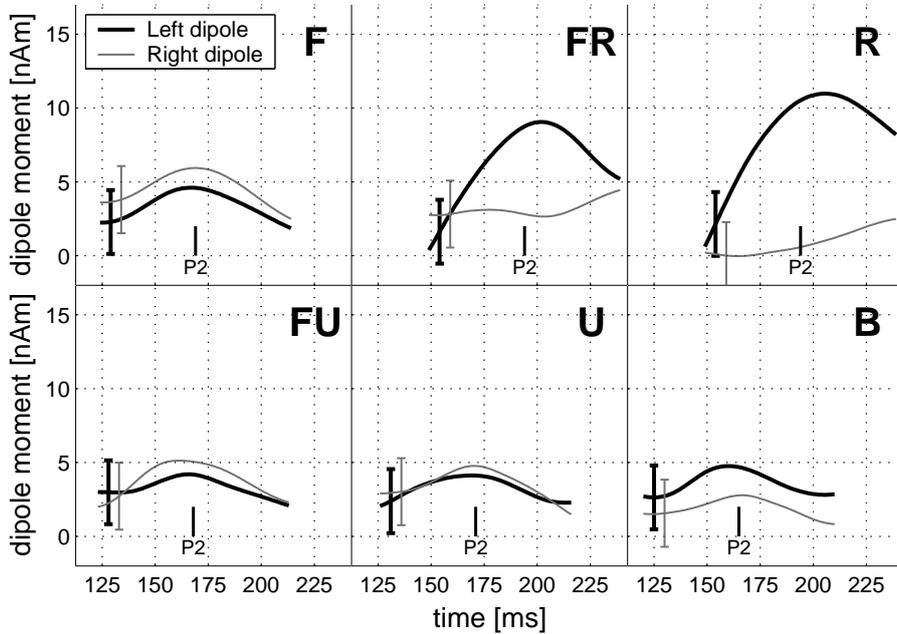


Figure 4.6: LAEP dipole moment magnitudes of two constrained fixed dipoles, dipole fits averaged across subjects. Each subplot represents one stimulus direction, indicated by the text in the top right corner. Vertical bars denote latency of wave P2, error bars the 95% confidence regions. Thick lines show the dipole in the left hemisphere, thin lines the one in the right hemisphere.

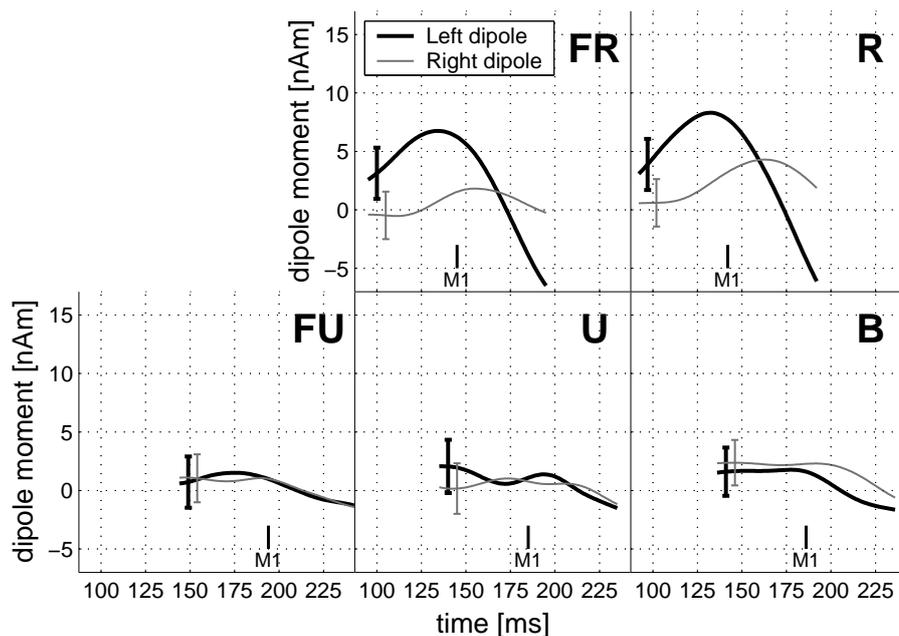


Figure 4.7: DC dipole moment magnitudes of two constrained fixed dipoles, dipole fits averaged across subjects. Each subplot represents one stimulus direction, indicated by the text in the top right corner. Vertical bars denote the mean latency of the MMN component (M1, averaged over frontal and parieto-frontal electrodes), error bars the 95% confidence regions. Thick black lines show the dipole in the left hemisphere, thin gray lines the one in the right hemisphere.

4.4 Discussion

Long-latency multi-channel recordings were performed for six virtual sound directions in an oddball paradigm. The virtual acoustics technique allowed for stimulation with externalized sound objects with azimuth and elevation cues. Single epochs were recorded and analyzed offline allowing for an improved averaging technique as well as for an estimation of the residual noise on a single-sweep basis (Riedel et al., 2001). The stimulus paradigm permitted evaluation of both LAEP and MMN components.

Stimuli off the midsagittal plane produced higher N1–P2 and M1–M2 amplitude differences than stimuli that were on it. For the N1–P2-complex, this increase was monotonic with azimuth, i.e., condition R yielded higher amplitude differences than condition FR. Both lateralized conditions resulted in increased latencies for components N1 and P2, but decreased latencies for components M1 and M2 when compared to the conditions on the midsagittal plane. All directions on the midsagittal plane showed very similar amplitudes and latencies, with one exception: The up-direction produced an M2-latency similar to that of both lateralized stim-

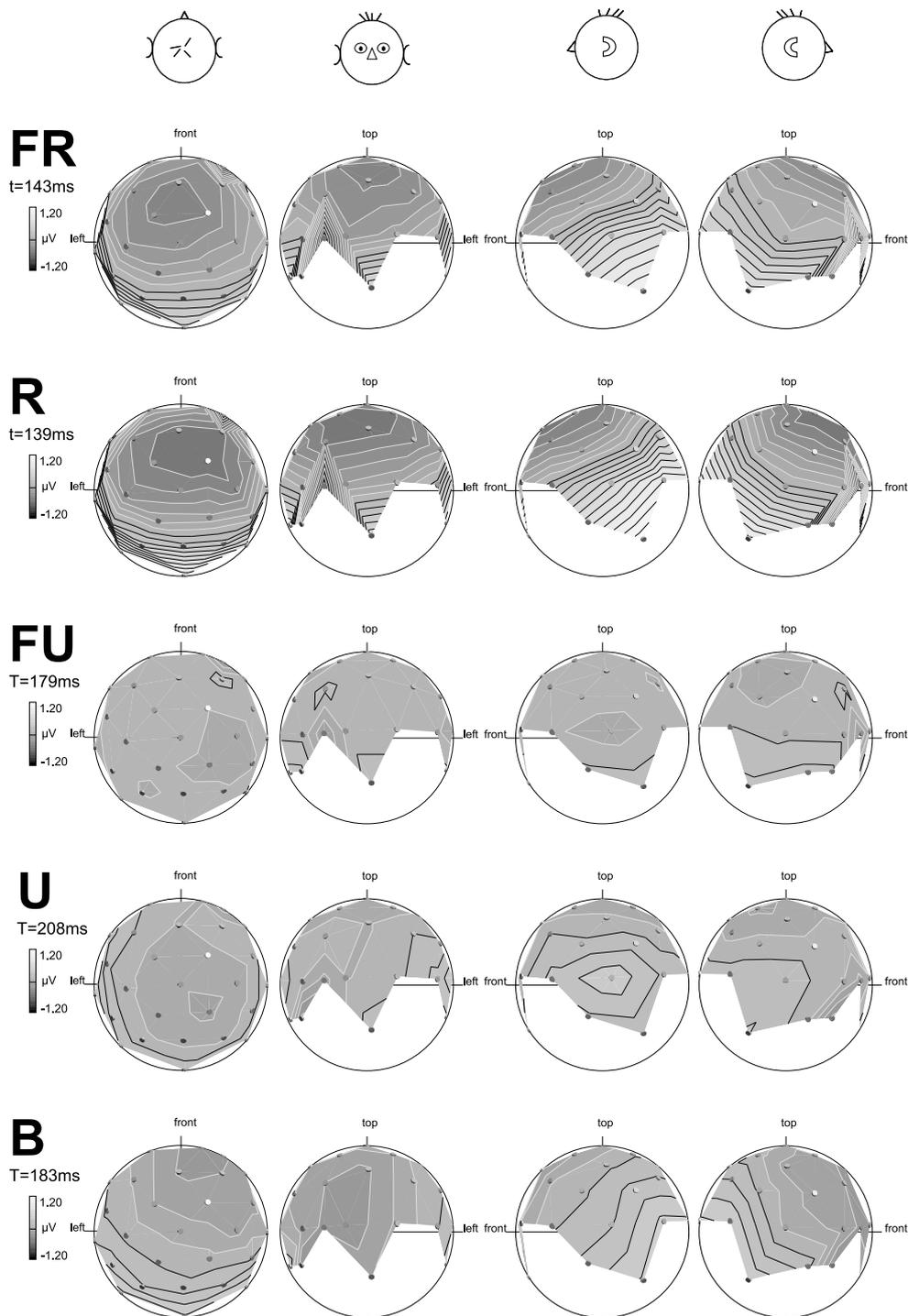


Figure 4.8: Topographic voltage maps for the DC data, waveform mean over subjects. Each row represents one stimulus direction, indicated by its label on the left. For each condition, maps were plotted at the latency of the MMN component (wave M1). Each condition is plotted in four perspectives (columns, from left to right): Top view, front view, view from right, view from left. Voltage increases from black (negative) to white (positive). Mean electrode positions, averaged across all subjects, are drawn as gray circles.

ulus conditions. However, this discriminability between the up-direction and the other conditions along the midsagittal plane could not be found in the dipole source analysis. Therefore, the LAEP and MMN components showed mostly a prominent coding of sound azimuth, but not of sound elevation. This finding is not compatible with a hypothetical ‘localization detector’ that would treat localization differences in elevation similar to differences in azimuth.

[McEvoy et al. \(1991\)](#) employed an ITD reversal paradigm and reported enhanced N1 and P2 amplitudes as ITD was increased. However, significant latency effects were not found. Possibly, in the present study the conjunction of ITD and ILD in the virtual acoustics paradigm might have caused the pronounced latency effects.

The source analysis indicated generators in the areas of the superior temporal cortices for both LAEP wave P2 and DC wave M1. Previous studies found no systematic differences in cortical source location, at least not as a function of ITD ([McEvoy et al., 1993](#)). Therefore, the issue of stimulus direction dependent dipole locations was not addressed in this study. Instead, dipole locations were restricted to be fixed for all conditions to stabilize the fits.

In accordance with the amplitude and latency effects found in the present study, the dipole fits for both LAEP and DC are significantly influenced only by the lateralized conditions. Right lateralized stimuli resulted in a stronger left dipole, since the auditory pathway projects more strongly to the contralateral than ipsilateral auditory cortex. This dominance of contralateral activation is also supported by neurophysiological studies, indicating that in cat auditory cortex the majority of spatially selective neurons are contralaterally-selective, with fewer ipsilaterally- and medially-selective cells (e.g., [Reale and Brugge, 1990](#)).

Contralateral activation for lateralized stimuli was also reported in MEG studies for the N1m component (the MEG equivalent of the N1) ([McEvoy et al., 1993, 1994](#); [Palomäki et al., 2005](#)). E.g., in a very recent study by [Palomäki et al. \(2005\)](#), the variation of sound azimuth revealed systematic influences on magnitudes and latencies of the N1m generators in the auditory cortices: In comparison to ipsilateral stimulation, magnitudes were increased and latencies were decreased when the stimulus was lateralized contralateral to the dipole site. This spatial tuning was more pronounced when using virtual acoustics rather than just ITD, ILD, or a combination of both. Unfortunately, no elevation cues were studied in their work. In addition to the contralateral dominance, [Palomäki et al. \(2005\)](#) found the right

hemisphere to be more strongly activated than the left for all stimulus conditions. This hemispheric dominance was not apparent in the present study.

A number of studies have shown that MMN components can generally be elicited by infrequent changes in sound azimuth (Schröger and Wolff, 1996; Schröger and Eimer, 1996; Damaschke et al., 2000). These findings are corroborated by the results of the present study and suggest that the memory system underlying the MMN encodes sound azimuth. Some authors tried to disentangle the neural processes underlying binaural cues and suggested separate processing of ITD and ILD in the auditory cortex (Schröger, 1996; Ungan et al., 2001). The virtual acoustics paradigm utilized in the present study does not allow for such a separation, as ITD, ILD, and spectral cues are inextricably intermingled in the HRIRs. However, it was shown that MMN amplitudes obtained from deviants with a synergistic combination of ITD and ILD are larger than those acquired with changes in one lateralization cue only (Schröger, 1996; Damaschke et al., 2000). Thus, the inherent combinations of ITD and ILD in the HRIRs incorporated in this work can be assumed to have enhanced the reported MMN amplitudes.

There are few studies that also used virtual acoustics in order to produce externalized stimulus objects. In a study by Paavilainen et al. (1989), MMN were elicited irrespective of whether location changes were presented by earphones or in free-field. In the headphone condition, sinusoids with non-zero ITD or ILD were used as deviants, and the peak amplitudes of the MMN components increased monotonically with increasing ITD or ILD. In the corresponding free-field presentation, loudspeakers at azimuth angles of 10°, 45°, and 90° served as deviants and the MMN amplitudes showed no systematic dependence on azimuth. However, their recording booth was not anechoic and they only used tonal stimuli. Hence, wall reflections and standing waves may have perturbed the sound field. In addition, tonal stimuli did not allow for exploiting spectral cues for sound localization.

In an MEG study with virtual acoustics by Fujiki et al. (2002), the deviant conditions differed from the standard by $\pm 30^\circ$ in either azimuth or elevation. The latencies of both MMN components were found to be increased by about 30 ms for stimuli deviating in elevation compared to those deviating in azimuth. This is in accordance to the results of this study, although the latency increase tended to be higher (about 40-45 ms, compare Fig. 4.5). The dipoles reported by Fujiki et al. (2002) were slightly more separated in the left–right direction (by about 1-2 cm) and lay about

4 cm superior to the dipoles analyzed in this study. These deviations might be a consequence of the different methodologies (MEG vs. EEG) and differing vertical positions of the coordinate systems. Another difference between the studies is that [Fujiki et al. \(2002\)](#) used no hemispheric symmetry constraint for their fits and the dipoles were not restricted to have the same location for all conditions. Their right sources were localized more medial and anterior compared to the left sources, and the right source location varied slightly with the deviant condition (in contrast to the findings of [McEvoy et al., 1993](#), who used no hemispheric constraint either but found no ITD-dependent source location). For their later MMN component (their M2) they found a stronger activation in the right hemisphere for four subjects, a stronger activation in the left hemisphere for two subjects, and an equal activation for two subjects. A stronger activation of the dipoles in the right hemisphere was also observed by [Palomäki et al. \(2005\)](#). However, there are also MEG studies to lateralized sounds that reported no such inter-hemispheric difference ([McEvoy et al., 1993](#); [Sams et al., 1993](#)), just as in the present study.

In a localization study with saccadic eye movements, [Hofman and Van Opstal \(1998\)](#) reported an effect of stimulus duration on the accuracy of elevation localization. They stated that the localization performance in elevation deteriorates below a stimulus duration of about 10 ms for broadband noise bursts. On the other hand, performance improved with increasing durations up to about 80 ms and saturated for higher durations. The chirp used in this work had a duration of 10.4 ms and might therefore not be optimal for an accurate elevation localization. Future work could incorporate longer-duration broadband noise stimuli for evoking the LAEPs. These stimuli would increase the elevation localization performance and might help to reveal systematic influences of stimulus elevation on the generation of LAEPs.

As stated above, the results presented here indicate a coding of sound lateralization (rather than localization) at the level of the LAEP and MMN generator site. In a psychoacoustic localization study, [Makous and Middlebrooks \(1990\)](#) found that the vertical and horizontal localization errors were in the same order of magnitude. Their results were also consistent with results from spatial discrimination studies (i.e., measures of the minimum audible angle, [Heffner and Heffner, 1988](#); [von Wettschureck, 1973](#)). Therefore, in the present work the general discriminability between the frontal standard and the deviants can be expected to also be in the same order of magnitude for both variation of azimuth and elevation (with

exception of the back-positioned deviant, where front-back confusions might have occurred), still no effects of elevation on the potentials could be found. Nevertheless, the auditory system must combine information from interaural and spectral cues across frequency at some level in the nervous system. As reported in chapter 3, this frequency-integrating processing stage could not be found on more peripheral levels of the auditory system by means of AEP, either. Hence, the lacking dependence on elevational sound position might indicate either that the hypothesized ‘localization detector’ is to be found yet more centrally in the auditory system, or that the method of AEP is generally too insensitive to the direction dependent changes in spectral cues.

In summary, the results of this study indicate that the hypothesis of a frequency-integrating ‘localization detector’, must be rejected for the auditory processing stages under investigation here. Instead, the influence of stimulus laterality on the responses suggests an ITD/ILD map at the generation sites of LAEP and MMN. Consequently, a processing stage which integrates interaural and spectral cues either might be situated at even higher levels of the auditory pathway, or cannot be observed by the method of AEPs.

Chapter 5

Summary and perspectives

In the present thesis the influence of realistic spatial acoustic stimuli on the generation of auditory evoked potentials (AEPs) was investigated. An optimized chirp stimulus was used, yielding responses with an improved signal-to-noise ratio compared to a click. The hypotheses of an ‘externalization detector’ and a ‘localization detector’, providing enhanced and distinct neural activation for realistic combinations of ITD and ILD across frequency, was tested at various levels of the auditory pathway.

In chapter 2, some fundamental (monaural) properties of the chirp were elucidated. Responses to long-duration stimuli with embedded chirps and responses obtained with single chirps were virtually indistinguishable for levels up to about 40 dB HL. For these conditions, the across-frequency integration of neural activity does not depend on the acoustic context and behaves linearly. At higher stimulation levels (50 dB HL and above), differences between the responses obtained in the two stimulation paradigms occurred and are probably caused by effects of neural saturation. Furthermore, with regard to temporal processing, the influence of the within-train repetition rate on wave-V latency and amplitude was examined for the click and the chirp. While wave-V latency increased with the temporal position in the train for all within-train rates (47.6, 95.2, 250 Hz), wave-V amplitude was only affected at the highest within-train rate tested (250 Hz). Wave-V latency therefore probably reflects a more sensitive indicator of neural synchronization than wave-V amplitude. The experimental results of chapter 2 demonstrate the importance of cochlear processing for the formation of ABRs, with interesting implications for clinical applications: For studies evaluating wave-V amplitude, the stimulus-train paradigm allows higher mean stimulus rates than the traditional single-stimulus paradigm. The findings

not only provide constraints on future models of peripheral and brainstem processing, but also emphasize the advantage of the chirp over the click for the follow-up chapters.

Chapters 3 and 4 investigated the dependence of chirp-evoked multi-channel AEPs on the spatial location of a sound source. A sound presentation with virtual acoustics, based on individual head related impulse responses (HRIRs), provided realistic, externalized auditory objects that varied in both azimuth and elevation. Chapter 3 focused on ABRs and middle latency responses (MLRs), corresponding to neural generators in the brainstem and primary auditory cortices, respectively. Both binaural potentials and binaural difference potentials (BDs) were examined. Response amplitudes did not benefit from using the externalized stimuli when compared to the diotic reference stimulus. For the ABR, the experimental findings showed a prominent dependence on stimulus laterality, i.e., the azimuthal distance from the midsagittal plane. This effect was prominent in binaural potentials, BDs, and the moment trajectories of a single rotating dipole that served as a source model. The sound elevation did not affect the ABRs. For the MLRs, stimulus laterality was only reflected in the latencies of the BDs. The source analysis of two fixed dipoles exhibited a greater activation of the dipole contralateral to the azimuthal position of the sound source compared to the ipsilateral dipole. With respect to elevational cues, the MLR right dipole source showed a slightly stronger activation for stimuli off the horizontal plane.

Chapter 4 extended the experimental findings of the preceding chapter highlighting the binaural processes at more central stages of the auditory pathway. Late auditory evoked potentials (LAEPs) were recorded in response to a subset of stimulus conditions used in chapter 3. The classical oddball paradigm was used, in which a frequent standard from the front was randomly substituted by five rare deviants with other sound locations. The analysis of LAEP components and dipolar sources revealed a strong influence of the stimulus laterality. Varying the elevation of the sound did not exhibit any systematic effects. By subtracting the response to the standard from the responses to the deviants, differences curves (DCs) according to the mismatch negativity (MMN) paradigm were obtained. Amplitudes, latencies, and dipolar sources of the MMN components showed a dependence on stimulus azimuth, but not on elevation.

Overall, the results of chapters 3 and 4 indicate that sound azimuth is encoded at all levels of the auditory pathway up to the generators producing the LAEPs and the MMN components. Influences of sound elevation are only weakly and not

systematically reflected in the MLRs evoked by transient stimuli such as the chirp. Since in chapter 3 the response amplitudes to externalized stimuli are not higher than those to the diotic stimulus and in chapters 3 and 4 no influence of sound elevation on the potentials was found, the hypothesis of a specific ‘externalization detector’ or ‘localization detector’, integrating neural activity for realistic combinations of ITD, ILD, and spectral cues, must be rejected for the investigated stages of the auditory pathway.

Due to the chirp’s short duration, the localization performance along the midsagittal plane might not have been optimal in chapters 3 and 4. In future work, the use of longer-duration stimuli, such as broadband noise or continuous sequences of chirps, could enhance elevation localization performance and might help to reveal systematic influences of stimulus elevation on the generation of LAEPs.

In this study, the dipole locations obtained from the AEP source analysis are given relative to the spherical head model. Individual anatomical head images, as derived from magnetic resonance imaging (MRI), could provide a more realistic head model, absolute positions of dipolar sources in the brain, and an enhanced accuracy of the dipole fits. Still, even though AEPs provide an excellent time resolution, the spatial accuracy of the reconstructed sources is generally poor compared to MRI. To overcome this methodological disadvantage, a similar experimental setup as presented in this thesis could be used in functional MRI to elucidate the direction dependent activation of the auditory cortices.

Further experiments could give insight into the role of attention on LAEPs. Conventionally, LAEPs are recorded with subjects listening ‘passively’ to a sequence of auditory stimuli while watching movies or reading books, i.e., the attention is undirected. A psychoacoustic alternative forced-choice task, performed during the LAEP recording session, would focus the subject’s attention on the stimulus and would allow for interesting comparisons with results from psychoacoustic studies and from LAEP with unattended stimulation. The necessary modifications for the stimulus presentation software are currently being implemented.

In conclusion, the results of this thesis provide one more step towards understanding the mechanisms underlying the generation of monaural and binaural AEPs, with valuable implications for future research. The fact that no evidence for an ‘externalization/localization detector’ could be provided here suggests that the integration of interaural and spectral cues might take place at levels subsequent to the sources generating the LAEP wave P2 and the MMN component in humans.

Bibliography

- Abbas, P. J., Gorga, M. P., 1981. AP responses in forward-masking paradigms and their relationship to responses of auditory-nerve fibers. *J Acoust Soc Am* 69(2): 492–499. [9](#), [30](#)
- Arehole, S., Salvi, R. J., Saunders, S. S., Gratton, M. A., 1989. Evoked-response forward-masking functions in chinchillas with noise-induced permanent hearing loss. *Audiology* 28(2): 92–110. [29](#)
- Arehole, S., Salvi, R. J., Saunders, S. S., Hamernik, R. P., 1987. Evoked response 'forward masking' functions in chinchillas. *Hear Res* 30(1): 23–32. [30](#)
- Békésy, 1960. *Experiments in Hearing*. McGraw-Hill, New York. [7](#)
- Blauert, J., 1997. *Räumliches Hören, 2. Nachschrift, Neue Ergebnisse und Trends seit 1982*. S. Hirzel Verlag, Stuttgart. [35](#), [63](#)
- Brantberg, K., Hansson, H., Fransson, P., Rosenhall, U., 1999. The binaural interaction component in human ABR is stable within the 0- to 1-ms range of interaural time differences. *Audiol Neurootol* 4(2): 88–94. [36](#), [57](#)
- Burkard, R., Hecox, K., 1983. The effect of broadband noise on the human brainstem auditory evoked response. II. Frequency specificity. *J Acoust Soc Am* 74(4): 1214–1223. [28](#)
- Campen, L. E. v., Hall, J. W., Grantham, D. W., 1997. Human Offset Auditory Brainstem Response: Effects of Stimulus Acoustic Ringing and Rise-Fall Time. *Hear Res* 103: 35–46. [7](#), [28](#)
- Damaschke, J., Granzow, M., Riedel, H., Kollmeier, B., 2000. Zusammenhang zwischen der Lateralisation aufgrund interauraler Zeit- und Pegeldifferenzen und

- kortikal evozierten Potentialen (MMN). In: Sill, A., ed., Fortschritte der Akustik – DAGA 2000, pages 326–327. DEGA e.V., Oldenburg. [81](#)
- Dau, T., 2003. The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *J Acoust Soc Am* 113(2): 936–950. [8](#), [27](#), [31](#)
- Dau, T., Wegner, O., Mellert, V., Kollmeier, B., 2000. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J Acoust Soc Am* 107(3): 1530–1540. [7](#), [8](#), [11](#), [16](#), [19](#), [23](#), [27](#), [37](#), [38](#), [39](#), [65](#), [66](#)
- de Boer, E., 1980. Auditory physics. Physical principles in hearing theory I. *Phys Rep* 62: 87–174. [7](#)
- Debruyne, F., Forrez, G., 1982. On-Effect in Brainstem Electric Response Audiometry. *Otolaryngology* 44: 36–42. [7](#), [28](#)
- Don, M., Allen, A. R., Starr, A., 1977. Effect of click rate on the latency of auditory brain stem responses in humans. *Ann Otol Rhinol Laryngol* 86(2 pt. 1): 186–195. [28](#), [29](#), [30](#)
- Fobel, O., Dau, T., 2004. Searching for the optimal stimulus eliciting auditory brainstem responses in humans. *J Acoust Soc Am* 116(4): 2213–2222. [7](#), [23](#), [37](#), [65](#)
- Fujiki, N., Riederer, K. A., Jousmaki, V., Mäkelä, J. P., Hari, R., 2002. Human cortical representation of virtual auditory space: differences between sound azimuth and elevation. *Eur J Neurosci* 16(11): 2207–2213. [81](#), [82](#)
- Furst, M., Levine, R. A., McGaffigan, P. M., 1985. Click lateralization is related to the beta component of the dichotic brainstem auditory evoked potentials of human subjects. *J Acoust Soc Am* 78(5): 1644–1651. [36](#), [57](#)
- Gardner, W. G., Martin, K. D., 1995. HRTF measurements of a KEMAR. *J Acoust Soc Am* 97: 3907–3908. [36](#), [58](#)
- Genuit, K., 1984. Ein Modell zur Beschreibung von Außenohrübertragungseigenschaften. Dissertation, RWTH Aachen. [35](#), [63](#)

- Gerull, G., Mrowinski, D., 1984. Brain stem potentials evoked by binaural click stimuli with differences in interaural time and intensity. *Audiology* 23(3): 265–276. [36](#), [57](#)
- Gorga, M. P., Abbas, P. J., 1981. AP measurements of short-term adaptation in normal and in acoustically traumatized ears. *J Acoust Soc Am* 70(5): 1310–1321. [30](#)
- Gorga, M. P., Thornton, A. R., 1989. The Choice of Stimuli for ABR Measurements. *Ear Hear* 10: 217–230. [7](#), [28](#)
- Granzow, M., Riedel, H., Kollmeier, B., 2001. Single-sweep-based methods to improve the quality of auditory brain stem responses. Part I: Optimized linear filtering. *Z Audiol* 40(1): 32–44. [11](#), [42](#), [56](#)
- Harkins, S. W., McEvoy, T. M., Scott, M. L., 1979. Effects of interstimulus interval on latency of the brainstem auditory evoked potential. *Int J Neurosci* 10(1): 7–14. [28](#)
- Hecox, K., Squires, N., Galambos, R., 1976. Brainstem auditory evoked responses in man. I. Effect of stimulus rise–fall time and duration. *J Acoust Soc Am* 60(5): 1187–1192. [7](#), [28](#)
- Heffner, R. S., Heffner, H. E., 1988. Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure. *Hear Res* 36(2-3): 221–232. [82](#)
- Hofman, P. M., Van Opstal, A. J., 1998. Spectro-temporal factors in two-dimensional human sound localization. *J Acoust Soc Am* 103(5 Pt 1): 2634–2648. [59](#), [82](#)
- Hoth, S., 1986. Reliability of latency and amplitude values of auditory-evoked potentials. *Audiology* 25: 248–257. [43](#), [70](#)
- Ito, S., Hoke, M., Pantev, C., Lutkenhoner, B., 1988. Binaural interaction in brainstem auditory evoked potentials elicited by frequency-specific stimuli. *Hear Res* 35(1): 9–19. [36](#), [41](#), [42](#), [43](#)
- Jacobson, J. T., Novotny, G. M., Elliott, S., 1980. Clinical considerations in the interpretation of auditory brainstem response audiometry. *J Otolaryngol* 9(6): 493–504. [14](#)

- Jasper, H. H., 1957. The ten twenty electrode system of the international federation. *Electroenceph Clin Neurophys* 10: 371–375, appendix. [40](#), [68](#)
- Jewett, D. L., Martin, W. M., Sininger, Y. S., Gardi, J. N., Morris, J. H., r., Williston, J. S., Randolph, M. G., Chimento, T. C., Shah, D., 1987. The 3-channel Lissajous' trajectory of the auditory brain-stem response. *Electroencephalogr Clin Neurophysiol* 68(5): 323–414. [36](#), [58](#)
- Jewett, D. L., Romano, M. N., Williston, J. S., 1970. Human auditory evoked potentials: possible brain stem components detected on the scalp. *Science* 167(924): 1517–1518. [2](#), [7](#), [28](#), [35](#), [64](#)
- Jewett, D. L., Williston, J. S., 1971. Auditory-evoked far fields averaged from the scalp of humans. *Brain* 94(4): 681–696. [2](#)
- Jiang, Z. D., Wu, Y. Y., Zhang, L., 1991. Amplitude change with click rate in human brainstem auditory-evoked responses. *Audiology* 30(3): 173–182. [21](#), [28](#)
- Jones, S. J., Van der Poel, J. C., 1990. Binaural interaction in the brain-stem auditory evoked potential: evidence for a delay line coincidence detection mechanism. *Electroencephalogr Clin Neurophysiol* 77(3): 214–224. [36](#), [57](#)
- Junius, D., Dau, T., 2005. Influence of cochlear traveling wave and neural adaptation on auditory brainstem responses. *Hear Res* (in press). [5](#)
- Kaltenbach, J. A., Meleca, R. J., Falzarano, P. R., Myers, S. F., Simpson, T. H., 1993. Forward masking properties of neurons in the dorsal cochlear nucleus: possible role in the process of echo suppression. *Hear Res* 67(1-2): 35–44. [9](#), [30](#)
- Kiang, N. Y., 1975. Stimulus representation in the discharge patterns of auditory neurons. In: Tower, D., ed., *The Nervous System, Vol. 3: Human Communication and Its Disorders*, pages 81–95. Raven, New York. [7](#)
- Kiang, N. Y. S., Watanabe, T., Thomas, E. C., Clark, L. F., 1965. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. MIT Press, Cambridge. [7](#)
- King, C., McGee, T., Rubel, E. W., Nicol, T., Kraus, N., 1995. Acoustic features and acoustic changes are represented by different central pathways. *Hear Res* 85(1-2): 45–52. [2](#), [64](#)

- Kodera, K., Yamane, H., Suzuki, J., 1977. The Effect of Onset, Offset and Rise-Decay Times of Tone Bursts on Brainstem Response. *Scand Audiol* 6: 205–210. [7](#), [28](#)
- Kramer, S. J., Teas, D. C., 1982. Forward masking of auditory nerve (N1) and brainstem (wave V) responses in humans. *J Acoust Soc Am* 72(3): 795–803. [29](#)
- Lasky, R. E., 1991. The effects of rate and forward masking on human adult and newborn auditory evoked brainstem response thresholds. *Dev Psychobiol* 24(1): 51–64. [28](#)
- Leonard, C. M., Puranik, C., Kuldau, J. M., Lombardino, L. J., 1998. Normal variation in the frequency and location of human auditory cortex landmarks. Heschl's gyrus: where is it? *Cereb Cortex* 8(5): 397–406. [57](#)
- Levine, R. A., 1981. Binaural interaction in brainstem potentials of human subjects. *Ann Neurol* 9(4): 384–393. [36](#), [40](#), [56](#)
- Levitt, H., 1971. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49(2): Suppl 2:467+. [40](#)
- Makous, J. C., Middlebrooks, J. C., 1990. Two-dimensional sound localization by human listeners. *J Acoust Soc Am* 87(5): 2188–2200. [82](#)
- McEvoy, L., Hari, R., Imada, T., Sams, M., 1993. Human auditory cortical mechanisms of sound lateralization: II. Interaural time differences at sound onset. *Hear Res* 67(1-2): 98–109. [57](#), [65](#), [80](#), [82](#)
- McEvoy, L., Makela, J. P., Hamalainen, M., Hari, R., 1994. Effect of interaural time differences on middle-latency and late auditory evoked magnetic fields. *Hear Res* 78(2): 249–257. [80](#)
- McEvoy, L. K., Picton, T. W., Champagne, S. C., 1991. Effects of stimulus parameters on human evoked potentials to shifts in the lateralization of a noise. *Audiology* 30(5): 286–302. [64](#), [80](#)
- McPherson, D. L., Starr, A., 1993. Binaural interaction in auditory evoked potentials: brainstem, middle- and long-latency components. *Hear Res* 66(1): 91–98. [36](#)

- McPherson, D. L., Starr, A., 1995. Auditory time-intensity cues in the binaural interaction component of the auditory evoked potentials. *Hear Res* 89(1-2): 162–171. [36](#), [57](#), [58](#)
- Mehrgardt, S., Mellert, V., 1977. Transformation characteristics of the external human ear. *J Acoust Soc Am* 61(6): 1567–1576. [35](#), [63](#)
- Middlebrooks, J. C., Green, D. M., 1991. Sound localization by human listeners. *Annu Rev Psychol* 42: 135–159. [35](#), [63](#)
- Mills, A. W., 1958. On the minimum audible angle. *J Acoust Soc Am* 30: 237–246. [1](#)
- Mills, A. W., 1972. Auditory localization. In: Tobias, J. V., ed., *Foundations of Modern Auditory Theory - Volume II*, pages 301–345. Academic, New York. [1](#), [35](#), [63](#)
- Näätänen, R., 1985. Selective attention and stimulus processing: reflections in event-related potentials, magnetoencephalogram and regional cerebral blood flow. In: Posner, M. I., Marin, O. S. M., eds., *Attention and Performance, XI*, pages 355–373. Erlbaum, Englewood Cliffs, NJ. [64](#)
- Näätänen, R., Alho, K., 1995. Mismatch negativity—a unique measure of sensory processing in audition. *Int J Neurosci* 80(1-4): 317–337. [2](#), [64](#)
- Otten, J., 2001. Factors influencing acoustical localization. Ph.D. thesis, Carl von Ossietzky Universität Oldenburg. [38](#), [66](#)
- Paavilainen, P., Karlsson, M. L., Reinikainen, K., Näätänen, R., 1989. Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalogr Clin Neurophysiol* 73(2): 129–141. [64](#), [81](#)
- Palomäki, K., Tiitinen, H., Mäkinen, V., May, P., Alku, P., 2005. Spatial processing in human auditory cortex: The effects of 3D, ITD, and ILD stimulation techniques. *Cogn Brain Res* (in press). [57](#), [80](#), [82](#)
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., Griffiths, T. D., 2002. The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36(4): 767–776. [57](#)

- Picton, T. W., Hillyard, S. A., Krausz, H. I., Galambos, R., 1974. Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol* 36(2): 179–190. [2](#), [35](#), [63](#)
- Polyakov, A., Pratt, H., 2003a. The cumulative effect of high click rate on monaural and binaural processing in the human auditory brainstem. *Clin Neurophysiol* 114(2): 366–375. [29](#), [30](#)
- Polyakov, A., Pratt, H., 2003b. Electrophysiologic correlates of direction and elevation cues for sound localization in the human brainstem. *Int J Audiol* 42(3): 140–151. [36](#), [58](#), [59](#), [60](#)
- Polyakov, A., Pratt, H., 2003c. Electrophysiological correlates of azimuth and elevation cues for sound localization in human middle latency auditory evoked potentials. *Ear Hear* 24(2): 143–155. [36](#), [58](#), [59](#), [60](#)
- Pratt, H., Har'el, Z., Golos, E., 1983. Three-channel Lissajous' trajectory of human auditory brain stem evoked potentials. *Electroencephalogr Clin Neurophysiol* 56(6): 682–688. [36](#), [58](#)
- Pratt, H., Polyakov, A., Bleich, N., Mittelman, N., 2004. The combined effects of forward masking by noise and high click rate on monaural and binaural human auditory nerve and brainstem potentials. *Hear Res* 193(1-2): 83–94. [30](#)
- Pratt, H., Sohmer, H., 1976. Intensity and rate functions of cochlear and brainstem evoked responses to click stimuli in man. *Arch Otorhinolaryngol* 212(2): 85–92. [14](#), [28](#), [30](#)
- Reale, R. A., Brugge, J. F., 1990. Auditory cortical neurons are sensitive to static and continuously changing interaural phase cues. *J Neurophysiol* 64(4): 1247–1260. [56](#), [80](#)
- Riedel, H., Granzow, M., Kollmeier, B., 2001. Single-sweep-based methods to improve the quality of auditory brain stem responses. Part II: Averaging methods. *Z Audiol* 40(2): 62–85. [11](#), [42](#), [56](#), [69](#), [78](#)
- Riedel, H., Kollmeier, B., 2002a. Auditory brain stem responses evoked by lateralized clicks: is lateralization extracted in the human brain stem? *Hear Res* 163(1-2): 12–26. [36](#), [57](#)

- Riedel, H., Kollmeier, B., 2002b. Comparison of binaural auditory brainstem responses and the binaural difference potential evoked by chirps and clicks. *Hear Res* 169(1-2): 85–96. [37](#), [65](#)
- Riedel, H., Kollmeier, B., 2003. Dipole source analysis of auditory brain stem responses evoked by lateralized clicks. *Z Med Phys* 13(2): 75–83. [36](#), [57](#), [58](#)
- Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., Vaughan, H. G., J., 1995. The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review. *Ear Hear* 16(1): 52–67. [2](#), [64](#)
- Rupp, A., Uppenkamp, S., Gutschalk, A., Beucker, R., Patterson, R. D., Dau, T., Scherg, M., 2002. The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans(1). *Hear Res* 174(1-2): 19–31. [8](#), [37](#), [65](#)
- Sams, M., Hamalainen, M., Hari, R., McEvoy, L., 1993. Human auditory cortical mechanisms of sound lateralization: I. Interaural time differences within sound. *Hear Res* 67(1-2): 89–97. [57](#), [64](#), [82](#)
- Scherg, M., 1991. *Akustisch Evozierte Potentiale, Grundlagen - Entstehungsmechanismen - Quellenmodell*. Kohlhammer. [2](#), [35](#), [43](#), [63](#), [70](#)
- Scherg, M., Cramon, D. v., 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr Clin Neurophysiol* 62(1): 32–44. [64](#)
- Scherg, M., Cramon, D. v., 1990. Fundamentals of dipole source potential analysis. In: Grandori, F., Hoke, M., Romani, G., eds., *Auditory Evoked Magnetic Fields and Electric Potentials*, vol. 6 of *Advances in Audiology*, pages 40–69. Karger, Basel. [43](#), [70](#)
- Scherg, M., Vajsar, J., Picton, T., 1989. A Source Analysis of the Late Human Auditory Evoked Potentials. *J Cog Neurosci* 1(4): 336–355. [64](#)
- Schröger, E., 1996. Interaural time and level differences: integrated or separated processing? *Hear Res* 96(1-2): 191–198. [65](#), [81](#)
- Schröger, E., Eimer, M., 1996. Effects of lateralized cues on the processing of lateralized auditory stimuli. *Biol Psychol* 43(3): 203–226. [65](#), [81](#)

- Schröger, E., Wolff, C., 1996. Mismatch response of the human brain to changes in sound location. *Neuroreport* 7(18): 3005–3008. [65](#), [81](#)
- Scott, M., Harkins, S., 1978. Amplitude of the brain stem auditory evoked response: the effect of interstimulus interval. *Int J Neurosci* 8: 147–152. [28](#)
- Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Lüders, H., Nuwer, M., Picton, T. W., 1991. American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. *J Clin Neurophysiol* 8(2): 200–202. [40](#), [68](#)
- Shore, S. E., Nuttall, A. L., 1985. High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. *J Acoust Soc Am* 78(4): 1286–1295. [8](#)
- Smith, R. L., 1977. Short-term adaptation in single auditory nerve fibers: some poststimulatory effects. *J Neurophysiol* 40(5): 1098–1111. [9](#), [29](#), [30](#)
- Smith, R. L., Brachman, M. L., 1982. Adaptation in auditory-nerve fibers: a revised model. *Biol Cybern* 44(2): 107–120. [9](#)
- Teder-Sälejärvi, W. A., Hillyard, S. A., 1998. The gradient of spatial auditory attention in free field: an event-related potential study. *Percept Psychophys* 60(7): 1228–1242. [65](#)
- Thornton, A. R., Coleman, M. J., 1975. The adaptation of cochlear and brain-stem auditory evoked potentials in humans. *Electroencephalogr Clin Neurophysiol* 39(4): 399–406. [21](#), [29](#), [30](#)
- Ungan, P., Yagcioglu, S., Goksoy, C., 2001. Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. *Clin Neurophysiol* 112(3): 485–498. [81](#)
- von Wettschureck, R., 1973. Die absoluten Unterschiedsschwellen der Richtungswahrnehmung in der Medianebene beim natürlichen Hören, sowie beim Hören über ein Kunstkopf-Übertragungssystem. *Acustica* 28: 197–208. [82](#)
- Walton, J. P., Frisina, R. D., Meierhans, L. R., 1995. Sensorineural hearing loss alters recovery from short-term adaptation in the C57BL/6 mouse. *Hear Res* 88(1-2): 19–26. [9](#), [30](#)

- Wegner, O., Dau, T., 2002. Frequency specificity of chirp-evoked auditory brainstem responses. *J Acoust Soc Am* 111(3): 1318–1329. [7](#), [8](#), [23](#), [37](#), [65](#)
- Westerman, L. A., Smith, R. L., 1984. Rapid and short-term adaptation in auditory nerve responses. *Hear Res* 15(3): 249–260. [9](#), [30](#)
- Westerman, L. A., Smith, R. L., 1985. Rapid adaptation depends on the characteristic frequency of auditory nerve fibers. *Hear Res* 17(2): 197–198. [9](#)
- Westerman, L. A., Smith, R. L., 1987. Conservation of adapting components in auditory-nerve responses. *J Acoust Soc Am* 81(3): 680–691. [29](#)
- Wightman, F. L., Kistler, D. J., 2005. Measurement and validation of human HRTFs for use in hearing research. *Acustica united with Acta Acustica* 91: 429–439. [35](#), [63](#)
- Woodworth, R. S., 1938. *Experimental psychology*. Holt, New York. [35](#), [63](#)
- Wrege, K. S., Starr, A., 1981. Binaural interaction in human auditory brainstem evoked potentials. *Arch Neurol* 38(9): 572–580. [36](#), [57](#)

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Oldenburg, den 9. Juni 2005

Dirk Junius

Danksagung

An dieser Stelle möchte ich all denjenigen Personen danken, die mich während der Promotionszeit auf vielfältige Weise unterstützt und damit zum Gelingen der vorliegenden Arbeit beigetragen haben.

Zuerst möchte ich Herrn Prof. Dr. Dr. Birger Kollmeier danken, der mir diese Dissertation in der Arbeitsgruppe „Medizinische Physik“ unter hervorragenden Arbeitsbedingungen ermöglicht und mit seinem Interesse und vielen anregenden Diskussionen unterstützt hat.

Herrn Prof. Dr. Volker Mellert danke ich für sein Interesse an dieser Arbeit und für die freundliche Übernahme des Korreferats.

Mein ganz besonderer Dank gilt Dr. Helmut Riedel für seine großartige Betreuung während der gesamten Doktorarbeit, insbesondere für die vielen konstruktiven Diskussionen, die methodische und inhaltliche Unterstützung sowie Korrekturen an den Kapiteln 3 und 4.

Prof. Dr. Torsten Dau möchte ich für die hervorragende Betreuung und Unterstützung bei der Bearbeitung des Kapitels 2 auch über Landesgrenzen hinweg ganz herzlich danken.

Dr. Jörn Otten danke ich für sein Know-How im Bereich der virtuellen Akustik, viele fruchtbare Gespräche und die Bereitstellung der Außenohrimpulsantworten für die Kapitel 3 und 4.

Einen ganz besonderen Dank verdient Anita Gorges, die viele der zeitaufwändigen Messungen durchgeführt und mich stets gutgelaunt unterstützt hat. Ein riesiges Dankeschön gilt ebenso meinen Versuchspersonen Stephan Ewert, Oliver Fobel, Rainer Huber, Michael Kleinschmidt, Matthias Müller-Wehlau, Carsten Reckhardt, Helmut Riedel und Kirsten Wagener, die mit unermüdlicher Geduld auch noch mehr Messungen über sich ergehen ließen als ursprünglich geplant.

Für das optimale Arbeitsklima, die gute Zusammenarbeit und freundschaftliche Hilfe möchte ich mich auch bei allen übrigen Mitgliedern der Arbeitsgruppe „Medizinische Physik“ bedanken. Insbesondere danke ich Helge Lüddemann, Stefan Heise und Helmut Riedel für das gute Büroklima sowie Gerke Hoiting für manche Kaffeepause und den stetigen Import niederländischer Spezialitäten. Ferner möchte ich hiermit Susanne Garre und Ingrid Wusowski für ihre Hilfe bei allen „bürokratischen“ Angelegenheiten danken.

Meiner Familie gebührt ein ausdrückliches Dankeschön für die uneingeschränkte Unterstützung und Geduld. Insbesondere danke ich meiner Frau Mailin für ihren liebevollen Rückhalt und ihre stetige Motivation.

Lebenslauf

Geboren wurde ich am 18. Juli 1973 als Sohn von Walburg und Wolfgang Junius in Bremen.

Im Jahre 1979 wurde ich an der Grundschule Kirchweyhe eingeschult und wechselte nach der vierten Klasse 1983 an die Orientierungsstufe Lahausen. Von 1985 bis 1989 besuchte ich die Klassen 7 bis 10 des Gymnasiums in Syke. Danach verbrachte ich das Schuljahr August 1989 bis Juli 1990 als Austauschschüler in Redding, Kalifornien (USA). Wieder in Deutschland, absolvierte ich die gymnasiale Oberstufe am Gymnasium Syke und legte dort im Mai 1993 das Abitur ab.

Nach meinem Zivildienst am Kreiskrankenhaus Bassum von Juni 1993 bis August 1994 nahm ich das Physikstudium zum Wintersemester 1994/1995 an der Universität Bremen auf. Nach dem Vordiplom wechselte ich an die Universität Oldenburg, wo ich mein Studium im Mai 2000 mit dem Diplom abschloss.

Im Juni 2000 begann ich meine Promotion in der Arbeitsgruppe „Medizinische Physik“ bei Prof. Dr. Dr. Kollmeier. Von Juni 2000 bis Mai 2003 war ich Stipendiat des Europäischen Graduiertenkolleg „Neurosensorik“. Von Juni 2003 bis Dezember 2004 arbeitete ich im Sonderforschungsbereich 517 („Neuronale Grundlagen kognitiver Leistungen“), seit Januar 2005 im DFG-Projekt „Psychoakustisch und physiologisch motivierte Modellierung der Modulationswahrnehmung“ als wissenschaftlicher Mitarbeiter.