Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Physicist Anita Kult
born in Schorndorf
Oral examination: 16. February 2006

An Integrated Magnetoencephalographic and Functional Magnetic Resonance Imaging Study on Temporal Asymmetry Processing in the Human Auditory Cortex

Referees:
Prof. Dr. Hans Günter Dosch/Prof. Dr. Hans-Joachim Specht
Prof. Dr. Peter Bachert

With present methods the skull and the scalp are too much in the way, and we need some new physical method to read through them. In these days we may look with some confidence to the physicists to produce such an instrument, for it is just the sort of thing they can do.

E.D. Adrian: Brain Rhythms

Nature 1944, 153: 360-362

Die zeitliche Asymmetrie ist eine fundamentale Eigenschaft der Sprache und Musik. So besitzt ein Ton mit einem exponentiellen Abfall der Einhüllenden (*"damped tone"*) einen überwiegend perkussiven Charakter. Wird der gleiche Ton zeitlich gespiegelt (*"ramped tone"*), hat er eine tonale Komponente. Mit Hilfe der Magnetoenzephalographie und funktioneller Magnetresonanztomographie (fMRI) wurden die durch auditorische Stimulation evozierten kortikalen Quellen lokalisiert und charakterisiert. Weiterhin wurde die Grösse der neuromagnetischen Antwort und der Ausgeprägtheit des Tons korreliert. Beide Grössen steigen mit zunehmender Halbwertszeit (HLT) der Stimuli. Die Asymmetrie die als Amplitudenverhältnis bestimmt wurde zeigt für die 4 ms HLT den grössten Effekt. Dies entspricht der psychophysischen Daten. Mittels fMRI wurde die spezifische Aktivierung des Kontrast ramped–damped für 4 ms HLT im Planum temporale lokalisiert wobei in der rechten Hemisphȅre der stärkere Effekt beobachtet wurde. Die stärke der N100m Antwort und die empfundene Ausgeprägtheit des Tons korrelierten darüber hinaus sehr hoch mit der Vorhersage eines neurophysiologisch motivierten Tonhöhenextraktionsmodel.

Temporal asymmetry is a fundamental property of speech and music. When a "damped" exponential is used to modulate a sinusoid, it reduces the sound quality typically associated with the carrier. When the modulator is reversed in time, producing a "ramped" sound, the sound quality of the carrier is more salient. Magnetoencephalography was used in conjunction with functional magnetic resonance imaging (fMRI) to localize and characterize cortical sources evoked by temporally asymmetric sounds. Additionally, the relation between neurophysiological and psychometric data was explored. We show that the relationship between stimulus perception and neuromagnetic responses are observed in the N100m component of auditory evoked magnetic fields. In response to ramped and damped sinusoids N100m peak amplitudes increase with the stimulus half–life time (HLT). The asymmetry in terms of N100m magnitude ratio was maximal at modulation HLT of 4 ms and greatly reduced at 0.5 ms and 32 ms which parallels the psychophysical data. fMRI revealed substantial right lateralized activation of Planum temporale in a response to a difference contrast 'ramped-damped' of 4 ms HLT, where we localized N100m sources. Temporally asymmetric sounds were simulated using auditory image model. The extracted carrier salience correlated with both, the N100m magnitude and perceived carrier salience.

Acknowledgements

This thesis is a result of many invisible hands helping me. During my dissertation work, as during all preceding stages of my graduate training, I received help from many people. The full extent of their contributions to my work I am only now beginning to realize.
My special thanks goes to the leader of the Section of Biomagnetism Dr. Andre Rupp for giving me the opportunity to work in such an inspiring and refreshing environment such as heidelberger MEG lab, for his optimism, advices and ideas that he'll recognize reading this thesis.
I would like to express sincere gratitude to my supervisors, Prof. Hans Günter Dosch and Prof. Hans-Joachim Specht for efficient guidance and insightful comments.
Furthermore I would like to thank Prof. Peter Bachert for refereeing my thesis. Thereby he made it possible for me to earn the doctorate at the Faculty of Physics in Heidelberg.
A special word of thank goes to Dr. Daniel Pressnitzer. His unique way of steering my research by giving me subtle hints at the right moment helped a lot during a research part of this PhD.
I'm grateful to my colleagues from the Department of Neuroradiology for giving me the opportunity to perform fMRI-experiments on their scanner.
Moreover, I would like to thank Barbara, Esther, Heide, Steffen, Peter, Renate, Matthias, Sebastian, Johannes, Nicole and Alexander for the refreshing and open atmosphere in the lab.
My research interests and dreams might never have realized without support of my parents, Ivan and Amalija and my brother Krešimir. I'm deeply grateful for their unconditional support at all times.
And last but not least I would like to thank Stefan Siebert whom I luckily met during my PhD time. His love, patience, support, sense of humor, optimism and ability to cheer me up turned the last year of my research–despite the thesis writing period–into the most enjoyable time of my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

All movements, thoughts and perceptions are accompanied by specific brain activity. In fact, under normal circumstances, the brain is always active, generating enormously complex patterns of neuronal firing. We understand fairly well the cellular mechanisms responsible for the generation of action potentials and post-synaptic potentials, we know micro- and macro-anatomy of the nervous system, but we are far away from understanding of how thousands of neurons are working as a whole, being involved in the organization of the particular event.
Humans have the natural ability to segregate, identify and recognize sounds in a variety of situations - separated by a wall from the sound source, in a concert hall, within a noisy traffic environment or at a cocktail party. How and what kind of information does the brain actually receive from our auditory sensory organs; which features are critically important and which are redundant or even cause confusion in the recognition and identification process? These and many more questions are still open.
Temporal asymmetry is a fundamental property of human speech and music. Its representation in the auditory cortex is therefore essential for our understanding of brain mechanisms underlying speech and music processing. The term "temporal asymmetry" (Patterson, 1994a) is used to describe sounds with differing attacks and decays. It has been shown that temporal asymmetry can affect a stimulus' timbre (Cutting and Rosner, 1974; Rosen and Howell, 1981; Patterson, 1994a), its perceptual timing (Morton and Marcus, 1976; Vos and Rasch, 1981; Gordon, 1987), and even its pitch (Hartmann, 1978). In music, the temporal envelope also contributes significantly to the identification of instruments. When sounds of asymmetrically shaped temporal envelopes produced by instruments are played backwards humans often fail to recognize the instrument (Paquette and Peretz, 1997). The ASA set of "auditory demonstrations" (Houtsma et al.,

1987) includes a very compelling illustration of the effects of short-term temporal asymmetry on auditory perception (No. 29). A piece of music is played on the piano, and then it is repeated with the waves for the individual notes reversed in time. The melody and harmony are largely unaffected by the manipulation, but the instrument is perceived to change from a piano to a reed organ whose notes end in disruptive complex transients.

The existence of the demonstration shows that the effect of time reversal has been known for some time. Nevertheless, there was little research on the topic until Patterson (1994a, 1994b) initiated a systematic study of auditory temporal asymmetry. Patterson (1994a, 1994b) compared the perceptions of a sinusoid multiplied with a periodically decaying exponential function and a sinusoid multiplied with a periodically rising exponential function. These stimuli were termed 'damped' and 'ramped' sinusoids, respectively. These sounds have two different perceptual components: one is a continuous tone with the pitch of the carrier; the other is a drumming sound, produced by the repetitive stream of abrupt onsets (damped) or offsets (ramped) of the stimuli. The longer the half-life of the exponential modulator, the stronger the tonal component. Akeroyd and Patterson (1995) used sounds with similar envelopes, but the carrier was broadband noise rather than a sinusoid. They reported that the damped sounds was heard as a drum stuck by wire brushes. It did not have any hisslike quality, typical for the noise carrier. In contrast the ramped sound was heard as noise, with a hisslike quality, that was sharply cut off in time. These experiments clearly demonstrate the important role of the temporal envelope on sound quality perception. Damped and ramped sinusoids pose an interesting problem for the understanding of temporal processing in the auditory system as their long-term Fourier energy spectra are identical and they contain two different periodicities (carrier and envelope) that are perceived simultaneously. Traditional models of auditory perception can not explain these perceptual differences. Patterson and Irino (1998) applied an auditory image model (AIM) of auditory processing (Patterson et al., 1995) to explain the phenomena. In this model the sound is filtered using an auditory filterbank, which converts the sound wave into a simulation of the basilar membrane motion (BMM). Temporal asymmetry in the neural activity produced by the cochlea is further enhanced by the "strobed" temporal integration that converts the neural activity pattern into an auditory image. The temporal asymmetry in the auditory image is sufficient to explain the perceptual asymmetry that we ultimately hear. To find the neuromagnetic representation of the perceptual asymmetry we used magnetoencephalography.

Magnetoencephalography measures magnetic fields outside the head generated by the synchronized neural activity with the sub-millisecond temporal resolu-

tion, and reflects therefore the real-time signal processing in the brain. The magnetic signals generated by cortical currents are picked up with superconducting coils connected to SQUIDs (superconducting quantum interference devices), the ultrasensitive detectors of magnetic fields. The present state-of-the-art neuromagnetometers contain more than 300 SQUIDs in helmet-shaped arrays so that signals can be recorded simultaneously over the whole cortex. However, MEG lacks the spatial specificity and resolution required for many brain mapping applications since it relies on the inverse problem which does not have the unique solution.

Functional magnetic resonance imaging (fMRI) has revolutionized cognitive neuroscience over the past decade. It takes advantage of the coupling between neuronal activity and hemodynamics (the local control of blood flow and oxygenation) in the brain to allow the non-invasive localization and measurement of brain activity. FMRI has excellent spatial resolution, along with great flexibility, but its temporal resolution is limited and it is an indirect measure of neural activity mediated by the cerebrovascular system. In order to track brain activity evoked by different sensory or cognitive tasks it is necessary to estimate both, the time-course and location.

Accordingly, we have chosen to measure the cortical neuromagnetic and hemodynamic responses to spectrally matched pairs of damped and ramped sinusoids with a range of half-lives to determine whether they exhibit temporal asymmetry, and if so, the temporal dynamics and location of their sources. Additionally, we performed psychophysical experiments to investigate whether temporal asymmetry is maximal in the range 4–16 ms as reported before, and finally, we simulated temporally asymmetric sounds with the auditory image model to test if the sensation can be predicted with the model. Our aim was to investigate the relationship between the perception of temporally asymmetric sounds, their neuromagnetic and hemodynamic representation and auditory image model predictions. The following Chapter 2 provides an insight to the basics of sound and hearing. A brief summery of the physical, physiological and anatomical background about the complex hearing mechanism, necessary for further understanding, is given. Chapter 3 and 4 are the introduction to the physics of functional brain imaging methods, e.g., MEG and fMRI. Chapter 5 is dealing with the behavioral tests using paired comparison analysis. The algorithm of the Bradley-Terry-Luce (BTL) method is shortly described and used to obtain the relative carrier salience scale which leads to the quantification of the perceived asymmetry in terms of asymmetry index. In Chapters 7 and 8 the MEG and fMRI experimental design and the results are described respectively. The effects of stimulus envelope on the neuromagnetic and hemodynamic responses dependent on stimulus' half-life time,

temporal direction and modality (transient versus steady state stimuli/responses) were investigated. In Chapter 9 we demonstrated the relation between the perception of stimulus salience and its temporal envelope. Furthermore in Chapter 10 we present the simulation results using auditory image model. Here we show that the Auditory Image Model can predict the perceived carrier salience. Finally, in Chapter 11 we investigated the relation between the perception, neurophysiology and modeling.

The temporal asymmetry had been observed in single neurons (Pressnitzer et al., 2000, Neuert et al., 2001, Lu et al., 2001) but our results represent the first in vivo neurophysiological ensamble correlate of the phenomenon.

Chapter 2

Sound and Hearing

Sound is an acoustic wave that results when a vibrating source (such as human vocal cords) disturbs an elastic medium (such as air). When a sound wave reaches a listener's ear drum, the vibrations are transmitted to the inner ear (or cochlea), where mechanical displacements are converted to neural pulses that are sent to the brain and result in the sensation of sound.

2.1 Mathematics of the Pure Tone

The pure tone is the most elementary of all sounds. Mathematically it is a sine wave described by the equation

$$x(t) = Asin(\frac{2\pi t}{T} + \Phi) \tag{2.1}$$

where A is the amplitude, t is time, T is the period in seconds, and Φ is the phase in radians. Knowing that frequency f corresponds to $f = 1/T$ and angular frequency $\omega = 2\pi f$, we can write the sine wave in an alternative form

$$x(t) = Asin(\omega t + \Phi) \tag{2.2}$$

The pure tone is special because it consists of only a single frequency so any linear operation performed on this waveform (i.e. any filtering) leaves the wave shape unchanged. Because the human auditory system behaves much like a frequency analyzer, the pure tone occupies a place of enormous importance in hearing science.

2.2 Fourier Analysis and Spectral Representation

2.2.1 Fourier Transform

Most natural sounds are not sine waves. However, other waveforms can be represented as a superposition of sine waves. The mathematician Fourier proved that any continuous function could be produced as an infinite sum of sine and cosine waves. His result has far-reaching implications for the reproduction and synthesis of sound. In particular, a periodic signal *x(t)* with a fundamental frequency f_0 can be represented as a complex Fourier series

$$x(t) = \sum_{-\infty}^{\infty} X_n e^{in2\pi f_0 t} \quad \text{where} \quad X_n = \int_{period} x(t) e^{-in2\pi f_0 t} f_0 dt \tag{2.3}$$

and a finite-energy signal *x(t)* can be represented as a Fourier integral

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{i2\pi ft} \mathrm{d}f \quad \text{where} \quad X(f) = \int_{-\infty}^{\infty} x(t) e^{-in2\pi ft} dt \tag{2.4}$$

where *X(f)* is called the Fourier transform of *x(t)*.

2.2.2 Energy Spectrum and Power Spectrum

The structure of sound, in terms of its frequency components, is often represented by its magnitude spectrum, a plot of sound amplitude, energy or power as a function of frequency.
The squared magnitude of X gives the power for a periodic signal and the energy density for a finite-energy signal. This lets us speak about the power or energy of a signal in different frequency bands. It is common to refer to X as the spectrum of x. Examples of magnitude spectra are given in Fig 2.1

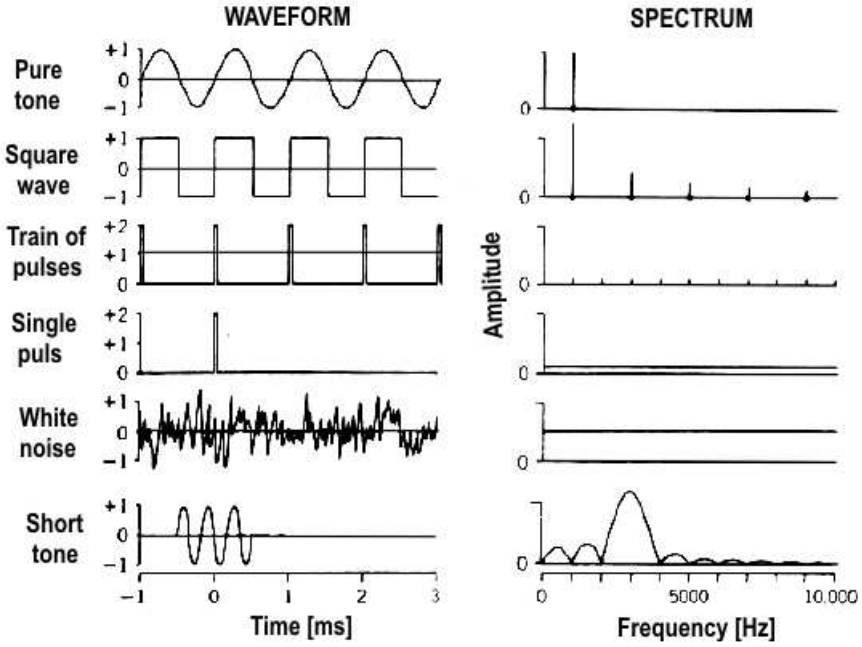

Figure 2.1: *Examples of Fourier magnitude spectra for various waveforms (adopted from Moore, 1989).*

2.3 Amplitude Modulation of the Sound

Amplitude modulation (AM) is a temporal feature of most natural acoustic signals. In AM the carrier's amplitude is varied so as to follow the magnitude of a modulating wave, while the carrier frequency remains unchanged. The idea and terminology of amplitude modulation (AM) owes much to radio engineering. The AM radio signal begins with a carrier of frequency ω_c. The amplitude of this signal is modulated by an audio signal, which is a program material to be broadcast. In hearing research we use audible instead of radio frequencies as carriers but the mathematics and terminology of AM signals remains the same. Suppose we wish to modulate a simple sine wave on a carrier wave. The equation for the carrier wave of frequency ω_c, taking its phase to be a reference phase of zero, is

$$c(t) = C\sin(\omega_c t). \tag{2.5}$$

The equation for the simple sine wave of frequency ω_m(the signal we wish to broadcast) is

$$m(t) = M\sin(\omega_m t + \phi), \tag{2.6}$$

with ϕ its phase offset relative to $c(t)$. Amplitude modulation is performed simply by adding $m(t)$ to C. The amplitude-modulated signal is then

$$x(t) = (C + Msin(\omega_m t + \phi))\sin(\omega_c t) \tag{2.7}$$

When $M = 1$ the wave is said to be 100% modulated. On the Fig. 2.2 the waveforms of two speech and two instrumental sounds are given.

2.4 Time Reversed Signals

For as long as there has been recorded sound, it has been possible for scientists to study the effects of reversing the direction of time by playing sounds backward.

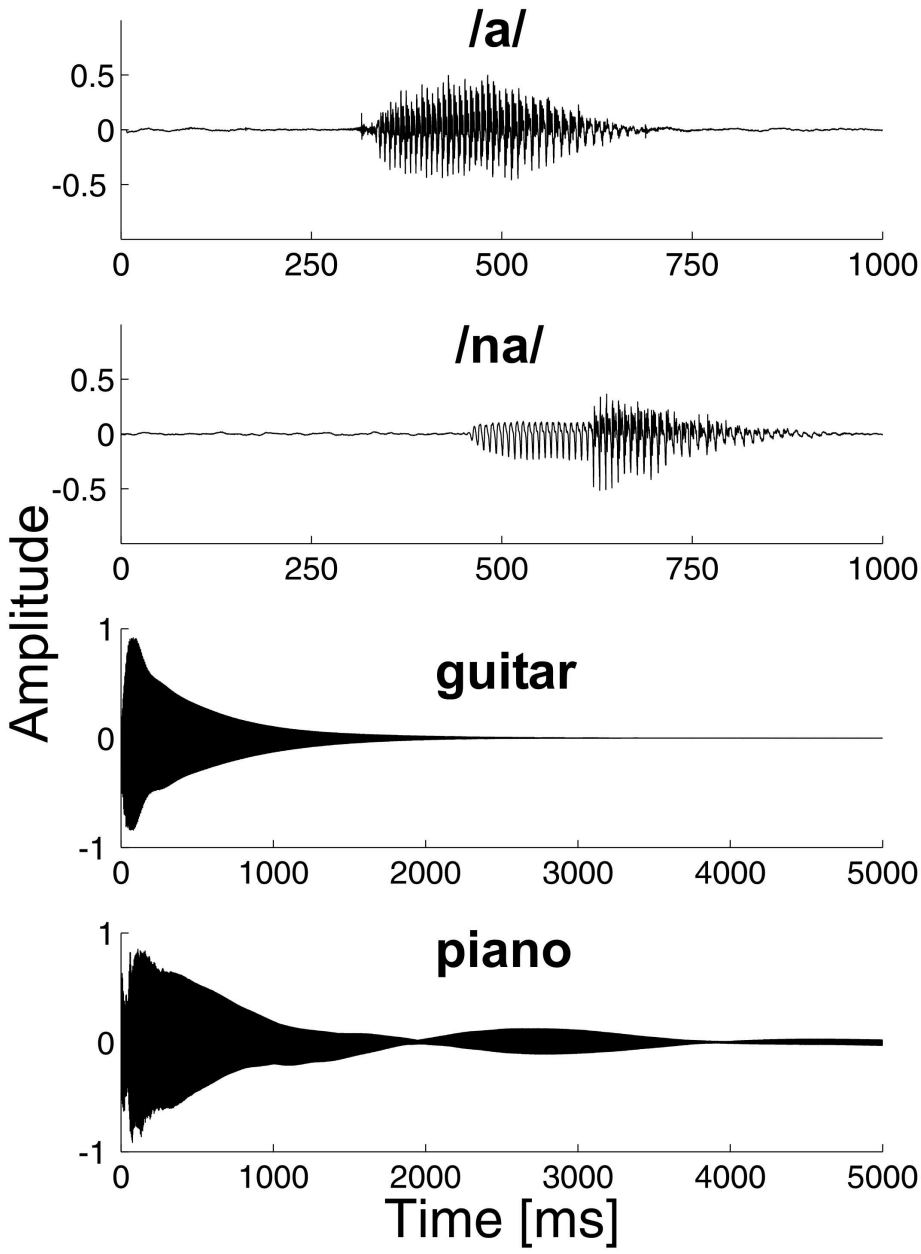

Figure 2.2: *Examples of amplitude modulation. Waveforms of the vocal /a/, syllable /na/, guitar and piano tones. The details of temporal envelopes provide important information in speech and music.*

We can learn a lot by listening the sounds played backward, especially about speech sounds, where the tendency to perceive normal speech categorically often obscures acoustical features.
A key point is that when a signal is played backward its spectrum remains unchanged. Let us consider a signal *x(t)* which extends from time $t = 0$ to time $t = T_D$. For the signal played backward we can write:

$$y(t) = x(T_D - t) \ (t < T_D). \tag{2.8}$$

The Fourier transform of *y(t)* is then

$$Y(\omega) = \int_{-\infty}^{\infty} dt e^{-i\omega t} x(T_D - t). \tag{2.9}$$

For $t' = T_D - t$ we get

$$Y(\omega) = -e^{-i\omega T_D} X(-\omega). \tag{2.10}$$

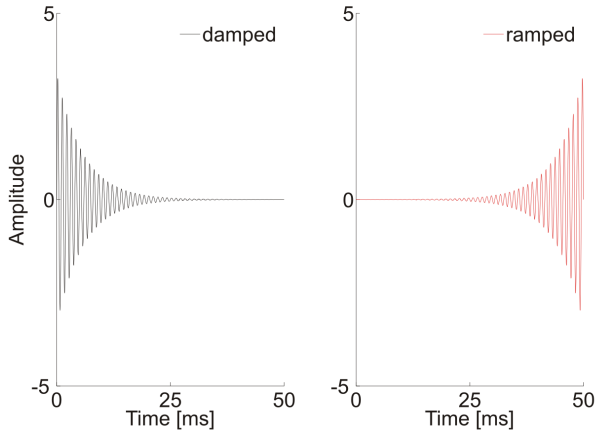

Figure 2.3: *A pair of time reversed sounds with the exponential envelope of 4 ms half-life time.*

Therefore, we have

$$|Y(\omega)|^2 = |X(\omega)|^2, \tag{2.11}$$

which says that the energy spectra of the signal and the time-reversed signal are the same.

As mentioned before, Patterson et al. (1994a,b) studied detailed perception of temporally asymmetric sounds using exponentially modulated sinusoids, which are our experimental stimuli, as well. On the Fig. 2.3, a pair of temporally asymmetric sounds, namely, damped (fast attack) and ramped (slow attack) 4 ms half-life time is shown. In a two-alternative, forced-choice experiment, Patterson's listeners were presented a ramped sinusoid in one interval and a damped sinusoid having the same or greater half-life in the other interval. In one version of the experiment, the listeners were asked to choose the interval containing the sound with the louder tonal component; in another version with the same stimuli, they were asked to choose the interval with the louder drumming component. Between trials, the half-life of the damped sinusoid was varied to determine the matching point, that is, the half-life of the damped sinusoid required to equate the probability of choosing either the ramped or damped stimulus as the one with the louder tonal or drumming component. The matching half-lives of the damped sounds were, on average, about 4 and 2.5 times larger than those of the ramped sounds, respectively.

2.5 Structure and Function of the Auditory System

2.5.1 The Ear

Classically the ear is divided into three sections: the outer ear, middle ear and inner ear as shown at the Fig. 2.4.

The Outer and Middle Ear

The outer ear is composed of the pinna and the auditory canal or meatus. Pinna modifies the incoming sound, particularly at high frequencies which is important for sound localization. The auditory canal acts as a closed tube resonator at one end, enhancing sounds in the range 2–5 kHz. Sound travels down the meatus and evokes the vibrations of the eardrum (or tympanic membrane). These vibrations

are transmitted through the middle ear by three bones called malleus (or hammer), incus (or anvil) and stapes (or stirrup) to the oval window, the port into the inner ear. The eardrum is some fifteen times larger than the oval window, giving an amplification of the pressure of about fifteen compared to the oval window alone. The major function of the middle ear is to ensure the efficient transfer of sound from the air to the fluids of cochlea.

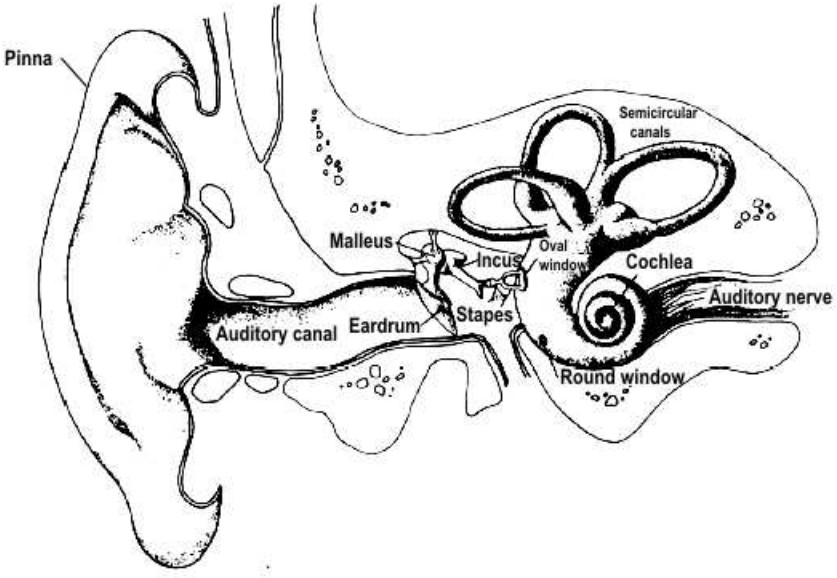

Figure 2.4: *Illustration of the structure of the peripheral auditory system showing the outer, middle and inner ear.*

The Inner Ear

The inner ear is embedded in the extremely hard temporal bone and is basically the same for all mammals. The snail-shaped cochlea of humans forms $2\frac{1}{2}$ turns, is filled with two electrochemically different fluids (perilymph and endolymph) and consists of three different channels of scalae, which run together from the base to the apex. The footplate of the stapes is in direct contact with the fluid in the scala vestibuli which is separated from the scala media only by the very thin and light Reissner's membrane. The fluid oscillations are transmitted to the basilar membrane which separates the scala media from the the scala tympani and supports the organ of Corti which contains about 3,000 inner and 16,000 outer haircells. They transform the mechanical oscillations in the inner ear into a signal (action potentials), that can be processed by the nervous system. The basilar membrane of humans has a total length of 32 mm and widens with a factor of about three from the base to the apex. The place theory of pitch perception suggests that pitch is determined by the place along this collection at which excitation occurs. High frequency sounds selectively vibrate the basilar membrane of the inner ear near the entrance port (the oval window). Lower frequencies travel further along the membrane before causing appreciable excitation of the membrane.

2.5.2 The Auditory Pathway

The auditory pathway (Figure 2.5) differs significantly from the visual and somatosensory pathways in that there is no large direct pathway from peripheral receptors to the cortex. Rather, information ultimately reaching the auditory cortex undergoes significant reorganization as it passes through the brainstem (Moore, 1994).
The auditory nerve (AN) enters the brain stem at the ponto-medullary junction and synapses on the ipsilateral cochlear nucleus (CN). It is made up of approximately 30,000 nerve fibers with afferent (sensory) portions projecting from cell bodies in the spiral ganglion and efferent axons coming from cells in the olivary complex. Organized into subdivisions, which can be identified by morphologically distinct cells, CN show the specific frequency response characteristics. Low frequency sound waves are processed in the ventral, rostral and lateral parts of each cochlear nucleus, while the dorsal, caudal and medial subdivisions respond to high frequency sound waves. From the CN, the majority of the auditory fibers project in the superior olivary complex (SOC), a group of nuclei, which is also

located in the medulla. Each of the superior olive receives projections from both the ipsylateral and contralateral CN. Because the contralateral projections are larger than the ipsilateral, the primary auditory cortex (AC) primarily receives information from the ear of the contralateral side of the head.
Arising from the SOC, fibers project through the lateral lemniscus (LL) tract to the inferior colliculus (IC) which is a component of the tectum of the dorsal midbrain. Studies in which the SOC is selectively damaged have demonstrated that it is essential in the localization of the source of a sound. Furthermore, it has been shown that lesions above the SOC (in the LL and IC) result in a loss of the ability to localize sound in the contralateral hemifield (Zatorre and Penhune, 2001). This along with the characteristics of the responses in the SOC support the concept that the SOC performs processing essential to determining the location of a sound. In addition, it seems that projections from the lateral olivary complex (LOC) to the IC provide a cross over point for information involved in source localization. This allows the segregation of information according to hemifield rather than ear source, much as the optic chiasm does in the visual system. Other types of information gathered from the ear (such as timber and frequency) however appear to cross over at different points or not cross over at all.
From the IC, two distinct pathways project to the ventral and dorsal regions of the medial geniculate nuclei (MGN) of the thalamus. While projections from the dorsal MGN reach the secondary auditory cortex, regions from the ventral part of the MGN project to the primary AC. The MGN also projects to subcortical structures in the frontal lobe as well as the dorsal amygdala and posterior neostriatum.

2.5.3 The Auditory Cortex

It is well established that the main human cortical auditory areas are located in the superior portion of the temporal lobe, including parts of the supratemporal plane and the supratemporal gyrus (STG). The organization of the human auditory cortex shows similarities with the auditory cortex of the macaque monkey, which has been parcelled into 12 areas organized in three regions: the core line region, containing the primary auditory cortex (PAC) and two more anterior areas, the rostral area (R) and the rostro–lateral area (RL), is surrounded by the belt region (itself subdividedminto as many as eight distinct areas) which in turn is boarded laterally by two areas forming the parabelt region in the STG (Pandya, 1995; Rauschecker, 1998a). Cytoarchitectonic studies in humans

(Galaburda and Sanides, 1980; Rademacher et al.; 1993; Pandya, 1995) have described a highly granular konio–cortex located in the first transverse gyrus of Heschl (H1), occupying about two-thirds of this structure. This area presumably forms the PAC, as also suggested by intracerebral recordings (Liégeois-Chauvel et al.; 1991). The first transverse gyrus is delimited anteriorly by the first temporal sulcus (TS1), and posteriorly by Heschl's sulcus (HS1). Although additional transverse gyri (H2, H3) might be present posteriorly to H1 (Rademacher et al., 1993; Penhune et al.,1996; Leonard et al., 1998), it has been suggested that the primary area remains within H1, occasionally extending to the posterior bank of HS1 (Rademacher et al., 1993). The core line is bounded medially by a more primitive area of prokonio–cortex located in the insular circular sulcus, called the 'root' area. On its other edges, the PAC is surrounded by several secondary belts of parakonio–cortex (associative areas), anteriorly toward the pole, laterally by the STG, and posteriorly by H2-H3 (if present) and by planum temporale (PT). Areas H1 and PT have been reported to be highly variable across individuals (Penhune et al., 1996; Leonard et al., 1998; Schneider et al., 2005).

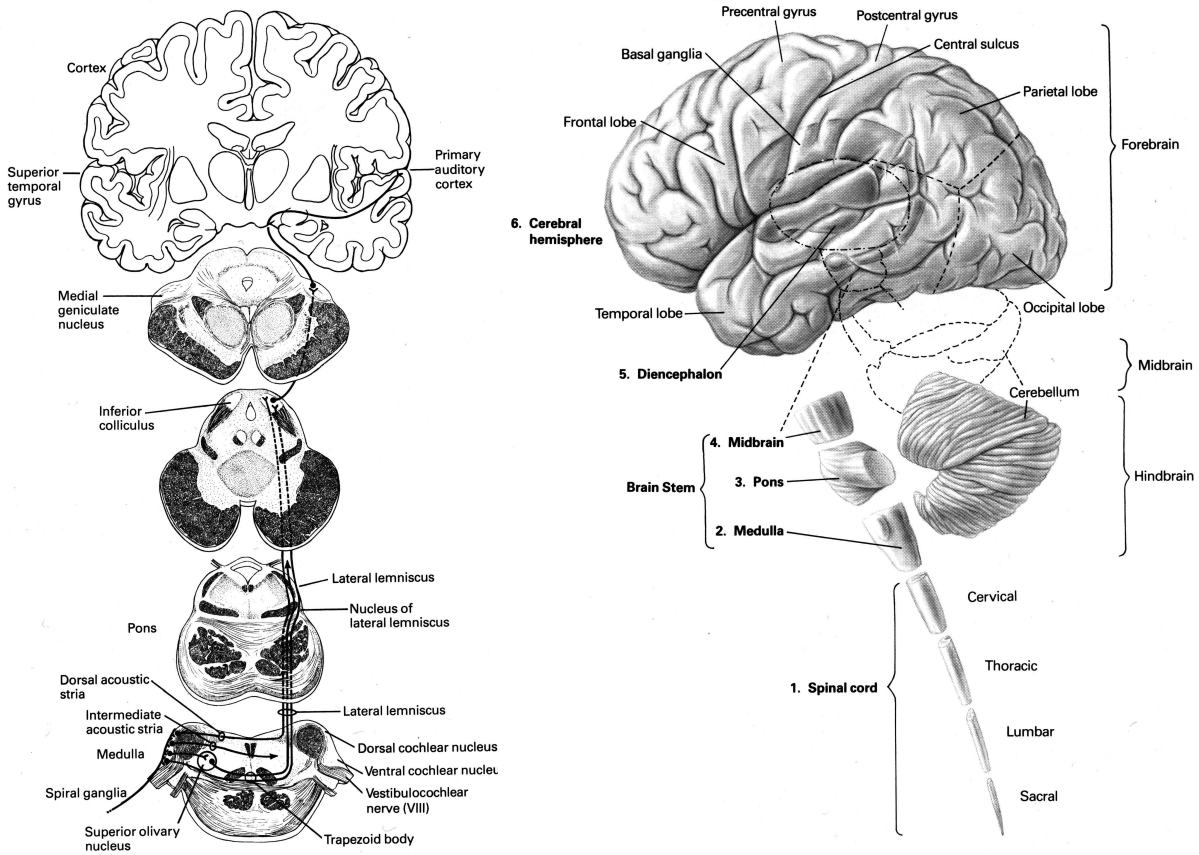

Figure 2.5: *Left: Schematic overview of the afferent connections in the central auditory pathway, starting from the brainstem (Superior olivary nucleus) up to the primary auditory cortex. Right: Lateral view of the human cortex. After Kandel and Schwartz (1985), page 406.*

Chapter 3

Magnetoencephalography

Magnetoencephalography (MEG) is a noninvasive technique used to measure magnetic fields generated by small intracellular electrical currents in neurons of the brain. The aim is to pick up the magnetic field at several locations outside the head and then to calculate the most probable source currents. Because of its submillisecond temporal accuracy, MEG represents a direct measure of the dynamics of evoked and spontaneous neural activity and provides the location of their sources.

3.1 Biological Sources

To understand the principles behind the MEG we should first consider brain electrophysiology. In MEG, the brain part of interest is the uppermost layer, the cerebral cortex, which is 2–4 mm thick and has a total surface area of about 2500 cm^2 folded to fit into the cranial cavity of the skull. The cortex consists mainly of neurons, connected among each other by a synapses. There are about 10^{10} neurons in the cortex connected by about 10^{14} synapses.
The most likely source of neural currents are the pyramidal cells in the thin cortical layers. As shown in Fig. 3.2 these cells are characterized by long apical dendrites.

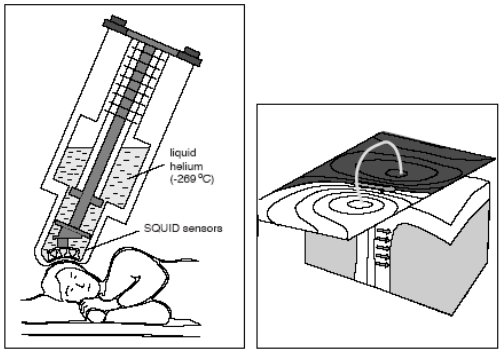

Figure 3.1: *MEG measurement setup. Left: The magnetic fields produced by neuronal currents are picked up with an array of superconducting sensors. Right: The most probable current configuration in the brain (small arrows) is calculated on the basis of the measured field pattern, in this case a current dipole model (large arrow) was used.*

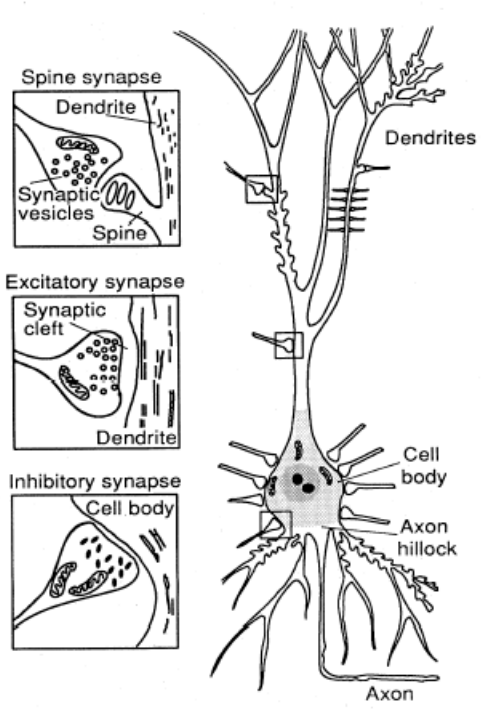

Figure 3.2: *Schematic illustration of a pyramidal neuron and three magnified synapses.The presynaptic action potential propagates along the axon. At the synaptic cleft, the liberated chemical transmitters diffuse to the postsynaptic membrane and initiate a postsynaptic action potential. Potential differences due to propagating action potentials, can be detected. (modified from Iversen, 1979).*

Extraordinarily small currents on the order of tens of picoamperes ($1\text{pA} = 10^{-12}\text{A}$) flow in these dendritic branches as a function of excitatory postsynaptic potentials (EPSPs). Consequently, the extracerebral magnetic field due to a single postsynaptic potential also is very weak: only of the order of attoteslas ($1\text{aT} = 10^{-18}\text{T}$). Many thousands of postsynaptic currents have to exist synchronously for the magnetic field to be strong enough to be detected even with the best of the instruments. Nominal calculations of neuronal density and cortical thickness suggest that the cortex has a macrocellular current density on the order of 100 nA/mm^2. If we assume the cortex is about 4 mm thick, then a small patch 5 mm x 5 mm would yield a net current of 10 nAm, consistent with empirical observations and invasive studies (Hämäläinen et al., 1993).

3.2 SQUID magnetometers

Figure 3.3 shows that magnetic signals from the human brain are extremely weak compared with ambient magnetic field variations and also compared with magnetic signals from other parts of the body.

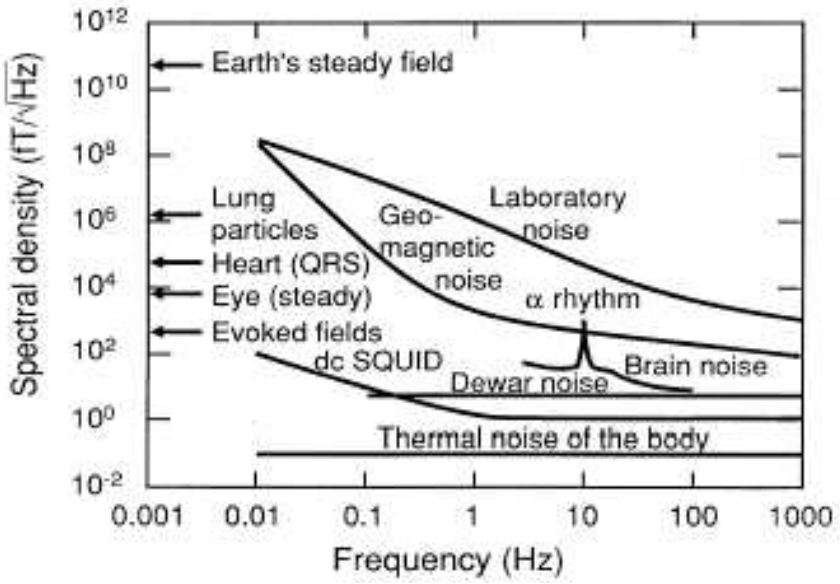

Figure 3.3: *Peak amplitudes (arrows) and spectral densities of fields due to typical biomagnetic and noise sources.*

The measurement of such a weak magnetic fields became possible through the development of very sensitive magnetometers, the superconducting quantum interference devices (SQUIDs). A SQUID magnetometer consists of a superconducting loop of wire (the detector coil) connected inductively to the SQUIDs loop. In the superconductor electrons form Cooper pairs which condense and are the carriers of the current in the superconductor. The SQUID measures the phase difference of the wave function of a Cooper pair at the junction. This phase difference is given by:

$$\Delta\phi = \frac{-iq}{\hbar c}\Phi, \tag{3.1}$$

where $\int_F B\mathrm{d}f$, the magnetic flux and $q = 2e$, the charge of a Cooper pair.
The phase difference 2π is reached for a magnetic flux quant with the flux Φ_m:

$$\Phi_m = \frac{\hbar c 2\pi}{2e} \approx 4 \cdot 10^{-7}\mathrm{Gscm}^2 = 4 \cdot 10^{-15}\mathrm{Wb}. \tag{3.2}$$

Phase differences $\Delta\Phi \ll \Phi_m$ lead to detectable interference effects at the junction. The sensitivity is indeed in the femto-Tesla region, as can be seen from the equation 3.2.

3.3 Quasistatic Approximation of Maxwell's Equations

Neuronal currents generate magnetic and electric fields according to Maxwell's equations. The neural current distribution can be conveniently described as the primary current, the "battery" in a resistive circuit comprising the head. The postsynaptic currents in the cortical pyramidal cells are the main primary currents giving rise to measurable MEG signals. The electric field **E** and the magnetic field **B**, induced by the total electric current density, **J**, can be solved from Maxwell's equations. The relevant frequency spectrum for electrophysiological signals in MEG and electroencephalography (EEG) is typically below 1 kHz, and most studies deal with frequencies between 0.1 and 100 Hz. Consequently, the physics of MEG and EEG can be described by the quasistatic approximation of Maxwell's equations. Thus inductive, capacitive and displacement effects can be neglected, and the true time-dependent terms in the field equations can be omitted. In the quasistatic approximation the Maxwell's equations read:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \nabla \times \mathbf{E} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} \end{aligned} \tag{3.3}$$

where ρ is the charge density and ϵ_0 and μ_0 are the permittivity and permeability of vacuum, respectively.
The quasistatic current flow $\mathbf{J}(\mathbf{r}')$ at location $\mathbf{r}'$ is therefore divergence free and can be related rather simply to the magnetic field at location r' through the Biot-Savart law,

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dv. \tag{3.4}$$

The total current density, $\mathbf{J}(\mathbf{r}')$, can be divided into two components, the passive volume or return current $\mathbf{J}^{\mathbf{V}}(\mathbf{r}')$ and the primary current $\mathbf{J}^{\mathbf{P}}(\mathbf{r}')$(see Fig. 3.4).

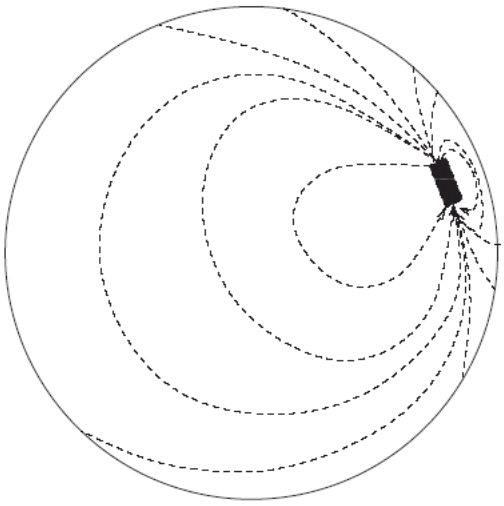

Figure 3.4: *Current distribution in the brain initiated by a primary current in the black box; the dashed lines are the stream lines of the volume current.*

We define the volume current as "passive" current that results from the effect of the electric field on extracellular charge carriers $\mathbf{J}^{\mathbf{V}}(\mathbf{r}') = \sigma(\mathbf{r}')\mathbf{E}(\mathbf{r}')$. From the quasistatic assumption, the electric field is simply the negative gradient of a potential, $E = -\nabla V$, and $\sigma(\mathbf{r}')$is the conductivity, which we will here assume to be isotropic. The primary current flow $\mathbf{J}^{\mathbf{P}}(\mathbf{r}')$ is the "driver" in the current flow

and can be defined at the macroscopic level as simply that portion that is not volume current,

$$\mathbf{J}(\mathbf{r}') = \mathbf{J}^{\mathbf{P}}(\mathbf{r}') + \mathbf{J}^{\mathbf{V}}(\mathbf{r}') = \mathbf{J}^{\mathbf{P}}(\mathbf{r}') - \sigma(\mathbf{r}')\nabla V(\mathbf{r}'). \tag{3.5}$$

The total current density may not have any volume currents (for instance, a closed loop of current), but every current density must have a primary current that generates the total current distribution. In MEG and EEG, we are most interested in the locations of the primary currents, since they represent the regions of active cell assemblies. The volume currents are necessary to "close the loop" to create divergence-free total current densities. We therefore adjust our forward models to emphasize the primary current distribution, but of course every model must still account for the effects of the volume currents. If we assume that the head consists of a set of contiguous regions each of constant isotropic conductivity, we can alter the Biot-Savart law above both to emphasize the primary current and to emphasize the boundary regions of the head. Using standard vector identities, the Biot-Savart law becomes

$$\mathbf{B}(\mathbf{r}) = \mathbf{B_0}(\mathbf{r}) + \frac{\mu_0}{4\pi}\sum_{ij}(\sigma_i - \sigma_j)\int\limits_{S_{ij}} V(\mathbf{r})\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \times d\mathbf{S'_{ij}}, \tag{3.6}$$

where the summation is over all boundaries, and $\mathbf{B_0}(\mathbf{r})$ is the magnetic field due to only the primary current. Typical macroscopic boundaries are the inner skull surface, outer skull surface, and scalp. This general equation states that the magnetic field can be calculated if we know the primary current distribution and the potential $V(\mathbf{r}')$ on all surfaces. We can create a similar equation for the potential on surface S_{ij},

$$(\sigma_i - \sigma_j)V(\mathbf{r}) = 2\sigma_0 V_0(\mathbf{r}) + \frac{1}{2\pi}\sum_{ij}(\sigma_i - \sigma_j)\int\limits_{S_{ij}} V(\mathbf{r}')\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \times d\mathbf{S'_{ij}}, \tag{3.7}$$

where $V_0(\mathbf{r})$ is the potential at $\mathbf{r}$ due to the primary current distribution.
The primary potential $V_0(\mathbf{r})$ is then used to solve (3.7) for the potential on all surfaces, and therefore solves the forward problem for EEG, which represents potential differences on the scalp. These surface potentials $V(\mathbf{r}')$ and the primary magnetic field $\mathbf{B_0}(\mathbf{r})$ are then used to solve (3.6) for the external magnetic fields.

3.4 Source Estimation

3.4.1 Current Dipole Model

Let us assume a small patch of activated cortex is centered at location $\mathbf{r_Q}$ and that our observation point $\mathbf{r}$ is some distance from this patch. If we insert this current source into the primary magnetic field model, the field can be well approximated as

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi}\mathbf{Q}\frac{\mathbf{r}-\mathbf{r_Q}}{|\mathbf{r}-\mathbf{r_Q}|^3}d\mathbf{r}' \tag{3.8}$$

where $\mathbf{Q} = \int \mathbf{J^P}(\mathbf{r}')d\mathbf{r}'$ is defined as equivalent current dipole. The current dipole is a straightforward extension of the better-known model of the paired-charges dipole in electrostatics. Though visually close to some simplistic geometrical representation of pyramidal cell assemblies, it is essential to bear in mind that we are dealing with different scales here. Neural macrocolumns typically extend at the millimeter scale, whereas the focal current dipole is a concept, or a metaphor, equivalent to the unidirectional activation of up to several square centimeters of gray matter. The current dipole model is the basis of EEG/MEG processing, since any arbitrary primary current density can always be broken down into small patches, each patch represented by an equivalent current dipole.
A problem arises, however, when too many of these patches are required to represent a single large patch. These larger patches may be more simply represented by a multipolar model, rather than many individual dipoles.

3.4.2 Head Models

Spherical volume conductor model

Given a specified primary current density, the general solution of the MEG forward model first requires solution of the EEG forward model, which in turn requires the solution of a non-trivial integral equation. If the head can be modeled as a single homogeneous sphere, then analytic solutions of 3.7 are well-known. If the head is a set of concentric homogeneous shells (e.g. brain, skull, and scalp), then the solution involves an infinite series expansion, but simpler approximations have recently been developed. Consider the case of a current dipole in a spherical head model (Fig. 3.5). If we examine the radial component of the magnetic field,

we can show that the contributions of the surface potentials in 3.6 vanish, and we are left with the remarkably simple formula,

$$\mathbf{B}_r(\mathbf{r}) = \frac{\mathbf{r}}{r}\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{\mathbf{r}}{r} \times \mathbf{r_Q}\frac{\mathbf{Q}}{|\mathbf{r} - \mathbf{r}'|^3}. \tag{3.9}$$

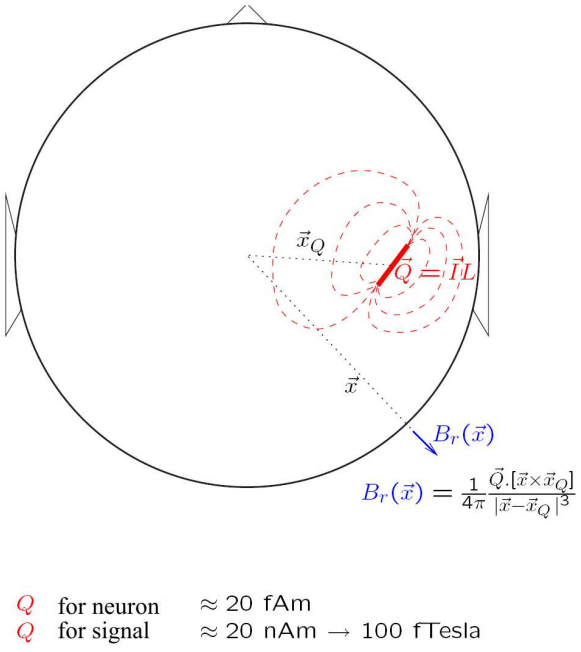

Figure 3.5: *Magnetic field of the single equivalent current dipole model in the spherical volume conductor.*

Thus these radial MEG measurements are a function of the primary current only, such that the volume currents do not contribute to this measurement. The conductivity profiles of the shells have become unimportant in the modeling due to their axial symmetry. However, non-radial measurements must include the effects of volume currents. Because of the simplicity of the radial model and the absence of the conductivity profile in the full field model, a mistaken assumption is sometimes made that MEG is insensitive to volume currents. A further interesting note is that if $\mathbf{Q}$ is radially oriented in the above equation, then $\mathbf{B_r}(\mathbf{r})$ is zero, and therefore so is the total magnetic field outside the sphere. In other words, a radially-oriented dipole inside a spherical conductor creates volume currents that identically cancel the primary field everywhere outside the sphere. However, this cancelation is extremely sensitive to the sphere assumption, and even a slight deviation from spherical allows a radial dipole to generate an external field.

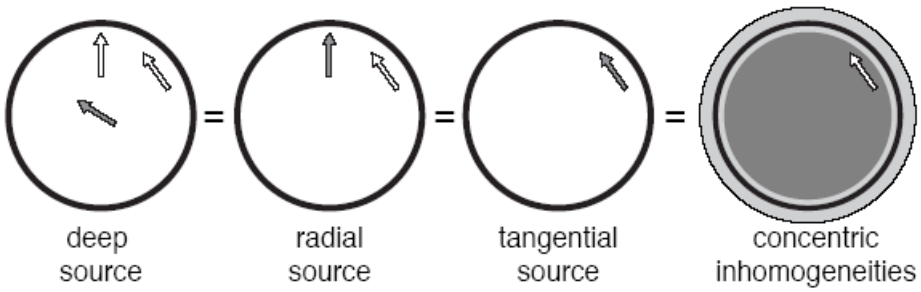

Figure 3.6: *Schematic presentation of the effects of deep, radial, and tangential currents on MEG signals detected outside a spherically symmetric conductor. In all situations the external magnetic field is identical because radial currents anywhere in the sphere do not produce any external magnetic field, sources exactly in the middle of the sphere are always radial, and because concentric inhomogeneities do not affect the magnetic field. EEG would see all these currents (tangential, radial, and deep ones) but would be affected by the electric inhomogeneities.*

3.5 The Inverse Problem

The problem to obtain potentials or fields from given sources is straightforward. Mathematically and physiologically much more interesting is the so called inverse problem, namely to obtain information on the primary from measured potentials or fields on the surface of the head. It has been shown already in 1853 by Helmholtz (Helmholz, 1853) that this problem has no unique solution.
That this is the case follows from the fact that there are non vanishing primary currents which leave no signal on the surface of the head. We have seen already in the previous section that in the spherically symmetric case produces an electric field, but is magnetically silent outside the volume. The case is illustrated in Figure 3.5, when $r \times r_Q$ is zero. Thus, a priori information needs to be incorporated into the analysis to make some assumptions regarding the location or the geometry of the source to derive a reliable solution. The activated neurons are most commonly described in terms of a "single-dipole- in-a-sphere" model. That is a single equivalent dipole, located in the center of a spherically symmetric homogenous medium. The solution of the inverse problem can be obtained with different methods: The Principal Component Analysis (PCA) calculates the autocorrelation function of the sampled signal matrix. Eigenvectors and eigenvalues of this matrix define the topographies, that contribute to the measured signal. Another method is to use the MUSIC-algorithm (multiple signal classification). One assumes, that the sources and the noise are orthogonal to each other and that the number of the sources is known. For a comparison of the advantages and disadvantages of these methods, see Hoechstetter (2001). The brain activity can also be modeled with a distributed source model, in which several current dipoles are densely spaced along the brain's gray matter. A problem occurs when the number of dipoles exceeds the number of sensors that record the field distribution. One has to find an underdetermined problem and needs to place constraints in order to find a solution. One way to overcome the problem is given with the "minimum norm solution", a solution for which the norm of the activities of all sources exhibits its minimum.
The auditory evoked fields (AEFs) have the immense advantage, that the pyramidal cell dendrites of the auditory cortex (the putative generators of the electrical sources) are arranged in a columnar fashion with an orientation normal to the cortical surface.

3.5.1 Lead Fields and Forward Fields

Discrete Fields

Regardless of the specifics of the head model, by electromagnetic superposition the forward model is linear, and we may write the relationship between the moment for a dipole at $\mathbf{r_Q}$ and the measurement at sensor location $\mathbf{r}$ as the inner product of a lead field vector $\mathbf{g}(\mathbf{r}, \mathbf{r_Q})$and the dipole moment $\mathbf{Q}$. We assume dipole moments of the 3x1 Cartesian form $\mathbf{Q} = [Q_x, Q_y, Q_z]^T$, where the individual components represent "elemental dipoles". By analogy, the extension of this model for the three magnetic dipole moments or nine magnetic quadrupole moments is straightforward, but for simplicity we will restrict the discussion to current dipoles. The canonical elemental dipoles form a basis for any dipole located at $\mathbf{r_Q}$ with any orientation and amplitude. This orthogonal dipole triplet is sometimes referred to as a regional source or rotating dipole which serves as a basic focal model for cortical activation within cortical folds in the neighborhood of $\mathbf{r_Q}$. The three components of the lead field vector $\mathbf{g}(\mathbf{r}, \mathbf{r_Q})$ are formed as the solution to either the magnetic or electric forward problem for each of these elemental dipoles. Explicit forms of $\mathbf{g}(\mathbf{r}, \mathbf{r_Q})$ are summarized in Mosher et al. (1999). For many dipoles (each representing a patch of cortex), we simply sum the individual contributions. An EEG or MEG measurement may be therefore represented as $m(\mathbf{r}) = \sum \mathbf{g}(\mathbf{r}, \mathbf{r_{Qi}}) \cdot \mathbf{Q_i}$ where $m(\mathbf{r})$ is the output of the single magnetometer. The measurements by an EEG or MEG array are made at N sensors, so we can write a matrix,

$$\mathbf{m} = \begin{bmatrix} m(\mathbf{r_1}) \\ \vdots \\ m(\mathbf{r_N}) \end{bmatrix} = \begin{bmatrix} \mathbf{g}^T(\mathbf{r_1}, \mathbf{r_{Q,1}}) & \dots & \mathbf{g}^T(\mathbf{r_1}, \mathbf{r_{Q,p}}) \\ \vdots & \ddots & \vdots \\ \mathbf{g}^T(\mathbf{r_N}, \mathbf{r_{Q,1}}) & \dots & \mathbf{g}^T(\mathbf{r_N}, \mathbf{r_{Q,p}}) \end{bmatrix} \begin{bmatrix} \mathbf{Q_1} \\ \vdots \\ \mathbf{Q_p} \end{bmatrix} = \mathbf{GQ}$$

where $\mathbf{G}$ is the "gain matrix" relating the set of p dipoles to the set of discrete sensor locations, and m is a generic set of N MEG or EEG measurements. Each column of $\mathbf{G}$ relates an elemental dipole to the array of sensor measurements and may be called the forward field, gain vector, or scalp topography, of the current dipole source. This forward field is effectively sampled by the discrete locations of the sensors. Alternatively, we may look along a row of $\mathbf{G}$ and observe that we are calculating the lead field vector for the same sensor location and many different dipole locations. For reasons of reciprocity, these discrete dipole locations are effectively samples of the lead field that would be generated if the sensors were used to generate currents in the brain.

Spatio-temporal Models

The above model is easily extended to multiple time samples. The most common model is to *add an each successive vector into a spatio-temporal data matrix.* The observed set of measurements over an N-sensor array for p sources can be expressed as a linear forward spatio-temporal model of the form $\mathbf{M} = \mathbf{GQ}$, where the observed forward field $\mathbf{M}$ (N-sensors x T-time samples) can be expressed in terms of the gain matrix $\mathbf{G}$ (N-sensors x $3p$-elemental dipoles) and a set of dipole moments $\mathbf{Q}$ ($3p$ x T). In this model, each row of the matrix $\mathbf{Q}$ represents the time series for an elemental dipole (e.g. x, y, or z directed). The three elemental dipoles at a fixed location represent a dipole of arbitrary orientation, but if the three time series associated with the elemental dipoles are allowed to have arbitrary time series, then the dipole appears to "rotate" in its position as a function of time. A common practice, therefore, in spatio-temporal modeling is to constrain each dipole to a "fixed" orientation by factoring out a unit orientation vector $\mathbf{u}$, such that time series for each dipole may be represented as

$$\begin{bmatrix} Q_x(1) & \dots & Q_x(T) \\ Q_y(1) & \dots & Q_y(T) \\ Q_z(1) & \dots & Q_z(T) \end{bmatrix} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} \begin{bmatrix} s(1) & \dots & s(T) \end{bmatrix} = \mathbf{us}^T \tag{3.10}$$

where $\mathbf{u}$ is constrained to have unity norm. For p sources and T time samples, all with fixed orientation, the spatiotemporal model can therefore be modified as

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} m(\mathbf{r_1}, 1) & \vdots & m(\mathbf{r_1}, T) \\ \vdots & \ddots & \vdots \\ m(\mathbf{r_N}, 1) & \vdots & m(\mathbf{r_N}, T) \end{bmatrix} = \begin{bmatrix} \mathbf{g}^T(\mathbf{r_1}, \mathbf{r_{Q,1}})\mathbf{u_1} & \dots & \mathbf{g}^T(\mathbf{r_1}, \mathbf{r_{Q,p}})\mathbf{u_p} \\ \vdots & \ddots & \vdots \\ \mathbf{g}^T(\mathbf{r_N}, \mathbf{r_{Q,1}})\mathbf{u_1} & \dots & \mathbf{g}^T(\mathbf{r_N}, \mathbf{r_{Q,p}})\mathbf{u_p} \end{bmatrix} \begin{bmatrix} \mathbf{s_1}^T \\ \vdots \\ \mathbf{s_p}^T \end{bmatrix} \\ &= \mathbf{A}(\{\mathbf{r_i}, \mathbf{u_i}\})\mathbf{S}^T \end{aligned} \tag{3.11}$$

The source parameters that define our forward model are the set of dipole locations and orientations $\{\mathbf{r_{Qi}}, \mathbf{u_i}\}$ for p sources, and the corresponding time series for each dipole, which are the columns of the time series matrix $\mathbf{S}$. Each column of $\mathbf{A}$ now corresponds to a single dipole at location $\mathbf{r_{Qi}}$ with orientation $\mathbf{u_i}$. Embedded in the lead field vector $\mathbf{g}(\mathbf{r_i}, \mathbf{r_{Qj}})$ are all of the important head model parameters that convert these dipolar sources into measurements.

3.6 MEG Hardware and Environment

The first biomagnetic measurements were performed with single-channel instruments (Cohen 1968; Cohen 1972). However, reliable localization of current sources requires mapping at several locations, which is time-consuming with only one channel. Besides, unique spatial features of, for example, ongoing background neuromagnetic signals cannot be studied. During the last fifteen years multichannel SQUID systems for biomagnetic measurements have developed to reliable commercial products. A state-of-the art multichannel MEG system comprises more than one hundred channels in a helmet-shaped array to record the magnetic field distribution all around the brain simultaneously. The dewar containing the sensors is attached to a gantry system, which allows easy positioning of the sensors above the subject's head. In our laboratory we use the whole head Neuromag 122 channels (Neuromag Ltd., Helsinki) device which comprises a total of 122 SQUID sensors, each containing two orthogonal planar gradiometers. The sensors are arranged into a helmet shaped array, which covers even the most peripheral areas of the brain, such as frontal and temporal lobes and the cerebellum. For additional rejection of external disturbances, MEG measurements are usually performed in a magnetically shielded room. Four different methods can be employed to construct such an enclosure: ferromagnetic shielding, eddy-current shielding, active compensation, and the recently introduced high-T_c superconducting shielding. Combinations of these techniques have been utilized in many experimental rooms (Cohen 1970a,b). A typical commercially available room utilized in biomagnetic measurements employs two layers of aluminum and ferromagnetic shielding. The inside floor area is usually 3 by 4 m, and the height is around 2.5 m. In addition to passive shielding, the external disturbances can be also canceled using active electronic circuits which are either independent of the actual MEG system or integrated with it. The most traditional active compensation system consists of flux-gate magnetometers and pairs of orthogonal Helmholtz coils positioned around the shielded room. The output of the field sensors is employed to control the current fed to the coils to cancel the detected noise.

3.7 Auditory Evoked Neuromagnetic Fields

The electromagnetic responses evoked from the human brain by auditory stimuli can be classified as transient, sustained or steady state. Transient responses are brief changes in the electromagnetic fields evoked by the onset or offset of a

stimulus or by some change in an ongoing stimulus. These responses are usually recorded in the time domain as a waveform with peaks and troughs occurring at latencies from one to several hundred milliseconds. Sustained potentials are continuous shifts in the recorded level that last through the duration of the stimulus. Steady state responses are responses to recurring stimuli that have settled so that they maintain a stable amplitude and phase relationship to the stimuli. These responses are often recorded in the frequency domain as peaks in a spectrum, with frequencies that are harmonically related to the stimulus rate.
Transient responses (see Figure 3.7 characterized by a polarity and either a location in a sequence of waves or a characteristic latency. For example, the N1m or N100m, the first prominent negative wave in the slow onset-response, has a characteristic peak latency of 100 ms (Näätänen and Picton, 1987). Auditory steady state responses can be recorded in response to a rapidly repeating stimulus. They are most prominent at frequencies of stimulation near 40 Hz, but they also occur at more rapid frequencies.

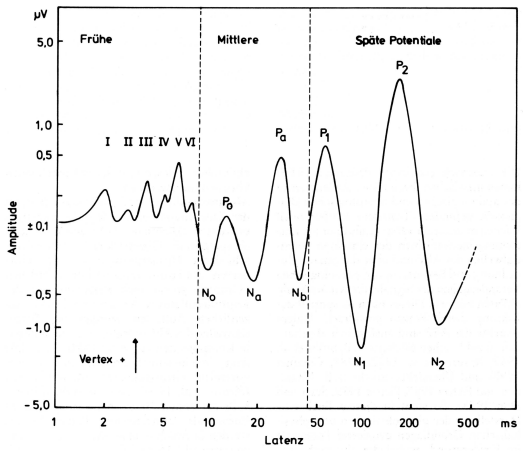

Figure 3.7: *Auditory evoked potentials as a response to a sound*

Chronology of Activation in the Human Auditory Cortex

An important issue of the investigation of the human auditory cortex is to detect precisely when and in which temporal order the cortical auditory areas are activated after stimulus onset. Electrophysiological studies (Picton, 1974; Wood and Wolpaw, 1982) have described that after stimulus onset the early brainstem evoked responses (occurring before 10 ms) are followed by different evoked components originating from the auditory cortex. The responses elicited between 10 and 70 ms after stimulus onset are termed middle latency components (MLCs),

whereas later components in the range 80-250 ms, are called long latency components (LLCs). MLCs elicited by clicks are usually described as a succession of three waves labeled as N19m (or Na) (latency: 19 ms), P30m (or Pa) (30 ms) and P50m (or Pb or P1) (50–70 ms). MLCs are followed by the N100m (or N1m) component having a latency around 100 ms and P200m component with the latency of around 200 ms (see Fig. 3.8)
The origin of the P30m is located in the posterior part of medial Heschl's Gyrus (Liégeois–Chauvel et al., 1994). There are general disagreement about the location of the generator of the P50m source. Reite et al. (1988) found the source of P50m to be localized in Planum temporale. Liégeois-Chauvel et al. (1994)

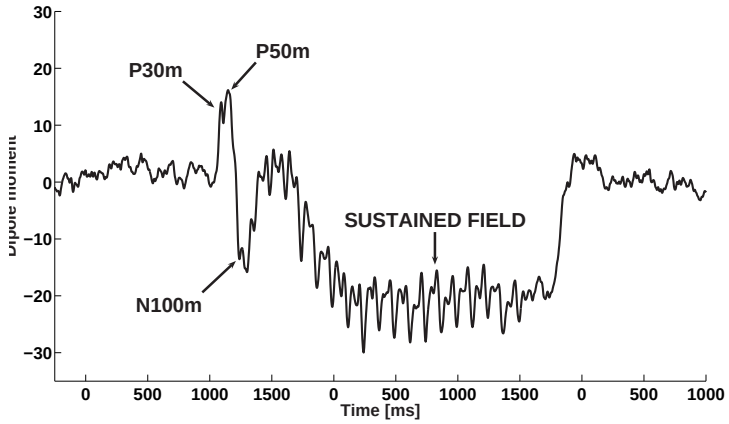

Figure 3.8: *Typical time course of auditory evoked responses is shown. The neuromagnetic responses were evoked by an exponentially amplitude modulated tone*

.

reported the generator of the P50m to be more lateral compared to the P30m source, which was validated in MEG measurements by Mäkelä et al. (1994), Scherg et al. (1989), and can be seen in Figure 3.9. The N100m component is composed of multiple, partially temporally overlapping, independent components (Näätänen and Picton, 1987; Lütkenhöner, 2001). Pantev et al. (1988) and Roberts and Poeppel (1996) showed tonotopical organization of the N100m generator. However Lütkenhöner argued, showing that the dipole location exhibited no significant frequency dependence at all. It has been shown that N100 peak amplitudes increase with the increasing stimulus duration (Kodera et al. 1979) but its latency remains the same (Jutsiniemi et al. 1989). Furthermore, the N100m increase with the increasing inter-stimulus interval (Polich et al 1988) and stimulus level (Biermann and Heil, 2000). The multiple sources of the N100m component are probably generated over a wide region of the supratemporal plane.

Magnetic field recordings by Lütkenhöner and Steinsträter (1998) suggested that the N100m arises from Planum temporale and that the P200m seems to have its center of activity in Heschl's Gyrus.

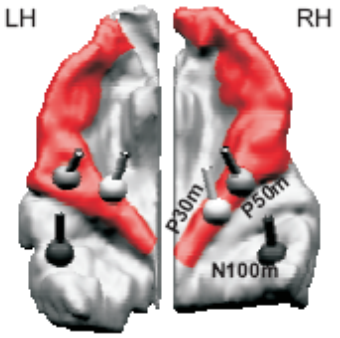

Figure 3.9: *Equivalent source dipoles of MEG components P30m, P50m and N100m fitted on the single components. The grand average locations of the corresponding dipoles are shown. Figure adopted from Schneider et al. (2004).*

Chapter 4

Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is an imaging technique used primarily in medical settings to produce high-quality images of the inside of the human body. MRI is based on the principles of nuclear magnetic resonance (NMR), a spectroscopic technique used to obtain microscopic chemical and physical information about molecules. The technique was called magnetic resonance imaging rather than nuclear magnetic resonance imaging (NMRI) because of the negative connotations associated with the word nuclear in the late 1970's. MRI started out as a tomographic technique, that is it produced an image of an NMR signal in a thin slice through the human body. Magnetic resonance phenomenon was discovered in 1946 by Felix Bloch and Edward Purcell, both of whom were awarded the Nobel Prize in 1952.

4.1 Physical Basics

Placing an object composed of magnetic interactive atoms in a static magnetic field, stimulating it with another oscillating one, excites a recognizable rotating magnetization in the object (see Figure 4.1). This phenomenon is based on an effect called nuclear magnetic resonance occurring in magnetic active atoms. Measurement of the induced magnetic field reveals information about the chemical composition of the object. In order to understand the macroscopic characteristics of a magnetized system our observations have to be taken down to the microscopic level of atoms and their spins. Consideration of their response to magnetic fields and their interaction in a whole system of atoms will illuminate the actually observed magnetization.

4.1.1 Nuclear Magnetic Resonance

Nuclei with odd atomic weight and/or atomic numbers, such as the nucleus of the hydrogen atom, possess an angular momentum I referred to as spin. γ, the gyromagnetic ratio, is a physical constant dependent on the nuclei involved and essentially proportional to q/m (=charge/mass). They are classified by their quantum spin number I, of which the value is a multiple of 1/2 (i.e. $I = \pm 0, \pm 1/2, \pm 1, \dots$). Nuclei with spin other than zero rotate around their own axis creating a magnetic moment μ, which is related to the angular moment by

$$\boldsymbol{\mu} = \gamma I \hbar, \tag{4.1}$$

($\hbar = h/2\pi$, where h is the Planck's constant (=6.6×10^{-34} Js). It is commonly given in MHz/T. In table 4.1 the values for some diagnostically relevant nuclei are listed.

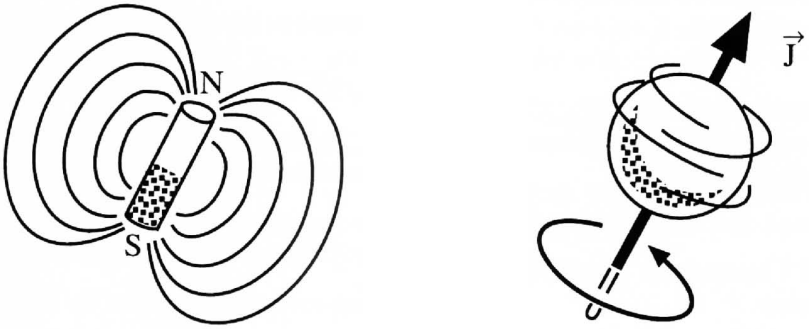

Figure 4.1: *Nuclear spin results in a microscopic magnetic field. Its direction and strength is represented by orientation and length of a vector, respectively (adapted from Liang and Lauterbur, 2000)*

While the magnitude of the magnetic moment μ, can be given by

$$\mu = |\boldsymbol{\mu}| = \frac{\gamma\sqrt{I(I+1)}}{2\pi}h, \tag{4.2}$$

its direction is completely random in the absence of external magnetic fields. No magnetic activation can be observed around a macroscopic object composed of nuclear magnetic active atoms. Exposing such objects to a strong magnetic field B_0 of strength B_0 causes the spin vectors to line up. By convention we assume

B_0 parallel to the z-direction of the laboratory frame:

$$B_0 = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}. \tag{4.3}$$

Due to the laws of quantum mechanics the magnetic moments show the discrete

Nucleus	Spin	Gyromagnetic Ratio (MHz/T)
^1H	1/2	42.58
^{13}C	1/2	10.71
^{19}F	1/2	40.05
^{31}P	1/2	11.26

Table 4.1: *Gyromagnetic ratio γ for some NMR active nuclei relevant in biomedical research and clinical applications.*

energy Eigenvalues E_m,

$$E_m = -\frac{\gamma h}{2\pi} B_0 m_I, \quad m_I = -I, -I+1, \ldots, I-1, I. \tag{4.4}$$

Thus, for a nucleus with $I = \frac{1}{2}$ two energy levels are allowed which result in a parallel or anti-parallel positioning of the moments (Figure 4.2). The energy gap is proportional to B_0:

$$E = \frac{\gamma h}{2\pi} B_0, \tag{4.5}$$

By emission or absorptions of photons with the energy $\Delta E = \frac{1}{2\pi} h\omega_0$ transitions between the two levels occur, so that by irradiation of an electromagnetic field with the frequency ω_0 corresponding to ΔE the spins can be forced to transit from the state of lower energy to the state of higher energy. This resonant transition is the principle of nuclear magnetic resonance. The average motion described by the magnetic moment of each nucleus in the object exposed to the magnetic field B_0, is a precession at angle α about the z-axis (nuclear precession). Driving force is the exerted torque T on the magnetic moment:

$$T = \boldsymbol{\mu} \times B_0 \tag{4.6}$$

The resulting change in the angular momentum is given by:

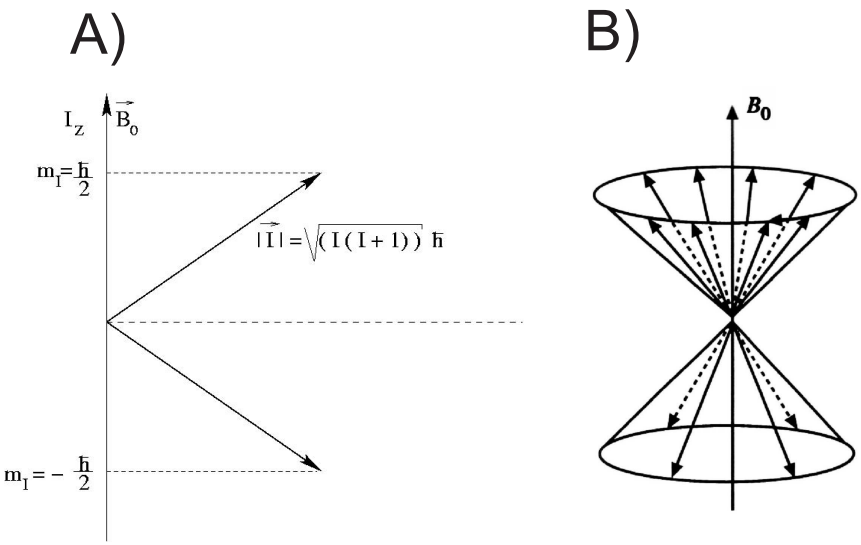

Figure 4.2: *Exposed to a static magnetic field B_0 the magnetic moment vectors of magnetic active nuclei align. (A) Two different energetic levels are possible for nuclei of spin 1/2. $\hbar$ denotes $h/2\pi$. (B) Distribution of the moments at an arbitrary time instant.*

$$\begin{aligned} \frac{dI}{dt} &= T = \boldsymbol{\mu} \times B_0 \\ \frac{d\boldsymbol{\mu}}{dt} &= \boldsymbol{\mu} \times \gamma B_0 = \boldsymbol{\mu} \times \boldsymbol{\omega_0} \end{aligned} \tag{4.7}$$

where the precession velocity $\boldsymbol{\omega_0}$ is known as Larmor frequency, and determined by the strength B_0 of the static magnetic field:

$$\boldsymbol{\omega_0} = \gamma B_0. \tag{4.8}$$

A snapshot of a number of spins will be a set of vectors spreading in two precessing cones (Fig. 4.2(B)). Establishment of full phase coherence among these spins is what is referred to as resonance.
The sum of all magnetic moments in the object result in an external bulk magnetization M pointing along the direction of the applied magnetic field B_0. It is directly proportional to the total number of spins N_s and the strength of B_0:

$$M = \sum_{i=1}^{N_s} \mu_i \tag{4.9}$$

Equation 4.7 remains valid, so that:

$$\frac{dM}{dt} = \boldsymbol{\omega_0} \times M = -\gamma \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix} \times M. \tag{4.10}$$

This system of differential equations

$$\begin{aligned} \dot{\mu_x} &= \omega_0 \mu_y \\ \dot{\mu_y} &= \omega_0 \mu_x \\ \dot{\mu_z} &= 0 \end{aligned} \tag{4.11}$$

is solved by a magnetic moment of the form

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_{xy} cos(\omega_0 t) \\ \mu_{xy} sin(\omega_0 t) \\ \mu_z \end{pmatrix} \tag{4.12}$$

Equation 4.8, the Larmor equation, is the physical basis for achieving nucleus specificity, since it reveals that the resonance frequency is linearly dependent on the magnetic field B_0 and the gyromagnetic ratio γ only. In practice, the different nuclei in a molecule are surrounded by orbiting electrons shielding B_0 to an extent depending on their positions in the molecule. This weakens or increases the NMR–effect. Accordingly, a specific spin system may have a range of different resonance frequencies, not only due to inhomogeneities of B_0 but whenever the object is subject to different chemical particles (chemical shift). Knowledge of these chemical shift frequencies and the corresponding spin densities is of great importance for the determination of chemical structures and is subject to NMR spectroscopy.

4.1.2 Excitation

To achieve resonance - that is to establish phase coherence among the spins - an external force is necessary. This force is provided by an oscillating magnetic field B_1(t). The resonance condition is that B_1(t) rotates with the same frequency as the precessing spins. Its effective field strength therefore has to vary according to

$$B_1(t) = B_1^{en}(t) e^{-i(\omega t + \phi)}. \tag{4.13}$$

ω denotes the excitation carrier frequency and ϕ the initial phase angle, which is assumed to be zero in the following. According to the resonance condition, ω must be chosen to equal the Larmor frequency of the atom aimed to be measured. The pulse envelope function B_1^{en}(t) determines the external shape of the applied field. It is responsible for exciting nuclei resonating in the desired frequency range. Due to the chemical shifts and inhomogeneities of B_0 mentioned above the design of an envelope function providing good selectivity is a difficult

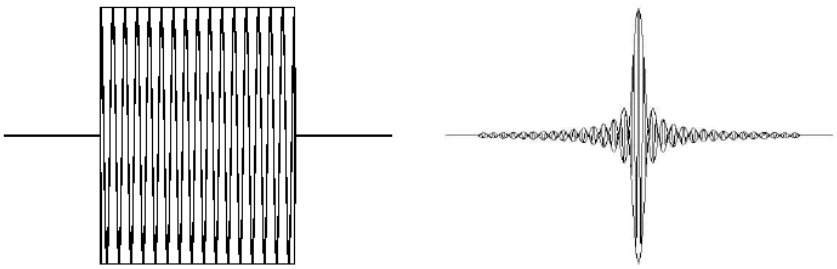

Figure 4.3: *Examples for RF pulses: boxcar-shaped (left) and sinc-shaped envelope function (right)*

task and still subject to research. Since $B_1(t)$ is only applied very shortly (a few micro or milliseconds) and oscillates in the radio frequency range it is commonly termed RF pulse (see Fig. 4.3). While the strength of the permanent field typically ranges from 0.1 T and 3 T the B_1(t) field is only of some mT strength. In the NMR system the RF pulse is generated by a transmitter coil placed around the object to be imaged.
If B_1 is applied perpendicular to the static field, the motion of the bulk magnetization M changes according to

$$\begin{aligned} \frac{dM}{dt} &= -\gamma M \times B(t) \\ B(t) &= \begin{pmatrix} B_1 cos(\omega t) \\ B_1 sin(\omega t) \\ B_0 \end{pmatrix} \end{aligned} \tag{4.14}$$

denotes the effective magnetic field vector. This means that during the RF pulse, while rotating with Larmor frequency, the axis of M is continuously receding from the z-axis with angular frequency $\boldsymbol{\omega_1} = -\gamma B_1$. In order to simplify visualization of the situation a rotating frame of reference is introduced, in which the transverse plane of the coordinate system is assumed to be rotating constantly at an angular frequency ω. To distinguish the new setting from the conventional stationary (laboratory) frame, the three axis are denoted by x', y' and z'. Correspondingly we refer to their orthogonal unit directional vectors as i', j' and k'. In this revolving system the motion of M can be shown as a simple precession of M about the x'-axis with angular frequency ω_1 as shown on the Fig. 4.4:

$$\frac{dM'}{dt} = -\gamma B' \times M' = \boldsymbol{\omega'} \times M' \tag{4.15}$$

where

$$B' = \begin{pmatrix} cos(\omega t) & sin(\omega t) & 0 \\ -sin(\omega t) & cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} B = \begin{pmatrix} B_1 \\ 0 \\ B_0 \end{pmatrix}. \tag{4.16}$$

The inclination angle $\theta(t)$ of the motion of M' is obviously:

$$\theta(t) = \omega_1 t = -\gamma B_1 t \tag{4.17}$$

and therefore dependent on the duration and magnitude of the RF pulse. In practice the pulse width (duration) is chosen according to the desired frequency, while varying its power (magnitude) results in the desired flip angle, e.g.

$$\begin{aligned} B_1\delta(t) = -\frac{\pi}{2\gamma} \rightarrow \phi = \frac{\pi}{2} \\ B_1\delta(t) = -\frac{\pi}{\gamma} \rightarrow \phi = \pi \end{aligned} \tag{4.18}$$

Since image formation is technically not possible performing the measurement

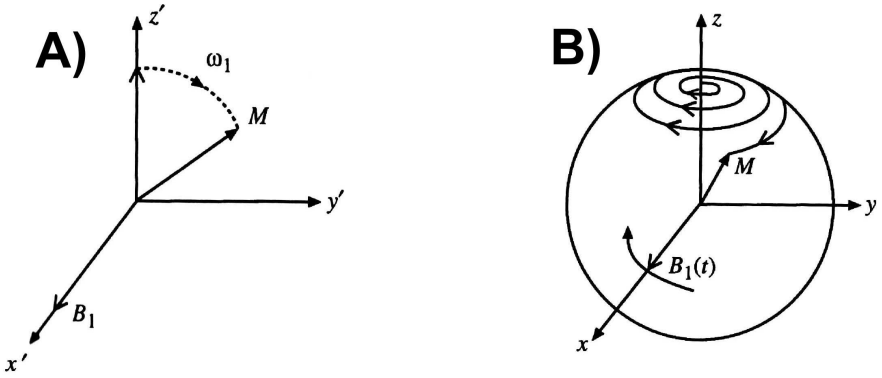

Figure 4.4: *Motion of the bulk magnetization vector M in the presence of a rotating RF field as observed in (A) the rotating frame, and (B) the laboratory frame (adapted from Liang and Lauterbur, 2000).*

while permanently sending RF-pulses knowledge of the evolution of the spin behavior after having switched of the RF pulse, that is the relaxation process, is of importance.

4.1.3 Relaxation

After a magnetized spin system has been perturbed it will return to its initial state of thermal equilibrium once the external force has been removed. The process

of mediating an externally forced magnetization or the decay of an internally remaining magnetization is called relaxation. In our case the relaxation process equals the return of the bulk magnetization vector M to its original position along the z-axis. The recovery of the longitudinal magnetization M_z is referred to as longitudinal or spin lattice relaxation, while destruction of the transversal component M_{xy}, that has been created by flipping M, is called transversal or spin-spin-relaxation. As these forces have the opposite effect of an RF pulse and proceed faster the larger the deviation of the spin system is from its equilibrium, we conclude:

$$\begin{aligned} \frac{dM_z}{dt} &= -\frac{M_z - M_0}{T_1}, \\ \frac{dM_{xy}}{dt} &= -\frac{M_{xy}}{T_2}, \end{aligned} \tag{4.19}$$

where
M_z is (magnetization component in z-direction),
M_0 (equilibrium component in z-direction),
M_{xy} (transversal magnetization component),
T_1 (time constant (longitudinal relaxation)), and
T_2 (time constant (transversal relaxation)).
Solving these differential equations leads to a build up of M_z and a decay of M_{xy} by following time courses:

$$M_z(t) = M_0(1 - e^{-t/T_1}), \tag{4.20}$$

$$M_{xy}(t) = M_0 e^{-t/T_2}. \tag{4.21}$$

Figures 4.5 to 4.8 schematically depict the relaxation processes and their time courses. It is important to note that these two relaxation processes take place independently, though simultaneously. T_1 relaxation occurs due to energy exchanges with the lattice while T_2 relaxation is a mere signal loss due to a randomization of the spin orientations. The dephasing of the spins in turn is dependent on a variety of other physical effects occurring during the experiment, so that the reduction of M_{xy} is completed in a shorter period than under assumption of no disturbance. This period is denoted as T_2^* . The differences in relaxation times T_1 and T_2 between different tissue types give rise to image contrast in anatomical MRI. In general $T_1 \geq T_2$.

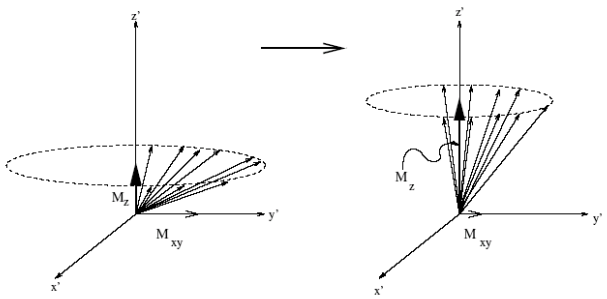

Figure 4.5: *Spin-Lattice Relaxation: Return of magnetization from the xy-plane towards M_z. (At the same time spin-spin relaxation takes place. This is indicated by a dephasing of the spins.)*

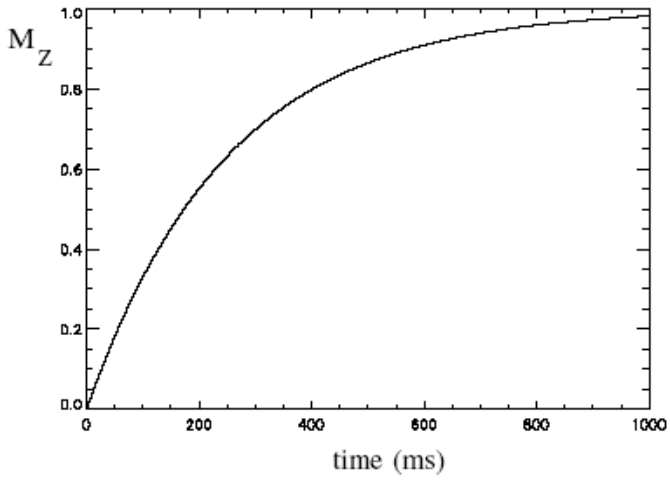

Figure 4.6: *T_1-relaxation curve in the recovery of the magnetization vector after the RF pulse has been switched off (ideal case).*

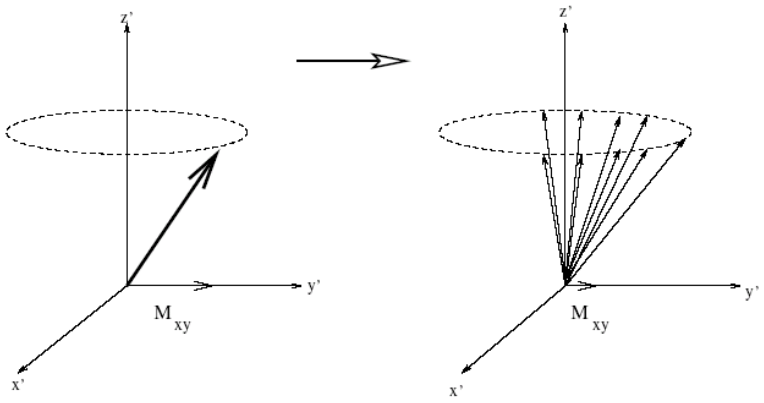

Figure 4.7: *Spin-Spin Relaxation. Shortly after the RF-pulse has been applied all spins are in phase with one another and the rotating frame (left). Dephasing causes the transversal magnetization M_{xy} to decrease (right).*

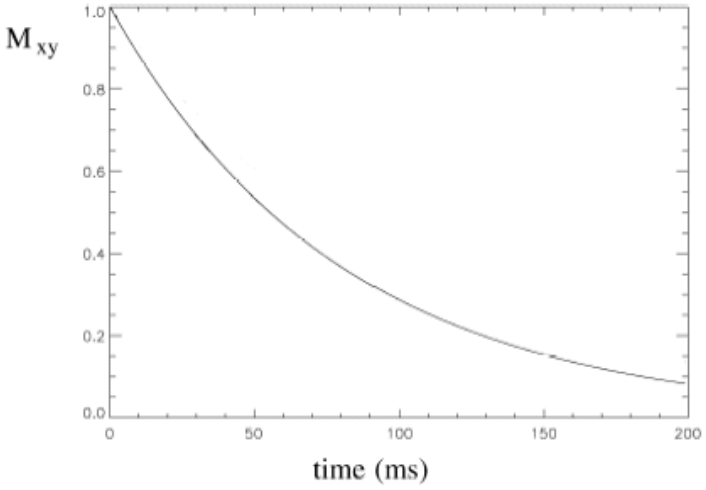

Figure 4.8: *Schematic time course of transversal magnetization M_{xy} during the return of the spin system to its equilibrium. The period in which M_{xy} gradually vanishes is referred to as T_2-relaxation.*

Considering the effects of relaxation in the equation of motion yields the famous Bloch equation:

$$\frac{dM}{dt} = \gamma M \times \mathbf{B}(t) - \frac{M_x\mathbf{i} + M_y\mathbf{j}}{T_2} - \frac{(M_z - M_0)\mathbf{k}}{T_1}. \tag{4.22}$$

It describes the trajectory of the magnetization vector M after cessation of the RF pulse. We remark that during the time M spirals back to the z-axis (free precession period) its magnitude is not preserved.

4.1.4 Signal Generation

After excitation of the spin system measurable signals have to be evoked from the rotating magnetic field. This is done by introducing a conducting loop (receiver coil). According to the Faraday law of induction, time varying magnetic flux through a conducting loop induces an electromagnetic force (voltage). Its strength equals the rate at which the flux is changing.

Free Induction Decay - FID

Consider free precession after a 90° RF pulse. Assuming that M is tipped right into the x-direction, the following decay of the magnetization in the resting frame of coordinates writes:

$$\begin{aligned} M_x &= M_0 e^{-t/T_2}\cos(\omega_0 t), \\ M_y &= M_0 e^{-t/T_2}\sin(\omega_0 t), \\ M_z &= M_0(1 - e^{-t/T_1}). \end{aligned} \tag{4.23}$$

Due to the orientation of M only the transversal component of the magnetization can give rise to a change of the magnetic flux in the pick-up coil. The induced signal $f(t)$ will therefore be of according shape, corresponding to M_x in Equation 4.23 as shown in the figure below (Fig. 4.9). The free induction decay (FID) is practically always mixed up with the RF frequency, so that merely the envelope function is measured. In addition, since stimulating and precession frequency are never perfectly identical, interference patterns will be observed.

Spin–Echo

Another essential way of MR signal generation is the Spin-Echo or Hahn pulse sequence (Hahn, 1950), which accomplishes a reversal of dephasing effects caused

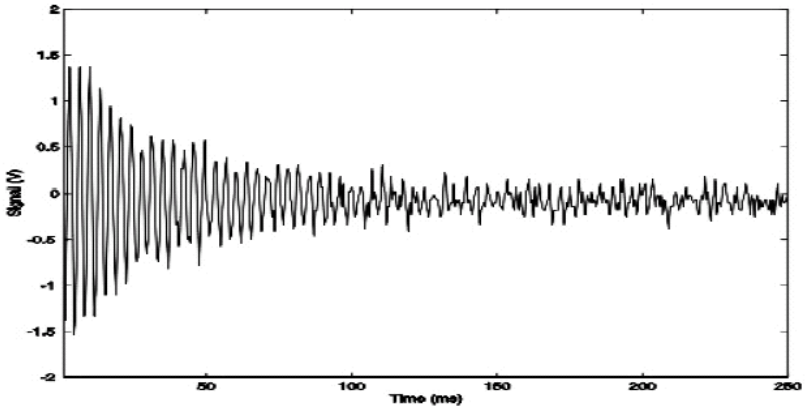

Figure 4.9: *Schematic Free Induction Decay (FID) of the nuclear magnetisation after a 90° RF pulse.*

by magnetic field inhomogeneities (pseudo relaxations). Echo signals can be produced by either multiple RF pulses or by magnetic field gradient reversal. Signals of the former type are called RF echoes, the latter gradient echoes.
Basically, after the 90° pulse, a second, a 180° pulse, is triggered. To gain the desired effect this refocusing pulse must be set at a time TE/2, where TE is known as the echo time (Fig. 3.2(A)). As previously described, the 90° pulse generates an FID signal. We assume the resulting magnetization vector M to be aligned with the y'-axis. Due to the inevitable variations in the applied magnetic field B_0, the magnetization vectors contributing to the bulk magnetization start to dephase. In parts of the sample where the magnetic field strength is is slightly lower than B_0, the nuclei will accordingly precess slightly slower than ω_0. Similar conditions apply for nuclei in areas affected by a slightly greater magnetic field strength than B_0. In the rotating frame of reference this shows in a magnetization vector fanning out (Fig. 4.10). As an effect the transverse component M_{xy} decays more rapidly than expected and gives rise to an envelope function of the FID signal proportional to e^{-t/T_2^*}, rather than $e^{-t/T2}$.

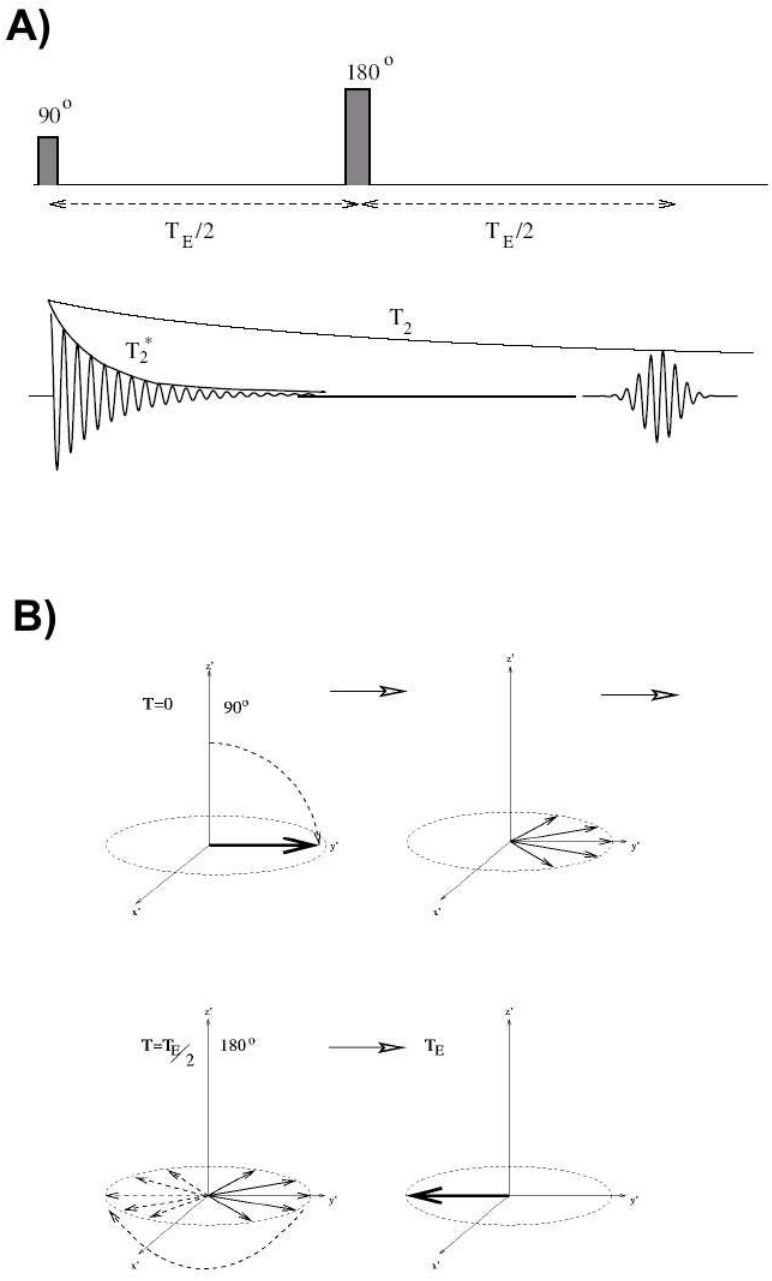

Figure 4.10: *Spin-Echo: (A) Pulse sequence (B) Spin behavior: After application of a* $90°$ *pulse the spins dephase. An additional 180° pulse accomplishes refocusing.*

The proximate refocusing pulse of 180° causes the whole fan to flip about the x'-axis. As a result all the precession vectors now precess towards the y'-axis and are aligned at a time TE after the initial FID pulse. They have been rephased. M reaches a maximum, corresponding to the center of the spin echo signal (Fig. 4.10(A)). Afterwards the spins dephase and the signal rapidly decays again.
By spin-echo imaging the T_2-value of an unknown sample can be estimated by performing a number of spin echo experiments with different TE values. The echo maximum is plotted versus TE and performing a least square fit to a function of the form $S = S_0 e^{-TE/T2}$ is performed. The calculated T_2 values are important information for adaptation of further imaging procedures involving the concerned tissues. Besides for T_2 measurements, spin-echo sequences are commonly applied in standard measurement techniques since they are as a result of the refocusing pulse T_2^* insensitive.

4.2 Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) allows visualization and localization of brain activity on the basis of serial, fast acquired MR images. During imaging the subject is presented with specific stimulation (or task) assumed to be correlated to activation of specific brain areas. Stimulation is organized in one of the following ways. In the investigation with on-off paradigms the subject has to listen to the presented sounds, perform a task etc. in consecutive time periods alternating with rest periods (block design). Experiments of event related design require a single action of the subject only according to a stimulus given at certain time instances. Both study designs may be arranged in equispaced or random time patterns.
In order to trace the spin history during the task from which the ongoing physiological processes can be recovered, contrast agencies (Aoki et al., 2002) can be used. Due to the varying rate of blood oxygenation as a physiological response to brain activation, the successive images show variations in their voxel intensities according to the stimulus presented during the experiment (Blood Oxygenation Level Dependent (BOLD) effect (Ogawa et al.,1990; Kwong et al., 1992; Ogawa et al., 1993). The physical basics of MR imaging and its application, fMRI, are given in detail in the following sections.

4.2.1 BOLD-Effect

The oxygen consumption of active neurons is higher than of neurons in the resting state. This causes the oxygen level in the surroundings of the neuron to drop and more oxygen is delivered by increased blood flow. In fact the blood flow over-compensates for the drop of oxygen which is why active cortical regions show an increase of concentration of oxygenated blood. Recognition of those concentration in- or decreases by a series of MR images is possible due to the fact the the magnetic state of blood depends on the degree of oxygenation. With deoxygenation of the hemoglobin the iron at the center of the heme (Fig. 4.11) becomes paramagnetic. Therefore deoxygenated blood is more paramagnetic than oxygenated blood (Pauling and Coryell, 1936) and creates local magnetic field gradients. If the magnetic field varies over a voxel the spins within this voxel dephase faster and T_2^* decreases. The slightly diamagnetic oxyhemoglobin distorts the magnetic field much less and T_2^* rises. As a consequence, the NMR-signal from oxygenated voxels is higher than the signal originating from deoxygenated voxels and an imaging sequence providing high T_2^* -contrast due to long echo times, such as echo planar imaging (see Appendix B), can reveal these signal changes (Springer et al., 1999).

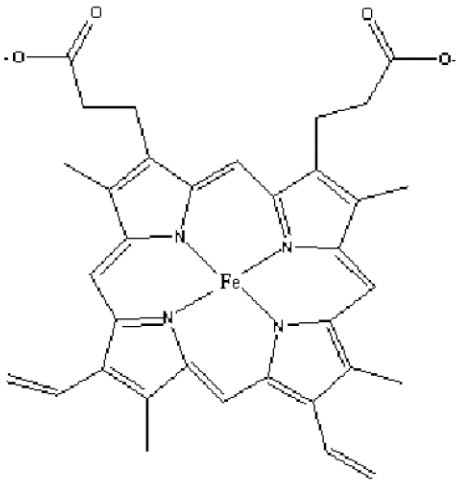

Figure 4.11: *Chemical structure of the heme (schematic). Depending on whether an oxygen atom is coordinatively bound to the iron or not, the heme is diamagnetic or paramagnetic, respectively.*

The fact that brain function is localized (sometimes in various networks) allows for mapping regions of cortical activity by measuring the hemodynamic changes that accompany neuronal activities. The spatial resolution of fMRI is accurate

enough to provide information about the location of activation. This was first demonstrated during a simple visual perception stimulus (Kwong et al., 1992) and has since been used to map many different kinds of activation in cortical areas as well as in deeper brain structures. Although fMRI is gaining increased attention in various fields such as clinical and neuroscience applications it is still limited by a relatively low signal-to-noise ratio (SNR), as well as the variable hemodynamic response delay as a consequence of the indirect detection of brain activity via BOLD contrast and blood perfusion changes. In addition, the applied EPI sequences employed to provide fast imaging suffer more than other techniques from distortion and signal loss arising from magnetic field inhomogeneities included by the brain. These often derive from susceptibility differences between brain and air and are therefore not amenable to correction by improved shimming (adjustment of the static field homogeneity with a set of coils). The latter difficulties reinforce the call for improvement of measurement and precise reconstruction techniques as well as development of sophisticated data analysis tools.

4.2.2 Auditory Stimulation and fMRI

An undesirable aspect of present-day MRI is the high-level sounds produced by the scanner and associated equipment. These unwanted sounds, or acoustic noise, pose particular difficulties for functional fMRI studies that measure brain activation in response to sound stimuli. For example, the background noise can mask the stimuli (Shah et al., 1999; Eden et al., 1999), and the noise itself can produce brain activity that is not related to the intended stimuli (Bandettini et al., 1998; Ulmer et al., 1998; Talavage et al., 1999; Edmister et al., 1999). If the noise can be heard, then the auditory system presumably is in a different state than during quiet conditions more typical of physiological or psychophysical experiments on hearing. Earmuffs or earplugs are commonly used to reduce noise levels heard by subjects (e.g., Savoy et al. 1999), but they are insufficient to achieve acceptably quiet conditions (Ravicz et al., 2000).
Acoustic noise in most imaging environments arises from various sources. Continuous noise can originate from ancillary equipment located in the room that houses the imager. This equipment often includes a pump for liquid helium used to supercool the imagers permanent magnet, a fan for supplying ventilation to the patient, and the air-handling equipment for the imager room. The highest-level noise, however, is produced whenever an image is acquired by the gradient coil switches which poses a challenge for performing auditory tasks in the MR

environment.

4.2.3 The "sparse temporal sampling" method

The scanner noise can (1) interfere with the auditory stimulation (Bandettini et al., 1998; Shah et al., 1999; Hall et al., 2000), (2) result in masking of the auditory cortical response depending on the frequency of the MR acquisition, (3) lead to activation of the auditory cortex itself depending on the effective repetition time (TR) for MR acquisitions (Bandettini et al., 1998; Ulmer et al., 1998), and (4) causes discrepancies in the attentional demands if frequency and intensity of the MR scanner noise differ between studies. A few imaging designs have been proposed to overcome these interferences (Edmister et al., 1999; Hall et al., 1999, 2000) using the sparse temporal sampling method. In this method, single volumes of brain images are acquired at the end of the stimulus and baseline conditions. To optimize detection of the activation without sampling the whole hemodynamic response function, images should be taken near to the maxima and/or minima of the hemodynamic response during the experimental cycle. Hall et al. (1999) compared continuous scanning with a sparse temporal sampling method, using an effective TR of 14 s and revealed a greater MR signal change for acquisitions with a long TR. In addition, it was found that recording scanner noise and playing it back to the subject could lead to a signal change of up to 1.5% primary and secondary auditory cortex. Hall et al. (1999, 2000) showed further that the activation peaked 4-5 s after stimulus onset and decayed after an additional 5-8 s. Others have shown a decrease in the spatial spread and lower z-scores of the activated auditory regions depending on the duration of the MR scanner noise (Shah et al., 1999). Overall, these studies showed, that sparse imaging is advantageous in auditory experiments since it avoids the four aforementioned problems of auditory stimulation in the presence of scanner background noise.

4.2.4 Data Analysis

Preprocessing in fMRI

The purpose of preprocessing in fMRI is to correct for non-task related variability in experimental data. These approaches are usually performed without any consideration for the experimental design and therefore are called preprocessing. The preprocessing steps seek to remove, rather than model data variability. One important term in fMRI analysis language is the signal-to-noise-ratio (SNR), which

is the quotient between task-related variability and non-task-related variability. The goal of the preprocessing is to remove as much non-task-related variability as possible in order to obtain a high SNR value. The sensitivity of an fMRI analysis is determined by the amount of residual noise (non-task related variability) in the image series. The following sections describe four basic steps of preprocessing (although the number of steps depends on the applied experimental design).

Motion Correction Motion correction corrects the data for movements of the head; it controls for motion-related variance. Small movements of approximately 1 mm occur (Turner et al., 1997; Frackowiak et al., 1997) even if the subjects head is fixated and the subject tries explicitly not to move. In addition, there are movements related to physiological factors, such as cardiac and pulmonary parameters. These movements can e.g. alter the spin history as a result to the change of position relative to the gradients of the scanner. Moreover, often it is impossible to eliminate the voxels in the activation pattern that are activated by an artifact that has arisen as a result of motion. The functional volumes are aligned in space by rigid-body transformations. The software program SPM2 (http://www.fil.ion.ucl.ac.uk/spm/software/spm2/) employs a least-squares approach and a six parameter (rigid body) spatial translation (Friston et al. 1996). In this approach, the first image in a time-series serves as a reference scan against which all subsequent scans are aligned. Motion correction has several limitations, including artifact-related limitations, which can result in a loss of data around the edges of the image volume. Distortions in fMRI images can also pose a problem because distortions may be dependent upon position in field, rather than position in the head. Therefore the rigid body model does not model these types of distortions.

Coregistration In order to coregister the anatomical and functional images, the software program in SPM2 requires three steps. First, it uses simultaneous affine registrations between each image and a template image of the same modality (functional or anatomical). Next, SPM2 separates the images into white and gray matter (segmentation). Finally, white and gray matter partitions acquired during step two are simultaneously registered via the sum of squared difference. The fMRI literature is ambiguous as to whether functional and anatomical images should be coregistered. Coregistration may improve normalization and allow activation on anatomical images to be displayed. It also permits comparison across modalities in an individuals brain. Nevertheless, possible disadvantages of coregistration exist. For example, coregistration may severely distort functional data

and may diminish correspondence between functional and anatomical images.

Normalization This preprocessing step spatially (stereotactically) normalizes images into a standardized space, which is defined by some ideal model or template image. It can thus be described as an intersubject averaging method. The software program SPM conforms the Montreal Neurological Institute (MNI) space (a combination of many MRI scans on normal controls; all right-handed subjects) and approximates that of the space described in Talaraich and Tournoux atlas (1988). Generally, these algorithms work by minimizing the sum of squares differences between the acquired images and the template (for automated algorithms see e.g. Ashburner and Friston, 1997). One major advantage of the normalization is that it allows the generalization of results to a larger population, thereby enabling averaging across subjects and improving comparisons with other studies. Normalization also provides coordinate space to report the results. One disadvantage of normalization can be cited in its potential to reduce the spatial resolution resulting from differences in interindividual brain organization and thus may reduce the activation strength by averaging subjects.

Spatial Smoothing Spatial smoothing convolves images with an isotropic Gaussian kernel (defined by its full width at half maximum (FWHM)) to potentially increase SNR, to compensate for residual between-subject variability after normalization, and to approximate a random field for statistical purposes (see multiple comparisons below). Issues such as reduced spatial resolution, which relate to smoothing, are discussed in detail in Friston et al., 2000.

Statistical Analysis of fMRI Data Several methods based on the General Linear Model (GLM) can be employed to assess brain activity in fMRI studies. The basic concept of the GLM is, that it treats the data as a linear combination of model functions plus noise. Typically, these methods result in a statistical parametric map. In order to obtain these parametric maps, each of these methods entails comparable successive steps. First, the response (signal) will be modeled at each voxel in the brain by a GLM. Subsequently, a hypothesis regarding the models parameters will be tested. The observed statistic map will then be represented at a given level (threshold) according to the point of distribution of the statistics (uncorrected level) or the field distribution of the statistics (corrected levels for local maxima). A correction to the significance of the t-statistics is suggested, which account for the multiple comparisons in the image (see multiple

comparisons below). Afterwards, the statistical parametric map can be visualized showing e.g. the t-value for each single voxel with a corresponding color on an image of the brain. Overall, these methods are voxel-by-voxel hypothesis testing approaches, which reliably identify regions showing a significant effect of interest. The GLM model used can refer to a single subject, one group of subjects or multiple groups of subjects, which can represent different subjects (e.g. musicians and non-musicians) or the same subjects (e.g. prior to training versus after training of an auditory task). The modeling part is univariate and thus separated for each voxel and usually each subject as well. For a simple t-test, the GLM is also utilized in order to account for covariates as well as temporal autocorrelation. Multi-subject fMRI experiments can also be performed using a GLM framework with different forms depending on the approach taken, e.g. fixed or random subject analysis. Fixed-effects model uses data from all subjects to construct a statistical test. This method allows interference to the analyzed subject sample. A random-effects model accounts for inter-subject variance in analysis and permits interference to population from which subjects are drawn, which is essential for group comparisons.

Multiple Comparisons Most statistics packages for functional imaging data create statistical parametric maps (see above), which have a value for a certain statistics (e.g. t-statistics) at each voxel in the brain. This results from the statistical test performed voxel-wise between states (experimental tasks) within one subject or over a number of subjects. One major problem arises, however, if more than one test is made, the collective alpha value is greater than the single test alpha (overall Type–I error increases). One option is to adjust the alpha value of the individual tests in order to maintain an overall alpha value at an acceptable level (controls overall for Type–I error; known as Bonferroni correction). However, conventional multiple comparison methods (e.g. Bonferroni) may overcompensate resulting in very strict significance values for even small data sets. This problem might be explained by the fact that this approach is not appropriate for correlated data. If a data set contains correlated data points, then the effective number of statistical tests may be greatly reduced and most fMRI data has significant correlations. Several approaches are used in fMRI analysis software programs. Some of the primary examples include an approach using the random field theory, which provides false positive rate for fMRI data based upon the smoothness of the data, a randomization-based analysis across replications, and applying a small volume correction, e.g. a Region of interest (ROI) analysis. A ROI can only be used if an a priori hypothesis has been made.

4.2.5 FMRI Hardware and Environment

In order to perform MRI experiments, it is necessary to have a magnetic resonance imager (see Fig. 4.12). An integrated MRI system consists of a magnet, gradient and shim coil(s) (often integrated), a console, radiofrequency (RF) and gradient amplifiers, and RF coils.

Super-conductive magnet Fundamental to all magnetic resonance devices is the need for an intense magnetic field that does not vary with time and that has very nearly the same field strength throughout the anatomical region to be examined. Typical magnetic fields for human research range between 1.5 T (i.e., proton resonance frequency: 64 MHz) and 7.0 T (i.e., 300 MHz). It is generally accepted that for the ^1H nucleus, signal-to-nose ratio (SNR) is linearly dependent on magnetic field strength; the higher field is in this respect the better. However, it should be noted that using a very high magnetic field requires overcoming several technical challenges such as stronger local magnetic inhomogeneities induced by tissue-air interfaces. To correct local inhomogeneous magnetic fields in the imaging volume, it is necessary to have room temperature shim coil that generate spatially non-uniform magnetic-fields that can be used to partially compensate the sample-induced non-uniformities. Higher magnetic fields require stronger shimming capability.

Gradient Coils A gradient coil set inserted into the magnet bore generates linear magnetic fields along x, y, and z directions. In order to acquire high-resolution images at high temporal resolution, fast slew rate (i.e., switching time) and high gradient strength are necessary. Since the gradient set with a smaller diameter will have higher gradient strength and faster slew rate with the same electronics, it is preferable to use dedicated smaller-diameter gradients for head-only research, rather than clinical body gradient sets. However, for a given gradient performance, the forces on the gradients increase with higher magnetic fields and consequently acoustic noises increase.

Console Console consists of receivers that detect, amplify, demodulate, and digitize the MR signals detected by the RF coil, and a set of electronics that can generate a pattern of RF and gradient pulses (which are sent to the amplifiers) for

the generation of the appropriate imaging signals from the sample. A computer controls both of these processes in the console. To obtain functional imaging, an MRI system must have high stability and high sensitivity because small differences between hundreds to thousands of images acquired during baseline (i.e. control) and stimulation periods are used to create an fMRI map. If MRI signals fluctuate 0.5 % peak-to-peak over a time in a given area, it is difficult to detect a less than 0.5 % signal change, which is commonly observed in fMRI.

Radio-frequency Amplifiers and Coils RF coils are used to transmit RF pulses for excitation of water and detect RF signals from water. Typically a smaller size coil is more sensitive, but its coverage of brain area is less. For whole brain studies, a homogenous head coil should be used, while a small surface coil can be used for localized brain studies. An alternative is to use an array of surface coils using multiple receivers to obtain volume coverage with surface coil sensitivity, albeit at the expense of significant increase in instrumentation complexity and signal processing. Sensitivity of surface coil depends on depth of cortical areas, and the surface of the cortex will be most sensitive. With multiple surface coil arrays, it is in principle also possible to obtain higher sensitivity in deeper regions.

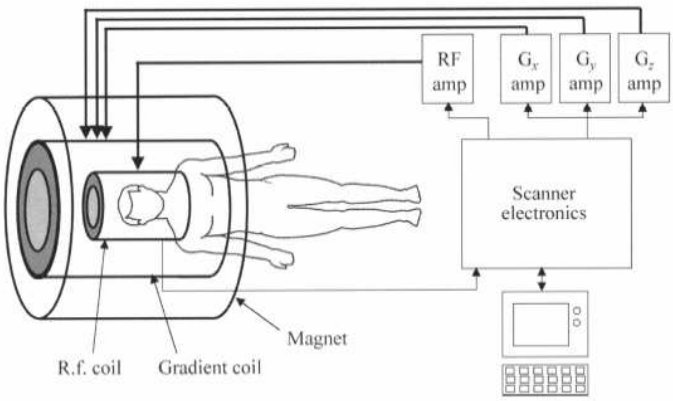

Figure 4.12: *Scheme of an MRI scanner (taken from Jezzard and Claire, 2001). A high static magnetic field is applied to get a sufficient net magnetic moment from the nuclear spins of the brain tissue. The excitation of these spins is then induced by signals from the radio frequency (RF) coil. The application of the three gradient coils allows for a 3D spatial reconstruction of the recorded resonances.*

Chapter 5

Psychometry

5.1 Paired Comparison Analysis

In a complete paired comparison experiment each subject is presented with $\frac{r(r-1)}{2}$ pairs of r items, sounds in our case, and asked to choose between two items at a time. To obtain a psychophysical scale from the binary paired comparison judgments we used the Bradley−Terry−Luce (BTL) method. The basic experimental unit is the comparison of two stimuli, A1 and A2, by a judge who must choose one of them. If more than two stimuli are under consideration every judge has to perform every possible paired comparison. The setup is called "balanced paired-comparison experiment" and for n stimuli and t judges the numbers of paired comparisons will be $t\binom{n}{2}$. The goal is to determine the probability that the stimulus A_i is chosen, $P(A_i)$ and to use this value to derive a relative scale of carrier salience. Carrier salience can be defined as a strength of tone sensation. In our experiments, the judges have to decide which sound has a more salient character ($A_i \succ A_j$). Ties or indifferent judgements are not permitted, so a listener must declare one of the two stimuli to be more salient. To avoid judging effects, depending on the order of representation, each pair was presented twice. With random variation, once in the order A_i, A_j and in reversed order A_j, A_i. The mean of both comparisons was used to get the raw data of the number of times each stimuli was judged by the t judges to be more salient than each of the

other stimuli. From these raw data, a square matrix A is formed as

$$A = \begin{pmatrix} - & a_{12} & \dots & a_{1n} \\ a_{21} & - & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & - \end{pmatrix} \tag{5.1}$$

where a_{ij} denotes the observed number of times, stimulus i was judged to be more salient than stimulus j. The total number of comparisons of the two stimuli is $t_{ij} = a_{ij} + a_{ji}$. Since a stimulus can not be compared to itself, the diagonal elements of the matrix are left vacant. From A, a probability matrix is constructed with elements $p_{ij} = \frac{a_{ij}}{tij}$. It is the probability that the stimulus A_i is preferred over stimulus $A_j (Ai \succ Aj)$. The summation of the symmetric elements is $p_{ij} + p_{ji} = 1$. The elements of the basic Bradley-Terry form of the probability matrix X can be derived with:

$$x_{ij} = \frac{p_{ij}}{1 - p_{ij}} = \frac{p_{ij}}{p_{ji}} \tag{5.2}$$

and analogously

$$x_{ji} = \frac{p_{ji}}{1 - p_{ji}} = \frac{p_{ji}}{p_{ij}}. \tag{5.3}$$

Using the natural logarithm on the single elements of X, the scale is transformed from a ratio scale to a difference scale. Finally, the values $s(i)$ of the single stimuli can be obtained by averaging over the columns of the transformed matrix X

$$s(i) = \frac{\sum_{n=1}^{\infty} \ln x_{ij}}{n}. \tag{5.4}$$

In the special case where all subjects judged the stimulus A_i to be more salient than stimulus $A_j (a_{ij} = t_{ij})$, the values in equation and in equation equal infinity. According to (David, 1988), who suggested to assume that 'half a subject' rated the pitch vice versa, the value of a_{ij} was corrected to $(a_{ij} = t_{ij} - \frac{1}{2})$.
When presented with two sounds, listeners do not compare their "true" carrier salience but rather two random variables centered on the real values. Therefore, different outcomes are expected for the same comparison between two sounds, across listeners but also with repetitions for a same listener. In order to test whether carrier salience differences were significant, they must be compared to the standard deviations of the results. These standard deviations are not readily available with the BTL method because judgments from all listeners are collapsed into proportions to build the scale. They were estimated by the bootstrap technique (Efron and Tibshirani, 1993). An empirical distribution can be generated by resampling with replacement from the data set. As a sample corresponds to

the data matrix for a given subject, each set can be thought of as a simulated subject group. For each set, the BTL analysis is performed anew, giving in the end a distribution of estimated carrier salience values for each stimulus. These distributions are than used to estimate the standard deviations. The standard deviations are finally used to test for the significance of the difference between any two stimuli at a given risk.

Chapter 6

Modeling temporal asymmetry in the auditory system

6.1 The Auditory Image Model

The inner ear and the auditory periphery convert sound to a set of neural signals that are relayed to the brainstem and brain. These neural signals form a neural code of the original sound. The neural circuits of the brainstem and brain further process this code to determine what and where are the objects that produced the sound. Our knowledge of how the peripheral auditory system operates and its consequences for auditory perception have advanced to the point where we can quantitatively describe how the neural code is formed from the original sound. Thus, the computational models of the peripheral auditory system are developed.

To explain perceptual asymmetry of temporally asymmetric sounds we used a time-domain model of auditory processing developed by Roy Patterson and colleagues (Patterson et al., 1995). Patterson's auditory image model (AIM) of peripheral auditory processing is constructed in three stages: an auditory filterbank that produces a representation of basilar membrane motion (BMM), a multi-channel, neural transduction mechanism that produces a representation of the neural activity pattern (NAP) in the auditory nerve, and a bank of strobed temporal integration units that produce the models representation of the auditory image that we hear in response to the sound as shown on the Figure 6.1.

The Spectral Analysis Stage

Spectral analysis is performed by a bank of auditory filters which converts a digitized wave into an array of filtered waves. The set of waves is AIM's representation of basilar membrane motion. The filters are distributed linearly along a frequency scale measured in Equivalent Rectangular Bandwidths (ERB's):

$$ERB = 24.7\,Hz + 0.108 f_0 \tag{6.1}$$

where f_0 is a central frequency. The ERB scale was proposed by Glasberg and Moore (1990) based on physiological research summarized in Greenwood (1990) and psychoacoustic research summarized in Patterson and Moore (1986). AIM provides a physiological auditory filter for generating the BMM–a non-linear, transmission-line filter (Giguere and Woodland, 1994a). A feedback circuit representing the fast motile response of the outer haircells generates level dependent basilar membrane motion (Giguere and Woodland, 1994a).
The upper panel of Figure 6.2 show the responses of the filterbank to the vowel in 'hat'. They have 75 channels covering the frequency range 100 to 6000 Hz (3.3 to 30.6 ERB's). In the high-frequency channels, the filters are broad and the glottal pulses generate impulse responses which decay relatively fast. In the low-frequency channels, the filters are narrow and so they resolve individual continuous harmonics. The rightward skew in the low-frequency channels is the 'phase lag,' or 'propagation delay,' of the cochlea, which arises because the narrower low-frequency filters respond more slowly to input.
.

The Neural Encoding Stage

The second stage of AIM simulates the mechanical/neural transduction process performed by the inner haircells. It converts the BMM into a neural activity pattern (NAP), which is AIM's representation of the afferent activity in the auditory nerve. The Meddis (1988) module simulates the operation of an individual inner haircell; specifically, it simulates the flow of neurotransmitter across three reservoirs that are postulated to exist in and around the haircell. The module reproduces important properties of single afferent fibres such as two-component time adaptation and phase-locking. The transmitter flow equations are solved using the wave-digital-filter algorithm described in Giguere and Woodland (1994a). There is one haircell simulator for each channel of the filterbank. The middle panel in Figure 6.2 show the NAPs obtained with adaptive thresholding and the

Meddis module in response to BMMs from the transmission line filterbank. The phase lag of the BMM is preserved in the NAP. The positive half-cycles of the BMM waves have been sharpened in time. Sharpening is also evident in the frequency dimension of the adaptive thresholding NAP.

The Temporal Integration Stage

Periodic sounds give rise to static, rather than oscillating perceptions, indicating that temporal integration is applied to the NAP in the production of our initial perception of a sound, our auditory image. Traditionally, auditory temporal integration is represented by a simple leaky integration process and AIM provides a bank of lowpass filters to enable the user to generate auditory spectra (Patterson, 1994a) and auditory spectrograms (Patterson et al., 1994b). However, the leaky integrator removes the phase-locked fine structure observed in the NAP, and this conflicts with perceptual data indicating that the fine structure plays an important role in determining sound quality and source identification (Patterson, 1994b; Patterson and Akeroyd, 1995). As a result, AIM includes a module which preserves much of the time-interval information in the NAP during temporal integration, and which produce a better representation of our auditory images. This is accomplished with strobed temporal integration (Patterson et al., 1994a,b). A bank of delay lines is used to form a buffer store for the NAP, one delay line per channel, and as the NAP proceeds along the buffer it decays linearly with time, at about 2.5%/ms. Each channel of the buffer is assigned a strobe unit which monitors activity in that channel looking for local maxima in the stream of NAP pulses.

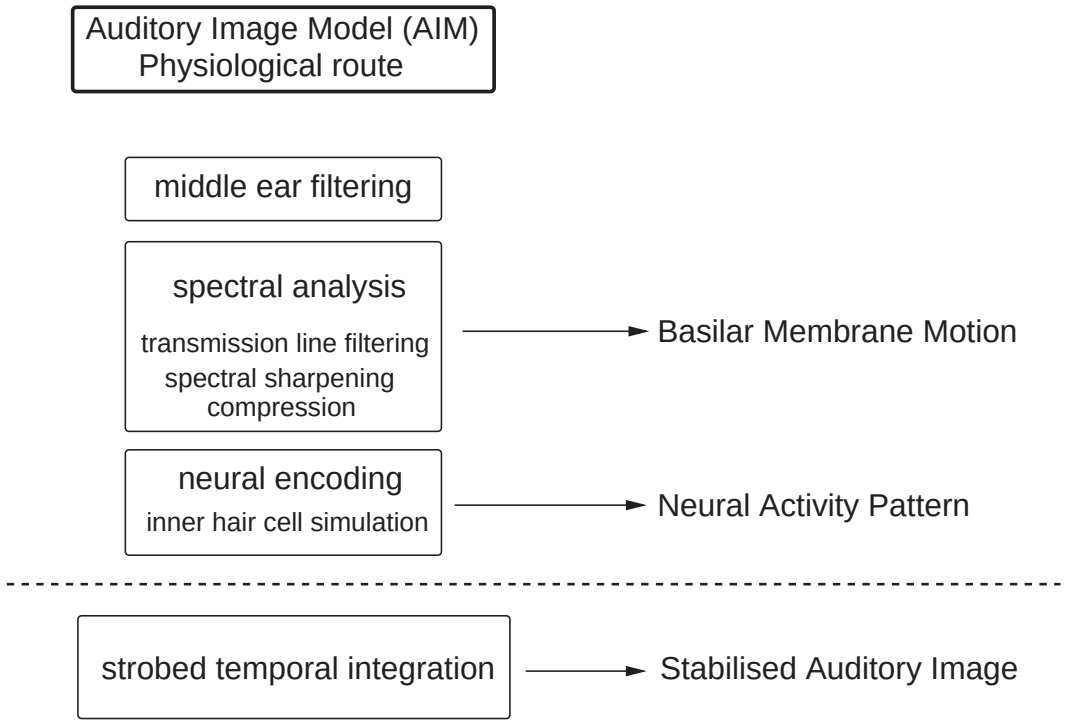

Figure 6.1: *The three stage structure of AIM.*

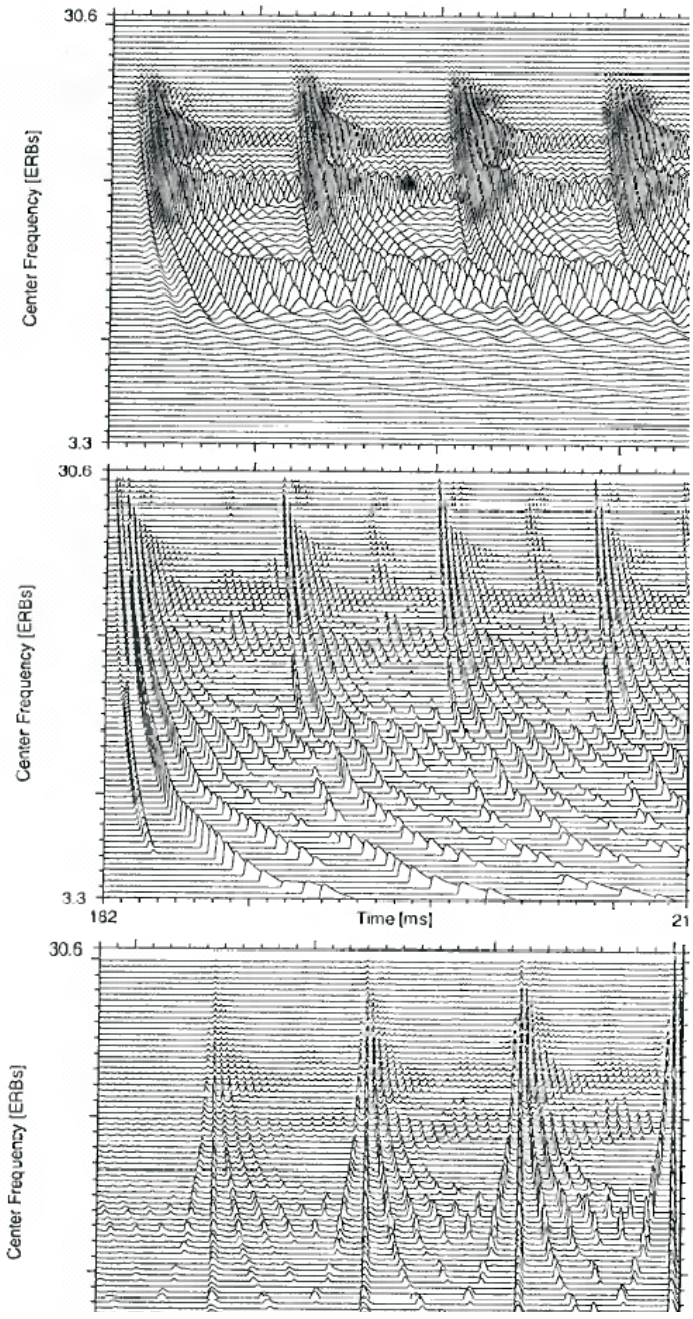

Figure 6.2: *Response of the model to the vowel in 'hat': (top) basilar membrane motion, (middle) neural activity pattern, and (bottom) auditory image (adapted from Patterson and Allerhand, 1995).*

When one is found, the unit initiates temporal integration in that channel; that is, it transfers a copy of the NAP at that instant to the corresponding channel of an image buffer and adds it point-for-point with whatever is already there. The local maximum itself is mapped to the 0-ms point in the image buffer. The quantized temporal integration of the periodic information generates a stabilized auditory image (SAI) as shown on the bottom panel of the Figure 6.2. It has been shown that the distance between 0 and the location of the first peak of the summed SAI corresponds to the perceived pitch (Ritter et al., 2005). We show that the height of this peak corresponds the perceived carrier salience (tonality). Periodic and quasi-periodic sounds cause regular strobing which leads to simulated auditory images that are static, or nearly static, and which have the same temporal resolution as the NAP. Dynamic sounds are represented as a sequence of auditory image frames. If the rate of change in a sound is not too rapid, as is diphthongs, features are seen to move smoothly as the sound proceeds, much as characters move smoothly in animated cartoons.

Applications

In hearing research, the AIM has been used to model phase perception (Patterson, 1987), and timbre perception (Patterson, 1994b), temporal pitch (Rupp et al., 2005, Ritter et al., 2005), as well as to simulate cochlear hearing loss (Giguere and Woodland, 1994b), and combination tones of cochlear origin (Giguere, Kunov, and Smoorenburg, 1995). In speech research, AIM has been used to explain syllabic stress (Allerhand et al., 1992), and as preprocessors for speech recognition systems (e.g. Patterson, Anderson, and Allerhand, 1994). In summary, the AIM provides a modular architecture for time- domain computational studies of peripheral auditory processing.

Chapter 7

MEG Experiments

7.1 Introduction

Rapid changes in spectral composition or amplitude of sounds constitute important information–bearing elements in audition. Neurons of the auditory system, including auditory cortex, respond preferentially to acoustic transients, such as the onsets of sounds. In this Chapter we investigated spatio-temporal dynamics of the sources evoked by presentation of temporally asymmetric sounds of five different half-life times, two modalities, e.g. transient sounds vs. steady state sounds, and fixed carrier frequency. We focus on three components of auditory evoked fields (AEFs), specifically the N100m, P50m and P30m for the transient modality, and N100m, P200m and sustained field (SF) for the steady state modality.

7.2 Materials and Methods

7.2.1 Subjects

Ten right-handed subjects (7 men, 3 women aged 22–44 years) with normal audiometric thresholds and without any history of audiological or neurological deficits participated in the study. Seven of the subjects participated in both experiments. Measurments were approved by a local ethics committee and conducted with an informed consent of each subject. In order to check the reproducibility of the results the whole paradigm was replicated for three subjects. The repetitions were carried out on different days.

7.2.2 Stimuli

The stimuli for the experiments were generated using 1000 Hz carrier and the amplitude envelope given:

$$E(t) = \frac{1}{\sqrt{hl}} e^{-t \ln(2)/hl} \tag{7.1}$$

The time t is the length of one cycle and was set to 50 ms to ensure that the discontinuity in the envelope at the end of each modulation cycle occurs at an upward-going zero-crossing of the carrier. hl is the half-life of the modulator, namely 0.5, 1, 4, 16 and 32 ms. To obtain approximately constant loudness for all conditions the amplitude was normalized by a factor proportional to the square root of the stimulus half life time, respectively (Patterson, 1994a). The stimuli were presented diotically at a stimulus intensity level of 65 dB SPL using ER-3 transducers (Etymotic Research, Inc.) with 90-cm plastic tubes and foam ear pieces. The sampling rate was set to 48 kHz. The order of the stimuli was randomized.

In the first experiment 600 single ramped/damped cycles of five different half life times with an interstimulus interval (ISI)(time between offset of the stimulus and onset of the next following stimulus) ranging from 300 to 350 ms were presented. In the second experiment 20 modulation periods were constructed adding up to a total duration of 1 s. ISI was set to 1.0–1.1 s. The session contained 120 repetitions for each condition. Waveforms and spectra of the stimuli are shown on the Figure 7.1, and paradigm designs for both experiments on the Figure 7.2.

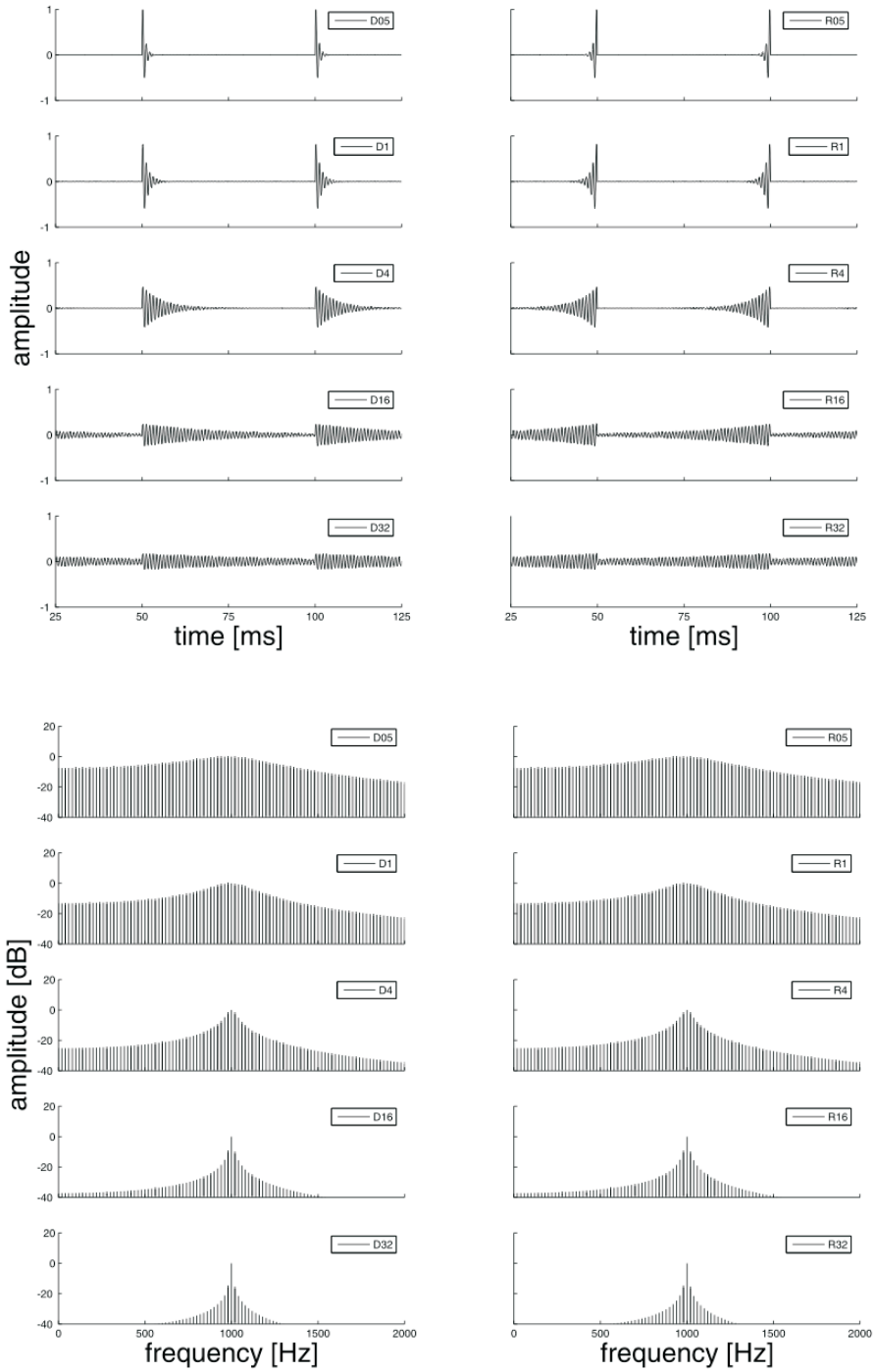

Figure 7.1: *Waveforms (above) and power spectra of damped and ramped tones. Damped sinusoids (D) are in the left column, ramped sinusoids (R) in the right column. HLTs used in the experiment are displayed: 0.5, 1, 4, 16 and 32 ms.*

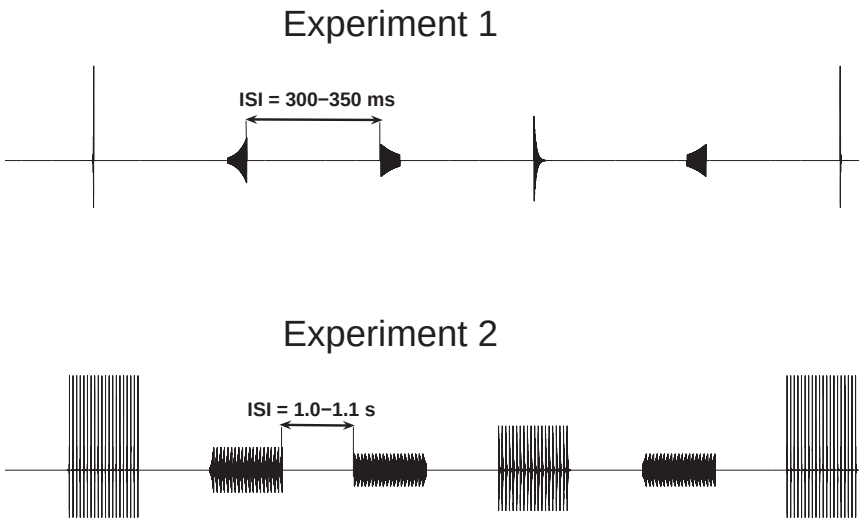

Figure 7.2: *Paradigm design for the experiments using transient (above) and steady state stimulation.*

7.2.3 Recording and Data Processing

Magnetic fields were acquired with a Neuromag 122 whole-head MEG system (Ahonen et al., 1993) inside of a magnetically shielded room (IMDECO, Switzerland). Subjects sat in an up-right position and watched a silent movie of their own choice. The sampling rate was 1000 Hz and a bandwidth ranging from 0.01 to 330 Hz. To increase signal-to-noise ratio neuromagnetic fields were averaged over an epoch of from 500 ms before to 500 ms (experiment 1) and 1200 ms (experiment 2) after tone onset. Offline averaging with artefact monitoring was performed using BESA5.1 software (MEGIS, Gräfelfing). Epochs containing signals exceeding an absolute level of 2000 fT and a gradient of 800 fT per sample were discarded automatically, resulting in about 5% rejection rate. The baseline was calculated over the 100-ms interval prior to tone onset.

7.2.4 Data Analysis

Spatio-temporal analysis (Scherg and von Cramon, 1985) of MEG data, using a two-dipole model (one dipole in each hemisphere), in a spherical volume con-

ductor aligned to the individual head surface was applied to the measured field distribution. The dipoles were freely fitted for each condition and each subject. The average location of all conditions was used as a spatial filter to derive the equivalent source waveforms for all conditions. The dipole orientations were fitted separately. No further constraints concerning dipole location, orientation or symmetry-conditions were applied. A principle component analysis (PCA) over an interval of 400–500 ms after the transition of the unfiltered auditory evoked responses was computed (Berg and Scherg, 1994) to compensate drifts and other low frequency artifacts due to the continuous stimulation. Estimates of the source parameters were accepted for further evaluation only if the goodness of fit of the field of the estimated source dipoles to the measured magnetic field was greater than 90 %. This analysis was applied to the 0.1–150 Hz band-pass filtered data and source coordinates of the waves P30m, P50m, N100m for single cycle paradigm, and N100m, P200m and the steady state responses for the steady state paradigm, were estimated.

Furthermore, we compared the areas of the sustained field for all conditions in a range between the beginning and the end zero–point of the sustained field as shown on the Figure 7.21.

To quantify the amplitude asymmetry of the pairs of ramped/damped responses an asymmetry index (AI) is defined as following: $AI = \frac{Ampl_{ramped} - Ampl_{damped}}{Ampl_{ramped} + Ampl_{damped}}$, and analogously for the areas of the sustained fields: $AI = \frac{Area_{ramped} - Area_{damped}}{Area_{ramped} + Area_{damped}}$. For comparison of the source waveforms obtained using source dipole model, individual source waveforms were averaged. 90% confidence intervals were determined using the Bootstraping-BC_a method (Efron and Tibshirani, 1993) with 1000 randomly drawn bootstrap samples of the data (MATLAB 7.1). This method allowed for calculation of confidence intervals without any prior assumptions about the distribution of the individual waveforms, e. g. a Gaussian distribution. Peak latencies and amplitudes were assessed by deriving the minimum and the corresponding value from the mean of the resamples. The critical t-intervals of latencies and amplitudes were computed using the resulting distribution of the single minima and the corresponding amplitude of each resample. Latency or amplitude differences of the peaks were considered significant when the t-intervals of two conditions did not overlap.

T1-weighted magnetic resonance images (MRI) were obtained from each listener on a 1.5-T, Siemens Symphony MRI-scanner. Scans were performed in 176 sagittal slices yielding an isotropic voxel size of 1 mm^3. Dipole locations were coregistered on individual scans and the positions were then transformed into the coordinate system of Talairach (Talairach and Tournoux, 1988) and projected

onto the probabilistic map provided by Schneider et al. (2004).

7.3 Results

7.3.1 Neuromagnetic Sources Evoked by Transient Stimulation

The duration of the evoked transient neuromagnetic responses generally lasted from 20 to about 300 ms as shown on the Fig. 7.3.

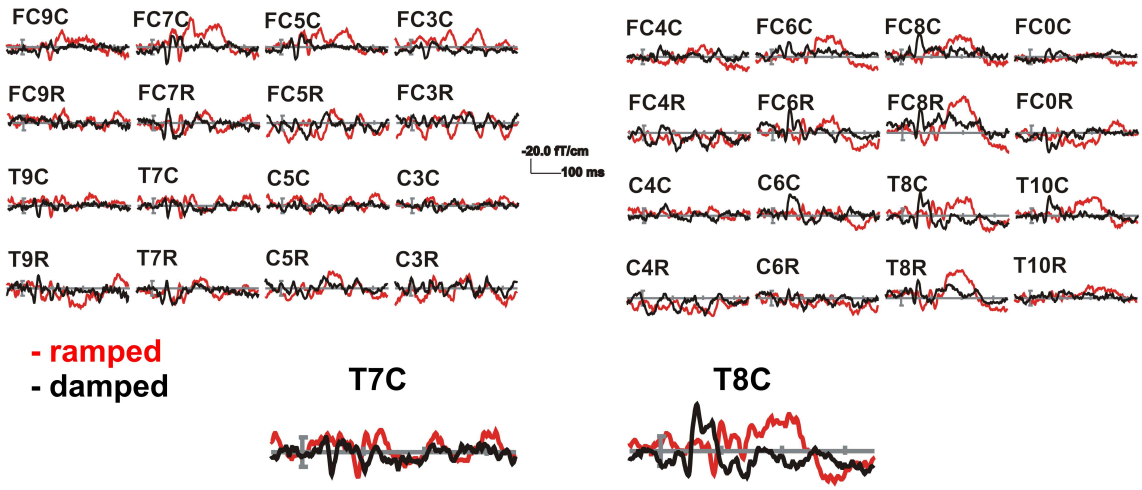

Figure 7.3: *A subset of averaged auditory evoked neuromagnetic responses of one subject to transient ramped and damped sounds of 4 ms HLT viewed from above and flattened onto a single plane. Circular (C) and radial (R) gradients located at the same coordinate are here separated for the sake of the clearness.*

While there was considerable variability between neuromagnetic responses evoked by different stimuli and for different subjects, the measured fields replicated well over time, within subjects. During the first 200 ms poststimulus we identified consistently three components: P30m, P50m and N100m for all ten stimulus conditions. All these components showed dipolar pattern topographies. Due to the envelope structure of the ramped sounds the latencies of the corresponding responses are delayed compared with their damped counterparts as shown on the Fig. 7.4.

On the figure 7.5 the grand average P30m source waveforms for left and right hemisphere are shown. In the table (see Appendix A) P30m amplitudes and

latencies for all conditions are given. P30m peak latency decreases with the stimulus half-life time (HLT). Amplitude change across the conditions, especially for the responses to the ramped stimuli. Here it is important to point out the difficulty of defining the "right" maximal amplitude for the early responses to the ramped sounds (especially for the ramped condition of the 4 ms HLT), since it is usually a point on the slope developing to the later P50m peak. The latency of the P30m peak do not change with the HLT for the responses to the damped stimuli. However, as mentioned before, due to envelope coding the latency of the P30m source waveforms in a response to ramped stimuli do change with the HLT. As shorter the HLT as longer the latency.

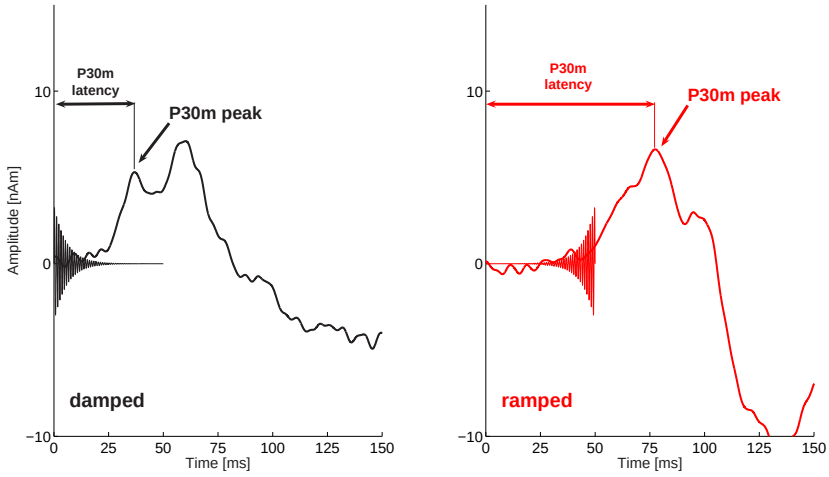

Figure 7.4: *An example of P30m evoked response to ramped and damped stimuli. Note the delayed response to the ramped stimuli due to the envelope coding*

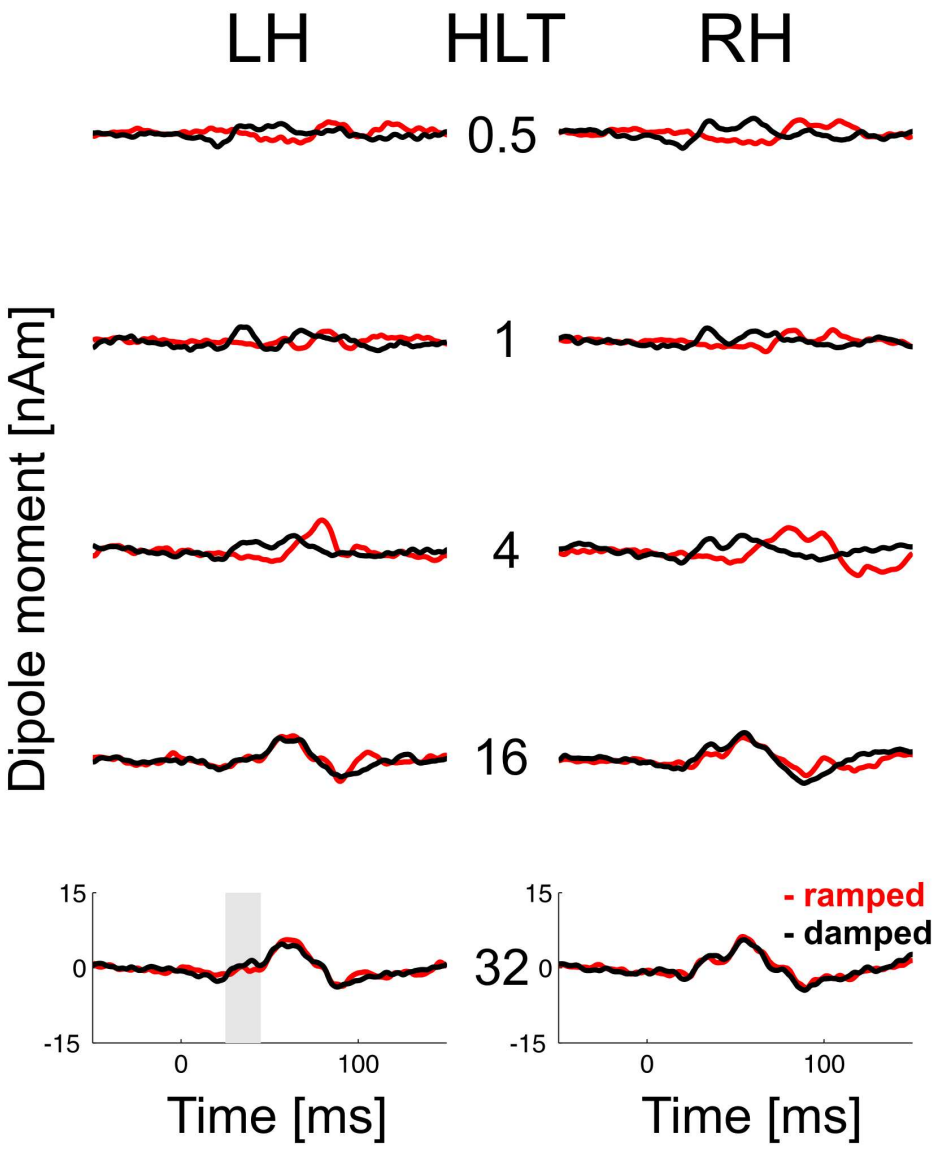

Figure 7.5: *P30m grand average source waveforms of both hemispheres for ramped and damped responses. Grey marked is a fit interval.*

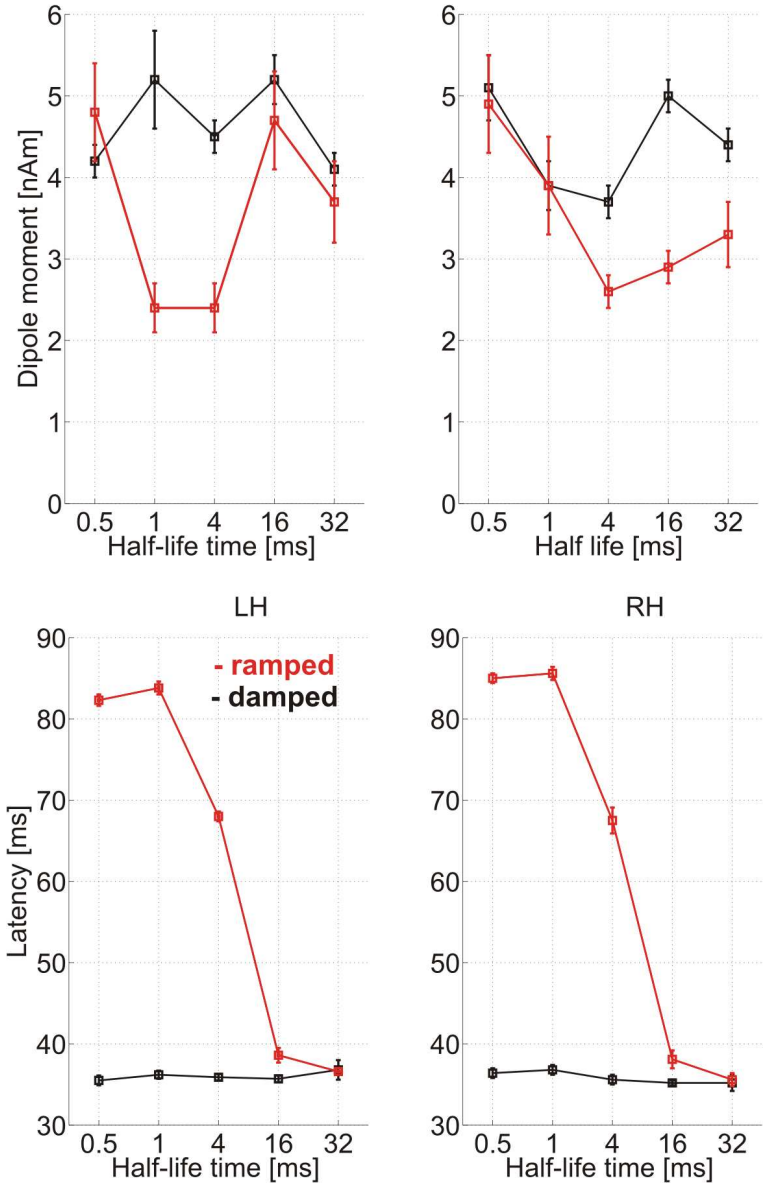

Figure 7.6: *P30m amplitudes (above) and latencies in the relation to the stimulus half-life time. P30m amplitudes differ significantly between conditions and hemispheres, especially for the ramped conditions. Latency difference between two types of stimuli (ramped vs. damped) is due to the specific envelope structure of the ramped stimuli.*

P30m amplitudes and latencies in a relation of the stimulus HLT are shown on the Figures 7.6, respectively. We localized the P30m sources bilaterally in the medial Heschl's gyrus, as shown on the Figure 7.23. Very similar observations are made for the P50m sources. They were localized in average in the lateral part of Heschl's gyrus, more lateral than the earlier P30m sources. A corresponding grand average source waveforms are shown at the figure 7.7 and the peak latencies and amplitudes in the table (see Appendix A) and on the Figure 7.8. Similar as before amplitudes do not differ significantly across the conditions and hemispheres. The latency delay of the ramped curves is again due to the envelope characteristics of the ramped stimuli.
The next processing stage is the N100m component. We localized it's sources bilaterally in average at the border between the lateral Heschl's gyrus (posterior duplication of Heschl's gyrus) and planum temporale. However, here we observe a high intersubject variability in the location of the sources. It can be partially explained by the high anatomical variability of planum temporale, the region considered to be one of the candidate sources of N100m component. The location of N100m sources projected to the 3D cortical surface reconstruction of one subject is shown on the Figure 7.12. On the Figure 7.9 the N100m source waveforms are shown. Here we observe the increase of amplitude as a function of HLT for all conditions (see Figure 7.10). In the right hemisphere the peak amplitudes are slightly higher than in the left hemisphere, but not statistically significant except for the 4 ms HLT ramped condition, where we get significant amplitude increase in the right hemisphere compared with the left hemisphere in terms of critical t-value. Furthermore, if we express the amplitude differences between ramped/damped pairs as an asymmetry index we get the plot shown on the Figure 7.16. In the left hemisphere the amplitude difference increase with the increasing HLT, whereas in the right hemisphere we get the clear peak at 4 ms HLT. The latencies behave as described earlier. The steeper the rise time, e. g. shorter HLT, the longer the latency as shown on the Figure 7.10. The table of amplitudes and corresponding latencies of N100m peak is given in Appendix A.

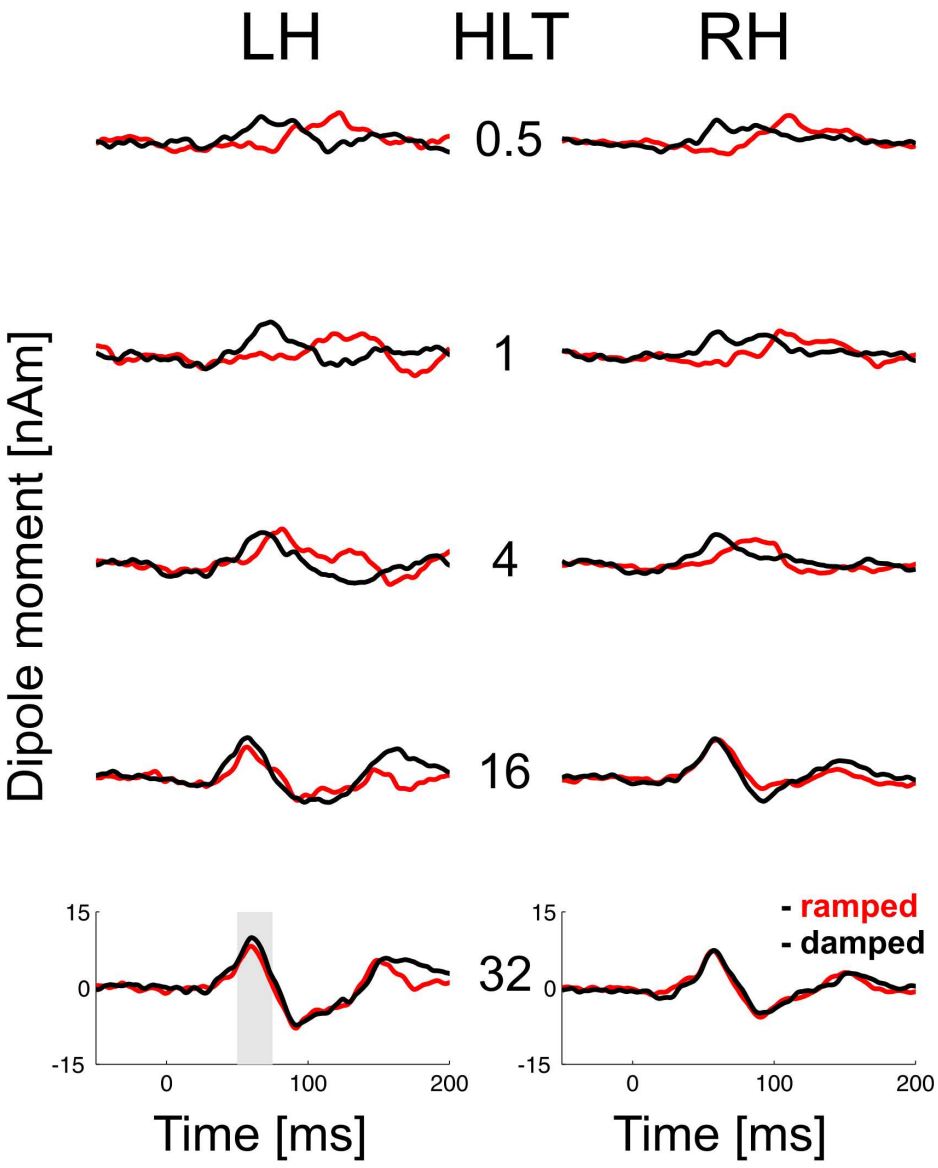

Figure 7.7: *P50m source waveforms. Grey marked is a fit interval.*

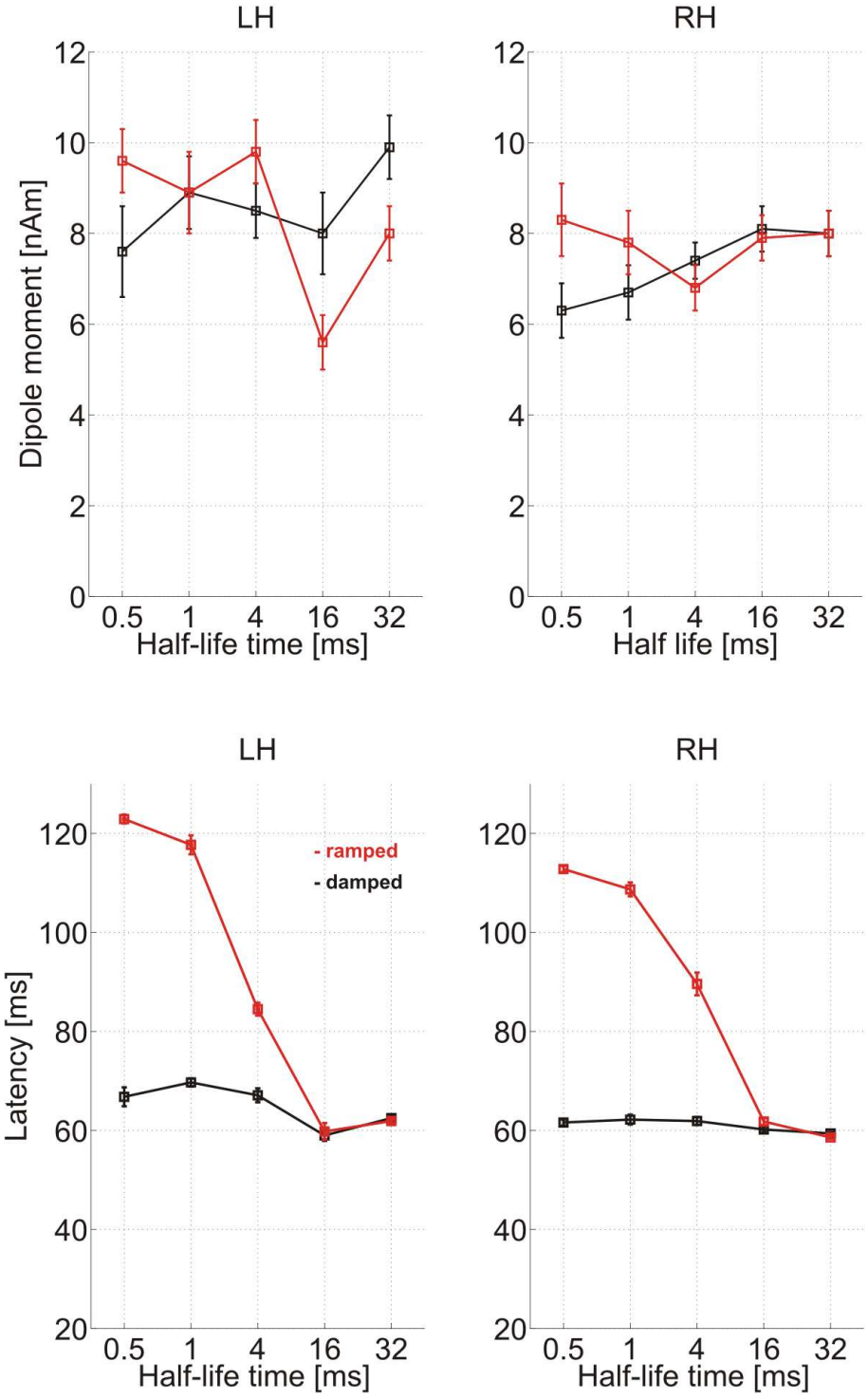

Figure 7.8: *P50m amplitudes and latencies in the relation to the stimulus half-life time. Amplitudes do not differ significantly across the conditions and hemispheres.*

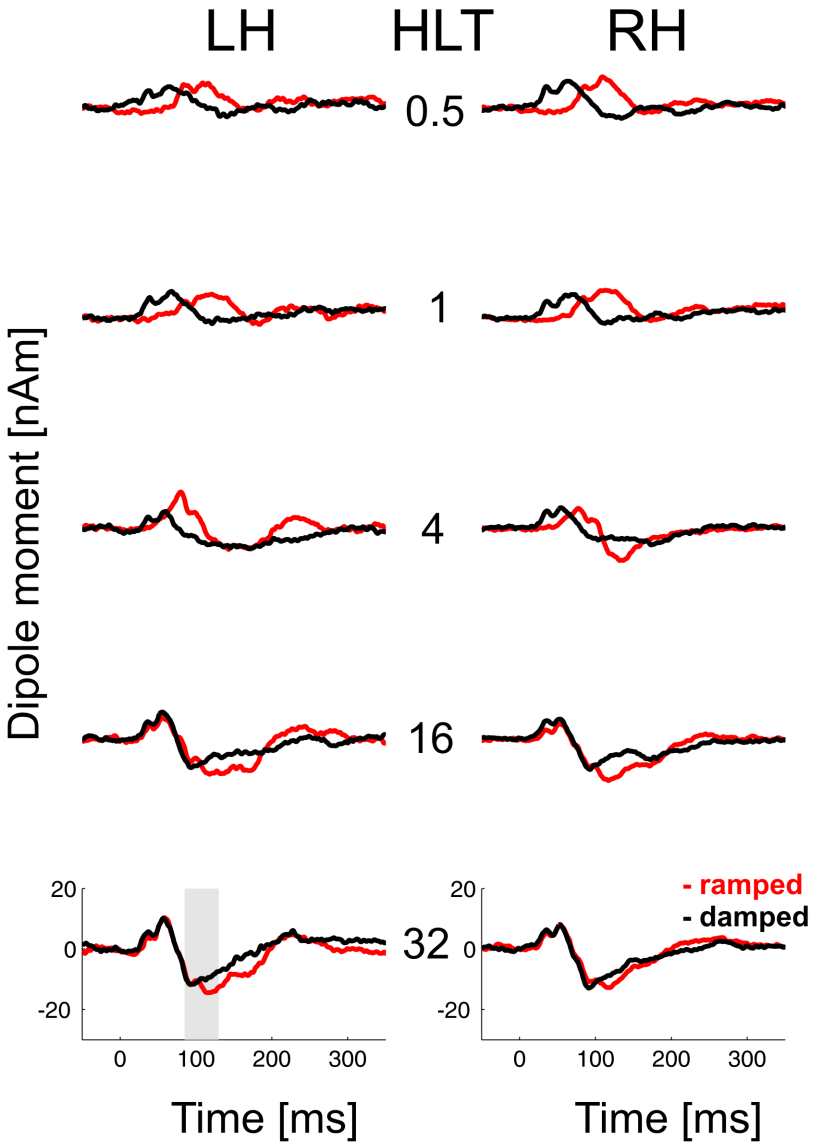

Figure 7.9: *N100m grand average source waveforms of both hemispheres for ramped and damped responses. Gray marked is a fit interval. Note the increase of the peak amplitude with the stimulus HLT.*

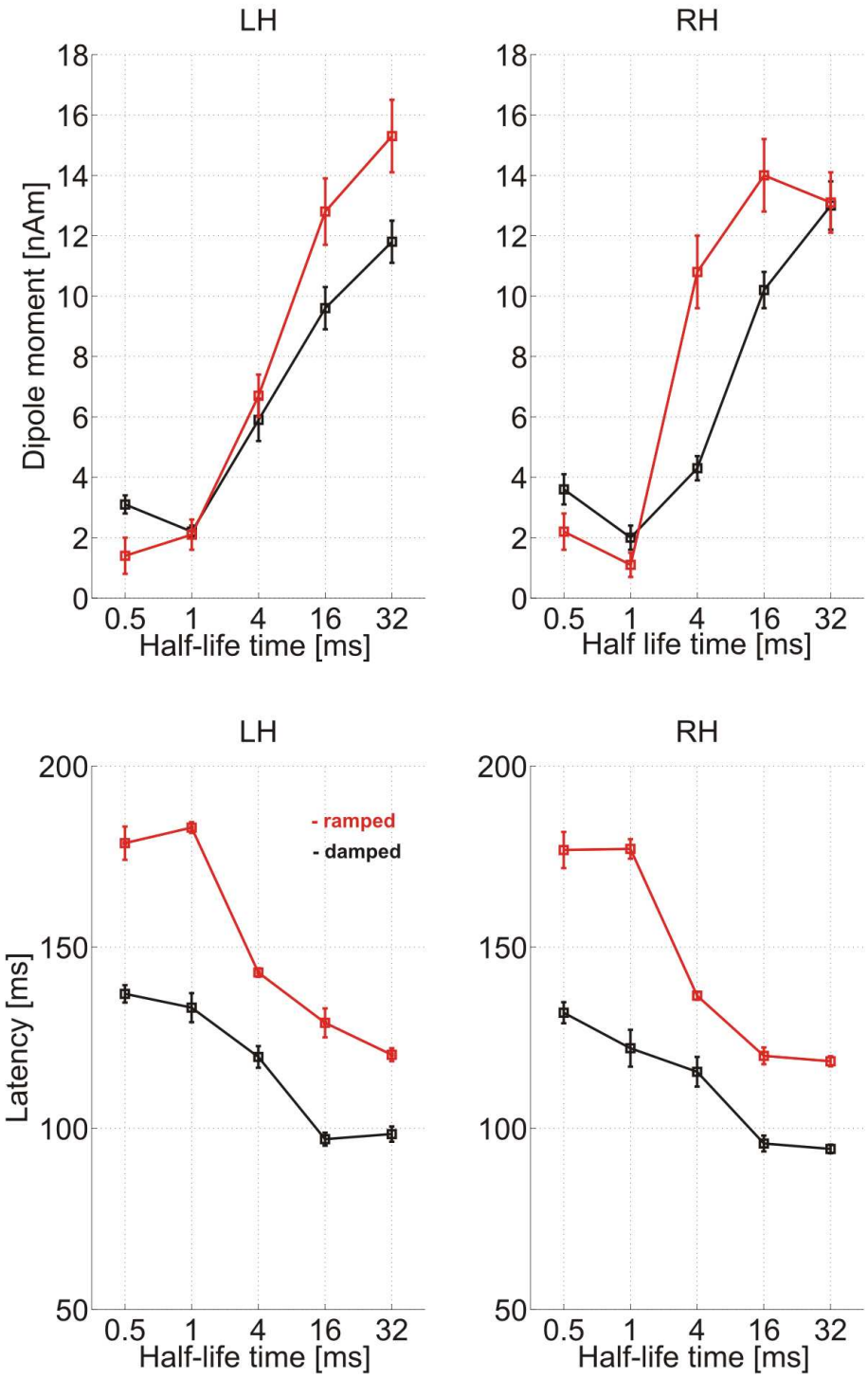

Figure 7.10: *N100m magnitudes and latencies in the relation to the stimulus half-life time. The amplitudes increase with the increasing HLT.*

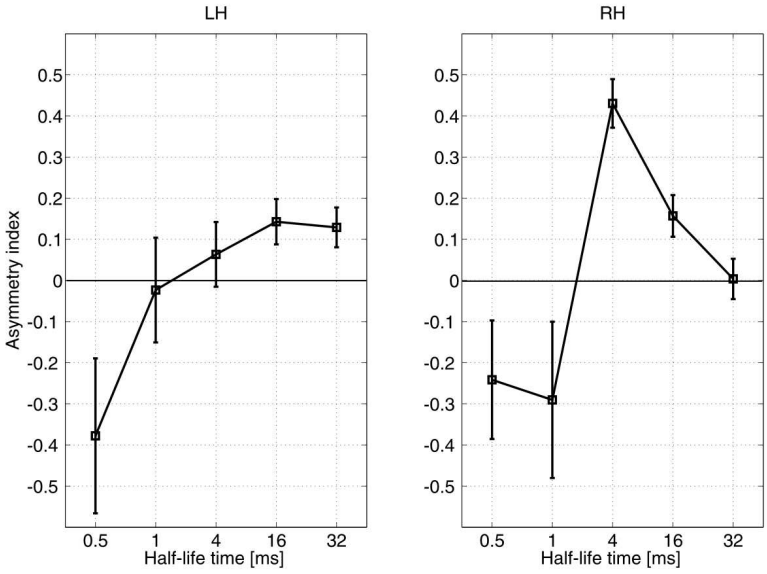

Figure 7.11: *N100m magnitude asymmetry index* $(AI = (Magn(ramp) - Magn(damp))/(Magn(ramp) + Magn(damp))$ *as a function of stimulus HLT. Note the peak at 4 ms HLT and the hemispherical asymmetries.*

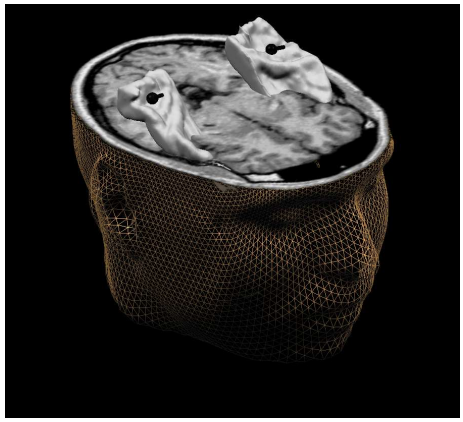

Figure 7.12: *Projection of N100m dipoles to the 3D head reconstruction obtained using anatomical MRI.*

7.3.2 Neuromagnetic Sources Evoked by Steady State Stimulation

The stimuli used in this experiment were sounds consisted of 20 single cycles from previous experiment. These stimuli evoked slow transient cortical responses and sustained and steady state fields as shown on the Figure 7.14. Because there were less response averages compared with the previous experiment the signal-to-noise ratios were suboptimal. Therefore we were not able to separate the early transients, e.g., P30m and P50m sources.

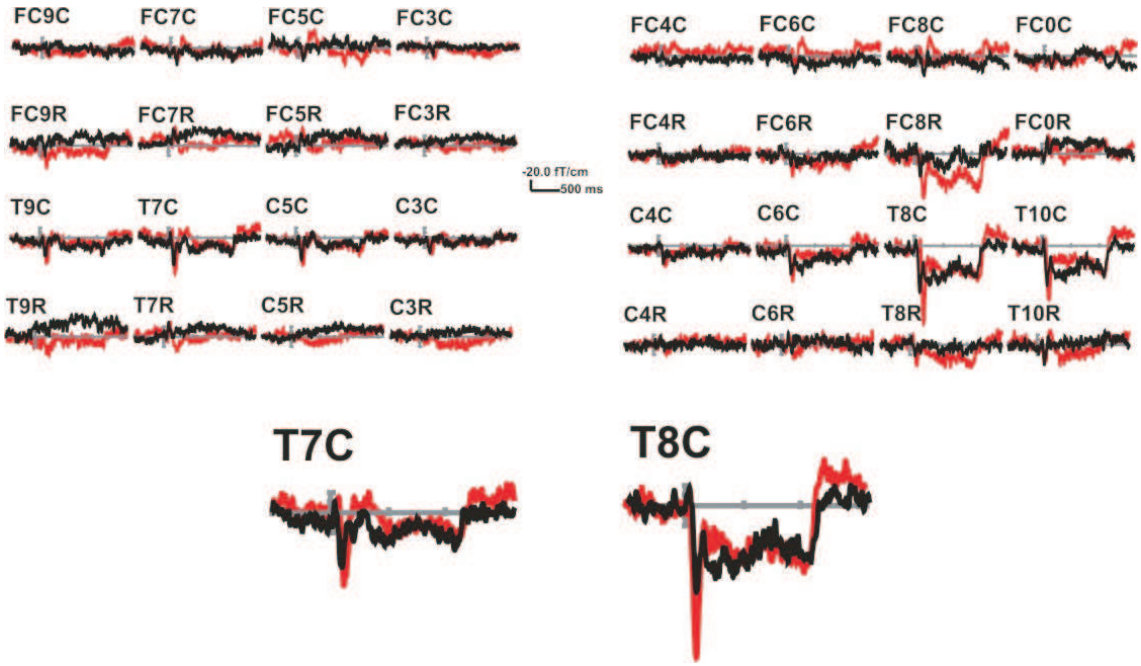

Figure 7.13: *Top view of the averaged auditory evoked neuromagnetic responses of one subject to steady state ramped and damped sounds of 4 ms HLT.*

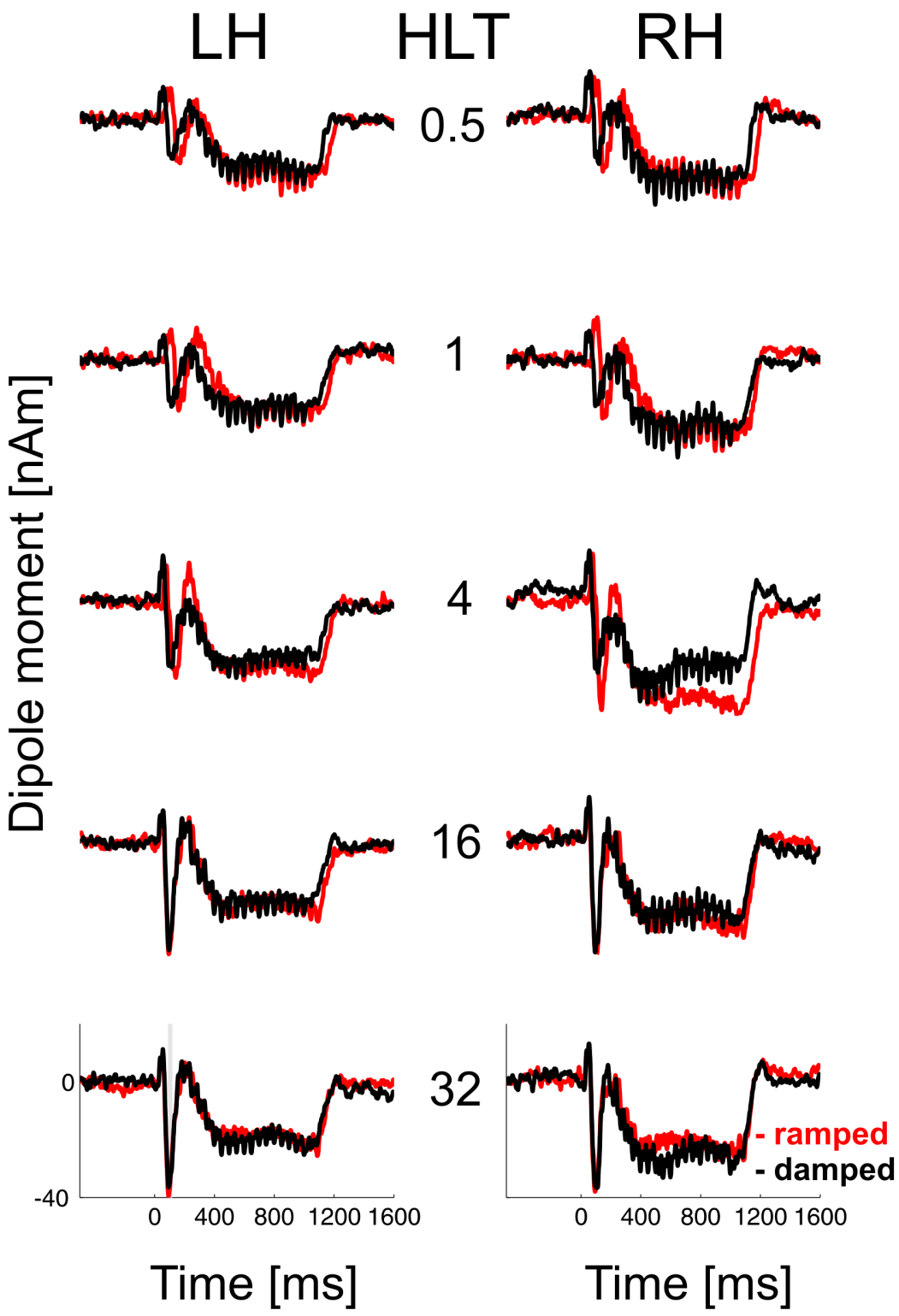

Figure 7.14: *N100m source waveforms evoked by steady state sounds. Note the increase of the amplitudes with the increasing HLT*

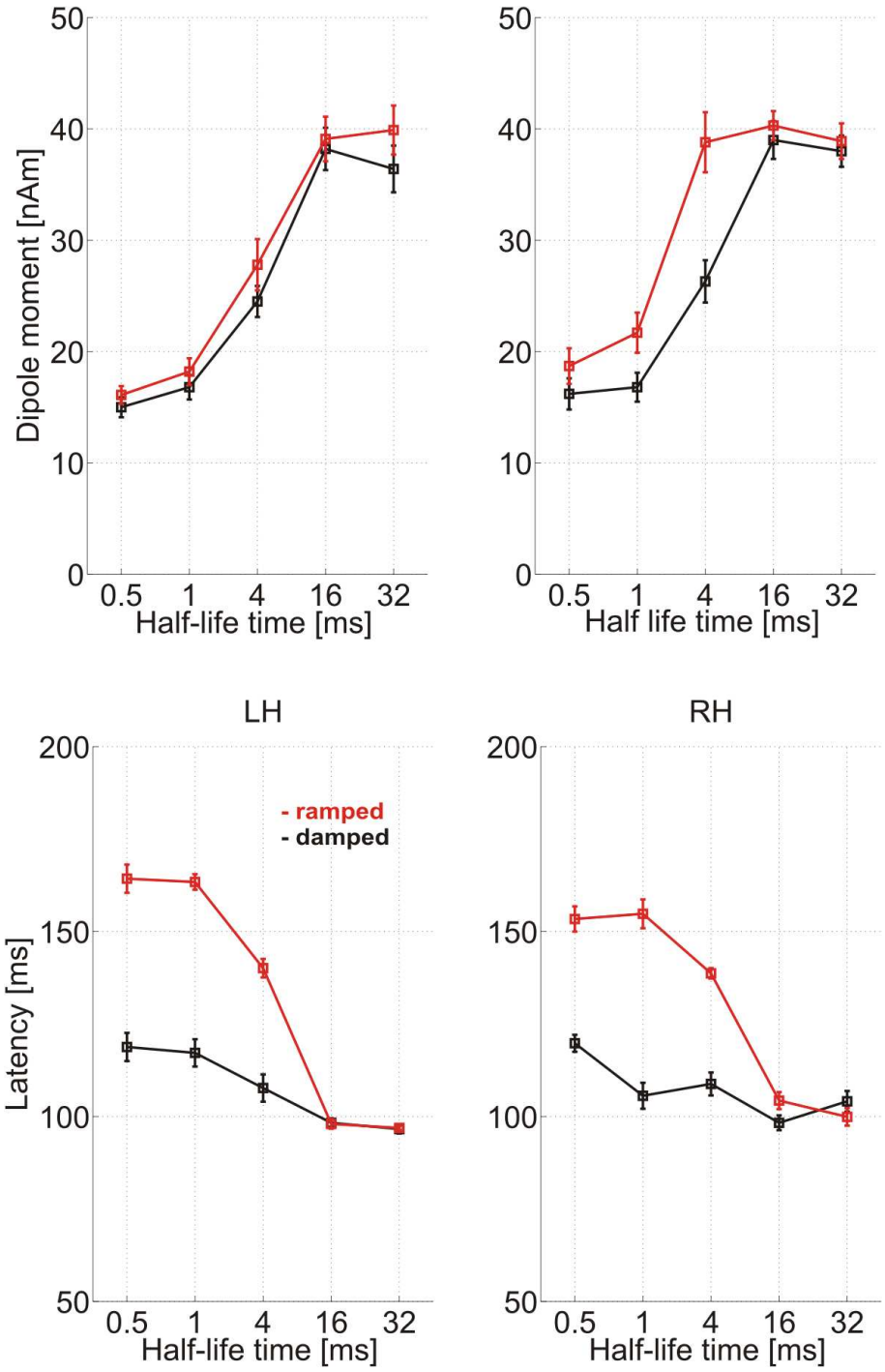

Figure 7.15: *N100m peak amplitudes and latencies in a relation to the half-life time of the stimuli. According to the bootstrap assessed t-intervals, the amplitudes evoked by presentation of the ramped stimuli of the 4 ms HLT was increased significantly compared to the amplitudes evoked by damped stimuli of the same HLT, and the right hemisphere amplitudes were significantly higher compared with the left hemisphere amplitudes .*

The first next prominent peak is N100m. N100m sources were in average localized in the Planum temporale. The projection onto the Talairach map is shown on the Figure 7.23. Here we observe similarly as in the previous experiment the increase of the peak amplitude with the increasing HLT for all conditions. N100m latencies decrease with the increasing HLT. Asymmetry index vary dependant of the stimulus HLT peaking at the 4 ms HLT condition. In the right hemisphere are again slightly higher amplitudes compared with the left hemisphere, but significantly just for the ramped 4 ms HLT condition. Corresponding values of amplitudes and latencies are given in Appendix A.

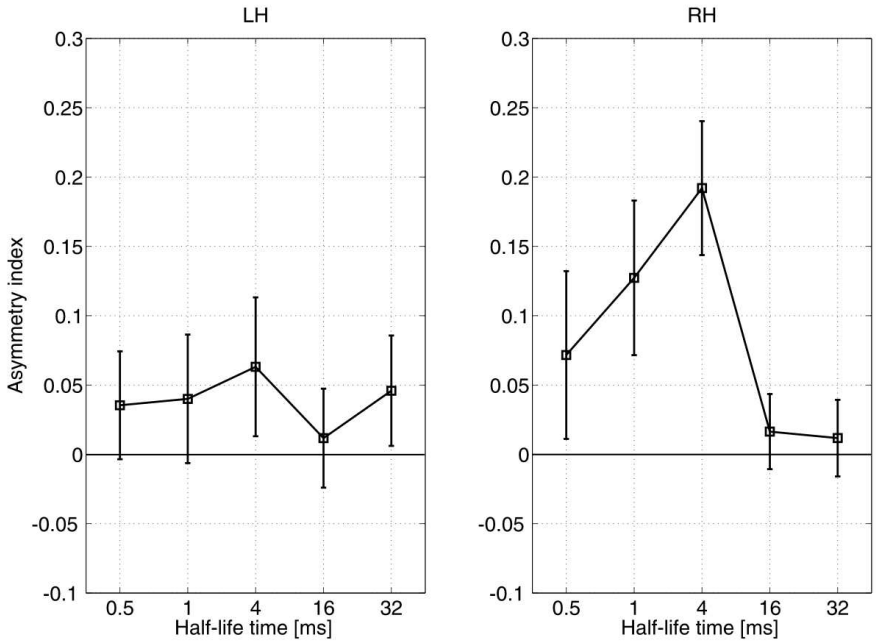

Figure 7.16: *N100m amplitude asymmetry index as a function of stimulus HLT. Note the peak at 4 ms HLT and the hemispherical asymmetries.*

Furthermore we localized P200m peak in the medial portion of Heschl's gyrus. To obtain peak amplitudes and latencies it was necessary to remove the sustained field using bandpass filter in a range between 4–100 Hz. The filters were applied after spatio-temporal analysis, because although they improve signal–to– ratio they mix up the topographical components of time-adjacent components and yields erroneous results (Yvert et al.,2001). On the Figure 7.17 P200m source waveforms are shown.

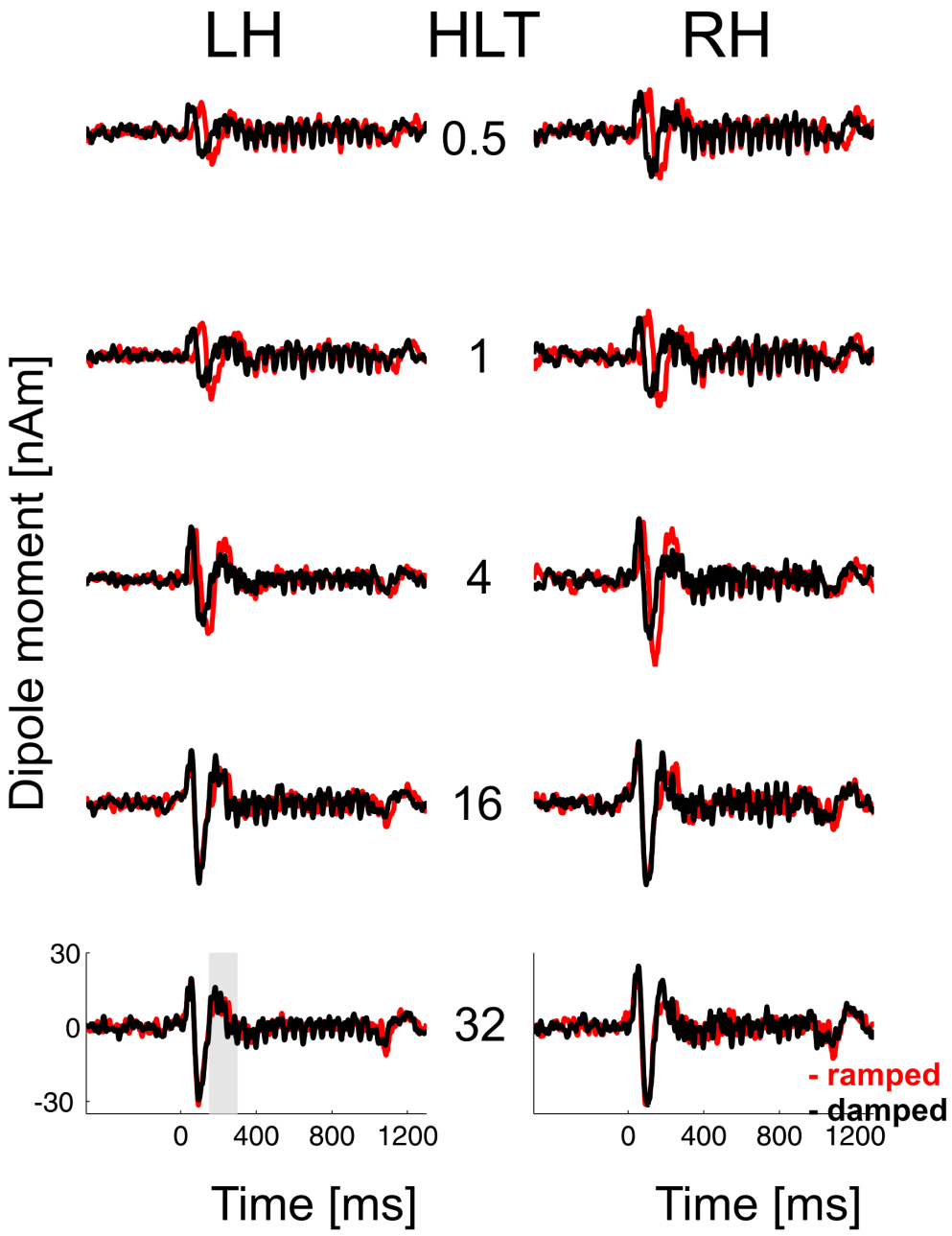

Figure 7.17: *P200m grand average source waveforms evoked by steady state sounds. Note the amplitude asymmetry for the 4 ms HLT condition.*

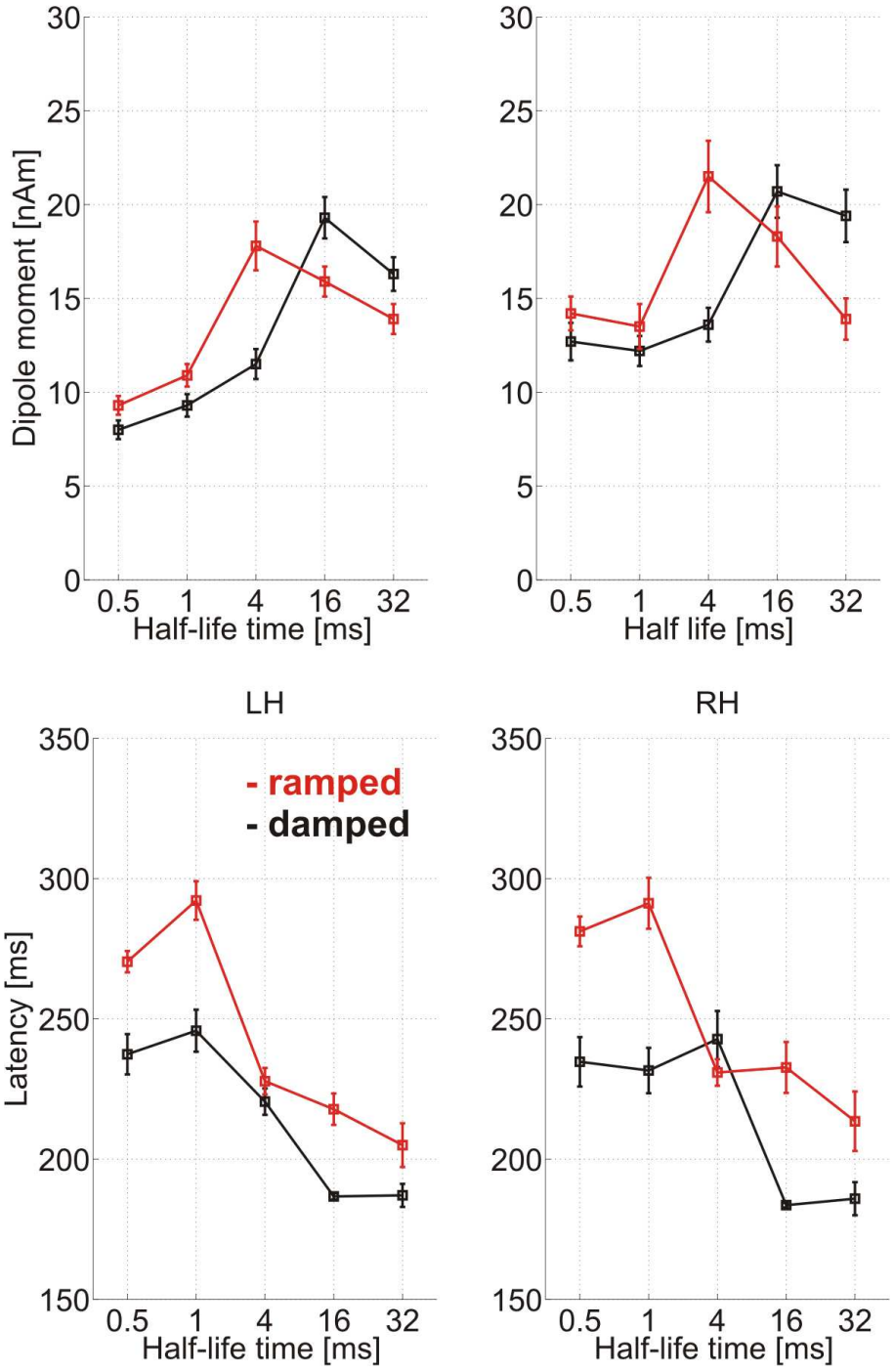

Figure 7.18: *P200m amplitudes and latencies expressed as a function of condition HLT.*

The P200m amplitudes do increase with the stimulus HLT similarly as N100m amplitudes. Latencies follow the envelope. The table of amplitudes and latencies is shown in Appendix A. If we express the amplitudes in a terms of asymmetry index we observe the bilateral peak at 4 ms HLT condition as shown on the Figure 7.19

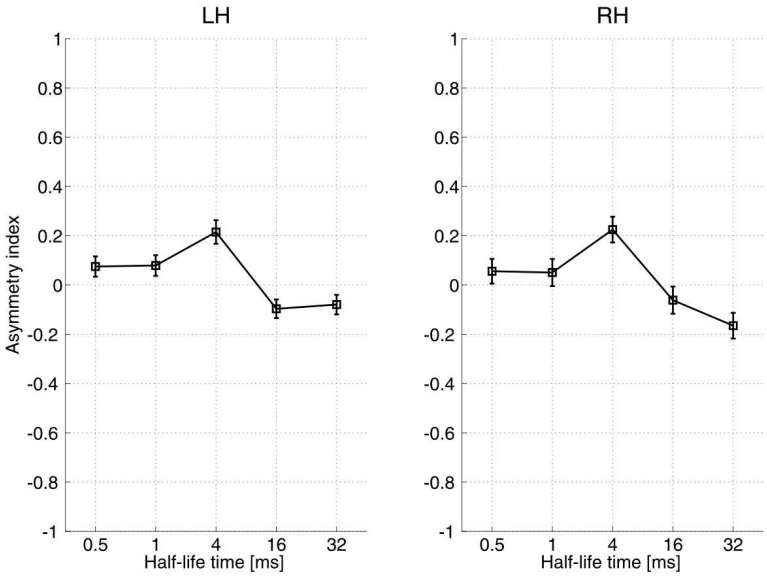

Figure 7.19: *P200m amplitude asymmetry index in a relation to the stimulus HLT. Note bilateral peak at 4 ms HLT condition.*

The final processing stage in our analysis were the sustained fields. We localized sustained fields in the medial portion of Heschl's gyrus, slightly more medial than P200m sources as shown on the Figure 7.23. Corresponding source waveforms are shown on the Figure 7.20. Since we wanted to test if the effect of temporal asymmetry affects the processing on the level of sustained/steady state field we decided to calculate the area under the curve as described in the Materials and Methods section of this Chapter. Furthermore to quantify the asymmetry we defined a new asymmetry index using areas instead of amplitudes as variables (see Figure 7.21). Here we observe a small peak for the 4 ms HLT condition in the right hemisphere but this effect was not significant according to the critical t-values obtained using Bootstrap (see Figure 7.22).

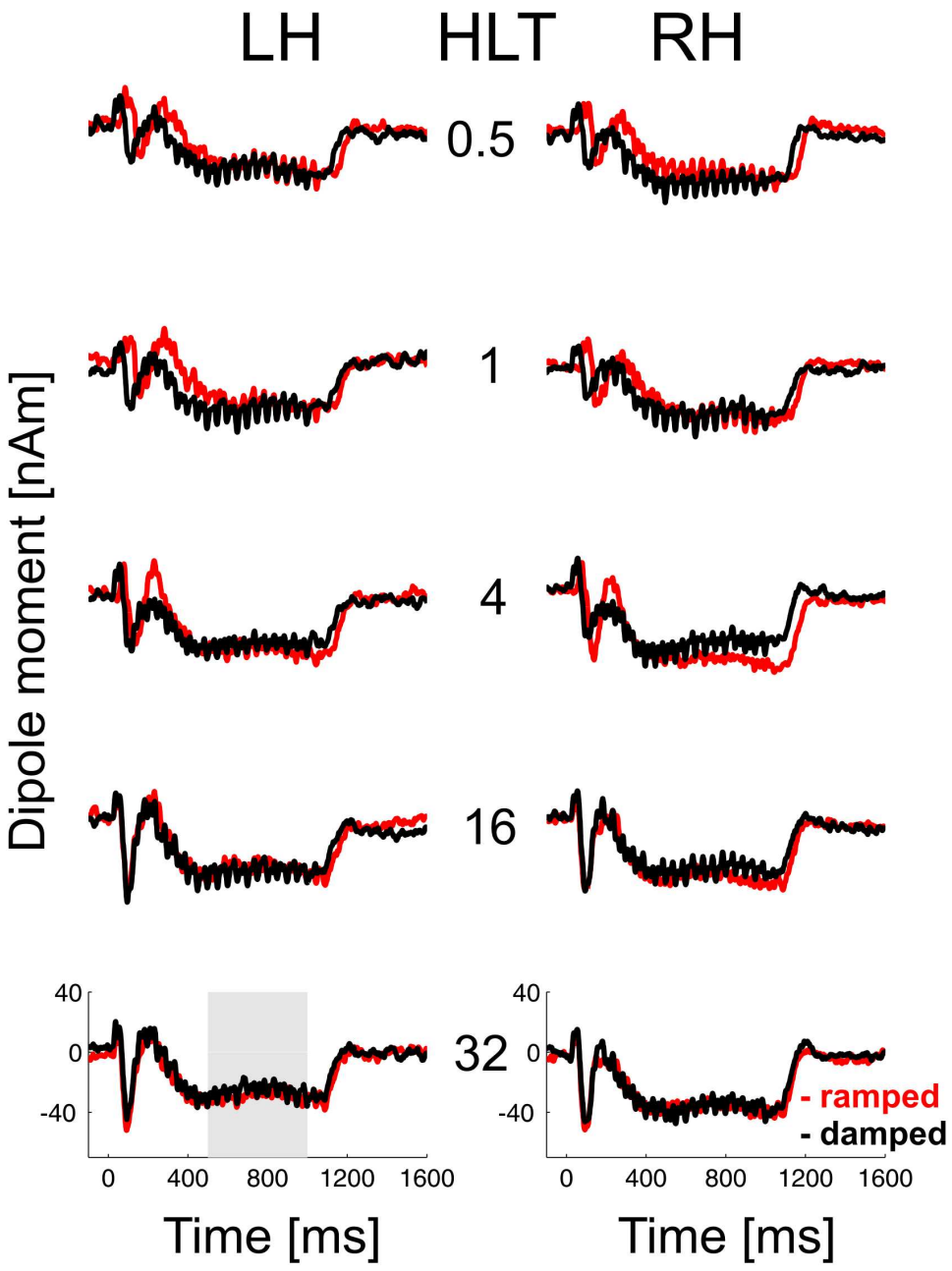

Figure 7.20: *Grand average sustained field source waveforms. Sustained field does not change with the stimulus HLT. There are no specific significant differences related to the conditions and hemispheres.*

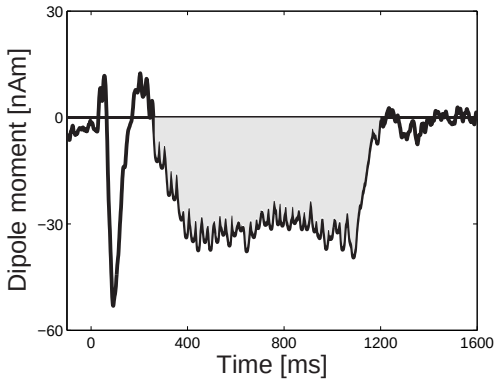

Figure 7.21: *The area of the sustained field was calculated between the zero-points of the sustained field as marked above.*

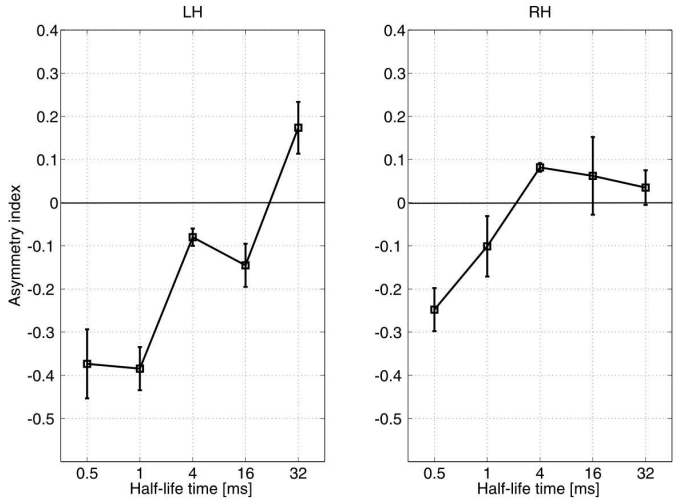

Figure 7.22: *Asymmetry index obtained using sustained field areas instead of amplitudes in a relation to the stimulus HLT.*

7.3.3 Transient vs. Steady State Stimulation

Temporally asymmetric transients evoke typically transient responses, e. g., P30m, P50m and N100m. We localized them and described their temporal dynamics. Steady state stimuli evoke additionally P200m component and sustained/steady state fields. The only component, in both experiments analyzed, was N100m. If we compare the morphology of the source waveforms of N100m evoked using transient stimulation with the ones evoked using steady state stimulation, we see that steady states evoke a single sharp peaks with well-defined maxima, whereas the transients develop slowly especially for the short stimuli HLTs. They often display one or two shoulders before the global maxima are reached. Furthermore comparison of amplitudes and latencies show that the amplitudes evoked using transient stimuli vary between 2 and 13 nAm for the damped sounds and 1 to 15 nAm for the ramped sounds. Their steady state counterparts reach the amplitudes of 15–38 nAm and 16–40 nAm respectively. In both experiments we observe the amplitude increase with the increasing stimulus HLT. Latencies in the first experiments were in the range 94–137 ms for the damped stimuli and 118–183 ms for the ramped stimuli. In the second experiment they are in the range between 96 and 118 ms for the damped stimuli and 96 and 164 ms for the ramped stimuli. Because of the specific envelope structure of the ramped stimuli the latencies are much longer compared with their damped counterparts. For both stimuli types and both modalities (transients and steady states) we observe the decrease of the latencies with the stimulus HLT. If we compare the asymmetry indices for both experiments, we observe the more pronounced asymmetry for the N100m sources evoked using single cycle stimuli (0.4 against 0.2 AI value for the right hemisphere). Here is important to note that the asymmetry is always more pronounced in the right hemisphere. In the following table are given the Talairach coordinates for all localized sources in both experiments. They are then projected onto the Talairach map obtained from Schneider et al. (2005).

Source	Talairach coordinates					
	Left Auditory Cortex			Right Auditory Cortex		
	x	y	z	x	y	z
Experiment 1						
P30m	-50 ± 6	-16 ± 9	18 ± 7	50 ± 7	-16 ± 4	17 ± 7
P50m	-50 ± 7	-21 ± 7	18 ± 8	54 ± 6	-17 ± 7	18 ± 10
N100m	-48 ± 10	-27 ± 6	9 ± 8	50 ± 6	-22 ± 10	9 ± 5
Experiment 2						
N100m	-52 ± 7	-24 ± 13	13 ± 8	49 ± 6	-25 ± 13	14 ± 12
P200m	-45 ± 8	-16 ± 9	9 ± 5	44 ± 10	-15 ± 9	10 ± 9
SSR	-45 ± 10	-19 ± 9	8 ± 5	44 ± 7	-17 ± 9	9 ± 8

Table 7.1: *Talairach coordinates of the auditory evoked sources using temporally asymmetric sounds.*

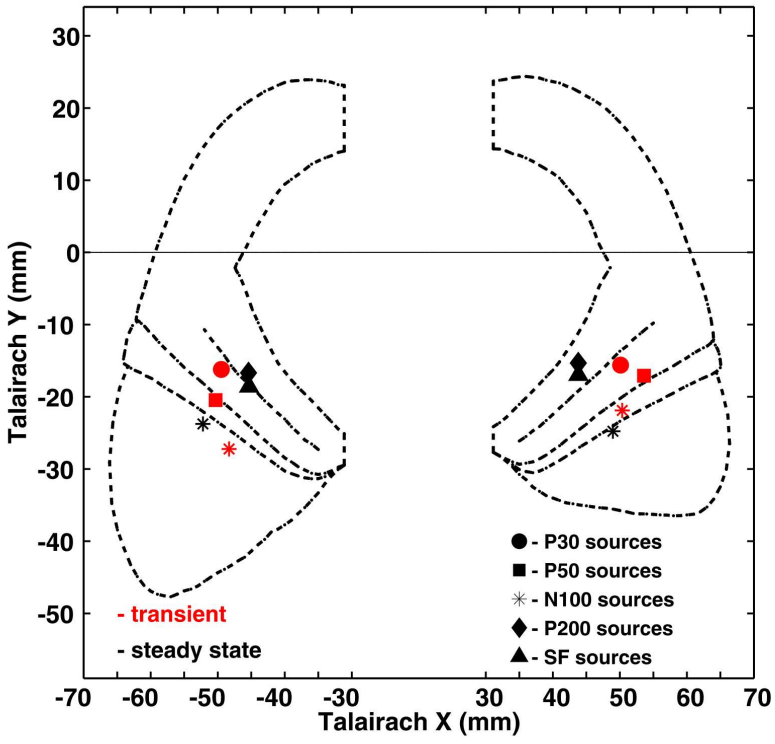

Figure 7.23: *Dipole locations in the space of Talairach and Tournoux (1988). The coordinates of the sulcal borders were provided by Schneider et al. (2005). The standard errors of mean are within the point size.*

Chapter 8

FMRI Experiments

8.1 Introduction

Since any MEG dipole analysis suffers from the non-uniqueness of the inverse problem, it is desirable to access and locate specific brain activation in a model-independent way. Functional magnetic resonance imaging allows the visualization of neuronal processes in the brain with a spatial resolution of several millimeters and a temporal resolution of a few seconds. The combination of anatomical and functional images from the same listeners is a unique tool to directly relate physiological processes to anatomical structures in the brain region of interest. Therefore we designed a functional magnetic resonance imaging (fMRI) experiment using sparse imaging to investigate the hemodynamic changes related to the processing of temporally asymmetric sounds.

8.2 Materials and Methods

8.2.1 Subjects

Seven right handed subjects (three females), aging 23–43 (mean age 32.5) with no history of peripheral or central hearing disorders participated in the experiments after having given informed consent.

8.2.2 FMRI Design and Parameters

Stimuli were steady state ramped and damped sounds (see subsection 5.1.2) of 6 s duration and three different HLTs, e.g. 0.5, 4 and 32 ms. A variation of Halls (Hall et al.,1999), method of sparse temporal sampling was used to examine MR signal changes related to a processing of temporally asymmetric sounds. Functional magnetic resonance imaging (fMRI) was performed on a Siemens Trio (Siemens, Erlangen, Germany) 3 Tesla whole-body MRI scanner, equipped with echo planar imaging (EPI) capabilities using the eight channel head coil for radio-frequency (RF) transmission and signal reception. A gradient-echo EPI-sequence with an effective repetition time (TR) of 9 s, an echo time (TE) of 30 ms and a matrix of 64x64 was used. Using a midsagittal scout image, a total of 12 axial slices (3.7x3.7x3.4 mm voxel size), parallel to the Silvian fissure, were acquired over 1s each 9 s. Initiation of the first set of 12 slices was triggered by a TTL pulse from a PC and all subsequent MR acquisitions were synchronized with stimulus presentation. A high resolution T1 and T2 weighted scans were acquired for each subject for anatomical coregistration. Using a variation of the sparse temporal sampling method, one set of 12 axial slices was acquired every 9 s to circumvent the scanner noise interferences. The stimulation started always 2 s after the acquisition to allow a partial recovery of the BOLD response. Stimuli were presented in a set of five of each type with the silence condition between two sets (see Fig. 8.1). 40 volumes were acquired for each condition with exception of the silence condition where 42 volumes were acquired.

FMRI Data Analysis

All structural and functional data were processed and analyzed using SPM2 ($http://www.fil.ion.ucl.ac.uk/spm$). Each BOLD and the structural time series was realigned to its first image and then normalized using affine nonlinear spatial transformations. All individual functional images were corregistred with the corresponding T1 images and normalized using standard SPM T1 template. The final preprocessing step, smoothing was performed using a Gaussian filter of 5mm (full width at half maximum). The location of the significant BOLD responses, given in MNI space, were transformed into Talairach coordinates. Fixed-effects analysis was carried out on individual (282 scans) and group data (1974 scans) using the SPM general linear model. The following contrasts were calculated:

- 'sound–silence' for all six conditions, e. g., ramped-damped pairs for 0.5, 4 and 32ms HLTs.

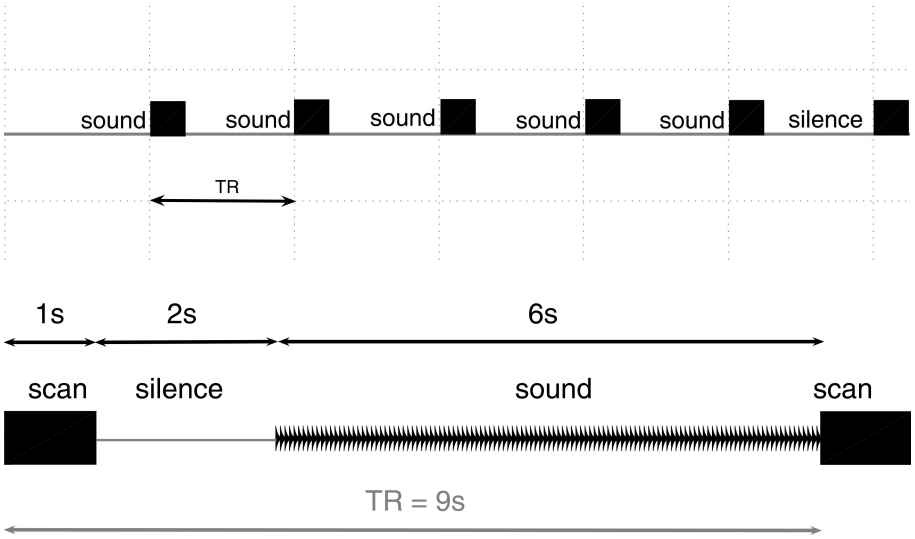

Figure 8.1: *Experimental paradigm design. The variation of sparse imaging allowed for investigating the BOLD signal in response to the stimuli rather than to the scanner noise. Below is the stimulation timing between two scans. Two seconds of silent interval between the acquisition and the stimulus provides the partial recovery of the BOLD response.*

- 'sound–sound' for all ramped-damped pairs.

The significance level for group activation was $p < 0.05$ corrected for multiple comparisons across the whole volume. This corresponded to a height threshold of $T > 4.95$. Thus, the group contrasts presented only show highly significant activations. In contrast, the presentation of the individual data is much less conservative as to illustrate the inter-individual variability. The significance level for the individual data was set to $p < 0.001$ (uncorrected), which corresponded to a height threshold of $T > 3.13$.

8.3 Results

The contrast 'sound-silence' revealed bilateral activation over auditory sensory areas. Figure 8.2 shows a group analysis for all listeners of the activation in a response to a sound, when the six sound conditions together are contrasted with the silence. Three main clusters spread over the Heschl's gyrus and planum temporale were localized. The most prominent cluster was in the medial portion of Heschl's gyrus. All Talairach coordinates are given in Table 8.1 and projected onto Talairach map in the Figure 8.3. Each of the six sounds, when contrasted separately with the silence produces similar patterns of activation with an exception of 4 ms HLT conditions. Individual data for all 'sound–silence' conditions of one subject are shown on the Figure 8.4.

Cluster	Talairach coordinates							
	Left Auditory Cortex				Right Auditory Cortex			
	x	*y*	*z*	*T-value*	*x*	*y*	*z*	*T-value*
mHG	-35	-22	8	19.95	40	-24	12	15.80
lHG	-48	-6	2	15.58	50	-18	4	10.55
PT	-45	-26	12	13.72	48	-28	8	13.91

Table 8.1: *Talairach coordinates of the six main clusters from the left and right auditory cortex localized using fMRI.*

Furthermore on the Figure 8.7 the 'ramped–damped' contrast for 4 ms HLT of all subjects is given. Individual data are presented with a much lower significance threshold to illustrate the inter-individual variability, rather than the very robust effects we are interested in when analyzing the whole group. The inter-individual

variability of the auditory responses might be largely explained by individual anatomical differences. The group data for 'ramped–damped' contrast of 4 ms HLT are shown in Fig. 8.7. Group analysis for ramped–damped contrast of 4 ms HLT revealed two clusters in right auditory cortex and one cluster in left auditory cortex. All Talairach coordinates are given in the Table 8.2 and projected onto the Talairach map on the Figure 8.6. There was no significant activation for 'sound–sound' difference condition of 0.5 and 32 ms HLT (neither ramped–damped nor opposite).

Cluster	Talairach coordinates							
	Left Auditory Cortex				Right Auditory Cortex			
	x	*y*	*z*	*T-value*	*x*	*y*	*z*	*T-value*
mHG	-35	-22	8	9.14	38	-28	8	8.7
PT	-	-	-	-	46	-34	16	11.13

Table 8.2: *Talairach coordinates of the three clusters from the left and right auditory cortex for the ramped–damped contrast of 4 ms HLT localized using fMRI.*

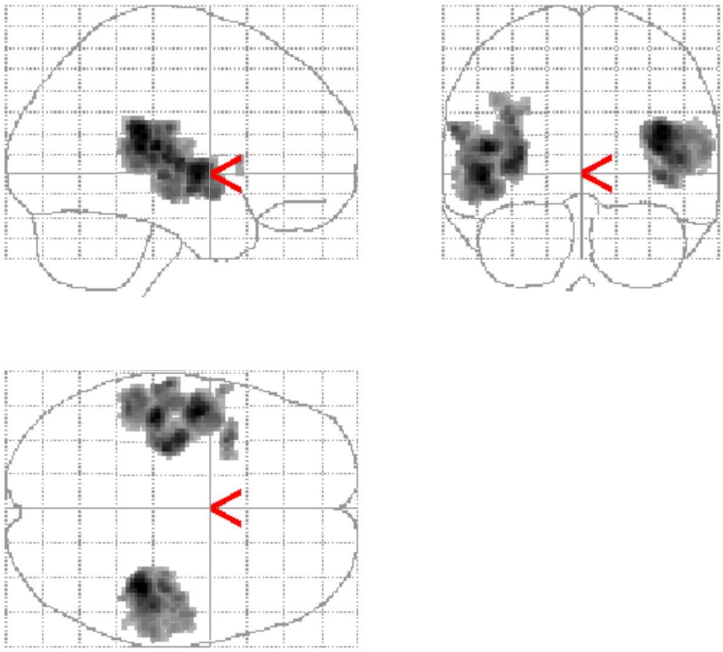

Figure 8.2: *Group analysis of activation in response to sound. All sound conditions together were contrasted with silence, ("glass brain" view, threshold for significance* $p < 0.05$ *corrected for multiple comparisons). Activation is largely confined to the temporal lobes bilaterally. The main peaks are in the Heschl's gyrus and planum temporale.*

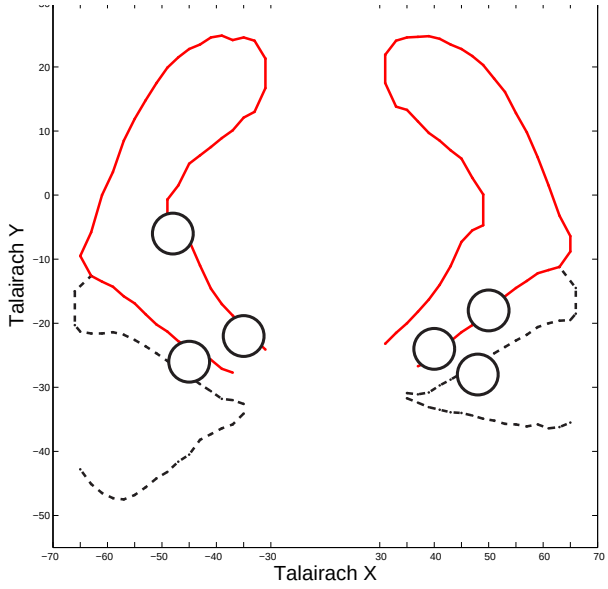

Figure 8.3: *Talairach coordinates of the six main clusters localized for all 'sound–silence' conditions projected onto the map adapted from Schneider et al. (2004). Their extension is roughly indicated by the radius of the circles.*

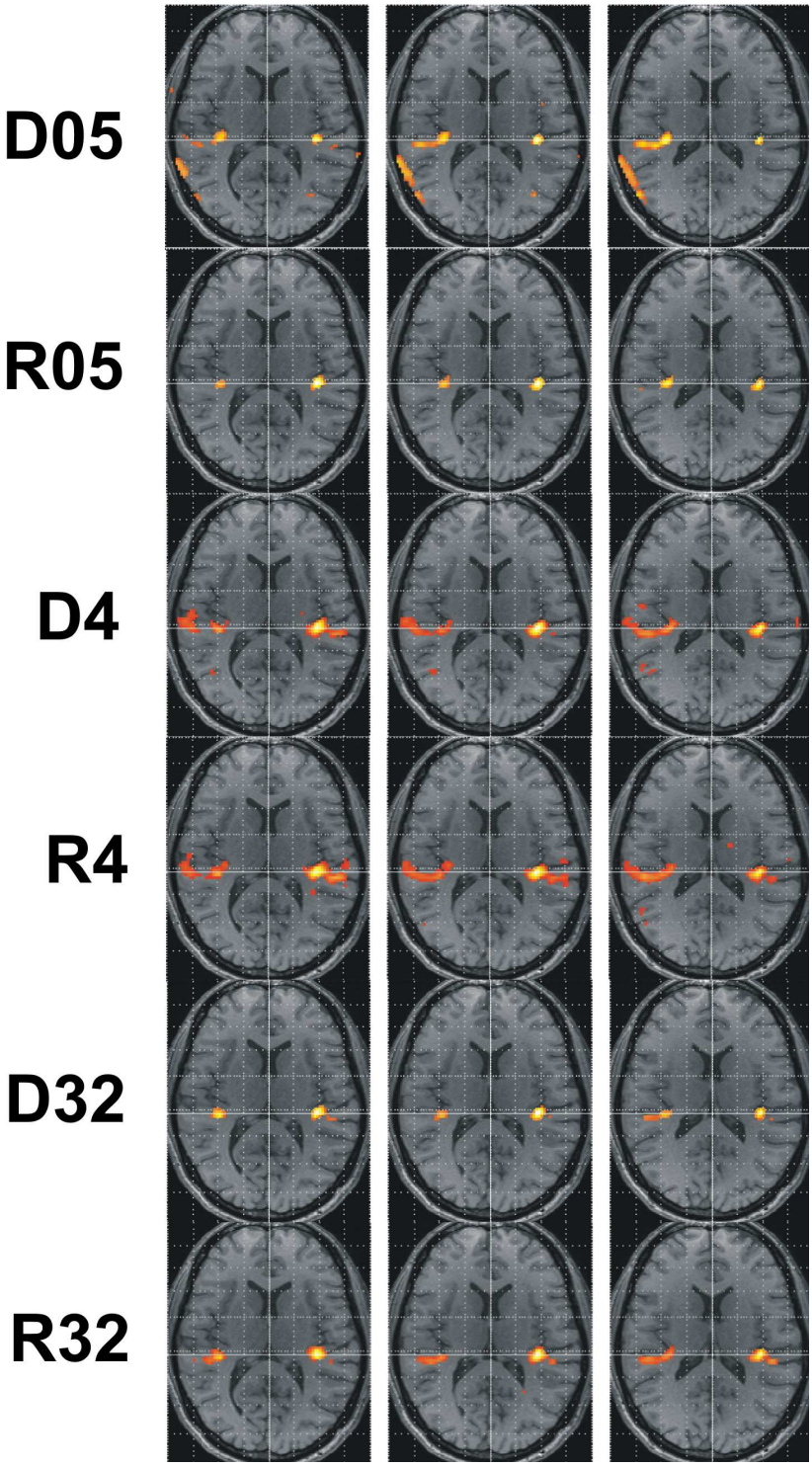

Figure 8.4: *Individual data of one subject for all 'sound–silence' conditions as projected onto the structural data of the high resolution anatomical MRI scan (in axial steps of 2mm starting from z = -2.*

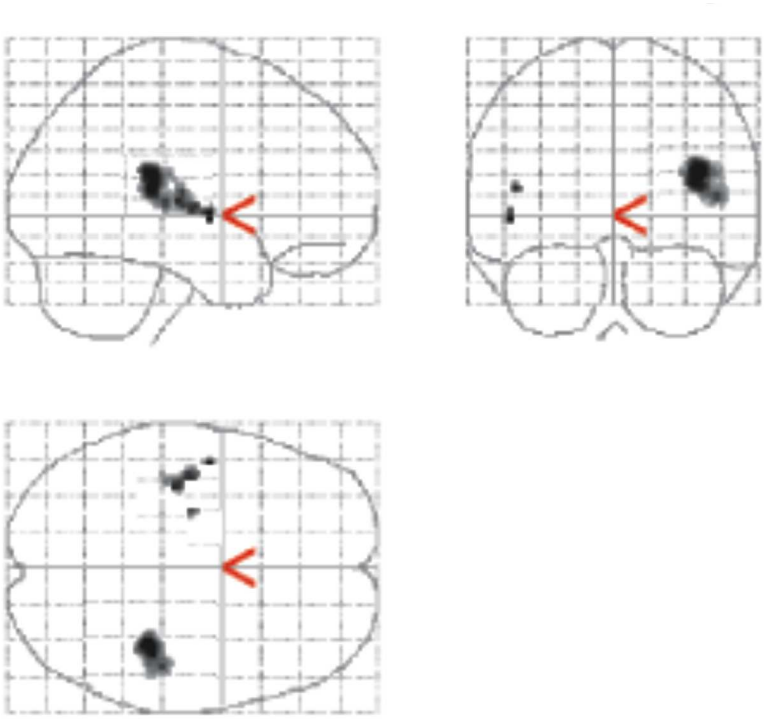

Figure 8.5: *Group Data: Contrast between ramped and damped conditions of 4 ms HLT ("glass brain" view, threshold for significance $p < 0.05$ corrected for multiple comparisons). Note the hemispherical asymmetries of the activity pattern.*

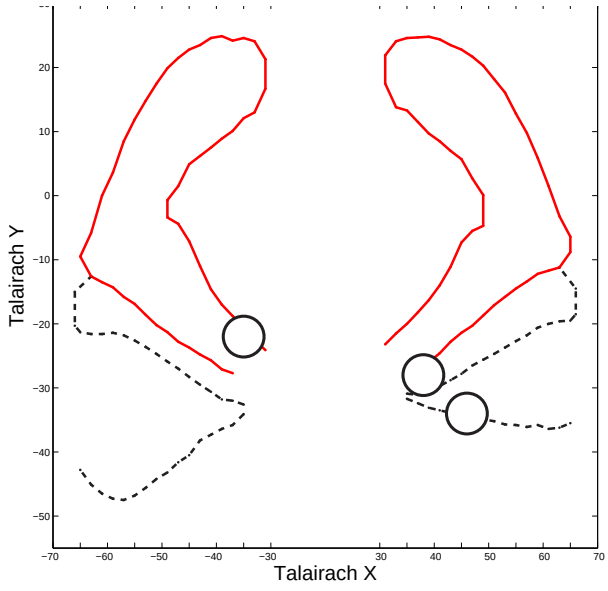

Figure 8.6: *Talairach coordinates of the significant clusters localized for ramped–damped contrast of 4 ms HLT projected onto the map adapted from Schneider et al., (2004). Their extension is roughly indicated by the radius of the circles.*

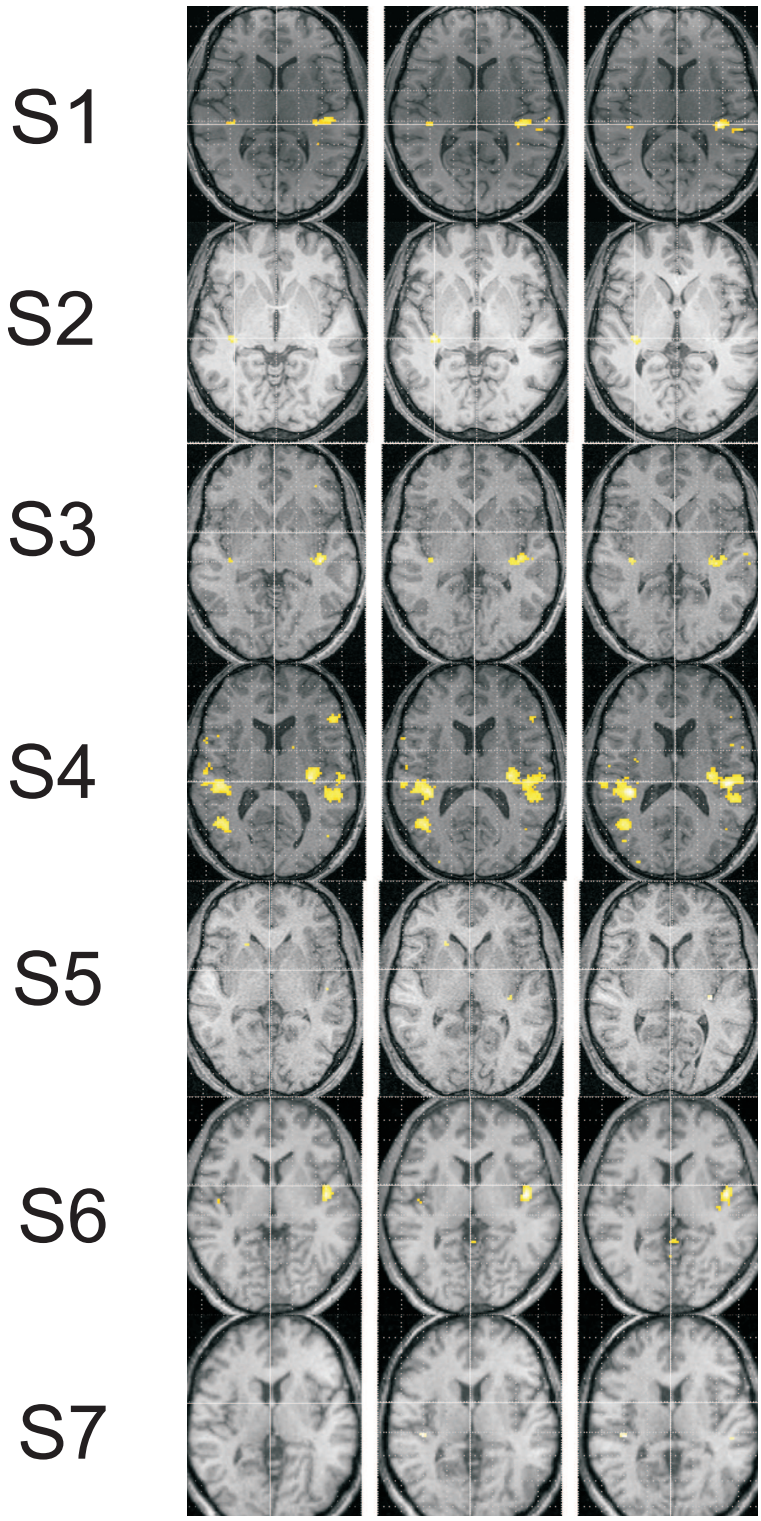

Figure 8.7: *Individual Data: Contrast between ramped and damped conditions of 4 ms HLT. In contrast to the presentation of the group data, here a less conservative statistic was used to illustrate the inter-individual variability. An uncorrected significance level of* $p < 0.001$ *was applied, which corresponded to a height threshold of* $T > 3.13$.

Chapter 9

Perception of Carrier Salience

9.1 Introduction

Subjective evaluation of sound quality is often conducted using direct scaling methods such as magnitude estimation or category scaling. These methods are attractive at first glance because they are fast and readily yield numerical values. However, the validity of scales thus obtained relies on many implicit assumptions, one of them being the unidimensionality of the underlying sensation continuum. Another assumption is the subjects ability to map their sensation magnitude onto a scale. Paired comparisons, on the other hand, require simple comparative judgments, and thereby eliminate response biases due to scale usage. In addition, from paired comparison judgments, the unidimensionality of the underlying sensation (e. g. the overall preference) can be verified. Finally, a ratio scale can be derived by means of probabilistic choice models such as the Bradley-Terry-Luce (BTL) model. To investigate the carrier salience extraction from temporally asymmetric sounds we designed a simple two forced choice task where the subjects judged which of two presented tones is more tonal.

9.2 Materials and Methods

9.2.1 Subjects

Twelve right-handed subjects (7 men, 5 women aged 21–64 years) with normal audiometric thresholds and without any history of audiological or neurological deficits participated in the study. Eight of the subjects participated in both experiments. They all had self-reported normal hearing.

9.2.2 Procedure

Psychoacoustic measurements of the paired comparison task were carried out using the same temporally asymmetric sounds as in the MEG experiment (see Subsection 5.2.2 for the description)with an inter-stimulus interval of 500 ms. In the first experiment single cycle stimuli, and in the second experiment steady state stimuli were compared. Sounds were delivered through the Sennheiser HD 520 II headphones at 65 dB. Listeners were instructed to decide which of the presented signals had more character of a musical tone. The stimuli were presented in one block for each experiment. In each block, all 45 pairs of nonidentical stimuli were presented in both orders in a randomized fashion, giving a total of 90 trials per block. For each trial, listeners were asked to decide which sound was more tonal (two alternate forced choice task). No feedback was given. After a few practice trials (5 to 10), the blocks were run just once. A scale for the relative 'carrier salience' was derived from the results of the paired comparison experiment, using the Bradley-Terry-Luce (BTL) method (David, 1988). According to the description of the algorithm (see chapter 4), the carrier salience of the temporally asymmetric stimuli can be ordered according to a linear scale. Additionally the asymmetry index is defined as $AI = \frac{CS_{ramped} - CS_{damped}}{CS_{ramped} + CS_{damped}}$ where CS denotes carrier salience.

9.3 Results

The results for the two series of stimuli are presented in Fig. 11.4 in which the carrier salience values are plotted as a function of stimulus HLT. Obviously, greater HLT produces higher carrier salience. We observe this consistence for both ramped and damped sounds and for both types of stimuli applied, transients and steady states. Ramped tones are judged as more salient than their damped counterparts. However, approaching the extreme HLT values (0.5 and 32 ms HLT) we observe the overlap of the curves which suggest the lower discrimination ability for these conditions. Asymmetry appears to be preserved independent of stimulus modality (transient vs. steady state). The clear peak at 4 ms HLT condition suggests the highest perceptual difference for this stimuli pair, which confirmed previous results (Patterson, 1994a, 1994b). Another feature of the data to note is the relatively small size of the standard errors, indicating a strong agreement across subjects.

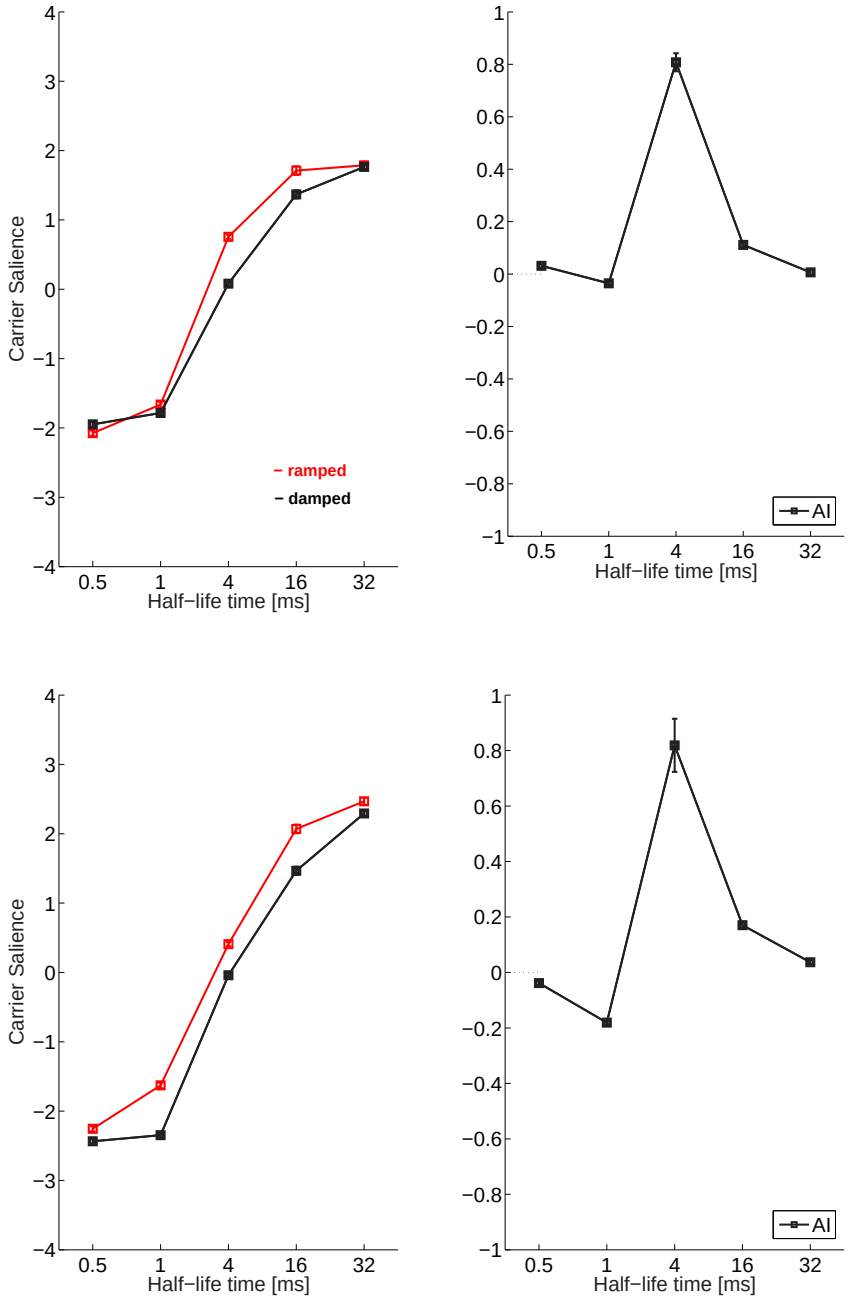

Figure 9.1: *The carrier salience values estimated by the BTL method and the standard errors estimated by bootstrap are plotted as a function of the stimulus HLT for transient (above) and steady state sounds. To quantify perceptional asymmetry we defined an asymmetry index as shown on the right side of the figure. Note the peak at 4 ms HLT condition for both sound modalities, as well as the similarity of the curves.*

Chapter 10

Modeling Temporal Asymmetry

10.1 AIM Simulations

Since the spectral models of auditory processing can not explain the perceptual asymmetry of temporally asymmetric sounds we simulate them with the Auditory Image Model (AIM) (Patterson et al., 1995). We show that AIM is able to predict the perceived carrier salience. The exact explanation of how the model extracts the carrier salience can be found in Chapter 6. AIM is based on the simulation of the spectral analysis performed along the basilar membrane using a bank of auditory filters: the neural transduction process of the inner hair cells and the primary auditory fibers is simulated in each of the frequency channels defined by the filters. In the next stage of processing, the time-interval information of each frequency channel is evaluated by strobed temporal integration. The array of these histograms, the 'auditory image' is shown for damped and ramped sounds of 4 ms half-life time (see Figures 10.2 and 10.1). Respectively, the abscissa represents the time interval of the image buffer. It decays over time and is limited to 32 ms. The activity of each frequency channel is summed up and results in the summed stabilized auditory image (SAI). We show that the first peak in the summed SAI highly correlates with the perceived carrier salience (tonality).
The simulations with AIM were carried out using the following parameters:

- transmission line filterbank,
- Meddis Haircells,
- medium firing rate,
- 100 channels

- internal sampling rate: 48000 Hz,
- for all SAIs the frame-to-frame analysis window is set to an interval of 1 ms.

10.2 Results

SAI's were computed on a 1 ms basis and concatenated the summery SAIs to each other, resulting in a 90 ms SAI-contour for all stimuli. On the figures 10.2 and 10.1 we show the buffer development for the damped and ramped sounds of 4 ms HLT taken at the point of the same amplitude hight, e. g., at 10 ms of the damped sound and 40 ms of the ramped sound. If we observe the temporal development of the buffer we see that the ramped sound excites much more channels and develops longer than its damped counterpart. Figures 10.2 and 10.3 show the development of the first peak of the summed SAI. The behavior of this peak can be summerized as follows: in the conditions with very short HLTs (0.5 and 1 ms) there is an almost flat contour for all stimuli with slight differences at the onset of the response. In the 4 ms HLT condition the ramped sounds shows a steep increase of the hight of the peak that indicates a specific carrier salience extraction. Longer HLTs (16 and 32 ms)show that damped sounds also elicit an increase of the its height. However, this peak is always lower compared to the ramped sounds of equal HLT. If we compare the simulations obtained using transient and steady-state sounds we notice very similar development of the buffer. It appears that the first cycle of the sound brings enough information for carrier salience extraction.

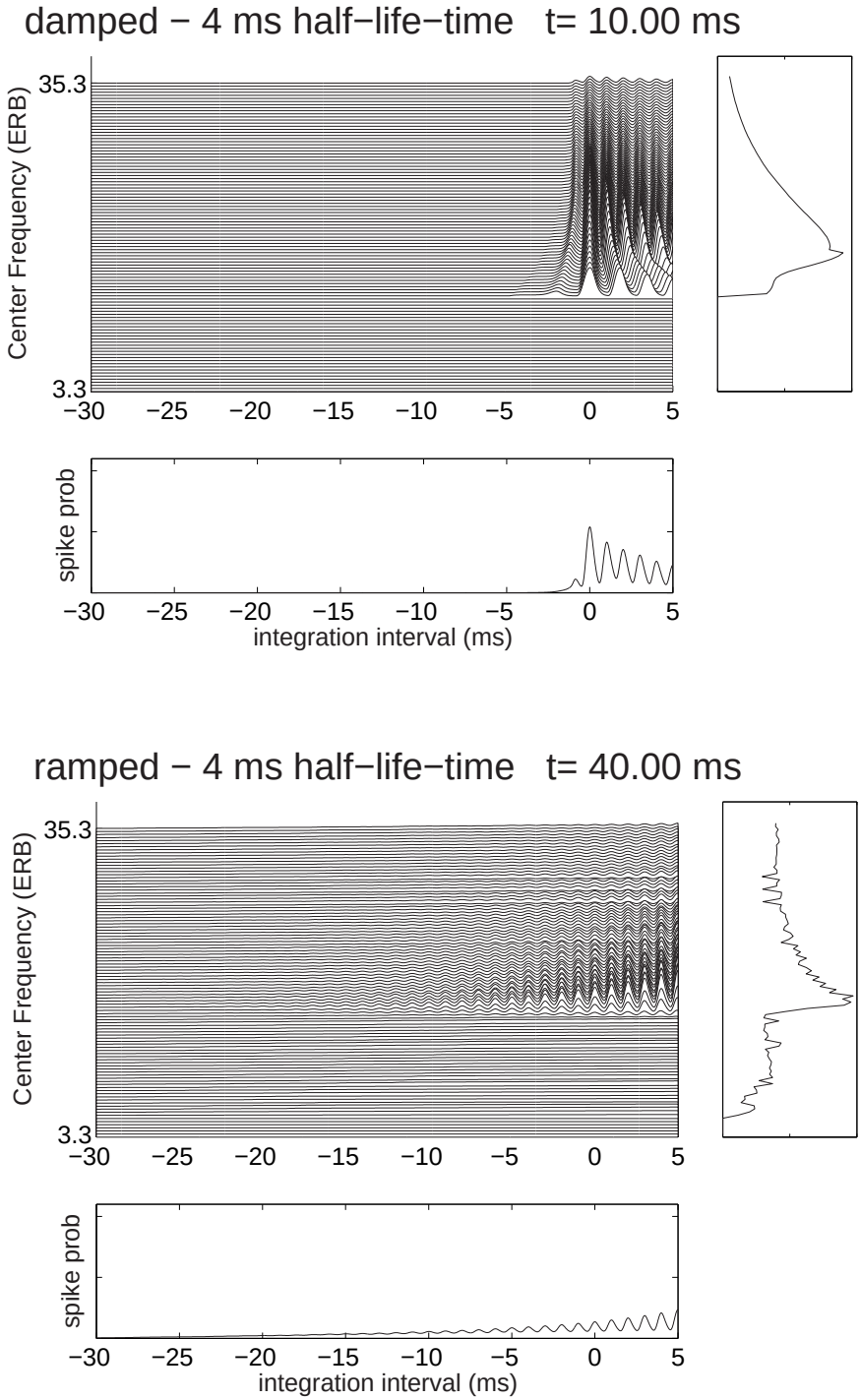

Figure 10.1: *Stabilized Auditory Images for steady state damped (above) and ramped sounds of 4 ms HLT taken at the time point of the same envelope hight. Note the difference in the buffer development. The height of the first peak at time predicts the perceived carrier salience*

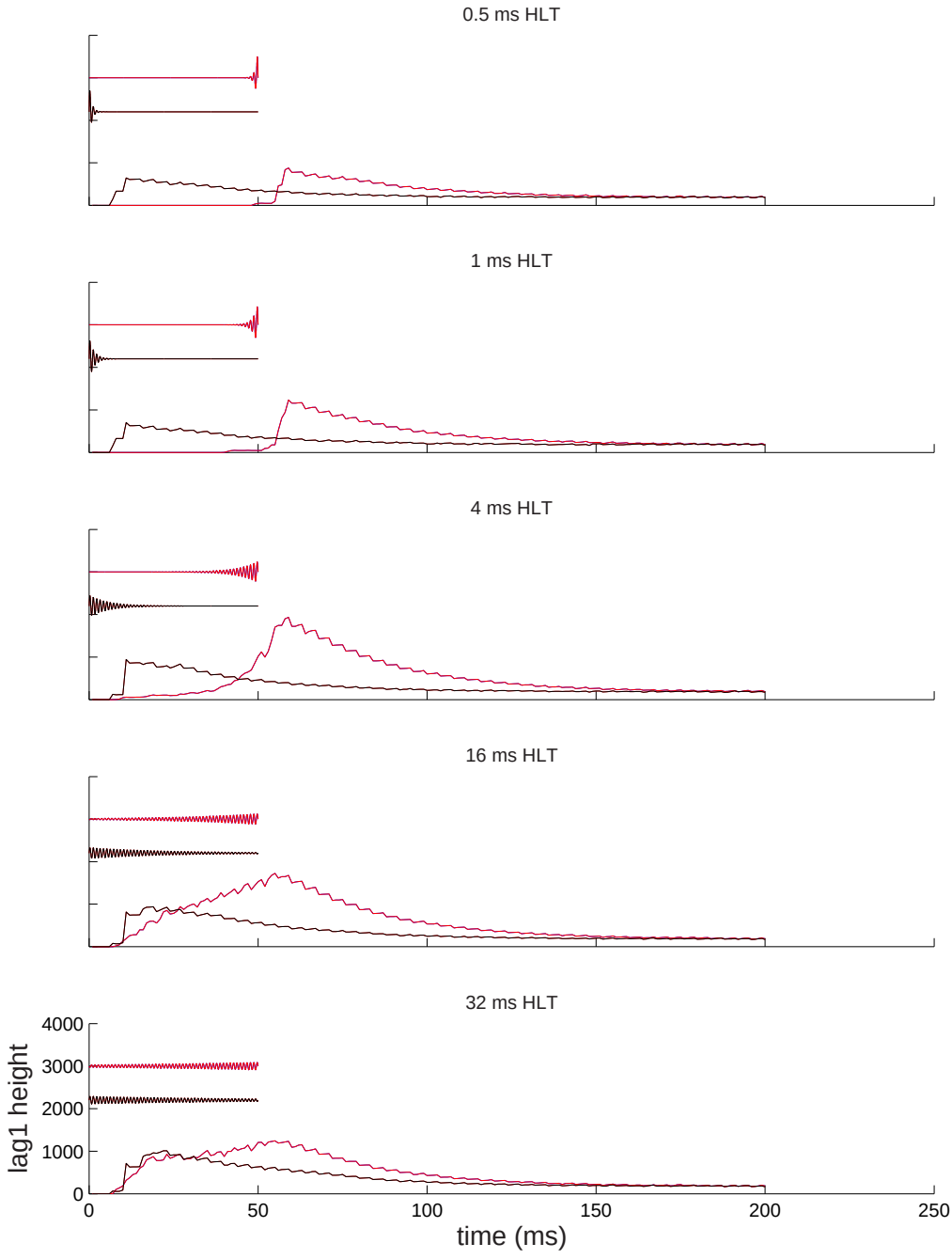

Figure 10.2: *Development of the first peak for the transient sounds. The ramped sounds evoke always higher first peak compared with its damped pair.*

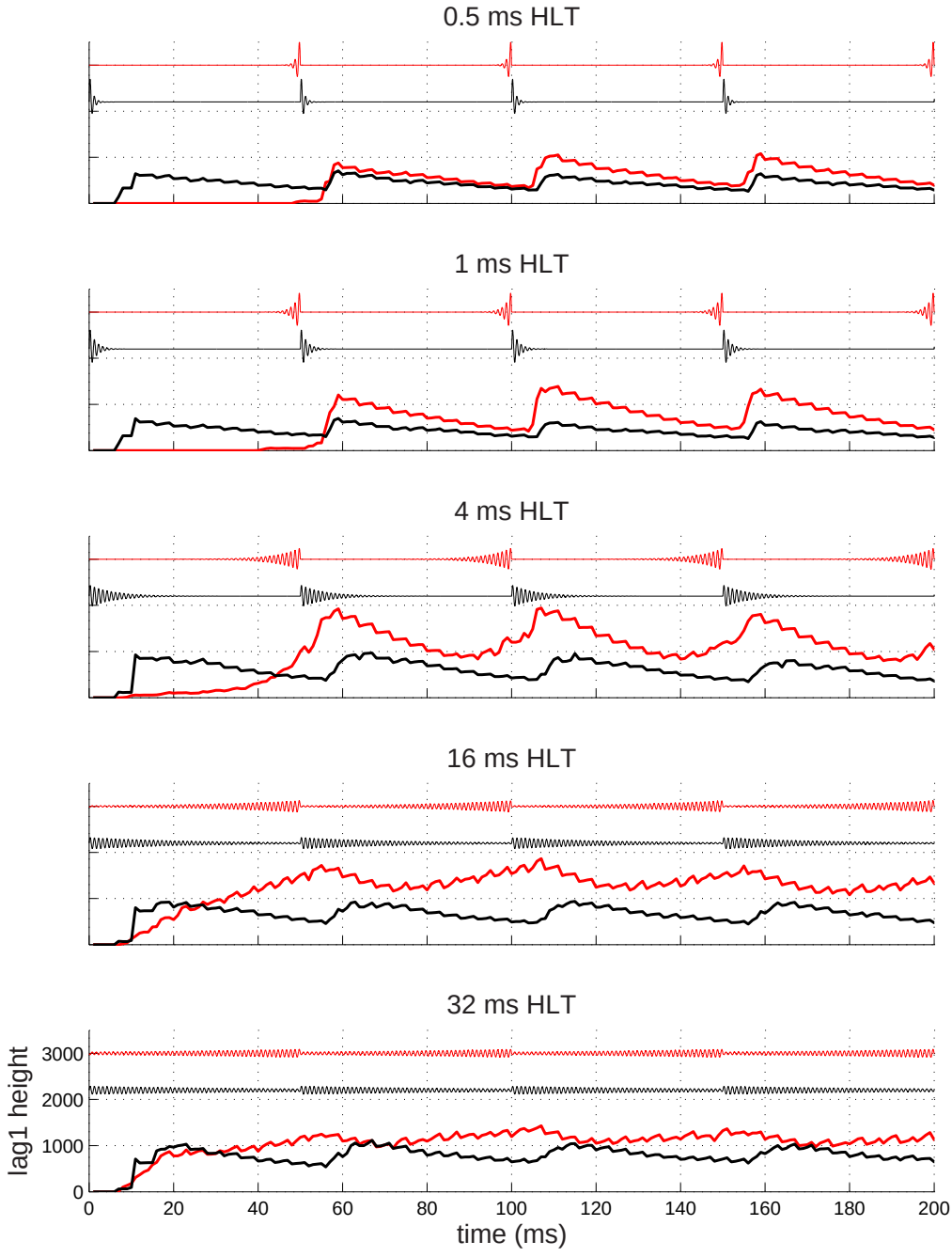

Figure 10.3: *Development of the first peak for the steady state sounds. The ramped sounds evoke always higher first peak compared with its damped pair.*

Chapter 11

Relation Between Perception, Neuromagnetic Responses, Hemodynamics and Modeling

Investigation of sound processing in the human auditory system combines the knowledge collected from findings of anatomy and physiology of the central auditory system, perception of the sound revealed by psychoacoustic measurements and predictions obtained using auditory models. Schematic we can represent the pattern by a triangle as shown in the Figure 11.1.
Here we demonstrate the relation between the perception of the temporally asymmetric sounds, its neurophysiological representation and the model predictions. We do not separate ramped from damped sounds to increase the data sample.

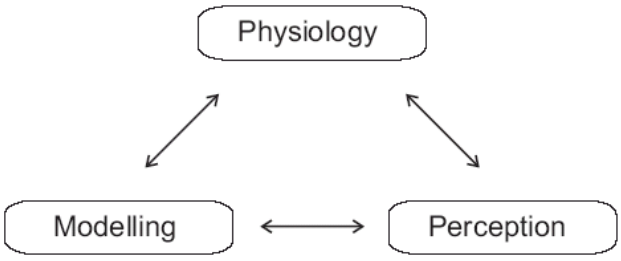

Figure 11.1: *The triangle in the research of human perception. Experimental evidence from anatomy and physiology combined with psychoacoustic results (perception) must correspond with models and theory (adapted from Ritter, 2004).*

Neuromagnetic Responses vs. Psychophysics

On the Figure 11.2 the correlations between the relative perceived carrier salience of the transient and steady state sounds and the corresponding evoked N100m magnitudes for both hemispheres are shown. The psychoacoustic results agree very well with the findings of the N100m response magnitudes in the MEG-experiment. The magnitude changes of N100m highly correlated with the tonality obtained using paired comparison analysis for both transients (r=0.96, p=0.001 for the left hemisphere and $r = 0.94$, $p = 0.001$ for the right hemisphere magnitudes) and steady states ($r = 0.98$, $p = 0.001$ for the left hemisphere and $r = 0.95$, $p = 0.001$ for the right hemisphere magnitudes).

Neuromagnetic Responses vs. Modeling

Here we investigated the correlation between the mean ridge hight for the first 90 ms of SAI summary image which represents the predicted measure of the perceived carrier salience, and neuromagnetic N100m source magnitudes again for transient and steady state sounds (see Figure 11.3. Simulated data on the temporal asymmetry processing on the level of cochlear nucleus revealed a high correlation with the N100m magnitudes for both transients($r = 0.86$, $p = 0.001$ for the left hemisphere and $r = 0.88$, $p = 0.001$ in the right hemisphere), and steady states ($r = 0.89$, $p = 0.001$ in the left hemisphere, and $r = 0.86$, $p = 0.001$ in the right hemisphere).

Psychophysics vs. Modeling

Finally we wanted to find out if our model predicts the perceived carrier salience to close the triangle. Simulated data correlated highly with the perceived carrier salience of the transients($r = 0.83$, $p = 0.001$ and steady states $r = 0.89$, $p = 0.001$) as shown on the Figure 11.4. That means that auditory image model is able to predict the perceived carrier salience changes across the HLTs and conditions, that is the perceptional asymmetry.

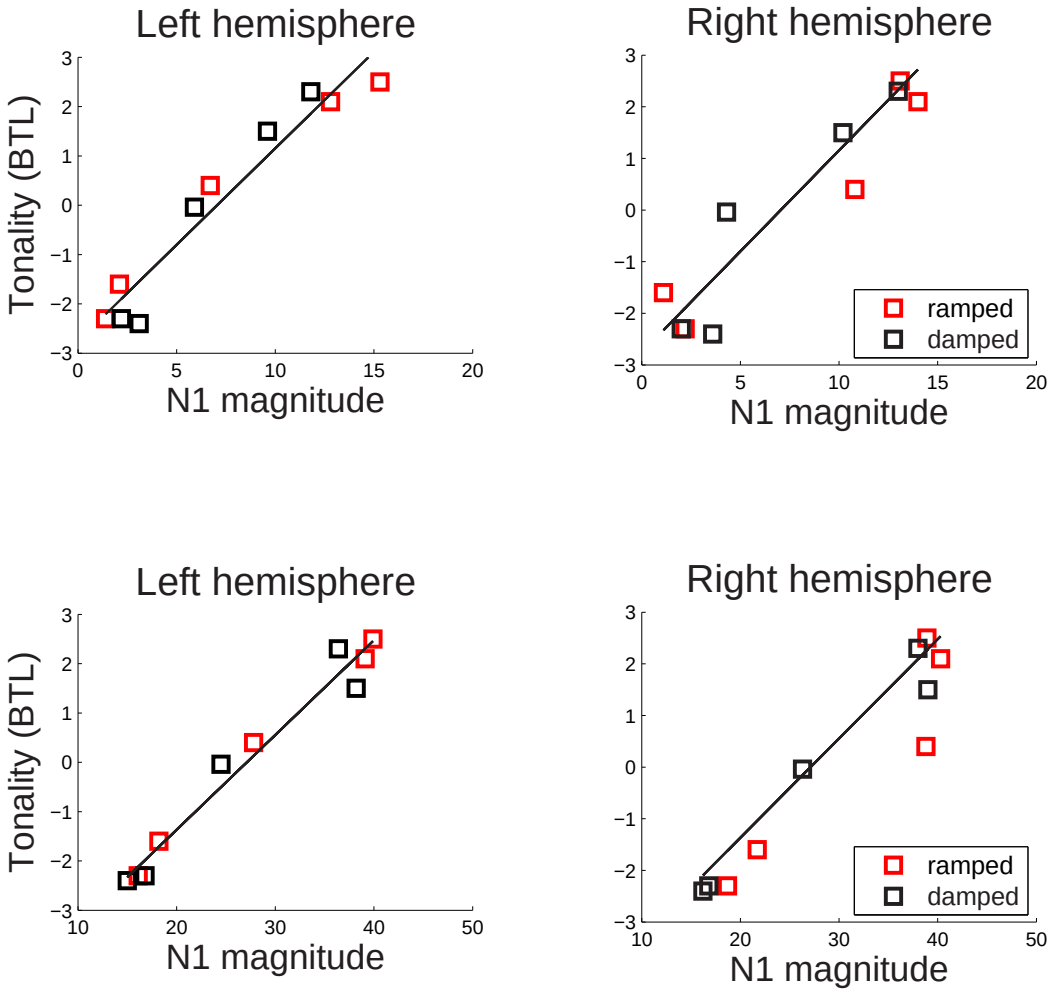

Figure 11.2: *Correlation between the magnitudes of N100m sources evoked by transient (above)(r=0.96, p=0.001 for the left hemisphere and* $r = 0.94$, $p = 0.001$ *for the right hemisphere magnitudes) and steady state (*$r = 0.98$, $p = 0.001$ *for the left hemisphere and* $r = 0.95$, $p = 0.001$ *for the right hemisphere magnitudes) temporally asymmetric stimuli and the relative perceived carrier salience of the corresponding sounds.*

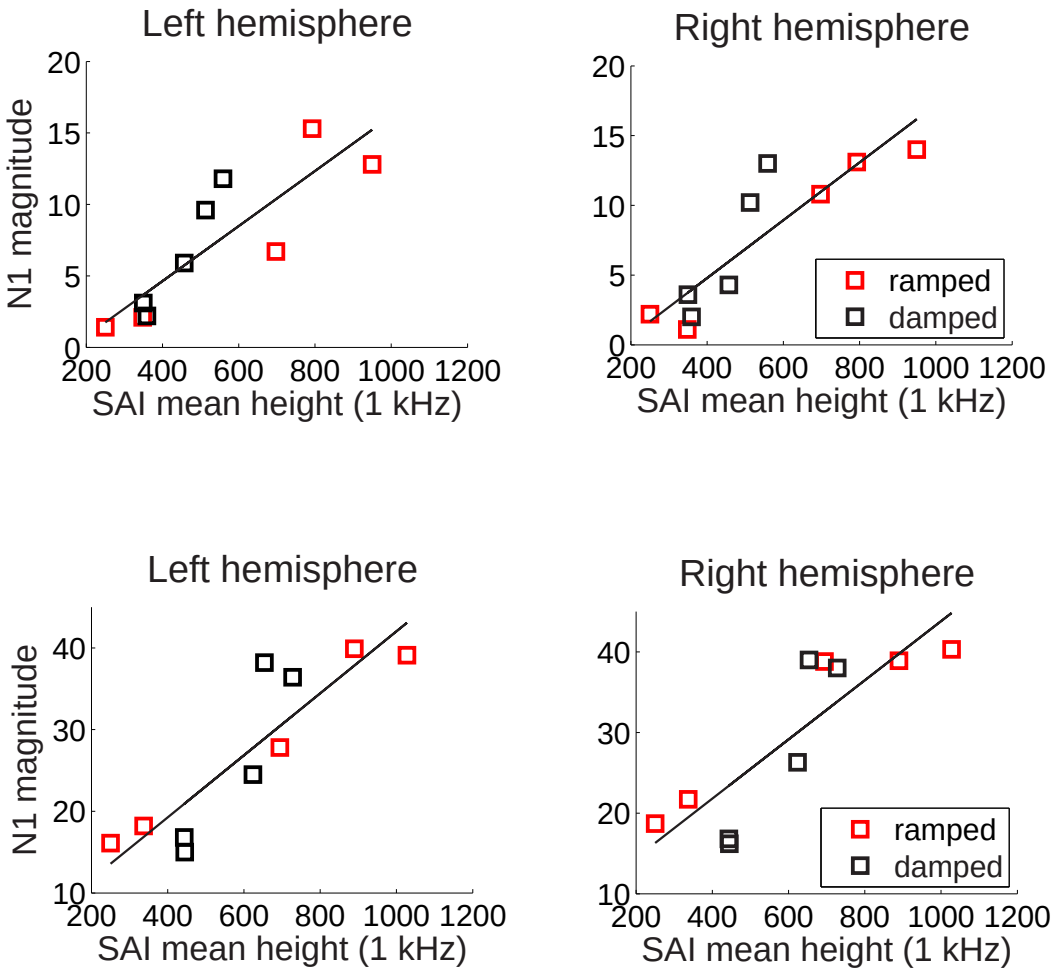

Figure 11.3: *Correlation between the mean ridge hight for the first 90 ms of SAI summary image which represents the predicted measure of the perceived carrier salience, and neuromagnetic N100m source magnitudes for transient (above)*($r = 0.86$, $p = 0.001$ *for the left hemisphere and* $r = 0.88$, $p = 0.001$ *in the right hemisphere) and steady state sounds* ($r = 0.89$, $p = 0.001$ *in the left hemisphere, and* $r = 0.86$, $p = 0.001$ *in the right hemisphere)*

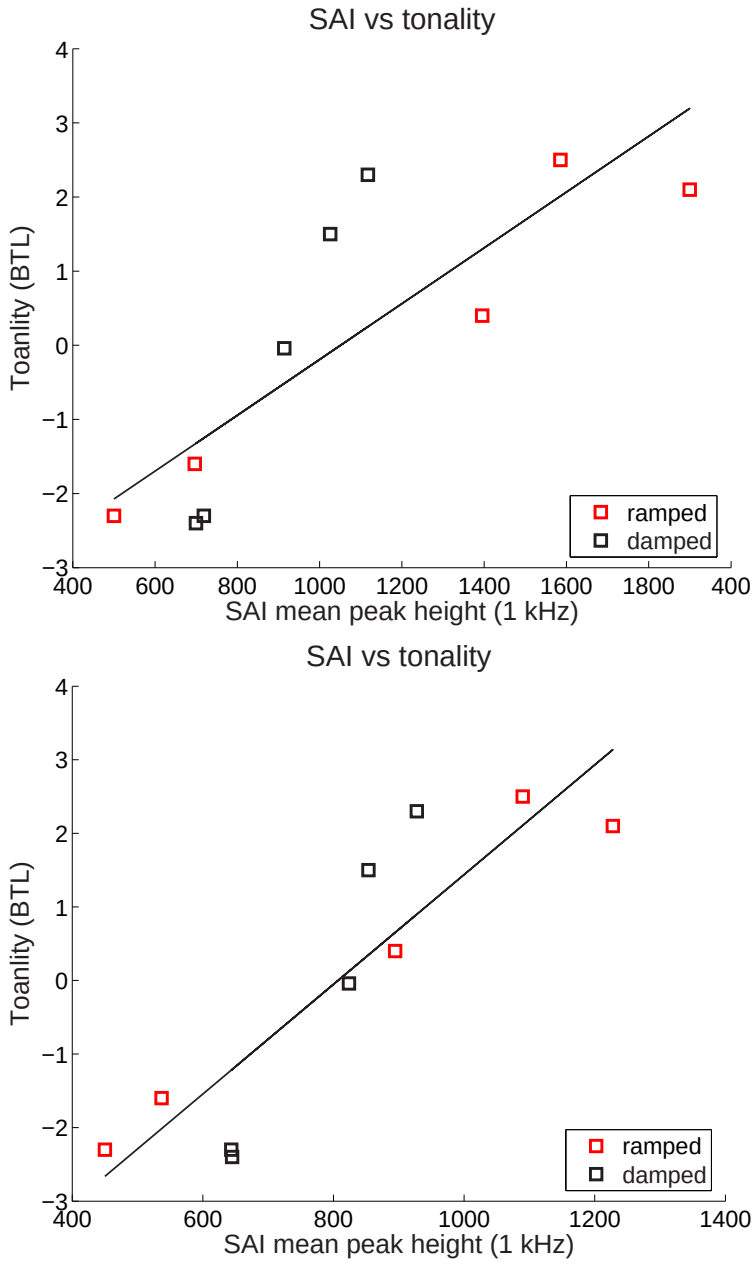

Figure 11.4: *Simulated data correlated highly with the perceived carrier salience of the transients*($r = 0.83$, $p = 0.001$) *and steady states* ($r = 0.89$, $p = 0.001$)

Hemodynamics vs. Neurophysiology

With regard to temporal resolution, the MEG has a big advantage. Ranging in the order of 1 ms its sensitivity is about three orders of magnitude higher than that of the fMRI. This is a consequence not of any technical limit but a principal one resulting from the time course of the hemodynamic response function. However, the main advantage of fMRI over MEG, is that fMRI is not plagued by an inverse problem. fMRI data are model-independent so that, for instance, any significant BOLD response found in a group analysis actually depicts enhanced activation, whereas the mean of different individual MEG models does not necessarily do so. However, we must not forget that fMRI analysis crucially rely on the warping of brains to enable the comparison of functional with structural data. On the Figure 11.5 the Talairach map is given with the projected neuromagnetic sources and main fMRI clusters for all 'sound–silence' conditions. The fMRI clusters partially overlap the MEG sources. However, medial fMRI clusters are much more medial compared with their MEG counterparts (e.g. P30m sources). The contrast 'ramped–damped' of 4 ms HLT revealed right hemisphere activity lateralization which might correspond to the increased asymmetry index values of N100m source amplitudes for 4 ms HLT condition.

Conclusion

We found a strong evidence for the existence of carrier salience extraction mechanism on the level of N100m component. N100m magnitude increase with the increasing stimulus HLT for both transient and steady state stimuli. Parallel we get changes in perceived carrier salience dependent on stimulus HLT. Carrier salience values increase with the increasing stimulus HLT. N100m magnitude differences between ramped and damped sounds are the highest for the 4 ms HLT condition. Perceptual asymmetry which is the most prominent for the 4 ms HLT condition. Our simulation study supports these findings. The mean ridge of the stabilized auditory image correlates highly with both, N100m magnitude and perceived carrier salience. That means that auditory image model is able to predict the perceptual asymmetry and corresponding neurophysiological asymmetry. Additional support for our hypothesis about the existence of special carrier salience extraction mechanism comes from our fMRI experiment. We localized region posterior to Heschl's gyrus as a result of difference contrast 'ramped–damped' of 4 ms HLT. This is the area where we localized N100m sources using MEG, where the highest magnitude differences for 4 ms HLT condition occur. This is additionally an approval of very good spatial resolution of magnetoencephalography.

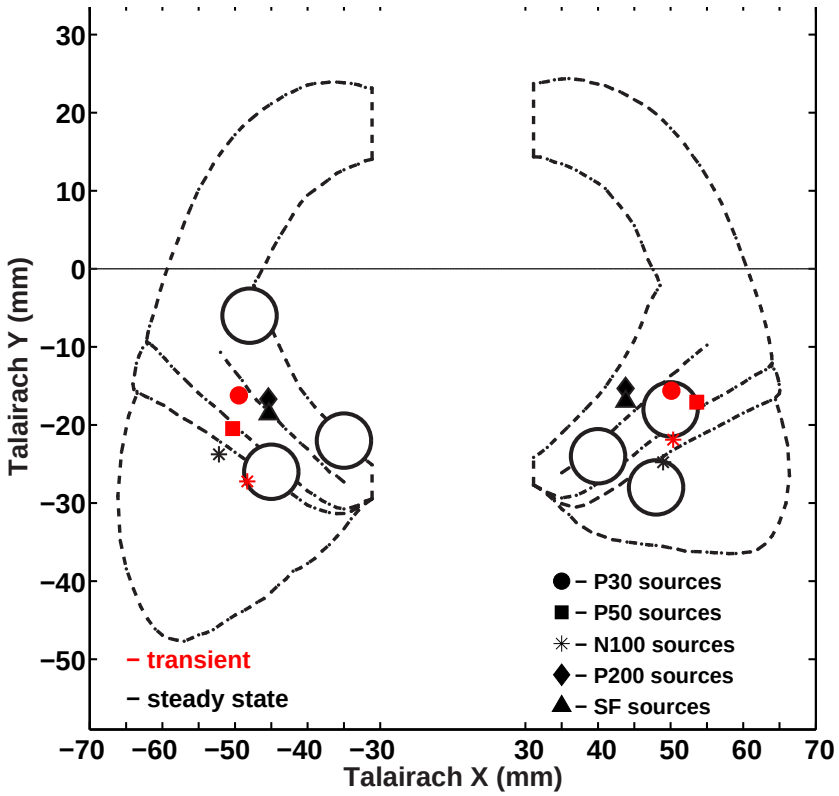

Figure 11.5: *Talairach coordinates of the neuromagnetic sources localized using MEG and fMRI obtained main clusters for all 'sound–silence' conditions.*

Chapter 12

Disscussion and Conclusions

This thesis attempted to determine special aspects of the mechanisms on temporal asymmetry processing in human hearing. Particularly, the relation between behavioral measurements and corresponding auditory evoked neuromagnetic responses and hemodynamic changes were examined. Magnetoencephalography (MEG) was used in conjunction with functional magnetic resonance imaging (fMRI) to localize and characterize cortical sources evoked by temporally asymmetric sounds. Additionally, the relations between neurophysiological and psychometric data were explored. Neuromagnetic responses and psychoacoustic results were compared with simulations of the temporal asymmetry processing based on the auditory image model provided by Patterson et al. (1995).

Neural responses to ramped and damped sinusoids have been studied in some subcortical and cortical structures (Fay et al. 1996; Pressnitzer et al. 2000; Neuert et al. 2001, Lu et al; 2001). Here, we review and compare the degree of asymmetry at successive stages. The representation of temporal asymmetry has been observed in the auditory nerve (AN) in goldfish (Fay et al. 1996). Fay showed that both ramped and damped sinusoids resulted in strong responses and that firing patterns to the ramped sinusoid were temporally different from those to the damped sinusoids. Wang and Sachs (1993) investigated the processing of the single-formant sounds, similar to the damped sinusoids used in the present study, and showed that they produced strong discharges in all types of AN fibers in cats. Pressnitzer et al. (2000) reported that onset and chopper neurons showed more pronounced response asymmetry to ramped and damped sinusoids than did the primary-like neurons, and response asymmetry appeared to be greater in some ventral cochlear nucleus (VCN) neurons than in AN. The observations on the level of inferior coliculus (IC) also showed various degrees

of response asymmetry (Neuert et al. 2001). In AN, VCN, and IC, the total spike count elicited by ramped sinusoids were reported to be generally greater than those by damped sinusoids (Fay et al. 1996; Pressnitzer et al. 2000; Neuert et al. 2001). Cortical neurons responded nearly exclusively to either ramped or damped sinusoids (Lu et al. 2001).

In the current study spatio-temporal analysis of auditory evoked transient and steady state neuromagnetic fields revealed the effects of temporal asymmetry at the level of primary and secondary auditory cortex. The resulting morphology of the P30m and P50m source waveforms depended strongly on the temporal features of the envelope which was reflected through the latency shift of the responses evoked by the ramped stimuli. This is in line with a study from Biermann and Heil (2000) that the peak latencies of the P50m and N100m components decrease with the rise time which in our case represents half-life time of the stimuli. The longer the stimulus HLT the shorter the latencies. However, P30m and P50m amplitudes remained unchanged across the half-life times, conditions and hemispheres. Therefore they represent the processing stage where the envelope coding takes place. The next processing stage that we analyzed was N100m component. When analyzing the N100m component we have to keep in mind other experimental parameters beside the dynamic properties of the stimuli, e.g., rise time and level, that influence the morphology of the source waveforms. It has been shown that N100 peak amplitudes increase with the increasing stimulus duration (Kodera et al. 1979) but its latency remains the same (Jutsiniemi et al. 1989). Furthermore, the N100m increase with the increasing inter-stimulus interval (Polich et al 1988) and stimulus level (Biermann and Heil, 2000). To account on that, we kept the level approximately constant for all conditions normalizing the amplitude by a factor proportional to the square root of the stimulus half life time, respectively (Patterson, 1994a). Source analysis of the N100m component for both stimulus modalities, e.g., transient and steady state stimuli, revealed the increase of the peak amplitude with the stimulus HLT for all conditions and both hemispheres. Ramped stimuli evoked always slightly higher amplitudes than damped, but the increase was significant only for the the 4 ms HLT condition. The amplitudes were also slightly higher in the right hemisphere compared with the left one, but the increase was significant only for the 4 ms HLT ramped condition. Furthermore, the N100m amplitudes evoked by steady state sounds were twice as large as the amplitudes evoked by transient stimuli. Knowing the influence of interstimulus interval and stimulus duration on the morphology of the source waveforms this finding does not surprise.

In his behavioral study Patterson (1995) showed that the temporally asymmetric stimuli become more tonal with the increasing HLT. We hypothesized that the change of N100m magnitude might be related with the change in the perception of stimulus tonality, e.g., carrier salience. Therefore we designed a simple two alternative forced choice task where the subjects judged which of the two presented stimuli is more tonal. Using the Bradley-Terry-Luce method we derived the relative carrier salience scale for all stimuli. Indeed, the stimulus tonality was judged as more prominent for the envelopes exhibiting a higher HLT. Relating the salience values for each stimulus with the corresponding N100m magnitudes for transients and steady states in both hemispheres (see Figure 11.4) we got a confirmation of our hypothesis. N100m magnitudes were highly correlated for both, transients (r=0.96, p=0.001 for the left hemisphere and $r = 0.94$, $p = 0.001$ for the right hemisphere amplitudes) and steady states ($r = 0.98$, $p = 0.001$ for the left hemisphere and $r = 0.95$, $p = 0.001$ for the right hemisphere amplitudes) with the carrier salience values.

Time-domain computational models of auditory processing which begin with some form of auditory filterbank reflect inherent temporal asymmetry. Patterson's auditory image model (AIM) of peripheral auditory processing is constructed in three stages: an auditory filterbank that produces a representation of basilar membrane motion (BMM), a multi-channel, neural transduction mechanism that produces a representation of the neural activity pattern (NAP) in the auditory nerve, and a bank of strobed temporal integration units that produce the models representation of the auditory image that we hear in response to the sound. We simulated temporally asymmetric sounds with the auditory image model to test if the sensation can be predicted with the model. We chose the mean of the ridge for the first 90 ms of SAI summary image as the relevant factor to extract carrier salience. In the 4 ms HLT condition the ramped sounds show a steep increase of the hight of the first peak in the summed SAI that indicates a specific carrier salience extraction. Longer HLTs (16 and 32 ms) show that damped sounds also elicit an increase of the first peak height. If we compare the simulations obtained using transient and steady-state sounds we notice a very similar development of the buffer which proves that the first cycle of the temporally asymmetric sounds carries enough perceptual information concerning carrier salience extraction. Simulated data on the temporal asymmetry processing on the level of cochlear nucleus in terms of mean ridge hight for the first 90 ms of SAI summary image revealed a high correlation with the N100m amplitudes for both transients ($r = 0.86$, $p = 0.001$ for the left hemisphere and $r = 0.88$, $p = 0.001$ in the right hemisphere), and steady states ($r = 0.89$, $p = 0.001$ in the

left hemisphere, and $r = 0.86$, $p = 0.001$ in the right hemisphere) and with the perceived carrier salience ($r = 0.83$, $p = 0.001$ for the transients and $r = 0.89$, $p = 0.001$, for the steady states).

To quantify the perceptual asymmetry as well as N100m magnitude asymmetry we defined an asymmetry index which was originally proposed by Patterson (1995). In Patterson's two-alternative, forced-choice experiment, listeners were presented a ramped sinusoid in one interval and a damped sinusoid having the same or greater half-life in the other interval. In one version of the experimental run, listeners were asked to choose the interval containing the sound with the louder tonal component. In another version with the same stimuli, they were asked to choose the interval with the louder drumming component. Between trials, the half-life of the damped sinusoid was varied to determine the matching point, that is, the half-life of the damped sinusoid required to equate the probability of choosing either the ramped or damped stimulus as the one with the louder tonal or drumming component. His asymmetry factor was defined as the average distance in logarithmic units between the matching half-life and the equal half-life. Having different behavioral task we defined the asymmetry index as $AI = \frac{CS_{ramped} - CS_{damped}}{CS_{ramped} + CS_{damped}}$, where CS represents relative carrier salience. For N100m magnitude asymmetry we use the same calculus with N100m magnitude instead carrier salience as a variable. We show that perceptual as well as neurophysiological asymmetry occurs for the 4 ms HLT condition for both transients and steady states. Peak magnitude asymmetry for the spectrally matched pairs for the N100m amplitudes has its maximum for the 4 ms HLT condition, and is more pronounced in the right hemisphere for both types of the stimuli (transients and steady states). Furthermore, the N100m magnitude asymmetry is slightly higher for transient stimuli than for the steady states. Amplitude asymmetry has further been observed bilaterally on the level of P200m stage evoked by steady state stimulation but because of high intersubject variability in the morphology of the source waveforms this component wasn't further analyzed. The P30m, P50m components and sustained field revealed no asymmetry concerning their amplitudes, e.g., areas in the case of SF (see subsection 5.1.4).

Functional magnetic resonance imaging revealed strong bilateral activation in both auditory cortices for all 'sound–silence' conditions. We localized six clusters which roughly correspond to the localized sources using MEG. However, our main finding using fMRI, concerns the asymmetry in the activation pattern that is observed only for the difference contrast "ramped–damped" for the stimuli pair of 4 ms HLT in the region posterior to Heschl's gyrus, e.g., between Heschl's Gyrus

and Planum temporale. This is the region where we localized the N100m sources using MEG, which is sensitive to carrier salience changes. This confirms, a very high spatial resolution of MEG. Additionally, we observe right activity lateralization which is in line with our findings on N100m responses. This might be due to the greater involvement of right auditory cortical areas in spectral processing as shown by Zatorre and collaborators (2000). Zatorre found a right hemisphere specialization in a task that required the judgement of the direction of pitch changes.

Our results confirmed previous findings (Pressnitzer et al., 2000; Neuert et al., 2001 and Lu et al., 2001) on temporal asymmetry encoding and paralleled psychoacoustic results (Patterson, 1995). On the Figure 12.1 we show the previous findings concerned temporal asymmetry encoding together with present findings. All data are shown in terms of asymmetry indices. Our results are marked in red. The maximal asymmetry for the 4 ms HLT condition was confirmed by all studies. However, the asymmetry for other HLTs appears to be more variable, especially in our study. We can offer a few possible explanations. First, N100m develops slowly for the low HLTs, e.g., 0.5 and 1 ms conditions, and often does not reach a clear maximum but rather displays one or two shoulders before the global maxima is reached. This makes the peak identification difficult and therefore a larger magnitude asymmetry index results. Second, concerning psychometric curve, in a paired comparison experiment it is difficult to judge the tonal character of the sound, especially for the sounds of short HLTs. The tonal character of such sounds is suppressed due to the drumming component which occurs at the same time. This might explain the different shape of the psychometric curve compared with this obtained by Patterson.

Our study extended previous findings to the transient sounds, and showed that psychophysical extraction of carrier saliency can be explained by the N100m magnitude changes. Simulations showed that it is possible to predict the perception with an auditory image model. fMRI revealed the position of the N100m generator and its lateralization. Thus, we found the strong evidence that the N100m is not just a simple consequence of the onset encoding but the neurophysiological mechanism specialized for the extraction of the carrier salience which might be the basis for the perceptual asymmetry.

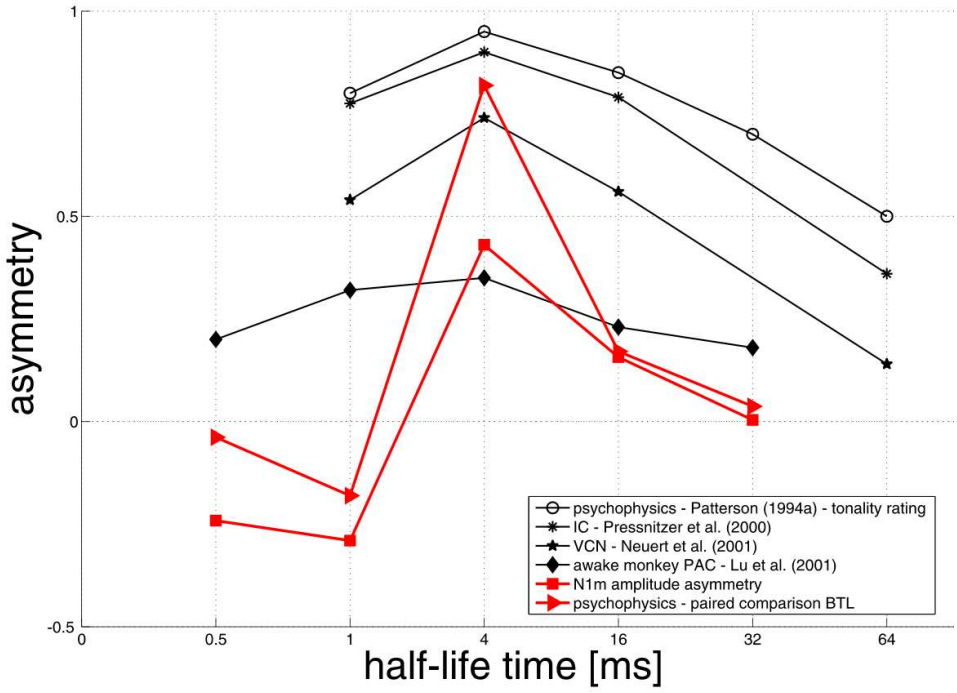

Figure 12.1: *Comparison between asymmetry preference of ventral cochlear nucleus (Pressnitzer et al., 2001), inferior coliculus (Neuert et al., 2001) and cortical neurons (Lu et al., 2001), human psychophysical performance in discriminating the ramped and damped sinusoids (Patterson, 1994a), N100m magnitude asymmetry and psychophysical perceptual asymmetry obtained using BTL method.*

Appendix A

Tables of Amplitudes and Latencies

Condition	P30m peak amplitudes and latencies			
	Left Auditory Cortex		Right Auditory Cortex	
	Amplitude [nAm]	Latency [ms]	Amplitude [nAm]	Latency [ms]
d05	4.2 ± 0.2	35.5 ± 0.6	5.1± 0.4	36.4 ± 0.6
r05	4.8 ± 0.6	82.3 ± 0.7	4.9 ± 0.6	85.0 ± 0.6
d1	5.2 ± 0.6	36.2 ± 0.5	3.9 ± 0.3	36.8 ± 0.6
r1	2.4 ± 0.3	83.8± 0.8	3.9± 0.6	85.6 ± 0.8
d4	4.5 ± 0.2	35.9 ± 0.4	3.7 ± 0.2	35.6 ± 0.6
r4	2.4 ± 0.3	68.0 ± 0.6	2.6 ± 0.2	67.5 ± 1.6
d16	5.2 ± 0.3	35.7 ± 0.4	5.0 ± 0.2	35.2 ± 0.3
r16	4.7 ± 0.6	38.6 ± 0.9	2.9 ± 0.2	38.1 ± 1.1
d32	4.1 ± 0.2	36.8 ± 1.2	4.4 ± 0.2	35.2 ± 1.0
r32	3.7 ± 0.8	36.6 ± 0.2	3.3 ± 0.4	35.6 ± 0.8

Table A.1: *P30m peak amplitudes and latencies with the standard error depicted.*

Condition	P50m peak amplitudes and latencies			
	Left Auditory Cortex		Right Auditory Cortex	
	Amplitude [nAm]	Latency [ms]	Amplitude [nAm]	Latency [ms]
d05	7.6 ± 1.0	66.8 ± 1.9	6.3± 0.6	61.6 ± 0.6
r05	9.6 ± 0.7	122.9 ± 0.9	8.3 ± 0.8	112.8 ± 0.6
d1	8.9 ± 0.8	69.7 ± 0.8	6.7 ± 0.6	62.2 ± 1.0
r1	8.9 ± 0.9	117.7± 1.9	7.8± 0.7	108.7 ± 1.4
d4	8.5 ± 0.6	67.1 ± 1.4	7.4 ± 0.4	61.9 ± 0.8
r4	9.8 ± 0.7	84.5 ± 1.3	6.8 ± 0.5	89.6 ± 2.3
d16	8.0 ± 0.9	59.0 ± 1.1	8.1 ± 0.5	60.2 ± 0.5
r16	5.6 ± 0.6	59.8 ± 1.7	7.9 ± 0.5	61.8 ± 0.7
d32	9.9 ± 0.7	62.5 ± 0.5	8.0 ± 0.5	59.4 ± 0.3
r32	8.0 ± 0.6	61.9 ± 0.6	8.0 ± 0.5	58.6 ± 0.3

Table A.2: *P50m peak amplitudes and latencies with the standard error depicted.*

Condition	N100m peak amplitudes and latencies			
	Left Auditory Cortex		Right Auditory Cortex	
	Amplitude [nAm]	Latency [ms]	Amplitude [nAm]	Latency [ms]
d05	-3.1 ± 0.3	137.1 ± 2.4	-3.6± 0.5	131.9 ± 2.9
r05	-1.4 ± 0.6	178.7 ± 4.6	-2.2 ± 0.6	176.8 ± 5.0
d1	-2.2 ± 0.2	133.3 ± 4.0	-2.0 ± 0.4	122.1 ± 5.1
r1	-2.1 ± 0.5	183.0± 1.5	-1.1± 0.4	177.1 ± 2.7
d4	-5.9 ± 0.7	119.7 ± 3.0	-4.3 ± 0.4	115.6 ± 4.1
r4	-6.7 ± 0.7	143.0 ± 1.2	-10.8 ± 1.2	136.6 ± 1.1
d16	-9.6 ± 0.7	97.0 ± 1.8	-10.2 ± 0.6	95.8 ± 2.2
r16	-12.8 ± 1.1	129.1 ± 4.0	-14.7 ± 1.2	120.0 ± 2.3
d32	-11.8 ± 0.7	98.4 ± 2.1	-13.0 ± 0.8	94.3 ± 1.2
r32	-15.3 ± 1.2	120.3 ± 1.8	-13.1 ± 1.0	118.5 ±1.4

Table A.3: *N100m peak amplitudes and latencies for the transient stimuli with the standard error depicted.*

Condition	N100m peak amplitudes and latencies			
	Left Auditory Cortex		Right Auditory Cortex	
	Amplitude [nAm]	Latency [ms]	Amplitude [nAm]	Latency [ms]
d05	-15.0 ± 0.9	118.8 ± 3.8	-16.2± 1.4	119.8 ± 2.3
r05	-16.1 ± 0.8	164.3 ± 3.8	-18.7 ± 1.6	153.4 ± 3.4
d1	-16.8 ± 1.1	117.2 ± 3.7	-16.8 ± 1.3	105.6 ± 3.5
r1	-18.2 ± 1.2	163.4± 2.1	-21.7± 1.8	154.8 ± 3.9
d4	-24.5 ± 1.4	107.7 ± 3.7	-26.3 ± 1.9	108.8 ± 3.1
r4	-27.8 ± 2.3	140.1 ± 2.5	-38.8 ± 2.7	138.7 ± 1.4
d16	-38.2 ± 1.9	98.3 ± 1.2	-39.0 ± 1.7	98.3 ± 2.0
r16	-39.1 ± 2.0	98.0 ± 1.3	-40.3 ± 1.3	104.3 ± 2.3
d32	-36.4 ± 2.1	96.6 ± 1.2	-38.0 ± 1.4	104.1 ± 2.8
r32	-39.9 ± 2.2	96.0 ± 0.6	-38.9 ± 1.6	99.9 ±2.4

Table A.4: *N100m peak amplitudes and latencies for the transient stimuli with the standard error depicted.*

Condition	P200m peak amplitudes and latencies			
	Left Auditory Cortex		Right Auditory Cortex	
	Amplitude [nAm]	Latency [ms]	Amplitude [nAm]	Latency [ms]
d05	8.0 ± 0.5	237.4 ± 7.2	12.7± 1.0	234.7 ± 8.8
r05	9.3 ± 0.5	270.4 ± 3.8	14.2 ± 0.9	281.2 ± 5.3
d1	9.3 ± 0.6	245.8 ± 7.5	12.2 ± 0.8	231.6 ± 8.1
r1	10.9 ± 0.6	292.2± 6.9	13.5± 1.2	291.2 ± 9.1
d4	11.5 ± 0.8	220.5 ± 4.7	13.6 ± 0.9	242.8 ± 10.0
r4	17.8 ± 1.3	227.8 ± 4.7	21.5 ± 1.9	230.9 ± 4.7
d16	19.3 ± 1.1	186.7 ± 1.1	20.7 ± 1.4	183.6 ± 0.8
r16	15.9 ± 0.8	217.8 ± 5.6	18.3 ± 1.6	232.7 ± 9.1
d32	16.3 ± 0.9	187.1 ± 4.1	19.4 ± 1.4	185.9 ± 5.9
r32	13.9 ± 0.8	205.0 ± 7.8	19.3 ± 1.1	213.5 ±10.6

Table A.5: *P200m peak amplitudes and latencies for the transient stimuli with the standard error depicted.*

Appendix B

MR Signal Localization and Imaging

We recall that the magnitude of an MR-signal depends on the number of resonating nuclei. Since the objects we deal with are usually not homogeneous, the amplitude of the signal will not directly correspond to the measured nucleus density. Consequently, identification of the signal components contributed by different parts of the object, becomes necessary. Basically two types of localization methods exist: selective excitation or reception and spatial encoding, which distinguishes frequency encoding and phase encoding. Central to both is the use of a gradient field, which by convention points along the z-direction. Modern systems provide three orthogonal gradients the shapes and forms of which can be adjusted independently.

B.1 MR Signal Localization

B.1.1 Selective Excitation

Selective Excitation allows NMR excitation limited to a chosen slice in the sample and is achieved by applying the 90° RF pulse simultaneously with a gradient field perpendicular to the desired slice. We remind that an RF pulse can only be frequency selective, so that spins at different locations will be excited in the same way if they resonate at the same frequency. Through augmentation of the $\mathbf{B_0}$ by the gradient field of strength G_z, the resonance frequency becomes position dependent:

$$\omega z = \gamma(B_0 + zG_z), \tag{B.1}$$

supposing that we wish to image a slice in the xy-plane. The width of the slice Δz is given by

$$\Delta z = \Delta\omega/\gamma G_z, \tag{B.2}$$

where $\Delta\omega$ is the bandwidth of the frequencies contained within the selective pulse. We can see from Equation B.2 that the slice thickness can be varied by either decreasing the spectral bandwidth $\Delta\omega$ or increasing the magnitude of the selection gradient G_z.

As the spread of frequencies close to the fundamental resonance frequency ω_0 of the magnet should be narrow, the RF pulse must be amplitude modulated. For small flip angles the slice profile in the frequency domain is given by the Fourier transform of the RF pulse envelope. In the case of a boxcar shaped slice profile one has to find the pulse function

$$B_1(t) \propto \int_{-\infty}^{\infty} \text{boxcar}(f) e^{-i2\pi ft} \mathrm{d}f. \tag{B.3}$$

Applying the scaling property of the Fourier Transform

$$h(bt) = \longleftrightarrow \frac{1}{|b|}\hat{h}(\frac{f}{b}) \tag{B.4}$$

to the boxcar-function

$$\frac{1}{a}\text{boxcar}(\frac{f}{a}) \longleftrightarrow \text{sinc}(\pi at) = \frac{\sin(\pi at)}{\pi at}, \tag{B.5}$$

results in

$$B_1(t) \propto c\Delta f \text{sinc}(\pi\Delta f t) \exp^{-2\pi i f_c t}, \tag{B.6}$$

where f_c the resonance frequency at the center of the slice. The envelope is given by

$$B_1^{env}(t) = c\Delta f \text{sinc}(\pi\Delta f t) \tag{B.7}$$

and the resonance condition by

$$\omega = 2\pi f_c \omega_0 + \gamma G_z z_0, \tag{B.8}$$

where c is a constant, determining the flip angle. Choosing the start of the pulse to be at $t - t_p/2$ we obtain a pulse symmetric to $t_p/2$:

$$B_1^{env}(t) = c\Delta f \text{sinc}[\pi\Delta f (t - \frac{t_p}{2})]. \tag{B.9}$$

Note, that the sinc pulse of infinite duration is impractical for an MR experiment. Truncation of the pulse in time domain:

$$B_1^{env}(t) = c\,\text{sinc}[\pi\Delta f (t - \frac{t_p}{2})] \cdot \text{boxcar}(t - \frac{t_p}{2}), \tag{B.10}$$

leads to the necessity of deconvolution in the frequency domain:

$$\hat{B}_1^{env}(f) = \text{boxcar}\frac{f - f_c}{\Delta f} \cdot \text{sinc}(\pi t_p(f - f)). \tag{B.11}$$

Figure B.1 depicts the selective excitation pulse sequence in a common pulse timing diagram, showing the RF pulse and selection gradient as a function of time. (The selection gradient is followed by a negative gradient pulse in order to rephase the spins across the slice, e.g. remove the gradient.)
The signal obtained by selective excitation now yield information on a specific section of the object. Still, to obtain an image, i.e. a two dimensional picture, two further signal encoding techniques are necessary.

B.1.2 Frequency Encoding

As previously mentioned the NMR resonant frequency is directly proportional to the strength of the main magnetic field. If we apply a gradient field, e.g. of magnitude G_x in x direction the variation of the resonance frequency ω with position x writes:

$$\omega(x) = \gamma(B_0 + xG_x). \tag{B.12}$$

In frequency encoding the gradient is switched on after excitation of the spins. The NMR signal is then observed in presence of the gradient. This is essential for frequency encoding. Suppose a homogeneous sample in the scanner at the position $x = 0$. The magnetic field at the object will be $\mathbf{B_0}$ and the angular frequency of the signal $\omega_0 = \gamma B_0$. If we move the object to x = -25 cm the local magnetic field will be weaker than B_0 due to the superimposed gradient field, resulting in a frequency less than ω_0.

Similarly at position $x = +25$ cm the signal frequency will be greater than ω_0. The spatial information is now encoded in the frequencies of the signal emitted:

$$S(t) = \int_{-\infty}^{\infty} \rho(x) e^{-i\gamma(B_o + xG_x)t} \mathrm{d}x \tag{B.13}$$

and can therefore be recovered by a Fourier transform (see Figure B.2). $\rho(x)$ denotes the projection of the proton density at location x. In practice the signal is demodulated with the frequency $\omega_0 = \gamma B_0$. The remaining signal is

$$S(t) = \int_{-\infty}^{\infty} \rho(x) e^{-i\gamma x G_x t} \mathrm{d}x. \tag{B.14}$$

We refer to the gradient field employed for frequency encoding as readout gradient and point out that it must not necessarily coincide with one of the frame-axis.

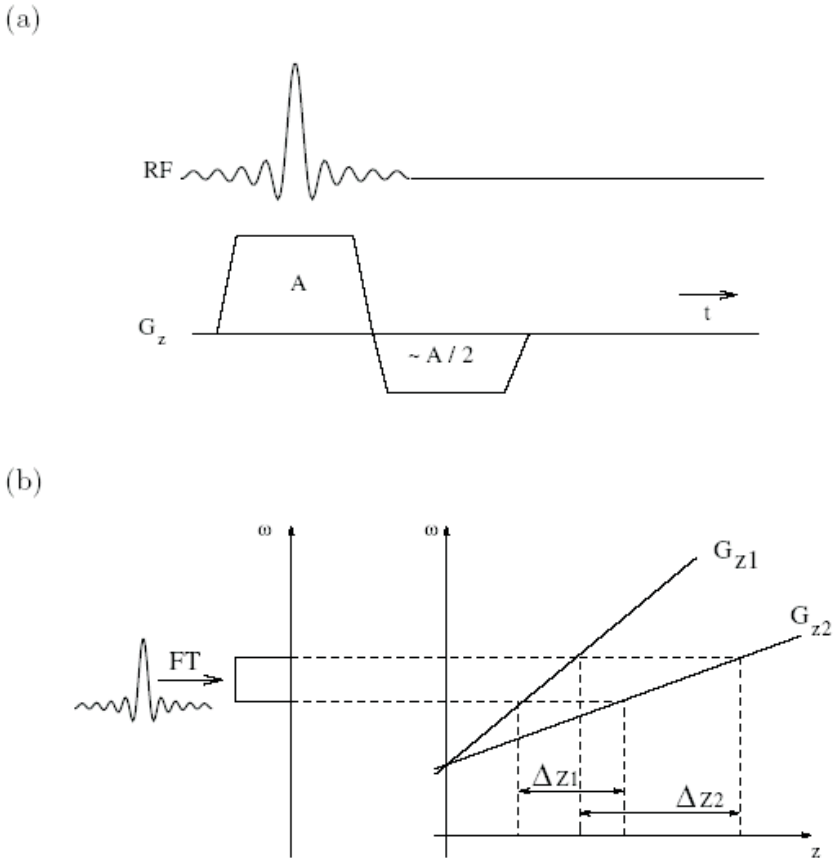

Figure B.1: *(a) Selective Pulse Excitation Sequence: By applying a shaped excitation pulse in presence of a selection gradient G_z only the spins of a thin slice in the sample will contribute to the signal. The gradient is afterwards reversed for reasons of rephasing. (b) Effect of different gradient strength G_z on slice thickness.*

As a consequence of frequency encoding, after an FID-pulse, the spins will dephase during data acquisition. It is therefore preferable to record the signal as a spin echo. Apart from using a refocusing RF pulse, it can be accomplished by applying a negative gradient before switching on the readout gradient. As an effect the spins are already dephased when starting the measurement and rephase in presence of the frequency encoding gradient, yielding an echo signal (gradient echo) as shown on the Figure B.3.

Now, when applying selective excitation and frequency encoding during the measurement, what one gets is a projection of the chosen slice along the direction of the applied gradient. In order to obtain information in two dimensions another method called phase encoding is added.

B.1.3 Phase Encoding

Assume that a gradient field similar to that of a frequency encoding gradient field is now applied in y-direction after the 90°-RF pulse, but before the readout takes place. Immediately after the pulse the magnetisation is in phase everywhere. While the gradient remains switched on the spins start to dephase relative to each other as a result of their different precession velocity. That means, that as data acquisition progresses the amount of dephasing advances. When the gradient is switched off the spins are still dephased, but do not dephase any further. They remain stationary in the rotating frame of reference.
The phase of the magnetisation is given by:

$$\phi(y) = \gamma y \int_0^T G_y \mathrm{d}t \tag{B.15}$$

where T denotes the period during which the phase encoding gradient of magnitude G_y is switched on. Equation B.15 implies that the dephasing rate also increases with the magnitude of the gradient. The dependency of the phase of the signal on the distribution of the nuclei along the gradient is called phase encoding. When applying frequency encoding and phase encoding at the same time the signal originating from an excited slice at position z_0 is modified to:

$$S(t) \propto \int_{Slice} \rho(x, y, z_0) e^{-i \int \omega(x,y,t) \mathrm{d}t} \mathrm{d}x \mathrm{d}y. \tag{B.16}$$

The phase of the signal is a function of the gradients applied between excitation and the moment of sampling

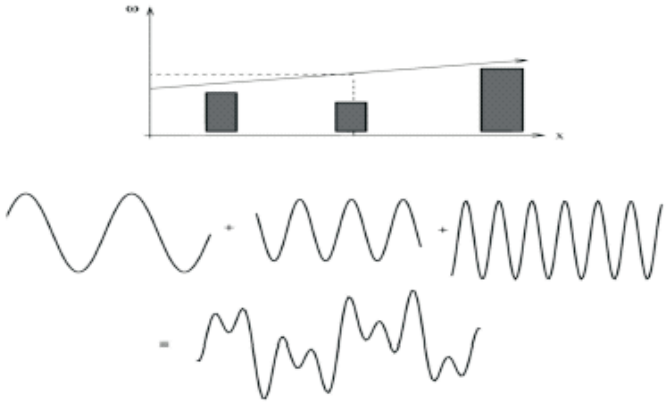

Figure B.2: *(a) Frequency encoding: Magnitude and frequency of the individual signals coincide with proton density and position of the sample, respectively, adding up to the actually observed signal.*

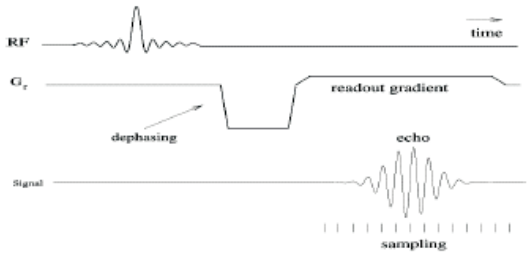

Figure B.3: *(Gradient Echo: produced by inverse gradient field.*

$$\int \omega(x,y,t)\mathrm{d}t = \gamma x \int G_x(t)dt + \gamma y \int G_y(t)\mathrm{d}t. \tag{B.17}$$

This concludes the technical part on the possibilities of NMR-signal generation suitable for image formation. Their usage for imaging as done in practice is given in the following. The term imaging roughly concerns two fields: data assessment and image reconstruction. In the first step encoding mechanisms are combined and the measured signals are digitally recorded. In the second step these signals afford proper decoding according to the measurement technique applied to gain an image.

B.2 Imaging

NMR imaging techniques comprise two main categories: reconstruction from projections and reconstruction by Fourier transform (FT). The simplest of all NMR imaging methods is accomplished by using a single field gradient yielding a one dimensional projection of the sample. Reconstruction of a 2D image is then possible from a number of projections at different angles (Radon,1917). The algorithms involved were inherently adopted from X-ray computed tomography, which was being enthusiastically developed in the early days of MR imaging. In practice, projection reconstruction is unduly sensitive to B_0 inhomogeneities and patient movement and has now largely fallen out of favour in standard MR and given way to the versatile Fourier transform methods. In the following we therefore focus on imaging methods using Fourier transform to reconstruct the images. For reconstruction from projection we refer to the extensive bibliography of Computed Tomography of which we cite only a few important providing an overview and further reading references (Hermann, 1980). A section on an essential tool of Fourier imaging, the k-space, serves as an introduction.

B.2.1 k-Space

The essential feature to Fourier imaging is that it comprises all techniques discussed in the previous chapters, namely: slice selection and spin excitation, a readout gradient, during which the signal is collected and a phase encoding gradient between excitation and the readout periods. The imaging sequence is then called Spin-Warp-Sequence (Edelstein and Hutchison, 1980). For simpler FT imaging the three orthogonal gradients are separate functions. For more sophisticated sequences though, this is not the case. In particular, the functions of

readout and phase encoding can be mixed and involve all three gradients. To enable description of these complex imaging schemes and at the same time justify image reconstruction by Fourier transform a notation known as k-space has been established (Twieg, 1983). It is obtained by simple variable substitutions.
Recall the frequency encoded signal S(t) :

$$\begin{aligned} S(t) &= \int_{-\infty}^{\infty} \rho(x) e^{-i\gamma(B_0 + xG_x)t} \mathrm{d}x \\ &= e^{-i\omega_0 t} \int_{-\infty}^{\infty} \rho(x) e^{-i\gamma x G_x t} \mathrm{d}x, \end{aligned} \tag{B.18}$$

where $\rho(x)$ represents the spin distribution at position x of the readout gradient. Removing the carrier signal $e^{-i\omega_0 t}$ yealds:

$$S(t) = e^{-i\omega_0 t} \int_{-\infty}^{\infty} \rho(x) e^{-i\gamma x G_x t} \mathrm{d}x. \tag{B.19}$$

By substituting

$$k_x = \begin{cases} \gamma/2\pi G_x t & \text{for FID signals} \\ \gamma/2\pi G_x (t - TE) & \text{for echo signals} \end{cases} \tag{B.20}$$

we obtain following relationship to the Fourier transform:

$$S(k_x) = \int_{-\infty}^{\infty} \rho(x) e^{-2\pi i k_x x} \mathrm{d}x. \tag{B.21}$$

Similar relations hold for the y and z components. Considering that the gradients are not necessarily constant throughout the measuring time this suggests:

$$S(k_x) = \int_{-\infty}^{\infty} \rho(x) e^{-2\pi i \boldsymbol{k} x} \mathrm{d}x, \tag{B.22}$$

where $x \in \mathbb{R}^3$ and

$$\boldsymbol{k} = \gamma \int_0^T \mathbf{G}(t) \mathrm{d}t. \tag{B.23}$$

This formula represents the point $\mathbf{k} = (k_x, k_y, k_z)$ reached by the spin system at time T under influence of the gradient system $\mathbf{G} = (G_x, G_y, G_z)$, assuming that the spins are all in phase at $T = 0$. If this is not the case the phase distribution at time $T = 0$ can be represented by an additional constant term k_0 on the right hand side of the equation above. By applying this formula to a field gradient sequence a path can be traced out in k-space. Figure B.5 depicts an example for a two dimensional image acquisition. As k-space directly corresponds to the Fourier space, in practice, the sequences are chosen according to the frequencies preferred to be imaged.

B.2.2 Fast Imaging

The term Echo Planar Imaging (EPI) refers to a class of high-speed imaging methods that collect a complete set of two-dimensional data during the free induction decay of a single excitation pulse. Thus EPI has become a synonym for single shot imaging.
The original sequence introduced by Mansfield (1977) involves the usage of two frequency encoding gradients during the readout period: a constant gradient and a rapidly alternating gradient. The alternating gradient produces a series of gradient echoes during the free precession period. This results in a coverage of k-space by a "zig-zag" trajectory.
Simpler in concept and much more widely used is a variant called Blipped Echo Planar Imaging (BEPI). It has been investigated by Johnson et al. (1983). Sequence and track through k-space are given in Figure B.6. Individual gradient echos are phase encoded using a series of blipped G_y pulses resulting in a rectilinear trajectory.
The number of k-space lines that can be scanned in a single excitation is restricted to the inevitable T_2-decay of the signal and the magnitude and switching rate of the readout gradient. This is particularly true when one wants to obtain high resolution images. For this reason the measurement must usually be repeated with the size of the initial negative blip changed slightly. This renders a vertical offset in the track so to fill the intermediate k-space lines.
The need to repeat the BEPI sequence to obtain enough data tends to defeat the idea of singleshot As a consequence various techniques have been developed to circumvent this problem. Due to the recent technical improvements several variations of single shot EPI have become viable.

B.2.3 Reconstruction

We recall that the measured signal is a frequency and phase encoded information on the nucleus density $\rho(x)$, of the measured object and can be recovered by a (inverse) Fourier transform. Thereby, the third dimension, representing the slice thickness, is often neglected. Naturally the usage of the well-known Fast Fourier Transform (FFT) is preferable, so that attention has to be paid to the frequencies scanned in k-space as well as to the spacing of the sampling points. The number of data samples N are required to be of a power of 2 and must also be equally distributed in the (k_x, k_y)-plane.
Accomplishing the Fourier Transform is not the only task in reconstructing an image. Due to technical and experimental limitations as well as physiological

phenomena a series of artefacts is introduced. If their origin can be determined and evolution course retained a removal might be possible by post processing. Typical technical deficiencies are e.g. inhomogeneities in the static magnetic field B_0 or inaccuracies in the switching gradients. Susceptibility artefacts at tissue boundaries, heart motion, breathing as well as patient motion artefacts are much harder to control. As they are more or less omnipresent in in-vivo MRI measurements their correction is of even greater interest. Finally, the images can not only be used for imaging anatomical structures but also to show up physiological changes. This can be accomplished by observing the signal fluctuations in each voxel of a sequence of MR images. The method is referred to as functional MRI (fMRI).

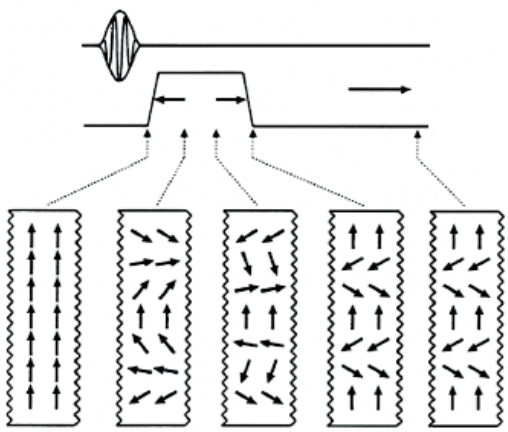

Figure B.4: *Phase Encoding: Behaviour of the spins before, during and after applying the phase encoding gradient. In this diagram it is assumed to be directed vertically.*

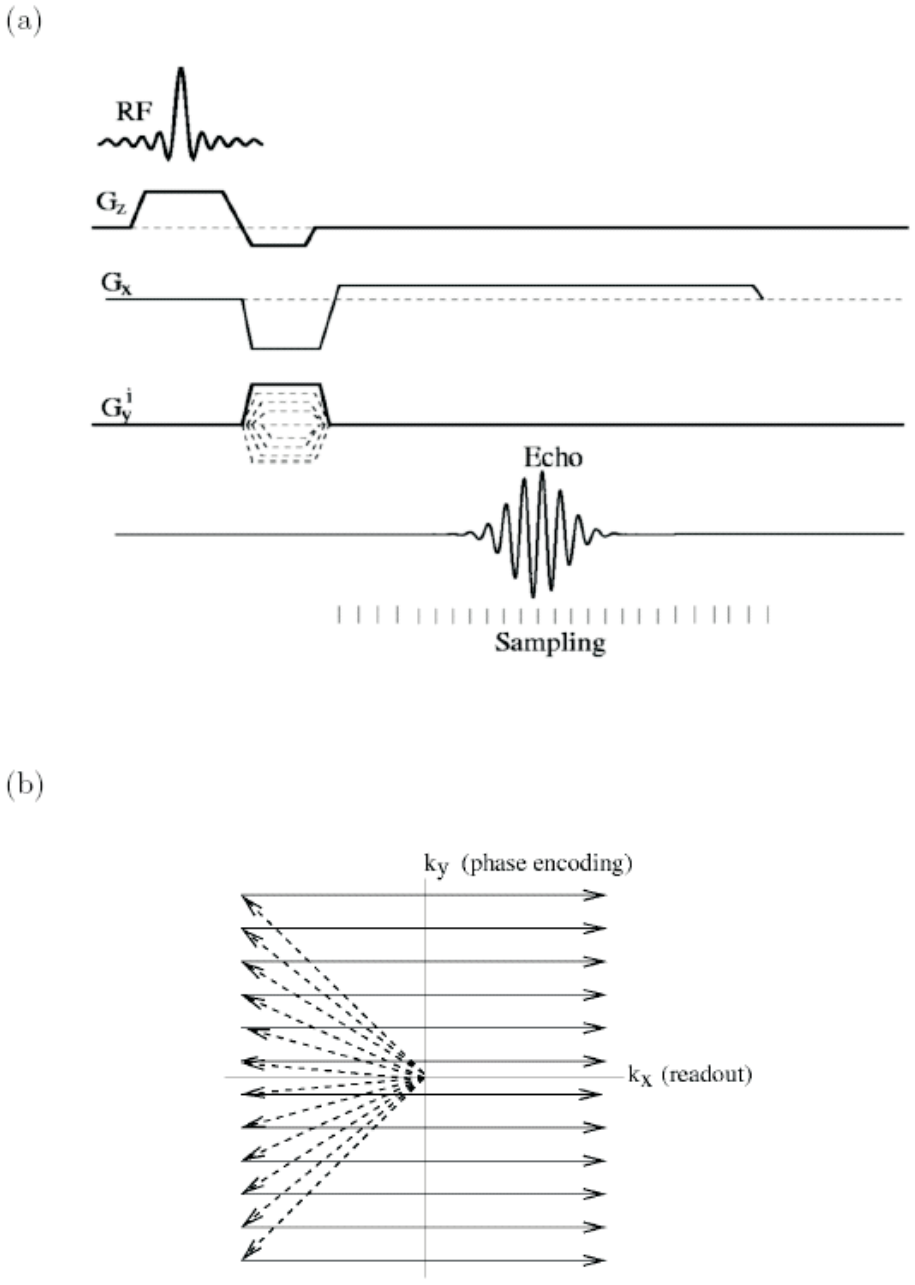

Figure B.5: *(a) 2D sequence employing slice selection gradient G_s, readout gradient G_r and phase encoding gradient G_p at the same time (Spin Warp). (b) according sampling trace in k-space (adapted from Rauscher, 2002)*

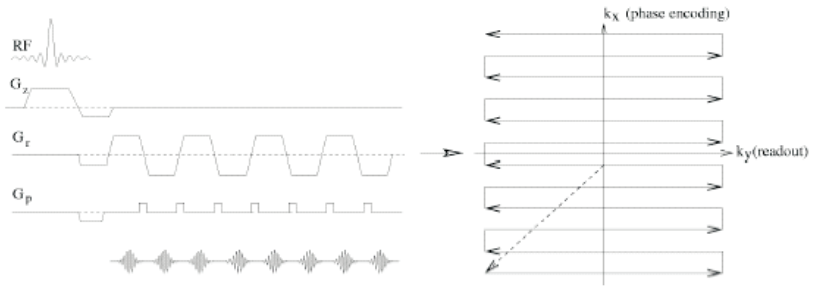

Figure B.6: *Blipped Echo Planar Imaging Sequence. The dotted line suggests the track after a slight change of the blips (adapted from Rauscher, 2002.*

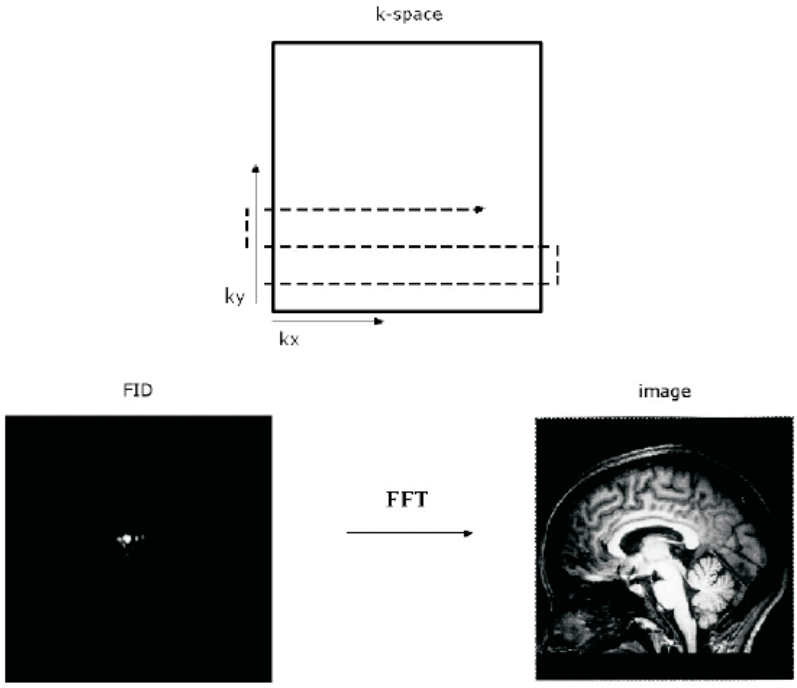

Figure B.7: *Image reconstruction: The acquired data in k-space are Fourier transformed to gain an image representing the density distribution of the measured nuclei.*

[12pt]book graphicx graphicx [body=6.0in, 8.2in,left=1.25in,right=1.25in]geometry amsmath,amssymb rotating [sectionbib]natbib setspace array caption2

Bibliography

Ahonen, A. I., Hämäläinen, M. S., Kajola, M. J., Knuutila, J.E.T., Laine, P. P., Lounasmaa, O. V., Parkkonen, L. T., Simola, J. T., Tesche, C. D. 1993. 122-channel SQUID instrument for investigating the magnetic signals from the human brain. *Physica Scripta* 49, 198-205.

Allerhand, M., Butterfield, S., Cutler, A., and Patterson, R.D., 1992. Assessing syllable strength via an auditory model, Proc. *Inst. Acoust.* 14, 297-304.

Aoki, I., Tanaka, C., Takegami, T., Ebisu, T., Umeda, M., Fukunaga, M., Fukuda, K., Silva, A., Koretsky, A., and Naruse, S, 2002. Dynamic activityinduced manganesedependent contrast magnetic resonance imaging (DAIM MRI). *Magnetic Resonance in Medicine*, 48(6), 927-933.

Akeroyd, M.A., Patterson, R.D., 1995. Discrimination of wideband noises modulated by a temporally asymmetric function. *J. Acoust. Soc. Am.* 98, 2466-2474.

Ashburner, J. and Friston, K., 1997. Multimodal image coregistration and partitioning a unified framework. *Neuroimage*, 6(3), 209-217.

Bandettini, P. A., Jesmanowicz, A., Van Kylen, J., Birn, R.M. and Hyde, J.S., 1998. Functional MRI of brain activation induced by scanner acoustic noise. *Magnetic Resonance in Medicine*, 39, 410-416.

Biermann, S., Heil, P., 2000. Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. *J. Neurophysiol.*, 84, 2426-2439.

Cutting, J.E. and Rosner, B.S., 1974. Categories and boundaries in speech and music. *Perception and Psychophysics*, 16, 564-570.

Cohen D., 1968. Magnetoencephalography, evidence of magnetic fields produced by alpha-rhythm currents. *Science* 161, 784-6.

Cohen D., 1970a. Large-volume conventional magnetic shields. *Rev. Phys. Appl.*, 5, 53-58.

Cohen D., 1970b. Low-field room built at high field magnet lab. *Phys. Today*, 23, 56-57.

Cohen D., 1972. Magnetoencephalography: Detection of brain's electric activity with a superconducting magnetometer. *Science*, 175(4022), 664-666.

Edelstein W. A. and Hutchison J. M. S., 1980. Spin warp NMR imaging and applications to human whole-body imaging. *Physics in medicine and biology*, 25, 751-756.

Eden, G.F., Joseph, J.E., Brown, H.E. and Zeffiro, T.A., 1999. Utilizing hemodynamic delay and dispersion to detect MRI signal change without auditory interference: the behavior interleaved gradients technique. *Magn. Reson. Med.*, 41, 13-20.

Edmister, W.B., Talavage, T.M., Ledden, P.J. and Weisskopff, R.M., 1999. Improved Auditory cortex imaging using clustered volume acquisitions. *Hum. Brain Mapp.*, 7, 89-97.

Efron, B. and Tibshirani, R. J., 1993. *An Introduction to the Bootstrap.* Chapman and Hall, New York.

Fay, R.R., Chronopoulos, M., Patterson, R.D., 1996. The sound of a sinusoid: Perception and neural representations in the goldsh (Carassius auratus). *Aud. Neurosci.* 2, 377-392.

Fay, R. R., Yost, W. A., and Coombs, S., 1983. Psychophysics and neurophysiology of repetition noise processing in a vertebrate auditory system. *Hearing Res.*, 12, 31-55.

Franckowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J. and Mazziotta, J.C., 1997. *Human brain function* San Diego: Academic Press.

Friston, K.J., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D. and Frackowiak, R.S.J., 1995a. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.*, 2: 189-210.

Galaburda, A., and Sanides, F., 1980. Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.* 190: 597-610.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S . and Turner R., 1995b. Analysis of fMRI time-series revisited. *Neuroimage*, 2, 45-53.

Friston, K. J., Williams, S., Howard, R., Franckowiak, R. S. and Turner, R., 1996. Movement-related effects in fMRI time-series. *Magn. Reson. Med.*, 35 (3), 346-355.

Friston, K. J., Price, C. J, Buechel, C. and Franckowiak, R. S., 1997. A taxonomy of study design. IN Franckowiak, R.S.J., Friston, K. J., Frith, C. D., Dolan, R. J. and Mazziotta, J. C. (Eds.), *Human brain function San Diego: Academic Press*, 141-159.

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouquette, S. and Poline, J., 2000. To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *Neuroimage*, 12(2), 196-208.

Galaburda A.M. and Sanides F., 1980. Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol*, 190, 597-610.

Gordon, J. W., 1987. The perceptual attack time of musical tones. *J. Acoust. Soc. Am.* 82, 88-105.

Giguère, C. and Woodland, P. C., 1994. A computational model of the auditory periphery for speech and hearing research. I. Ascending paths. *J. Acoust. Soc. Am.*, 95, 331-342.

Giguere, C., Kunov, H., and Smoorenburg, G.F., 1995. Computational modelling of psycho-acoustic combination tones and distortion-product otoacoustic emissions, *15th Int. Cong. on Acoustics, Trondheim (Norway), 26-30 June*

Giguère, C. and Woodland, P. C., 1994. A computational model of the auditory periphery for speech and hearing research. I. Descending paths. *J. Acoust. Soc. Am.*, 95, 343-349.

Glasberg, D. D. and Moore, B. C. J., 1990. Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47, 103-138.

Greenwood, D. D., 1990. A cochlear frequency-position function for several species-29 years later. *J. Acoust. Am.*, 87, 2592-2605.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V., 1993. Magnetoencephalography – theory, instrumentation, and applications to

noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2), 413-497.

Hahn, E., 1950. Spin echoes. *Physical Reviews*, 80, 580-594.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.P., Gurney, E.M. and Bowtell, R.W., 1999. Sparse temporal sampling in auditory fMRI. *Hum. Brain Mapp.*, 7, 213-223.

Hall, D.A., Summerfield, A.Q., Goncalves, M.S., Foster, J.R., Palmer, A.R. and Bowtell, R.W. (2000). Time-course of the auditory BOLD response to scanner noise. *Magn. Res. Med.*, 43, 601-606.

Hartmann, W. M., 1978. The effect of amplitude envelope on the pitch of sine wave tones. *J. Acoust. Soc. Am.* 63, 1105-1113.

Helmholtz, H. L. F. von, 1853. Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. *Annalen der Physik und Chemie,*, 89, 211-233, 353-377.

Herman, G. T., 1980. *Image Reconstruction from Projections - The Fundamentals of Computerized Tomography. Academic Press*

Hoechstetter, K., 2001. *Magnetic Source Imaging of Tactile Evoked Activity in the Human Secondary Somatosensory Cortex.* PhD Thesis, Faculty of Physics, Heidelberg.

Houtsma, A. J. M., Rossing, T. D., and Wagenaars, W. M., 1987. Effects of echoes in *Auditory Demonstrations Institute for Perception Research IPO*, Eindhoven, The Netherlands.

Iversen, L. L., 1979. The chemistry of the brain. Sci. Am. 241, 118-129.

Jezzard, P., and Clare, S., 2001. Principles of nuclear magnetic resonance and MRI. In P. Jezzard, P. M. Matthews, and S. M. Smith, editors,*Functional MRI: An Introduction to Methods, Oxford University Press, Oxford*, 67-92.

Johnson G., Hutchison J. M. S., Redpath T. W., and Eastwood L.M., 1983. Improvements in performance time for simultaneous three dimensional NMR imaging. *Journal of Magnetic Resonance*, 54, 374-384.

Joutsieniemi S.-L., Hari R., and Vilkman V., 1989. Cerebral magnetic responses to noise bursts and pauses of different durations. *Audiology* 28, 325-333.

Kandel, E. R. and Schwartz, J. H., 1985. *Principles of neural sciences.*, Elsevier, New York.

Kodera K., Hink R. F., Yamada O., and Suzuki J. I. 1979. Effects of rise time on simultaneously recorded auditory-evoked potentials from the early, middle and late ranges. *Audiology*, 18, 395-402.

Kwong, K. K., Beliveau, J. W., Chester, D. A., Goldberg, I.E., Weisskoff, R. M., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *In Proceedings of the Natural Academic Sciences*, 80, 5675-5679.

Leonard, C. M., Puranik, C., Kuldau, J. M., and 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, 397-406.

Liégeois-Chauvel, C., Musolino, A. and Chauvel, P., 1991. Localization of the primary auditory area in man. *Brain*, 114, 139-153.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P. and Chauvel, P., 1994. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of middle latency components. *Electroencephal. Clin. Neurophysiol.*, 92, 204-214.

Lu, T., Liang, L., Wang, X., 2001. Neural representations of tempo- rally asymmetric stimuli in the auditory cortex of awake primates. *J. Neurophysiol.*, 85, 2364-2380.

Lu, T., Liang, L., Wang, X., 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neurosci.*, 4, 1131-1138

Luce, D., 1959. *Individual choice behavior.* John Wiley and Sons, New York.

Lütkenhöner, B., and Steinsträter, O., 1998. High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol. Neurootol.*, 3, 191-213.

Lütkenhöner, B., Lammertmann, C., and Knecht, S., 2001. Latency of Auditory Evoked Field Deflection N100m Ruled by Pitch or Spectrum? *Audiol. Neurootol.*,6, 263-278.

Mäkelä, J. P., Hämäläinen, M., Hari, R., and McEvoy, L., 1994. Whole-head mapping of middle-latency auditory evoked magnetic fields. *Electroencephalogr. clin. Neurophysiol.*, 92, 414-421.

Mansfield. P., 1977. Multi-planar image formation using NMR spin echos. *Journal of Physics C: Solid State Physics*, 1977.

Meddis, R., 1988. Simulation of auditory-neural transduction: Further studies *J. Acoust. Soc. Am.*, 83, 1056-1063.

Moore J.K., 1994. The human brainstem auditory pathway. *In: Jackler. RK, Brackmann DE, editors. Neurotology. St. Louis: Mosby.* 117. 305

Morton, J., Marcus, S.M. and Frankish, C.R., 1976. Perceptual centers (P-centers), *Psychological Review* 83, 405-408.

Mosher J.C., Leahy R.M., Lewis P.S., 1999. EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng.*, 46(3), 245-259.

Näätänen, R. and Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24, 375-425.

Neuert V., Pressnitzer D., Patterson R.D., Winter I.M., 2001. The responses of single units in the inferior colliculus of the guinea pig to damped and ramped sinusoids. *Hearing Research*, 159(1-2), 36-52.

Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W, 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the Natural Academic Sciences*, 87, 9868-9872.

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., and Ugurbil, K., 1993. Functional brain mapping by blood oxygenation level dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysics Journal*, 64, 803-812.

Pandya D.N., 1995. Anatomy of the auditory cortex. *Rev Neurol (Paris)*, 151, 486-494.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., and Wittkowski, W., 1988. Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalogr. clin. Neurophysiol.*, 69, 160-170.

Paquette, C., Peretz, I., 1997. Role of familiarity in auditory discrimination of musical instruments: a laterality study. Cortex 33, 689696.

Patterson, R. D. and Moore, B. C. J., 1986. Auditory filters and excitation patterns as representations of frequency resolution.
In: *Frequency selectivity in Hearing edited by* B.C.J. Moore, Academic PressLimited, London, 123-177.

Patterson, R. D., 1987. A pulse ribbon model of monaural phase perception. *J. Acoust. Soc. Am.* 82, 1560-1586.

Patterson, R.D., Anderson, T., and Allerhand, M., 1994. The auditory image model as a preprocessor for spoken language,*in Proc. Third ICSLP, Yokohama, Japan*, 1395- 1398.

Patterson, R.D, 1994a. The sound of a sinusoid: Spectral models. *J. Acoust. Soc. Am.* 96, 1409-1418.

Patterson, R.D., 1994b. The sound of a sinusoid: Time-interval models. *J. Acoust. Soc. Am.* 96, 1419-1428.

Patterson, R. D., Allerhand, M. H., and Giguere, C., 1995. Time-domain modelling of peripheral auditory processing: A modular architecture and software platform. *J. Acoust. Soc. Am.* 98, 1890-1894.

Patterson, R. D., Handel, S., Yost, W. A., and Datta, A. J., 1996. The relative strength of tone and noise components in iterated rippled noise. *J. Acoust. Soc. Am.* 100: 3286-3294.

atterson, R. D., and Irino, T., 1998. Modeling temporal asymmetry in the auditory system, *J. Acoust. Soc. Am.* 104, 2967-2979.

Penhune V.B., Zatorre R.J., MacDonald J.D., Evans A.C., 1996. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb Cortex*, 6(5), 661-672.

Picton, T. W., Hillyard, S. A., Krausz, H. I., and Galambos, R., 1974. Human auditory evoked potentials. I: evaluatiuon of components. *Electroencephalogr. Clin. Neurophysiol.*, 36, 179-190.

Polich J., Aung M., Dalessio D. J., 1987. Long latency auditory evoked potentials: intensity, inter-stimulus interval, and habituation. *Pavlov J Biol Sci.*, 23(1), 35-40.

Pressnitzer, D., Patterson, R. D., and Krumbholz, K., 2001. The lower limit of melodic pitch. *J. Acoust. Soc. Am.*, 109, 2074-2084.

Rademacher J., Caviness V.S. Jr, Steinmetz H., Galaburda A.M., 1993. Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology. *Cereb. Cortex* 3(4), 313-329.

Radon, J. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, 1917. *Berichte der Sächsischen Akademie der Wissenschaften B*, 96, 262-279.

Rauschecker J.P., 1998a. Parallel processing in the auditory cortex of primates. *Audiol Neurootol*, 3, 86-103.

Rauscher A. 2002. Correction of geometric distortion in fast magnetic resonance imaging at 3 tesla. Masters thesis, Technical University of Vienna, TU Wien.

Ravitz, M. E., Melcher, J. R., and Kiang, N. Y.-S., 2000. Acoustic noise during functional magnetic resonance imaging. *J. Acoust. Soc. Am*, 108(4), 1683-1696.

Reite, M., Teale, P., Zimmermann, J. T., Davis, K., and Whalen, J., 1988. Source location of a 50 ms latency auditory evoked field component. *Electroencephalogr. clin. Neurophysiol.*, 92, 149-160.

Ritter, S., 2004. Magnetoencephalographic Study on the Representation of Huygens Noise in the Auditory Cortex of Humans. PhD Thesis, Faculty of Physics, Heidelberg.

Ritter, S., Dosch, H. G., Specht, H. J., and Rupp, A., 2005. Neuromagnetic responses reflect the pitch changes of regular interval sounds. *Neuroimage*, 27:3, 533-543.

Roberts, T. P. and Poeppel, D., 1996. Latency of auditory evoked M100 as a function of tone frequency. *Neuroreport*, 7, 1138-1140.

Rosen, S.M. and Howell, P., 1981. Plucks and bows are not categorically perceived, *Perception and Psychophysics* 30: 156-158.

Rupp, A., Uppenkamp, S., Bailes, J., Gutschalk, A., and Patterson, R., 2005. Time constants in temporal pitch extraction: a comparison of psychophysical and neuromagnetic data. In: *Auditory signal processing: physiology, psychoacoustics, and models, edited by D. Pressnitzer, A. de Cheveigne, S. McAdams, L. Collet, Springer, New York.*

Savoy, R. L., Ravicz, M. E., and Gollub, R, 1999. The Psychophysiological Laboratory in the Magnet: Stimulus Delivery, Response Recording, and Safety, *In Medical Radiology, Diagnostic Imaging and Radiation Oncology: Functional MRI, edited by C. Moonen and P. Bandettini Springer, Berlin*, 347-365.

Scherg, M., and von Cramon, D., 1986. Evoked dipole source potentials of the human auditory cortex. *Electroencephalogr. clin. Neurophysiol.*, 65, 344-360.

Scherg, M., Vajsar, J., and Picton, T. W., 1989. A source analysis of the human auditory evoked potentials. *J. Cognit. Neurosci.*, 1, 336-354.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., and Rupp, A., 2004. Tonotopy, Tonochrony and Pitch Mapping in Heschl's Gyrus of Musicians. *Proceedings of the 14th International Conference on Biomagnetism (BIOMAG),* 570.

Schneider P., Sluming V., Roberts N., Scherg M., Goebel R., Specht H.J., Dosch H.G., Bleeck S., Stippich C. and Rupp A., 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat Neurosci.*, 8(9), 1241-1247.

Shah, N.J., Jancke, L., Grosse-Ruyken, M.L. and Muller-Gartner, H.W., 1999. Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination. *J. Magn. Reson. Imaging*, 9, 19-25.

Springer,C. S., Patlak, C. S., Palyka, L. and Huang, W. 1999. Functional MRI, *Springer Verlag*, 9, 91-102.

Stufflebeam, S. M., Poeppel, D., Rowley, H. A., and Roberts, T. P., 1998. Peri-threshold encoding of stimulus frequency and intensity in the M100 latency. *Neuroreport*, 9, 91-94.

Talairach, P., Tournoux, J., 1988. *A Stereotactic Coplanar Atlas of the Human Brain.* Thieme, Stuttgart.

Talavage, T.M., Edmister, W.B., Ledden, P.J. and Weisskopff, R.M., 1998. Quantification of the impact of MRI scanner noise on auditory cortex. *In: Proceedings of the 6th Annual Meeting of ISMRM, Sydney, Australia, 1998*, 1502.

Twieg D. 1983. The k-trajectory formulation of the NMR imaging process with application in analysis and synthesis of imaging methods. *Medical Physics*, 10, 610-621.

Ulmer, J.L., Biswal, B.B., Yetkin, F.Z., Mark, L.P., Mathews, V.P., Prost, R.W., Estkowski, L.D., McAuliffe, T.L., Haughton, V.M. and Daniels, D. L., 1998. Cortical activation response to acoustic echo planar scanner noise. *J. Comput. Assist. Tomgr.*, 22, 111-119.

Vos, J. and Rasch, R., 1981. The perceptual onset of musical tones. *Perception and Psychophysics* 29, 323-335.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C., 2001. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb Cortex*, 11(5), 411-423.

Wood, C. C. and Wolpaw, J. R., 1982. Scalp distribution of human auditory evoked potentials. II. Evidence for overlapping sources and involvement of auditory cortex. *Electroencephalogr. Clin. Neurophysiol.*, 54, 25-38.

Zatorre, R.J. and Penhune, V.B., 2001. Spatial Localization after Excision of Human Auditory Cortex. *J. Neurosci.* 21, 6321-6328.