ANALYSIS OF THE STRUCTURE OF TIME–FREQUENCY INFORMATION IN ELECTROMAGNETIC BRAIN SIGNALS

Doctoral Dissertation

Ville Mäkinen

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Laboratory of Biomedical Engineering

ANALYSIS OF THE STRUCTURE OF TIME–FREQUENCY INFORMATION IN ELECTROMAGNETIC BRAIN SIGNALS

Doctoral Dissertation

Ville Mäkinen

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Laboratory of Biomedical Engineering

Teknillinen korkeakoulu
Teknillisen fysiikan ja matematiikan osasto
Lääketieteellisen tekniikan laboratorio

Author   Ville Mäkinen

Name of the dissertation

Analysis of the Structure of Time–Frequency Information in Electromagnetic Brain Signals

Date of manuscript	18.10.2005	Date of the dissertation	28.4.2006

☐ Monograph

☑ Article dissertation (summary + original articles)

Department	Department of Engineering Physics and Mathematics
Laboratory	Laboratory of Biomedical Engineering
Field of research	Signal processing and neuroscience
Opponent(s)	Professor Bernd Lütkenhöner
Supervisor	Acting Professor Risto Ilmoniemi
(Instructor)	Docent Patrick May and Docent Hannu Tiitinen

Abstract

This thesis encompasses methodological developments and experimental work aimed at revealing information contained in time, frequency, and time–frequency representations of electromagnetic, specifically magnetoencephalographic, brain signals.

The work can be divided into six endeavors. First, it was shown that sound slopes increasing in intensity from undetectable to audible elicit event-related responses (ERRs) that predict behavioral sound detection. This provides an opportunity to use non-invasive brain measures in hearing assessment. Second, the actively debated generation mechanism of ERRs was examined using novel analysis techniques, which showed that auditory stimulation did not result in phase reorganization of ongoing neural oscillations, and that processes additive to the oscillations accounted for the generation of ERRs. Third, the prerequisites for the use of continuous wavelet transform in the interrogation of event-related brain processes were established. Subsequently, it was found that auditory stimulation resulted in an intermittent dampening of ongoing oscillations. Fourth, information on the time-frequency structure of ERRs was used to reveal that, depending on measurement condition, amplitude differences in averaged ERRs were due to changes in temporal alignment or in amplitudes of the single-trial ERRs. Fifth, a method that exploits mutual information of spectral estimates obtained with several window lengths was introduced. It allows the removal of frequency-dependent noise slopes and the accentuation of spectral peaks. Finally, a two-dimensional statistical data representation was developed, wherein all frequency components of a signal are made directly comparable according to spectral distribution of their envelope modulations by using the fractal property of the wavelet transform. This representation reveals noise buried processes and describes their envelope behavior.

These examinations provide for two general conjectures. The stability of structures, or the level of stationarity, in a signal determines the appropriate analysis method and can be used as a measure to reveal processes that may not be observable with other available analysis approaches. The results also indicate that transient neural activity, reflected in ERRs, is a viable means of representing information in the human brain.

This thesis was prepared in collaboration between the following units:

Apperception & Cortical Dynamics (ACD), Department of Psychology, University of Helsinki, Finland

BioMag Laboratory, Engineering Centre, Helsinki University Central Hospital, Helsinki, Finland

Laboratory of Biomedical Engineering, Department of Engineering Physics and Mathematics, Helsinki University of Technology

Table of contents

List of publications

This thesis consists of an overview and of the following articles:

I   Mäkinen V, May P, Tiitinen H, 2004. Transient brain responses predict the temporal dynamics of sound detection in humans. Neuroimage 21, 701–706.

II   Mäkinen V, Tiitinen H, May P, 2005. Auditory event-related responses are generated independently of ongoing brain activity. Neuroimage 24, 961–968.

III   Mäkinen VT, Tiitinen H, May PJC, 2004. Auditory evoked responses are additive to brain oscillations. Neurol Clin Neurophysiol.

IV   Mäkinen VT, May PJC, Tiitinen H, 2004. Human auditory event-related processes in the time-frequency plane. Neuroreport 15, 1767–1771.

V   Tiitinen H, Mäkinen VT, Kičić D, May PJC, 2005. Averaged and single-trial brain responses in the assessment of human sound detection. Neuroreport 16, 545–548.

VI   Mäkinen VT, May PJC, Tiitinen H, 2005. The use of stationarity and nonstationarity in the detection and analysis of neural oscillations. Neuroimage 28, 389–400.

Author's contribution

The initial idea to all studies, except to V, was my own. In all studies, I prepared the stimuli, performed the main part of the measurements, and designed and decided most parts of the experimental setups. I determined and carried out all the data analyses in Studies II, III, IV, and VI, and most parts in Studies I and V. Aside from those used in source localization, the data analysis tools were of my own design and implementation. I drew the illustrations and wrote the initial manuscript versions for all the publications. The final form of the manuscript texts was, however, the result of a group effort with all authors providing invaluable contribution. During the thesis venture, I also had the privilege to participate in the work of several researchers and worked for the technical maintenance and development of the BioMag Laboratory of the Helsinki University Central Hospital, where all measurements were performed.

Acknowledgements

Abbreviations

ARMA	**Autoregressive moving average**
CCWW[†]	**Cycle-length convoluted weighted-wavelet**
CWT	**Continuous wavelet transform**
DOG2	**Second derivative of Gaussian wavelet**
DWT	**Discrete wavelet transform**
ECD	**Equivalent current dipole**
EEG	**Electroencephalography**
ERD	**Event-related desynchronization**
ERF	**Event-related field**
ERP	**Event-related potential**
ERR[†]	**Event-related response; comprises ERF and ERP**
ERS	**Event-related synchronization**
FFT	**Fast Fourier transform**
FSEM[†]	**Fractally scaled envelope modulation**
IIR	**Infinite impulse response**
MEG	**Magnetoencephalography**
MMN	**Mismatch negativity**
N1/N1m	**ERP/ERF deflection with peak latency of around 100 ms**
P1/P1m	**ERP/ERF deflection with peak latency of around 50 ms**
P2/P2m	**ERP/ERF deflection with peak latency of around 200 ms**
PRSE[†]	**Partition-referenced spectral estimate**
RSW[†]	**Response source weighting**
SNR	**Signal-to-noise ratio**
SSP	**Signal-space projection**
SQUID	**Superconducting quantum interference device**

[†] These abbreviations were introduced in the publications of this thesis.

PREFACE

About the research approach of the thesis: This thesis is concerned with revealing the time, frequency, and time–frequency information that is contained in electromagnetic signals measured from healthy, functioning human brains. Two main approaches exist for gathering new information on the operation of the human brain: one, developing and employing novel analysis methods; two, posing novel questions, typically in the form of new measurement paradigms. Both approaches were utilized in this thesis. The first approach involves typical engineering work, but there is no overarching theoretical framework in this thesis. The avoidance of rigid theoretical frameworks is intentional but also due to the signal being generated by a system of tremendous complexity. Therefore, modeling was not attempted, although it could be the road to more profound understanding. Instead, the main challenge was obtaining an accurate description of the neural signals buried in noise. The used data transforms provide informative descriptions, but these descriptions are ambiguous: Several explanations can account for any observed phenomenon. The underlying signals are, however, not abstract entities; with a combination or a battery of analysis methods, the actual constituents of the signals may be revealed, and in so doing, answers to some fundamental questions of this research field may be found.

A few matters of form and composition: Regarding statistics, let us consider an example of testing whether lottery wins differ significantly from zero (i.e. not having won at all). Taking ten consecutive jackpots (period 25.9.2004 – 14.1.2005) of the Finnish national lottery (seven out of seven correct, mean prize 1342000 €) and the corresponding lowest scores that yield money (four out of seven correct, mean prize 11.2 €), the significances are $p = 0.03$ and $p = 3.0*10^{-12}$ for the seven correct and four correct, respectively (two-tailed t-test). The jackpot is therefore just barely statistically significant, whereas the four correct is overwhelmingly significant. The reason for this is that the relative variance of the jackpot is large and that of the four correct is small. In science, nevertheless, statistical significance is often the only recognized measure of an effect and used without further consideration. Then again, most scientists would probably agree that only by understanding the properties of the data can the quantities derived from the data be correctly interpreted. Thus, according to a common practice, statistical values, as well as several other details, are left to the publications, and the emphasis of this thesis is instead placed on an attempt to provide prerequisites for intuition.

This thesis is written in American English. This preface also illustrates another aspect of the composition of this thesis: The text is structured using bolded outline titles. This somewhat uncommon layout is used to avoid the otherwise very numerous subsections. In addition, in this scheme it is natural to use the same outline titles, which, I hope, helps the reader to notice that the same aspects of different data are considered. In general, the thesis composition follows scientific convention: The concepts of the thesis title and their relevance to current work are systematically described in the introduction, which is followed by methods, results, and discussion. However, in order not to disconnect the description of questions from the answers, the methods and results of each study follow in succession. This manner of presentation indicates that the studies are not totally disjunct, but rather that each was performed to answer questions raised in the preceding thesis work.

1. INTRODUCTION

1.1 Signals

Some signal basics: Of the many definitions of signal, "a detectable physical quantity or impulse (as a voltage, current, or magnetic field strength) by which messages or information can be transmitted"[1] captures several points relevant for present examination. This thesis is concerned with time-series signals that reflect the operation of the human brain, but the used methods, *per se*, are general and suitable, for instance, for interrogating signals that are functions of some variable other than time.

The magnitude of most physical signal sources changes continuously, but all signals of the current work were transformed to a discrete digital format prior to analysis. The digital format allows efficient and precise manipulation, storage, and transmission of signals; it has been essential for the development of current information and communication technology (e.g. Smith, 2001). Values of digital signals are usually defined only at discrete intervals Δt (the inverse of sampling rate); the signals in this thesis can be presented in the form

$$x(n), \quad n = 1,2,3,\dots N, \tag{1}$$

where both n and $x(n)$ have integer values. The assumption is that the range of values that $x(n)$ can have is both sufficiently broad and sufficiently dense to accurately describe the examined phenomenon. The maximum duration of the Δt that allows perfect reconstruction of the continuous signal is governed by the well-known Nyquist theorem (formulated by Nyquist in 1928 and formally proved by Shannon in 1949). It states that the sampling rate needs to be higher than twice the bandwidth of the sampled signal. With the frequency band of interest usually containing the low frequency end, the measured signals need to be low-pass filtered prior to sampling in order to avoid aliasing of frequency components above half the sampling rate to the intended frequency band (e.g. Lyons, 2001). In this thesis work, the examined frequencies were at least an order of magnitude below the sampling rate to avoid practical problems such as the requirement to use a reconstruction filter[2] for obtaining an accurate shape of the original signal.

We further restrict our focus to one-dimensional signals. The thesis measurements, however, were performed using multiple separate sensors simultaneously, which yielded multidimensional signals. The methods examined and developed in this thesis are one-dimensional in the sense that they do not specifically exploit simultaneous information of signal properties from multiple sensors. The methods that do exploit simultaneous information include, for example, principal and independent component analysis, and signal source localization with respect to some *a priori* knowledge of the sensor configurations. These multidimensional methods have been actively developed specifically in the current brain research field (e.g. Baillet *et al.*, 2001; Darvas *et al.*, 2004; Vrba, 2002) and complement the methods that are examined in this thesis. In

---

[1] http://www.merriam-webster.com

[2] $$x(t) = \sum_{n=-\infty}^{\infty} \left(x(n) \frac{\sin(\pi(ft - n))}{\pi(ft - n)} \right),$$

where f is the sampling frequency. This equation is known as the Nyquist–Shannon interpolation formula.

accordance with the single dimension of the examined signals, all studies of this thesis deal with the auditory system, where the sensory input signal is essentially a one-dimensional time-series (for the use of binaural information by the human auditory system see e.g. Moore, 1995; Palomäki, 2005).

In measuring a physical process, the obtained values are partly attributable to the process of interest and partly to the measuring system, a variety of external sources, and possible subsidiary processes of the measured object. The first part is signal and the second part is noise. In practice, noise, not signal, is often the dominant component. Furthermore, it is typically the researcher's whim that decides what is noise and what is signal. There is also a terminological problem in that although the word signal is dedicated to the quantity of interest, it is also the word used for the entire measured quantity that contains the signal and the noise. An important related quantity is signal-to-noise ratio (SNR), which describes the amount of the desired signal and the amount of noise. Properties of the signal often determine how SNR is calculated, but a commonly used equation is

$$SNR = A / \sigma_\text{n}, \qquad\qquad [2]$$

where A is the amplitude of the signal, and σ_n is the standard deviation of noise. Improving SNR is often the main goal of signal processing and underlies much of this thesis work.

Information and structure in signals: Besides improving the SNR, we want to extract information from signals. According to information theory, the amount of information can be quantified as a reduction of uncertainty when one of the possible alternatives for an event occurs (Shannon, 1948, 1949). It is measured by the entropy function (H) defined by

$$H = -\sum_{k=1}^{K} P_k \log P_k \qquad\qquad [3]$$

for a source with K alternative outputs with probabilities P_k. This function is useful, for example, in describing how much we can compress data for storage: Space can be saved by encoding commonly occurring events with shorter code words than those that occur rarely (Huffman, 1952). Nevertheless, neither equation [3] nor those taking into account the correlations between signal values (joint probabilities; e.g. Gonzalez and Woods, 2002) allow us to evaluate the amount of relevant information in a signal or facilitate its identification. For example, consider signals obtained from two radios, one tuned and another not tuned to a transmission channel. The latter is theoretically the more informative signal: It is unpredictable (each sample is independent from the previous values), whereas there is a lot of predictability and uneven probability distributions in radio programs. In practice, however, both in radio and in brain recordings, the part of the signal that has some predictability, or structure, yields information; the rest is considered noise.

Signal structure (e.g. sinusoidal wave shapes) may be visible in one signal representation but not in another, and finding the descriptive representations is one of the tasks of this thesis. There are methods that find in someway optimal signal representation (e.g. based on minimization of entropy, Coifman and Wickerhauser, 1992), but in general, there is no explicit reason to assume that these would best reveal the signal structures of interest, and they typically do not work well with low SNR. One general step towards obtaining relevant information can, however, be identified: A representation of a brain signal can be cumbersome if not impossible to interpret as such, and therefore the task is to obtain a reference representation or a baseline level against which the data elements that convey information can be identified.

1.2 Electromagnetic brain signals

Signaling in the brain: The structural details of the human brain are known to a considerable degree (e.g. Kandel *et al.*, 2000), but this knowledge has provided only limited insight into the operation of the human brain; it is the signals in the brain, not the brain matter, that matter. Uncovering the principles of information processing in the brain can be approached from the same direction as the brain receives information, through the pathway of sensory systems. In all sensory systems, the information is encoded into electrical activity of some of the $10^{\sim11}$ neurons of the brain. The electrical signals are transferred between neurons using chemical transmitters (with the exception of gap junctions), but around 99.998% of the distance the signals travels in electrical form in the brain[3]. The importance of electrical activity for brain function is highlighted by non-invasive transcranial magnetic stimulation which can, for example, disturb cognitive functions (Rossi and Rossini, 2004), and direct invasive electrical stimulation can produce a variety of sensations including pleasure (Heath, 1972). The currently dominant view in neuroscience regarding short-time-scale information representation in the brain might tentatively be summarized in the following way: Information is manifested in the spatiotemporal pattern of the membrane potentials of neurons.

Measurable currents and neural activity: The cerebral cortex contains several topographically organized areas: Sensory stimuli produce neural activity that is constrained to specific patches of the cortex, and the physical stimulus features determine the location of the activity in an orderly fashion. Pyramidal neurons are abundant in the cortex and their apical dendrites have a preferentially perpendicular orientation to the surface of the cortex (e.g. Kandel *et al.*, 2000). These observations indicate that, for example, following sensory stimulation there are favorable conditions for local concentration of activity in the cortex that gives rise to a macroscopic net current.

The generation of the currents that can be non-invasively observed from the surface or outside of the head has been discussed for example by Baillet *et al.* (2001), Hämäläinen *et al.* (1993), and Schaul (1998). In brief, the postsynaptic potentials of neurons have dipolar current patterns that decay more slowly as function of distance than the quadrupolar patterns of the action potentials. The postsynaptic currents also last an order of magnitude longer than the action potentials. Hence, it is the temporally overlapping currents in the apical dendrites of the pyramidal cells that apparently give rise to a primary current, which underlies the non-invasively measurable electrical and magnetic signals (for neural details of the process see Murakami *et al.*, 2002, 2003).

The exact neural mechanisms underlying the measured signals are important in, for example, interpreting the effects of pharmacological agents on brain processes (e.g. Kähkönen *et al.*, 2005). However, it appears that in many cases one can assume that the signal obtained from a location reflects approximately linearly the level of activity (including both pre- and postsynaptic parts) in the underlying neural circuitry. This is because the activity in a cortical patch is mostly recurrent (Douglas *et al.*, 1995) and driven overwhelmingly by surrounding neurons (Shadlen and Movshon, 1999), whereby the post- and presynaptic activity are temporally overlapping in the patch. In addition, the relationship of the primary current $\vec{J}^{P}$ both to the potential difference and to the magnetic field is linear and described by

$$V_i = \int \mathcal{L}_i^E(\vec{r}) \cdot \vec{J}^P(\vec{r}) dv \qquad [4]$$

---

[3] This crude estimate is based on the approximate length of the pyramidal cells in the cerebral cortex (1 mm) and the width of the synaptic cleft (20 nm).

and

$$b_i = \int \vec{\mathcal{L}}_i^M(\vec{r}) \cdot \vec{J}^P(\vec{r}) dv, \qquad\qquad [5]$$

respectively (Hämäläinen *et al*., 1993), where $\mathcal{L}_i^E$ and $\mathcal{L}_i^M$ are the electric and magnetic lead fields that describe the sensitivity of the measurement sensor to a current source within a volume (Malmivuo and Plonsey, 1995).

Electro- and magnetoencephalography: In electroencephalography (EEG; for general reference see Niedermeyer and Lopes da Silva, 2004), the potential differences on the surface of the skull are measured. These typically have a range from a few to a few hundred microvolts (μV). The basic scheme of EEG is electrodes attached to the scalp, an amplifier stage, a data preprocessing system, and a data storage device (for details on instrumentation see e.g. Geddes and Baker, 1989; Virtanen, 1998). Several aspects of EEG systems determine the type of information that can be obtained. For example, the relevant frequency range of EEG can be from 0 to 1000 Hz, but is limited by the electrodes, amplifier, and used sampling rate. With its long history (first recording reported by Berger in 1929) and relatively simple instrumentation, EEG has an established role in clinical neuroscience (e.g. Wallace *et al*., 2001; Nuwer, 1998), but much remains unknown about the relationship of EEG signals and cognitive functions as well as about the exact mechanisms that generate the EEG signals.

The first measurements of the magnetic fields produced by the human brain were made with room temperature coils by Cohen in 1968, but the advent of magnetoencephalography (MEG, for general reference see Hämäläinen *et al*., 1993) was facilitated by the development of reliable superconducting quantum interference devices (SQUIDs, for review see Ryhänen *et al*., 1989). The SQUIDs currently used in MEG require a dewar construction, where cooling is provided by liquid helium. MEG devices are further placed in magnetically shielded rooms in order to attenuate the external fields that are typically several orders of magnitude greater than the neural signals, whose magnitudes are in the femto- to picotesla range. A separate flux transformer, referred to as pickup coil, provides effective coupling of the SQUID and the external magnetic field. The shape of the pickup coil determines the lead field, the profile of the source sensitivity. For example, planar gradiometers have focal lead fields and further have the practical advantage of showing maximal signal directly above a current dipole (Knuutila *et al*., 1993).

MEG and EEG signals are alike, for example, in reflecting the same primary current sources and in providing a millisecond time resolution. However, the interest is generally in the differences rather than in the similarities of these two methods. EEG provides portability and allows for continuous monitoring, whereas MEG requires less measurement preparation. MEG devices and their supporting electronics and facilities, then again, are much more complex and expensive than those for EEG. The development of MEG has mainly been driven by the desire to localize the sources of the neural signals: The magnetic permeability of the head is approximately equal to that of free space and hence has negligible effect on the magnetic fields, whereas the electric potentials measured with EEG are determined (distorted) by the inhomogeneous conductivity distribution of the head. In practice, EEG rarely matches the source localization capability of MEG.

Even the shortest texts comparing MEG and EEG typically never fail to mention that MEG sees tangential but is blind to radial sources, whereas EEG is sensitive to both. It is thus often assumed that MEG sees only around half of the cortical sources, which lie in the sulci. However, as briefly discussed by Mäkinen (2002) and considered in detail by Hillebrand and Barnes (2002), pronounced cortical activations that would be invisible with MEG are unlikely. This follows from the spatial span of the activation, the close proximity of the gyri to the MEG sensors, and because only around 5% of the cortex is closely radial. Hillebrand and Barnes concluded that source depth instead of source orientation is the critical factor in determining the source visibility in MEG. Nevertheless, even brainstem responses have been reported with MEG

(Parkkonen and Mäkelä, 2002). It may also be noted that while EEG is sensitive to tangential current sources, its sensitivity to them is lower than to radial sources.

Both MEG and EEG have a shortcoming regarding source localization. An infinite number of current configurations within a volume can give rise to the exact same electromagnetic pattern on the surface (shown already in 1853 by Helmholtz). Hence, the spatial resolution of MEG and EEG is dependent on assumptions that can be made of the current source and, subsequently, of the method chosen to tackle this inverse problem. The benchmark inverse solution is the equivalent current dipole (ECD, see e.g. Hämäläinen *et al.*, 1993). It is highly popular in MEG research because of the simplicity of the information that it provides: A neural source is accounted for with one or few ECDs, which again are completely determined by location, orientation, and strength. Other inverse solutions (e.g. Baillet *et al.*, 2001; Darvas *et al.*, 2004; Uutela *et al.*, 1999) tend to provide continuous current distributions of the whole volume of the head, which may be difficult to quantify.

When using MEG and EEG, one might do well in not focusing on the spatial information, as positron emission tomography and functional magnetic resonance imaging tend to be superior in this respect (e.g. Budinger, 1998; Turner *et al.*, 1998). Instead, MEG and EEG provide time-resolution beyond these methods, and the signal directly reflects the electrical activity of the neurons providing the possibility to draw causal relationships between the neural events and cognition. However, it may be noted that a poor spatial resolution of a method can also limit the time resolution: When multiple processes have overlapping (and jittering) time courses, the spatial summation may act as a low-pass filter in the time-domain. In this respect, the advantage of MEG and especially that of the planar gradiometers compared with EEG is the possibility to analyze focally generated signals, which is particularly important for the analysis of raw, low-SNR signals such as those examined at length in this thesis.

1.3 Time information in brain signals

Averaging: Deflections of the measured signal, produced by neural activity, with an approximately fixed time course to identifiable events (e.g. sensory stimuli) are termed event-related fields (ERFs) in MEG and event-related potentials (ERPs) in EEG. For convenience, both are referred in this thesis as event-related responses (ERRs) when no distinction between ERFs and ERPs is made. The term ERR is reserved to those event-related processes that can be observed with event time-locked (i.e. time-domain) averaging of the measured signal. That is, the neural responses need to be both approximately time- and phase-locked to the events in order to survive the averaging processes.

The main reason for employing the event time-locked averaging is the large extent of the lead field of an MEG or EEG sensor. That is, besides picking up the activity from the event-activated patch of the cortex, the sensor picks up other activity such as spontaneous activity of a large amplitude generated over large brain areas (Schaul, 1998). This background activity dwarfs the event-related activity. According to the traditional model (e.g. Hämäläinen *et al.*, 1993) the measurement is of the form *measured signal = ERR + uncorrelated noise*. In averaging, the uncorrelated background activity is attenuated so that the SNR gain G increases in the square root of the number of trials N_t (for the derivation of this results see e.g. Mäkinen, 2002), which in dB is

$$G \text{ (dB)} = 20*\log_{10}(N_t^{0.5}) = 10*\log_{10}(N_t). \qquad [6]$$

However, in contrast to this traditional model, there are an increasing number of studies (e.g. Başar, 1980; Gruber *et al.*, 2005; Hamada 2005; Jansen *et al.*, 2003; Karakaş *et al.*, 2000; Klimesch *et al.*, 2004; Makeig *et al.*, 2002; Sayers *et al.*, 1974) suggesting that ERRs are not additive

components to the background activity, but rather that ERRs result from stimulus-induced phase reorganization of the background activity. Determining which is accurate—this new hypothesis or the traditional view—is one of the goals of this thesis.

Regardless of the generation mechanism of the ERRs, it is an unrealistic assumption that the human brain produces completely invariant responses to a repeated stimulus. Consequently, several different approaches have been adopted to enable the examination of single trials in order to reveal the information lost in averaging (e.g. Lugger *et al.*, 1998; Parra *et al.*, 2002; Quian Quiroga and Garcia, 2003; Tang *et al.*, 2002; Woody, 1967), and this issue is also considered in this thesis. Another issue of averaging is that the typical number of trials used and needed in ERR studies ranges from a few dozen to a few hundred. The recording of 100 trials typically takes a few minutes and, according to [6], yields a 20 dB increase in SNR when compared to raw signal. However, if the reliable detection of a different ERR requires another 20 dB increase in the SNR, then presentation of 10000 trials and recording of few hundred minutes is required. This is already beyond what is usually considered feasible. Thus, MEG and EEG measurements are typically limited to examine ERRs within a rather narrow SNR range.

The auditory ERRs: A variety of ERRs have been reported from various sensory modalities under various conditions (e.g. Regan, 1989) and also within the auditory modality (e.g. Näätänen, 1992). In this thesis, the focus is on the robust, cortically generated ERRs that can be observed with very simple measure settings. When discrete sound stimuli are separated by around 1 second, the stimulus time-locked averaging of the MEG signal obtained above auditory brain areas yields a waveform similar to that shown in Figure 1. It has three clear deflections: the first downward deflection is the P1m (or P50m, where *m* stands for MEG response), the next deflection is the N1m (N100m), and the last is the P2m (P200m, previously considered an intrinsic part of the N1(m) but now with an increasingly independent

Fig. 1: Grand-averaged ERR to sine-tone stimulation obtained with MEG over temporal area of the right hemisphere. The black curve is low-pass filtered at 40 Hz and the grey curve is unfiltered.

status; Crowley and Colrain, 2004). All of these deflections are believed to be generated mainly in the cortex; Anurova (2005) and Renvall (2003) have recently reviewed the information of the pathways from the cochlea to the cortical sites generating the ERRs.

The N1(m) is the most conspicuous and the most examined of the auditory ERRs. Its main generators lie in the supratemporal auditory cortex (e.g. Godey *et al.*, 2001; Lütkenhöner and Steinsträter, 1998). Its behavior as a function of various stimulus parameters has been extensively examined (Näätänen and Picton, 1987; Roberts *et al.*, 2000), and sophisticated stimulus manipulations have shown its specificity to speech (e.g. Mäkelä *et al.*, 2002, 2004) and direction of sound source (Palomäki, 2005). Moreover, neural models of the N1(m) have begun to emerge, which account for its behavior as a function of stimulus parameters (May, 1999; May and Tiitinen, 2001A). Nevertheless, the relationship between the N1(m) and sound perception remains unclear.

Evidence supporting the linkage between perception and the N1(m) has been provided, for example, by studies of Parasuraman *et al.* (1980, 1982), where behaviorally detected noise-

masked sine tones elicited N1 responses, whereas no responses were observed with undetected tones. Sanders *et al.*, 2002 showed that a behaviorally measured capability to segment word onsets from a continuous sound stream was positively correlated with the amplitude of the N1. In light of recent results (May, 1999; May and Tiitinen, 2004) suggesting that the mismatch negativity (MMN) response[4] is the result of amplitude and latency modulation of the N1(m), the evidence indicating that the MMN indexes perceptual sound discrimination (Sams *et al.*, 1985; Tiitinen *et al.*, 1994) can also be seen to support the linkage between the N1(m) and perception. However, conflicting evidence has also been reported. For example, the latencies of the N1 and behavioral reaction times to sound detection have been observed to display different slopes as a function of sound intensity (Jaskowski *et al.*, 1994). Eddins and Peterson (1999) observed that as sound intensity level decreases, the N1 responses vanish at a level where the sounds could still be behaviorally detected. Also, the amplitude behavior of the N1 does not appear to reflect the perceptual loudness of the stimuli (Näätänen and Picton, 1987). The mixed nature of the evidence, however, does not show that the N1(m) is somehow intangible, but rather implies that our capability to observe the effects of stimulation is not matched by our knowledge of the methodological issues (e.g. effects of averaging, source localization, filtering) and of the cognitive processes involved. Thus, it may be possible to find isolated conditions where an accurate relationship between perception and N1m-like ERRs may be established.

1.4 Frequency information in brain signals

About frequency, Fourier transform, and stationarity: Frequency is defined as the number of occurrences of a pattern in a unit time. For a process to be associated with a certain frequency, it should display approximately the same period over several cycles. That is, the rationale underlying the concept of frequency is that if a process displays the same repeating pattern, then describing the period of pattern is relevant, whereas time information is not very useful, as the process merely repeats the same pattern. In signal processing, the convention is to define frequency via sine and cosine functions that are the basis of Fourier transform. For a discrete signal $x(n)$ whose length is N, the discrete Fourier transform is written as:

$$X(m) = \sum_{n=0}^{N-1} x(n)\big[\cos(2\pi nm / N) - i\sin(2\pi nm / N)\big] = \sum_{n=0}^{N-1} x(n)e^{-i2\pi nm / N} \ . \qquad [7]$$

The Fourier transform provides a representation, or decomposition, of a signal in frequencies (*m*) that are sinusoidal functions of a certain phase (for details on Fourier transform see e.g. Bracewell, 1999). A signal with a constant period but non-sinusoidal shape is not represented by a single Fourier frequency (*m*), yet Fourier transform is extremely useful and well justified: The largest magnitude Fourier component will ordinarily be the one that has the period of the signal, the original signal shape will be recovered with the inverse transform of the Fourier components, and the sinusoid has several favorable properties as a basis function[5].

A complication of the Fourier basis functions is that in order for a signal to be precisely of a single frequency, it should be a stationary sinusoid whose span is from negative infinity to infinity. More generally, a process is held stationary if its probability distributions are time-invariant (e.g. Hayes, 1996). In several studies, it has been estimated that relatively short segments (e.g. ~1 s) of MEG or EEG data can be considered stationary (Angelidou *et al.*, 1992; Jansen *et*

---

[4] The MMN response is obtained by subtracting the ERR elicited by frequently occurring stimulus from the ERR elicited by an infrequent stimulus. It has peak latency in the range of 100–250 ms (Näätänen, 1992).

[5] Sine and cosine functions are, for example, smooth with smooth derivatives, depict the location of constant velocity object in one coordinate as it progresses along a circular path, and describe the behaviour of many simple physical objects (e.g. pendulum).

al., 1981; Tseng *et al.*, 1995). Nevertheless, with the measured finite length signals, one can only obtain an estimate of the stationarity of the signal, and it appears obvious that sensory events and cognitive processes will continuously produce unpredictable changes into the measured signal (see also Popivanov and Mineva, 1999). Hence, in brain research as well as in applied signal processing, stationarity is usually regarded as a relative concept, often emphasized by referring to the level of stationarity. In this thesis, the level of stationarity of an oscillatory process is defined to be inversely proportional to the magnitude and rate of changes in the properties (e.g. frequency and amplitude) of the oscillation.

Estimation of spectral power: The power spectrum describes signals in terms of power per unit of frequency and, unlike Fourier transform, it does not contain phase information. The power spectrum of a stochastic process is defined as the Fourier transform of its autocorrelation function (e.g. Hayes, 1996). As Fourier transform provides a frequency representation of a signal, a straightforward way to obtain a power spectrum is to calculate the magnitude-squared Fourier transform of the signal. This calculation method is referred to as the periodogram, which, however, provides only an estimate of the power spectrum, because it is based only on one realization of the process and in using only a segment of the data makes the implicit assumption that the signal is zero beyond the segment. To elaborate, the squared Fourier transform provides the frequency representation of a signal within a time window, whereas the goal of spectral estimation is to give a frequency description of the actual stochastic processes that generate the signal.

The variance of the periodogram estimate does not converge to zero with increasing length of the signal, because the number of estimated frequency points (bins, *m* in [7]) increases in unison with the increase in the number of samples (*N*). The length of the signal, however, determines the frequency resolution of the spectral estimation, which increases with increasing data segment length: The periodogram works like a bank of *N* band-pass filters with the filter shape being a sinc function[6], where the width of the main lobe is inversely related to the length of the data segment. Increasing the frequency resolution of spectral estimation, by increasing the length of the estimation window, may however not yield more accurate frequency information of brain processes. This is because the processes of the human brain are continuously evolving, and are unlikely to be confined to narrow frequency ranges. Using relatively short data windows in spectral estimation, then again, yields a high number of estimates, which provides for a smooth, high SNR average spectrum. The averaging of power spectra is referred to as incoherent averaging (Lyons, 2001), where the gain in SNR (G_{incoh}) is described by

$$G_{incoh}(\text{dB}) = 10 * \log_{10}(N_t^{0.5}).$$ [8]

Incoherent averaging increases the SNR more slowly as a function of the number of trials N_t than the time-domain averaging. A task of this thesis is to develop a technique for obtaining an optimal data window length for spectral estimation of an oscillatory process.

A wide range of methods has been developed to provide improvements to the periodogram in the task of spectral estimation. The improvements are obtained as a trade-off between different properties of the spectrum, or require that the processes generating the signal can be modeled. Windowing (e.g. Hamming, Hanning) the data prior to calculation of the Fourier transform effectively modifies the filter shape of the periodogram in a manner that reduces the side lobes but widens the main lobe. This results in a smoother estimate, but the frequency resolution in the estimate is reduced (Harris, 1978; Nuttal, 1981). Another technique is to calculate the periodograms utilizing overlapping segments of data and subsequently average the periodogram estimates (Welch, 1967). Again, a smoother estimate is obtained at the expense of

---

[6] $\operatorname{sinc}(t) = \begin{cases} 1 & \text{for } t = 0 \\ \dfrac{\sin(t)}{t} & \text{otherwise} \end{cases}$

frequency resolution. Of the more recent model-free (nonparametric) methods, the multitaper approach is notable; it is suggested to provide an optimal trade-off between resolution and variance (Thomson, 1982; Percival and Walden, 1993; Mitra and Pesaran, 1999). In engineering, one can typically make assumptions about the processes generating the signal. Autoregressive (AR) and autoregressive moving average (ARMA) are the most popular of the model-based (parametric) spectral estimation methods (Makhoul, 1975; for other model based spectral estimation methods, see e.g. Marple, 1987). With ARMA, a relatively small number of coefficients (typically 5–50) are used to form a linear model, or predictor, of a signal. While it is apparent that processes generating MEG or EEG signals are beyond accurate description as an ARMA process, these methods have also been used widely in the analysis of MEG and EEG (e.g. Tseng *et al.*, 1995; Pardey *et al.*, 1996), although their use has also been criticized (Mitra and Pesaran, 1999). In general, however, the advances in spectral estimation do not solve the problems of spectral estimation in MEG and EEG work, where the problems arise from the inherent properties of the measured signal.

The frequency scene of MEG and EEG: In frequency-domain analysis of MEG and EEG, ongoing (endogenous) oscillations are the processes of interest. The ongoing oscillations, or brain rhythms, are categorized according to frequency (< 4, 4–8, 8–13, 13–25 and 25–80 Hz oscillations are referred to as delta, theta, alpha, beta, and gamma, respectively; for oscillations with higher frequency than gamma see e.g. Ikeda *et al.*, 2005) and also increasingly according to their source location. For example, the prominent alpha rhythm, which was described already in the first reports of EEG (Berger, 1929; Adrian and Matthews, 1934), now refers specifically to the rhythm generated in the posterior, visual brain areas. A rhythm of approximately 10 Hz observed in the auditory brain areas is considered separate from the alpha and referred to as the tau rhythm (Tiihonen *et al.*, 1991; Lehtelä *et al.*, 1997).

It is a general observation that the resting brain produces more oscillatory activity than the vigilant brain (e.g. Hari and Salmelin, 1997). A suggested role of ongoing oscillations is idling (Kuhlman, 1978), which would be a form of readiness state. Recently, there has been a strong trend to find and to emphasize the role of oscillations in neural information processing. This has reached such an extent that there now is "the tantalizing conjecture that perception, memory, and even consciousness could result from synchronized networks" (Buzsáki and Draguhn, 2004), referring specifically to synchronized oscillatory activity. In line with this trend is the suggestion that modulation of ongoing oscillations underlies the generation of ERRs. In this thesis, the role of the oscillations is examined more cautiously.

When one examines the power spectrum of an alert person, it is not the spectral peaks of the brain rhythms, but rather a generally decreasing power with increasing frequency, that is likely to be the dominant feature (Freeman *et al.*, 2000, and Fig. 2). The decrease in power with increasing frequency approximately follows the inverse of frequency and is referred to as 1/f type of spectrum or pink noise (white noise would have level spectrum). This type of spectrum is
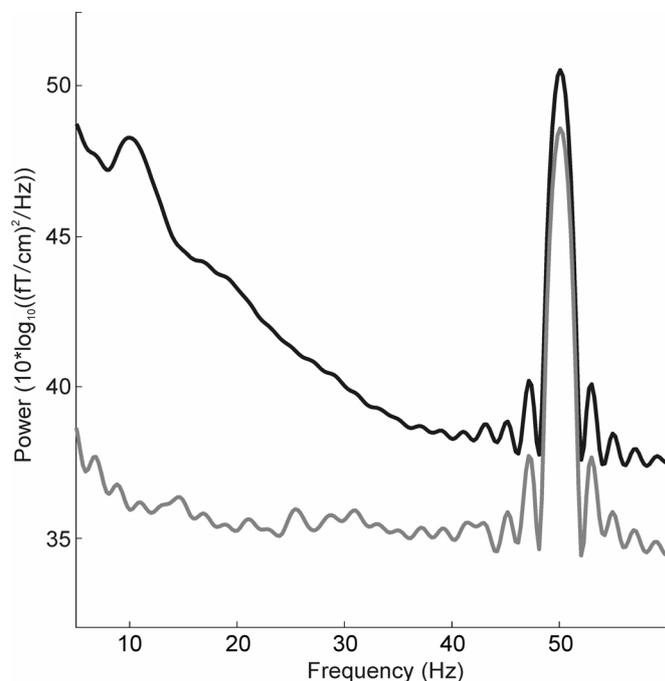
Fig. 2: Grand-averaged spectral estimate of MEG data from auditory brain areas (black curve) and an empty room spectrum (gray curve). The spectrum from subjects displays 1/f shape, a peak at 10 Hz, and both spectra display a prominent peak at 50 Hz produced by the mains power. The scale of power is equal to dB/Hz.

typical for a large variety of natural phenomena (e.g. Milotti, 2002).

Hence, the neural oscillations, when present, occur within a prominent noise profile, which has roughly $1/f$ shape but different noise characteristics can occur along the frequency axis. This noise profile presents several practical problems; for example, the statistical testing of the existence of a spectral peak is complicated because the absolute value at any frequency is likely to depend more on the underlying noise than on the oscillation peak. The noise profile, then again, varies in a manner that need not be correlated with the strength of the oscillation. The straightforward solution is to estimate the noise profile and subtract it (Buzsáki and Draguhn, 2004; Dumermuth and Molinari, 1987), but the shape of the profile can be complex and the neural oscillation can be spread along the noise slope. It is not trivial to determine what is noise and what is signal in the slope; this issue is engaged in this thesis.

1.5 Time–frequency information

The audible description of a signal: The aim of time–frequency analysis is to describe how frequency components of a signal, representing for example specific brain processes, behave over time. Besides signal processing, time–frequency analysis is also related to the sensory modality examined in this thesis. We do not hear time-evolution of sound pressure waves directly; instead, we hear the time-evolution of the magnitudes of the frequency components. This is a result of efficient time–frequency decomposition of the sound signal in the auditory system (see e.g. Lewicki, 2002). After the decomposition, information is integrated over time and frequency by the neural machinery in order to construct an auditory environment with identifiable objects (see e.g. Griffiths and Warren, 2004). In this thesis, time–frequency transforms are used to analyze and identify the auditory evoked and ongoing brain processes.

The problems of time–frequency estimation: To appreciate the task of estimating processes that make up a signal at any arbitrary time point (as characterized by frequency), let us begin by considering a simple one-component sine signal written in the form

$$x(n) = A \sin(2\pi f n + c), \qquad\qquad [9]$$

where A is the amplitude, f is the frequency and c sets the phase. This describes a stationary oscillation. Brain oscillations undergo amplitude modulations, and with all the brain areas receiving transient input from external and internal events, the phase of the oscillation cannot be expected to change linearly. Therefore, a one-component brain process is more appropriately written in the form

$$x(n) = A(n) \sin(G(n)), \qquad\qquad [10]$$

where $A(n)$ is the modulated amplitude and $G(n)$ is the function determining frequency and phase. Now the problem is that our measurements provide us with $x(n)$, and an infinite amount of pairs of $A(n)$ and $G(n)$ can generate the same $x(n)$. In addition, the actual measured signals contain noise and, without *a priori* knowledge, we have no way of determining what part of the signal is noise and what represents dynamic behavior of the examined phenomenon. Furthermore, the measured signals are often composed of multiple simultaneous components. Thus, time–frequency estimation can be considered a type of inverse problem (Cohen, 1995), unsolvable without *a priori* knowledge. A common approach in time–frequency analysis is to find a representation of a signal in terms of basis functions that may or may not accurately describe

the processes generating the signal. If the used basis functions are not descriptive (e.g. they assume too high a level of stationarity), the results can be misleading.

This inverse problem of time–frequency estimation is not often considered in MEG and EEG work, whereas the so-called time–frequency uncertainty is well known. This uncertainty principle occurs in several fields of mathematics (Selig, 2002) and is only distantly related to the Heisenberg uncertainty principle. In fact, in signal analysis, the uncertainty principle has been considered a misnomer (Cohen, 1995), as it only means that the more accurately we wish to localize the signal in time (frequency), the broader its frequency (time) range must be. The uncertainty is in the broadness of either time or frequency, and it is typically represented as

$$\sigma_{time}\sigma_{frquency} \geq \frac{1}{4\pi} \Leftrightarrow \sigma t \sigma \omega \geq \frac{1}{2},$$ [11]

where σ is the standard deviation (to one side, i.e. one-tailed) around the center frequency or time. The uncertainty is a signal property and not solvable with methodological advances.

The methods of time–frequency estimation: The most widely used of the current time–frequency methods in applied signal analysis (for mathematical approaches see e.g. Gröchenig, 2001) are the short-time Fourier transform, continuous and discrete wavelet transform including wavelet packets and matching pursuit, family of quadratic distributions developed from Wigner–Ville distributions, and time-varying (parametric) ARMA models (see e.g. Addison, 2002; Cohen, 1995; Durka, 2003; Mainardi *et al.*, 2002; Mallat, 1998; Thakor and Tong, 2004; and also Gabor, 1946). These methods have different advantages, and the properties of the examined signal and the goal of the examination determine the method best suited for the analysis.

In this thesis, the method of choice is continuous wavelet transform (CWT), defined by

$$CWT_n(s,b) = \frac{1}{\sqrt{s}}\sum_{n=0}^{N-1}x(n)\psi*\left(\frac{n-b}{s}\right).$$ [12]

ψ is the wavelet (a function localized in time and frequency with finite energy and zero mean), s is the scale of the wavelet (the center pass-band frequency) and b is the time point where the transform is calculated. The calculation of [12] is equal to an inner product and quantifies the similarity between wavelet and signal portion with which the wavelet has overlap. By employing a range of scale parameters s, the wavelet is dilated (and contracted) so that it corresponds to different frequencies, and by further calculating over points b, one obtains a time–frequency distribution of the data (see also Addison, 2002; Mallat, 1998; Torrence and Compo, 1998).

CWT is a redundant transform and not equally suited for computer applications (e.g. data compression) as discrete wavelet transform (DWT); however, for the type of data analysis performed in this thesis, CWT offers at least the following advantages:

- CWT has no requirements for the data from which it is calculated and does not require *a priori* knowledge of the signal (although such knowledge can be exploited in CWT analysis). Some methods such as time-varying ARMA models tend to require that the data properties change smoothly, preferably with a constant rate of change, and may not cope well with transient events such as ERRs. CWT does not suffer from the cross-terms (i.e. spurious, ghost components) that complicate the interpretation of Wigner–Ville distributions. Methods that find the best data description for each short data segment are best suited for analysis of high SNR data with few components, whereas CWT is suitable for analysis of data with multiple components and low SNR (typical properties of MEG and EEG data).
- CWT is a highly versatile and adaptable time–frequency transform. Wavelets are an open-ended group of functions facilitating, for instance, complex valued data descriptions as well as those with real part only; even a waveform that does not fulfill the expected

properties of wavelet can be used as a pseudo-wavelet (Qiu *et al.*, 1995). The trade-off between frequency and time resolution is nearly arbitrarily selectable even within the same wavelet family such as the Morlets (e.g. Addison, 2002). One can therefore take advantage of *a priori* knowledge of the signal structure by selecting a descriptive wavelet. The calculation of CWT is also very flexible; the scales s and time-points b can be selected arbitrarily, although if CWT is not calculated at every time point, it is, technically speaking, not a continuous wavelet transform.

- CWT has several favorable properties for highly detailed signal analysis. CWT is time-shift invariant, a necessity for accurately determining the latency of an event. DWT is not time-shift invariant (Durka, 2003) and has been recommended to be used in overcomplete form (Bradley and Wilson, 2004), thus making it almost equivalent to CWT. For determining the characteristic frequency of a process, CWT offers the possibility to use arbitrarily dense frequency spacing, a feature often underappreciated. The short-time Fourier transform has an inflexible, constant time–frequency trade-off that is usually suited only for analysis of a limited frequency range. Wavelets automatically stress high frequency resolution at low frequencies and high time-resolution at high frequencies, which corresponds to general distribution of relevant information in a signal. A point of considerable importance for the visualization and for the ease of interpretation is that CWT is based only on a single basis function.

- The computation of CWT is more laborious than that of DWT, but there is also an efficient algorithm for CWT. As the symbolic Fourier transforms of all the common wavelets are known, the arduous calculation of time-domain convolutions can be avoided by performing the calculations in the frequency-domain. First, a Fourier transform of the data is obtained (with fast Fourier transform, FFT). This is multiplied by the discretized Fourier transforms of the scaled wavelets, after which inverse Fourier transforms (with inverse FFT) are calculated from the products. This algorithm was used in this thesis.

On the time–frequency descriptions of MEG and EEG data: When is a time–frequency transform of a signal useful? For instance, for the averaged ERR of Figure 1, the time-domain description is the simplest and the easiest to interpret, and therefore in most cases the most informative. However, not all brain processes are phase- and time-locked to stimulation or suitable for examination with time-domain averaging. Time–frequency methods are well-suited to an examination of event-related, non-phase-locked changes in brain processes, where the SNR can be improved with averaging, utilizing absolute or squared values (Equation [8]). The prerequisites for effectively employing CWT in the analysis of event-related brain processes are considered in this thesis.

Event-related power changes have been under active investigation (e.g. Pfurtscheller and Lopes da Silva, 1999A; Schnitzler and Gross, 2005). The power reductions are termed event-related desynchronization (ERD) and power increases as event-related synchronization (ERS). The terminology alludes to the assumed underlying mechanisms, whereby ERD is the result of cortical activation and subsequent disruption of the ongoing rhythmic activity and ERS is typically regarded as the result of cortical deactivation and idling (Pfurtscheller and Lopes da Silva, 1999B). Both ERD and ERS may also easily result from methodological artifacts or misinterpretations (e.g. power increase of ERR thought of as ERS).

A CWT of raw, unaveraged MEG or EEG data theoretically describes the time-evolution of the processes in the signal, but in practice is likely to be of too low SNR to yield useful information. Thus, one needs a method of gathering information over time to increase the SNR. The statistical (i.e. average) behavior of a frequency band (a wavelet scale) can be acquired from CWT by first obtaining the envelope of a frequency band and then characterizing its properties, for instance, with spectral estimation (see also Linkenkaer-Hansen *et al.*, 2001, 2004). If the envelope behavior of a range of frequency bands could be characterized in a scale-free manner so that these characterizations were directly comparable, a compact two-dimensional data

representation would be obtained. By providing a statistical description of the temporal behavior of frequency components, this representation could be a powerful tool for identifying neural processes from noisy signals. Apparently, the development of such a data representation was a task of this thesis.

1.6 Objectives of the thesis

In the previous introductory sections, several elemental unanswered questions and possibilities for developing new methods were described. Below, they are encapsulated as six objectives for the thesis. Although they arise from MEG and EEG research, they also concern two more general topics: analysis of elusive signals buried in noise and information processing in the human brain. The connections between these objectives will, hopefully, become apparent in the following chapter and in Section 3.1.

Objective 1: To establish the stimulation conditions where ERRs describe and predict perceptual sound detection (Studies I and V).

Objective 2: To resolve whether ERRs are generated by phase reorganization of ongoing brain oscillations or by processes additive to ongoing oscillations (Studies II and III).

Objective 3: To obtain a conclusive mapping of auditory event-related processes in time, frequency, and spatial domains and to determine the prerequisites for the effective use of CWT in this task (Study IV).

Objective 4: To exploit the information on the time–frequency structure of auditory ERRs in order to determine whether the amplitude differences in averaged ERRs are due to changes in single-trial amplitudes or in their temporal alignment (Study V).

Objective 5: To develop a method that would remove the $1/f$ noise slope, characteristic for spectral estimates of MEG and EEG, thereby enabling the accentuation of the oscillatory processes from frequency data (Study VI).

Objective 6: To develop a compact, two-dimensional data representation where the envelope spectra between frequency components are directly comparable with one another (Study VI).

2. METHODS and RESULTS

Notes concerning all measurements: The measurements were performed in the BioMag Laboratory of the Helsinki University Central Hospital with the approval of the hospital's Ethical Committee. The subjects (nine in Studies II and V and ten in the other studies) were healthy human volunteers. In all recordings, a 306 sensor MEG device (Vectorview, Elekta Neuromag, Helsinki, Finland) was used. Auditory stimulation (frequency 750 Hz, duration 100 ms in Studies II and IV, 50 ms in Studies III and VI) was binaurally delivered with piezoelectric loudspeakers connected to the subjects via air tubes. A-weighted measurements of sound pressure level were employed to calibrate the stimulus output (80 dB in Studies II–IV and VI, empty room noise level was ~ 30 dB). When the subjects did not have a behavioral task, they either read a book (Study I) or watched silent films (Studies II–VI). The stimuli were presented at least 150 (Studies I and V), 400 (Study II), or 800 (Studies III, IV, and VI) times.

In Study I, the sampling rate was 468 Hz, whereas in the rest of the studies it was 601 Hz. An anti-alias low-pass filter was set to one-third of the sampling rate. Studies I and V had no high-pass filter, and the rest had a 0.03 Hz high-pass cut-off. In all studies, both raw data and online averages were collected. The source locations were quantified with ECDs employing sensor selections that covered each hemisphere and that consisted of both planar gradiometers and magnetometers. The analysis of raw data was performed in Matlab environment using data from the gradiometer sensors only. The direct relationship between the lead fields of the planar gradiometers and the source locations (Hämäläinen *et al.*, 1993; Knuutila *et al.*, 1993) provides for unambiguous data analysis. In addition, the magnetometer sensors rely on signal-space projection (SSP, Tesche *et al.*, 1995; Uusitalo and Ilmoniemi, 1997) to provide usable SNR; the calculation of SSP would have complicated the analysis.

In the analysis of raw data, band-pass limiting was performed with infinite impulse response (IIR) filters. IIR filters provide better frequency response with fewer computations than finite impulse response filters but have a non-linear phase response (e.g. Lyons, 2001). Therefore, two-way filtering was used: The signal is passed through the filter once, and then the filtered signal is reversed and run back through the filter again. This cancels the phase distortion of the first run and yields a zero phase distortion filtered signal. The Chebyshev type II filter was judged to provide the best trade-off between properties. It provides a flat pass-band, essential for the interpretation of the filtered signal, and a steeper roll-off from the pass-band to the stop-band (attenuation > 30 dB in the thesis studies) than, for example, Butterworth filters. Increasing the filter order improves the filter's performance, but the algorithm becomes unstable with too high a filter order resulting in, for example, pass-band ripple. The optimal filter order, or the highest order without pass-band ripple, was obtained through an iterative search.

2.1 Linking brain activity to perception via time information (Studies I and V)

Background (Study I): The relationship between ERRs and perceptual sound detection remains unestablished. One complication may be the theoretical approach of dividing ERRs to endogenous (non-obligatory) and exogenous responses (Donchin *et al.*, 1978). The brain response

reflecting perception should occur for all stimuli (i.e. be exogenous) that the subjects can reliably report perceiving. Yet, if perception is treated as a subjective, endogenous phenomenon, this theoretical framework appears problematic. Here, the research idea was to make a maximal distinction between perception and the acoustical structure of the stimuli (and the consequent obligatory ERR). To achieve this, sounds with intensity slowly increasing from inaudible to audible were used: The acoustic change is slow and smooth while the perception makes a quantum leap when the undetected sound becomes detectable.

The effect of sound slopes on auditory brain responses has been examined in numerous studies (for review see Loveless and Brunia, 1990; Phillips *et al.*, 2002) but it is difficult to draw a coherent picture from the obtained results. The sound properties (e.g. frequency) and stimulus presentation interval vary from study to study, the type of the sound slope (e.g. linear, logarithmically linear, or Hanning) and the onset or the offset intensity levels are often not reported. The noise level of the measurement environment and the level at which the subjects could detect the stimuli are almost never reported. ERRs are further often filtered with a high-pass filter, which can severely distort ERR shape with long sound slopes that are likely to elicit sustained brain activity (Picton *et al.*, 1978AB). In Study I, the sound slopes were longer and increased more slowly than those used in previous studies, but the difference between previous and current work is conceptual also: Previous studies have examined the effect of sound onset slopes on ERRs, whereas the main goal of Study I was to utilize the high time-resolution of MEG in relating the temporal evolution of brain processes to the timing of behaviorally measured sound detection.

Stimuli, measurements and data analysis (Study I): Sounds with durations of 1, 1.5, and 2 seconds were used. The intensity increased across the stimulus duration with a slope that is linear in the logarithmic decibel scale corresponding to perceptually linear slope (e.g. Moore, 1995). The intensity range (-10 to 60 dB) was set so that the midpoint corresponded with the average perceptual sound detection level (acquired with preliminary measurements using constant intensity tones). The sounds were presented in active and passive conditions with a random offset-to-onset interval of 0.5 to 4.5 seconds. The subjects' task in the active condition was to identify when the sounds became audible by pressing a response key.

With time-domain ERRs, determining the baseline is an unambiguous operation (but see also May and Tiitinen, 2001B) with a portion of the prestimulus signal (here -200 to 0 ms) providing the reference level. The analysis methods used here were those commonly applied in the analysis of ERRs, with the exception of selective averaging. That is, to explore whether the single-trial MEG responses would vary in a manner that is reflected in behavioral responses, the MEG responses were sorted according to reaction times, and those MEG trials corresponding to faster and slower half of each subject's reaction times were averaged separately. This is a straightforward application of temporal information from behavioral measures to signal analysis of MEG.

Results (Study I): It has been assumed that abrupt changes in sound energy are required for the elicitation of cortical ERRs (Clynes, 1969; Näätänen and Picton, 1987). Nevertheless, the smooth sound slopes elicited prominent ERRs with an abrupt onset (Fig. 3a), followed by sustained brain activity (tracking the intensity envelope of the stimulus) and an offset response. The abrupt onset was more pronounced in the active than in passive condition. The onset ERR was localized to approximately same areas that generate the N1m (whose location was obtained with 1000 Hz control stimuli).

The latency of the ERR accurately predicted the timing of the behavioral responses. The behavioral responses followed the onset peak with a constant delay of around 180 ms with all three sounds (correlation between the two time measures $r = 0.97$). Selective averaging showed that the onset ERRs preceding fast behavioral responses occurred earlier than those preceding the slow behavioral responses to the same stimuli. This latency difference between the fast and slow ERRs increased from around 50 to 100 ms as a function of decreasing sound slope and was found to be matched by increase in variation of the behavioral reaction times.

Thus, with the ERRs predicting both the behavioral reaction times as well as their trial-to-trial variation, Study I presents a strong case for the intimate linkage between sound perception and noninvasively measurable brain events. The reaction time variance can be explained with there being one constant variance component, possibly arising from sensorimotor areas (Hanes and Schall, 1996), and the rest of the variance arising from the operation of the auditory system. With longer slopes, there is a longer portion of the sound with intensity near the detection threshold, where it is susceptible to stochastic events in the nervous system as well as in the auditory environment. Thus, with longer slopes, there is more latency jitter in the brain events.

Finally, while the focus of this examination was specifically on the time information, the current data would have been poorly suited for frequency or time–frequency examinations. The ERR is a step-like entity (Fig. 3a) that is widespread in frequency. Yet, with the selective averaging, the ERRs were found to exhibit considerable latency jitter. This means that despite their step-like morphology, they fill a large portion of the time–frequency plane through being spread in both frequency and time, thus effectively concealing other event-related brain processes.
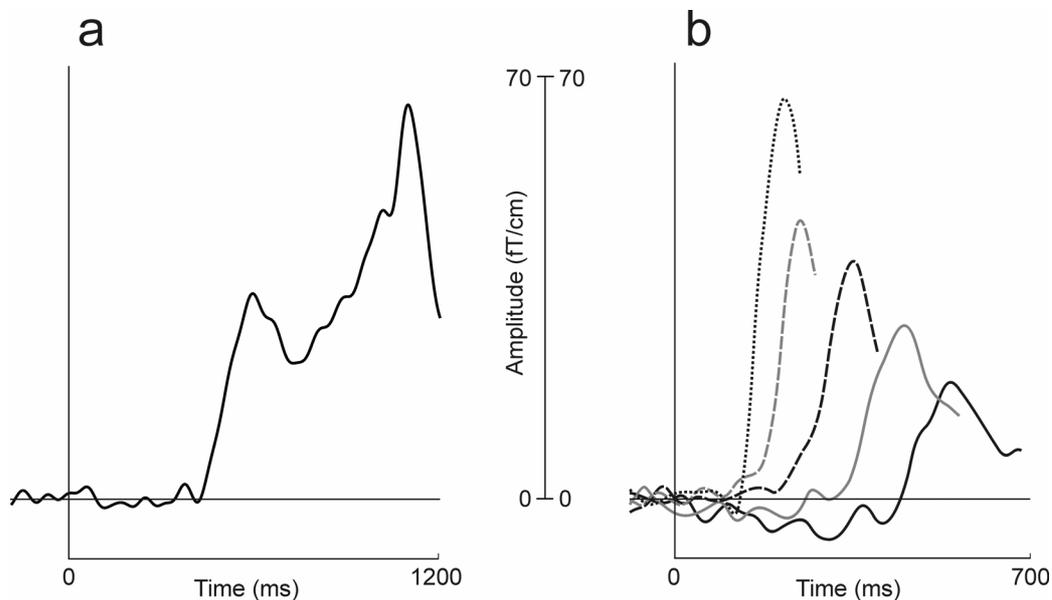
Fig 3: In (**a**) the grand-averaged ERR of Study I (1 s sound, active condition) displays an abrupt beginning followed by sustained activity and a transient offset response. In (**b**) the systematic behavior of the ERRs of Study V (active condition) as a function of sound slope is shown. The dotted black, dashed gray, dashed black, solid gray, and solid black display the ERRs to 125, 250, 500, 750 and 1000 ms sounds, respectively. Only the onsets of the ERRs are shown in order to make them more discernible (0–20 Hz pass-band).

Background (Study V): The results of Study I indicate that the onset of the ERR elicited by the sound slopes is an objective, non-invasive marker of perceptual sound detection. Thus, it holds a promise of non-invasive assessment of the functionality of the human auditory system. This would be specifically useful when examining hearing development in children and pre-speech infants, as well as when examining patients with compromised responsive capabilities. The first step towards these goals is the replication of the close correlation between the ERR latencies and behavioral responses. In addition, the optimization of this measurement paradigm may be initiated: The duration of the measurements should be minimized and the ERRs accentuated. A key parameter is sound duration; by using shorter stimuli, the measurements might be hastened and the response morphology clarified. Here, the shortest duration sound of Study I (1 s) was the

starting point with additional sound durations of 750, 500, 250, and 125 ms used. The offset-to-onset interval (0.5–2.5 s), as well as the intensity range (0 to 60 dB), was slightly reduced.

Results (Study V): The stimuli elicited prominent ERRs. As in Study I, the responses were of higher amplitude in the active condition, but unlike in Study I, the amplitudes decreased with increasing sound duration (Fig. 3b). One reason for this is that with the shortest sound durations the onset and offset response merged. Despite the changes in response morphology from Study I, the latencies of the ERRs were still highly correlated with behavioral reaction times ($r = 0.94$).

These results indicate that sound durations of around 500 ms and above are useful in providing a clear ERR peak while allowing for unambiguous separation of the onset and offset responses. The shortened stimulus presentation interval may, however, not be recommended as the amplitudes of the ERRs were reduced compared to Study I (with corresponding stimulus durations). The reduction of amplitude with decreasing stimulus interval is well documented for the N1(m) (e.g. Näätänen and Picton, 1987; but see also May and Tiitinen, 2001B).

In this study, the amplitudes of the averaged ERRs were found to depend on the task (attention) as well as on the sound slope. As observed in Study I, ERRs can exhibit considerable latency jitter and thus the amplitude differences in the averaged data can reflect amplitude differences in single trials but may also arise from changes in the accuracy of the alignment of the single-trial ERRs. Resolving this question requires examination of single-trial data. The analysis technique developed for this examination is derived from the lessons of Studies II–IV and is considered in Section 2.4.

2.2 Resolving the generation mechanism of event-related responses (Studies II and III)

Background: Ever since the advent of stimulus time-locked averaging (Dawson, 1951), it has been assumed that ERRs are the result of processes additive to the ongoing activity (*measured signal = ERR + uncorrelated noise*, referred here as the additive model of the generation of ERRs). However, an alternative hypothesis has emerged. Sensory stimuli could result in phase resetting of ongoing oscillations. The subsequent phase coherent state of oscillations would survive the averaging process and could thus underlie the ERRs. Here, this alternative is referred to as the organized oscillations model. It was first proposed by Sayers *et al.* in 1974, and since then several studies have supported this view (e.g. Başar, 1980; Gruber *et al.*, 2005; Hamada 2005; Jansen *et al.*, 2003; Karakaş *et al.*, 2000; Klimesch *et al.*, 2004; Makeig *et al.*, 2002).

If we consider this issue from the viewpoint of Study I, the organized oscillations model appears problematic. The ERRs of Study I are unipolar, and it is difficult to explain such a shape with organized oscillations model, whereas that shape would be the expected result of increased, additive activity. Moreover, in contrast to the results of Study I, the organized oscillations model appears to suggest that the ERRs do not reflect neural information processing. That is, if the information of sensory stimuli would be represented by phase-coherent states of ongoing oscillations, this entails the problem that such states also occur stochastically and without any stimuli or phase resetting. The information cannot be readily related to the actual phase resetting process either, because when the oscillations were already in the phase towards which they would otherwise be set, the brain processes would be uninfluenced by the stimulus. These points, along with certain apparent problems in previously used methods, served as the motivation to resolve this question. Studies II and III were designed to provide clear N1m responses for this examination, but simulations as well as theoretical considerations may be equally important for settling this issue.

In addition to methods used directly to resolve the generation mechanism of ERRs, a signal preprocessing method that combines data over sensors was introduced and employed in

Study II. The method, termed response source weighting (RSW), weights and sums raw data according to ERR amplitude distribution over sensors: Its operation is analogous to that of a neuron in an artificial neural network (see e.g. Haykin, 1999). RSW provided moderate improvements to SNR when compared to single-sensor data. Spatial filtering methods that are likely to be more effective are available (Hillebrand *et al.*, 2005).

Examining the presence of additive components: If the organized oscillations model underlies ERRs, then nothing is added to the signal, and the power of pre- and poststimulus unaveraged signals should not differ[7]. Such negative observations (Jansen *et al.*, 2003; Makeig *et al.*, 2002; Sayers *et al.*, 1974) can be explained, for example, with the following:

- If signal is not properly preprocessed (removing the mean of the signal and detrending) before applying spectral estimation methods, the resulting artifacts are likely to dwarf the possible small power changes produced by event-related brain processes.
- The length of the data segment used for examination strongly influences the results. For example, Sayers *et al.* (1974) used a 940 ms time window. As the auditory ERRs are mostly confined to a time range of 50–250 ms from stimulus onset (see Fig. 1), they cannot make a large contribution to the power of the whole data segment. Moreover, if a window function (e.g. Hamming, Hanning) is used with a long data segment, the activity occurring at the beginning or end of the segment will be strongly attenuated.
- When power is examined within relatively narrow frequency bands (obtained with band-pass filtering or with time–frequency methods), the band-limiting will inevitably spread the power of the possible additive components, and consequently, the prestimulus signal may not provide a valid baseline.
- The distribution of artifacts (e.g. eye blinks) may be different in pre- and poststimulus signals. For example, artifact rejection is often performed within the poststimulus time-window according to the expected duration of ERR. This can lead to artificially large prestimulus power with the poststimulus data being free of artifacts while the corresponding length of the prestimulus is not. To attain matched pairs, the artifact rejection should be performed with segments comprising both pre- and poststimulus data, after which the accepted segments can be divided into the two parts (the technique used in Studies II–IV).
- A different type of reason is that an event-related dampening of oscillatory activity could counter the power increases of additive activity.

In Study II, the estimation of spectral power on a 1–45 Hz band using a non-parametric multitaper spectral estimation method (Percival and Walden, 1993; Mitra and Pesaran, 1999) showed that the post-stimulus power was, on average, 20% greater than the prestimulus power (400 ms estimation windows). As the interest here is in the power of a broadband signal rather than in resolving the spectral components, the power can also be obtained in a straightforward manner by limiting the signal to the frequency band of interest and then calculating the magnitude-squared value of the samples (i.e. based directly on the definition of signal power). In Study III, this technique showed that on a 4–40 Hz band the poststimulus signal power was, on average, 27% greater than the prestimulus power (250 ms windows). When the averaged ERR of each subject was added to the corresponding unaveraged prestimulus signals, it was found to account for 90% of the power increases.

It has been observed that the trials with lowest poststimulus power are associated with low-amplitude ERRs (and high power with high amplitude). This has been taken to be inconsistent with the assumption that ERRs are the result of additive activity (Makeig *et al.*, 2002). However, when a waveform is added to a stochastic oscillation, the summation results in trials of high and low amplitude depending on whether the phase of the oscillation is in enforcing or dampening

---

[7] To be more precise, the phase towards which the phase resetting occurs determines whether there is a slight power increase or decrease; nevertheless, such an effect is minimal compared to the magnitude of the power increase if ERRs are the result of additive processes.

relation with the waveform. Hence, the result that low power trials correspond with low-amplitude ERRs and *vice versa* is not inconsistent with the additive model of ERRs but is the expected result in the lines of the well-established phenomenon of stochastic resonance (e.g. Moss *et al.*, 2004). In Study II, this expected result was obtained by categorizing the trials into three groups according to their power at poststimulus (several frequency bands used in the categorization) and then averaging the trials and observing the N1m amplitude. When the prestimulus power was used in the same sorting procedure, no effect on the N1m amplitude was observed. This appears inconsistent with the view that modulation of ongoing oscillations underlies the generation of ERRs.
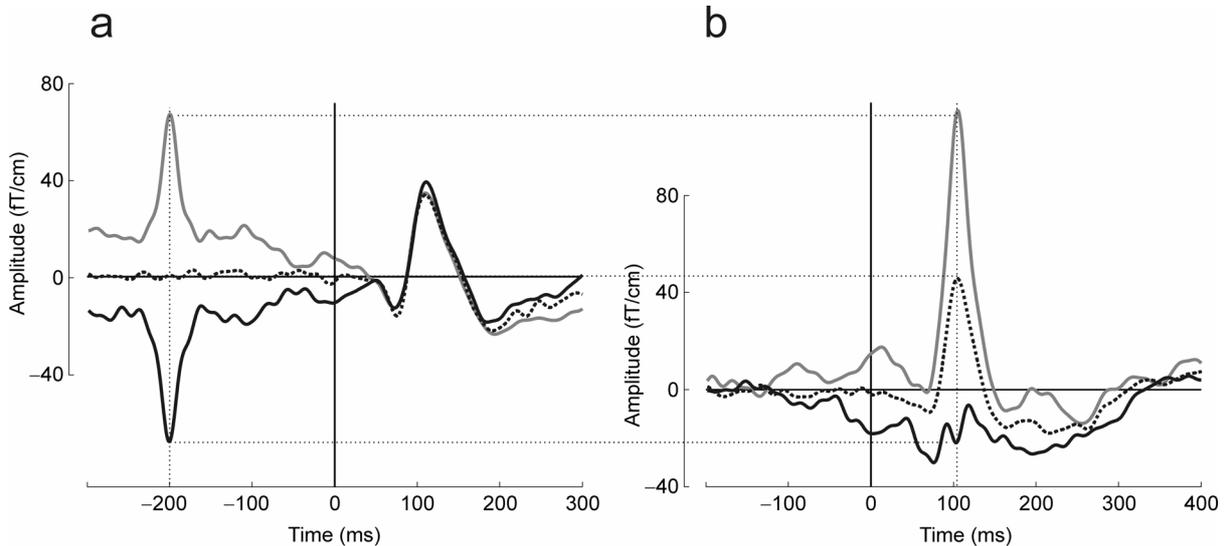
Fig. 4: Inter-trial amplitude variation used for examining the generation mechanism of ERRs. Single-trial, single sensor data (Study II, 1–45 Hz pass-band) was sorted into three groups according to the amplitude at 200 ms prestimulus (**a**) and at the peak latency of the N1m (**b**) and subsequently averaged (and grand-averaged). The variation of amplitudes is equal (the scales are equal) at the two examination points, which shows that ongoing activity is not in a coherent state at the latency of the N1m, but is biased by an additive constant-amplitude component. The variation is actually slightly increased at the latency of the N1m. The improved SNR provided by RSW was useful for examining this increase in more detail.

Examining phase organization: In some examined frequency bands, an inter-trial phase coherent state has been observed at the time of ERRs (Gruber *et al.*, 2005; Hamada 2005; Jansen *et al.*, 2003; Klimesch *et al.*, 2004; Makeig *et al.*, 2002). This has been interpreted as direct evidence that the neural oscillations are in an organized state at the latency of the ERRs (see also Penny *et al.*, 2002). However, an additive component that is phase-coherent over trials will produce an uneven phase distribution to all frequencies where it contains power (for elaboration, see simulations in Study II). Thus, it appears that with the signal decomposed to narrow bands, it is impossible to distinguish between phase modulation of ongoing activity and the presence of additive phase-coherent component. However, inter-trial phase coherence of ongoing oscillations, which results in a waveform in the averaged data, can equally appropriately be defined as inter-trial amplitude coherence; an organized state of ongoing oscillations is marked by reduction in inter-trial amplitude variance compared to baseline level. Adding a constant waveform to each trial does not effect the inter-trial amplitude variance. Thus, a necessary condition for the presence of inter-trial phase coherent state of ongoing oscillations is a reduction in the inter-trial amplitude variance of a broadband signal.

Here, the idea of the inter-trial amplitude variation estimation is illustrated in Figure 4, where the trials were sorted according to their amplitude into three groups and then averaged at an arbitrary time point (200 ms before stimulus onset) and at the peak latency of the N1m (difference between high and low amplitude signals shows the amplitude variation). The

variations are equal at the two time points; ongoing activity is not at an organized state at the latency of the N1m. The mean amplitudes at the two time points, however, are not equal and hence an approximately constant bias, an addition of the size of the N1m, has occurred at the single-trial level. In Studies II and III, the appropriate broadband examinations (1–45 Hz and 4–40 Hz, respectively) showed increases in the inter-trial amplitude variance at the latencies of the N1m, in contrast to the decreases expected from the organized oscillations model. These slight increases are explained by the trial-to-trial amplitude variance of the N1m.

2.3 The time–frequency composition of auditory event-related processes (Study IV)

Background: Studies II and III demonstrate that the auditory ERRs are not generated by phase resetting of ongoing oscillations. To attain a comprehensive view of auditory ERRs, their relationship to possible non-phase-locked event-related responses should also be established. This is a complicated task with commonly used moderate or high-frequency resolution time–frequency methods, which are sensitive not only to individual responses but also to complexes of responses, which makes it difficult to distinguish between sequential transient events and oscillatory processes.

Studies II and III highlight the independence of ERRs from the ongoing oscillations, but they do not imply that the ongoing oscillations are unaffected by sensory stimulation. In the somatosensory system, event-related power reductions (event-related desynchronization, ERD) are a well-established phenomenon (Pfurtscheller and Lopes da Silva, 1999A; Schnitzler and Gross, 2005). With auditory stimulation, the character of ERDs has remained unclear; power reductions have been observed, but they have been interpreted as reflecting some special circumstances or processes subsequent to the actual auditory processing (Bastiaansen *et al.*, 2001; Kaiser *et al.*, 2002; Krause, 1999; Sutoh *et al.*, 2000). Here, we sought to obtain a conclusive picture of the event-related processes in passive recording conditions with simple sine tone stimulation. If power reductions are observed, then they are an integral, unconditional part of the auditory event-related processes. Only crude source localization was performed (with a mapping of the observed processes to the MEG sensor grid employing planar gradiometers); the emphasis was on the time–frequency issues.

Details in the time–frequency plane: Two continuous wavelets are considered here. High time resolution was gained with the second derivative of a Gaussian (DOG2, Mexican hat) wavelet defined by

$$\psi_{DOG2}(t) = (1 - t^2)e^{-t^2/2}.$$ [13]

A Morlet wavelet with a wavenumber of six (Morlet-6, $\omega_0 = 6 \Leftrightarrow f_0 = 0.95$) provided a lower time resolution but a better frequency resolution. The Morlet is a complex sinusoid modulated by a Gaussian bell curve, defined by

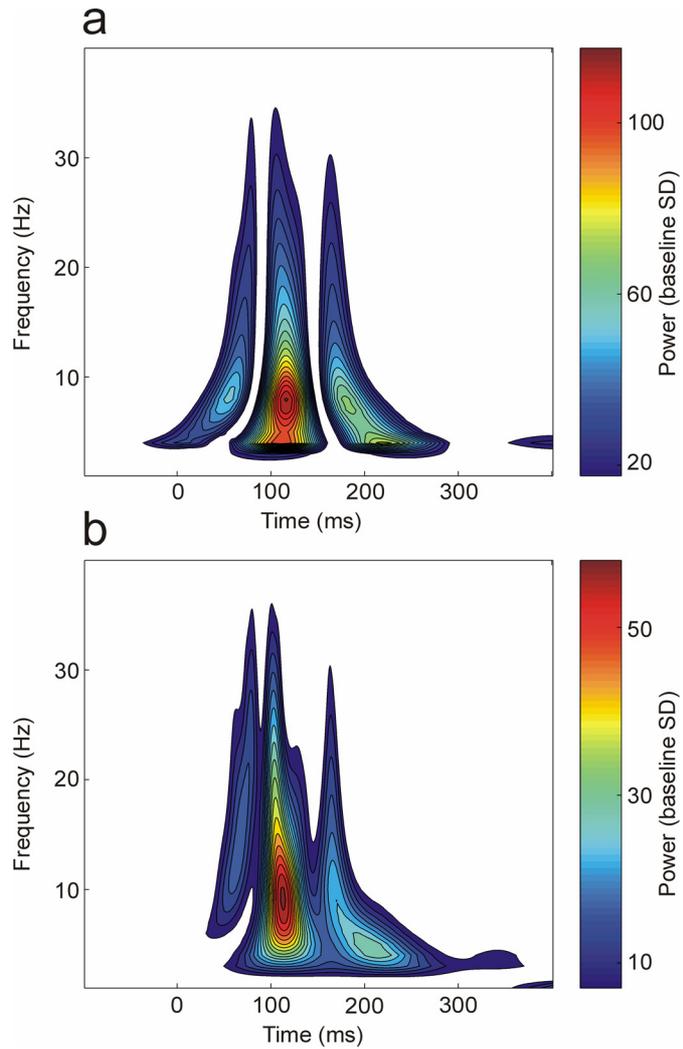$$\psi_{Morlet}(t) = \frac{1}{\pi^{1/4}} e^{i2\pi f_0 t} e^{-t^2/2}.$$ [14]

Both of these wavelets are just one instance of the family formed by a different order of derivates and different wavenumbers (for more details, see Addison, 2002). The Morlet wavelet and its

sibling Gabor wavelet, which consists of the real part only, are the most common continuous wavelets. The Morlet wavelet (with $f_0 \approx 1$) is usually considered to be of high time resolution, but a purpose of the present examination is to highlight the differences between CWTs obtained through Morlet-6 and those obtained through the truly high time resolution DOG2.

The time resolution (i.e. time spreading) of a wavelet can be demonstrated by calculating the CWT of an impulse response, which displays the time-domain shape of the wavelet. With 10 Hz DOG2 and Morlet wavelets, the impulse response is spread to around 100 and 200 ms on both sides of the latency of the impulse response. The time span of the wavelet gives rise to a computational problem at the edges of the data segment: The wavelet will continue to an area where there is no signal. This problem needs to be resolved somehow, for example by discarding data around the edges (see Addison, 2002 for other techniques).

In examining event-related processes, one is not typically interested in the general spectral content of the signal, but instead on what kind of power changes occur as a result of the used stimulation. The key point is that the processes can be identified within each used wavelet scale (the center pass-band frequency) with respect to a baseline level of that scale. Obtaining the baseline level entails several possibilities for artifacts that are often

Fig. 5: The DOG2 wavelet transforms of the averaged auditory ERRs (**a**, absolute values) and of the same but single trial data (**b**) are alike; the power increases visible in **b** are accounted for by the averaged ERRs in **a**. The time–frequency data is thresholded to depict only the activity that is clearly distinguished from noise. The data is from Study III.

not considered. Because of the time spreading of the CWTs, a poststimulus power increase (or decrease) is diffused in time so that it can influence the prestimulus signal. If the immediate prestimulus signal is used as a baseline, it can be on an elevated level. With this baseline, the post-stimulus power increases are reduced, and the real baseline level will appear as a power decrease (i.e. ERD). The impulse response simulation allows us to assess how far these effects would spread; the baseline needs to be obtained from a signal portion that is sufficiently separated from the stimulus onset. This separation distance increases as a function of decreasing frequency (increasing wavelength) and increasing frequency resolution of the used wavelet. With short stimulus intervals and low frequencies, it may be impossible to find an artifact-free baseline. Acquiring the baseline from a separate measurement without stimulation is also problematic, as the background activity cannot be assumed equal to that present during stimulation.

With a functional implementation, artifact-free baseline, and a sufficient number of epochs to yield a high SNR (Equation [8]), any deviation from the baseline level marks the event-related processes. With the time-series data mapped to two dimensions and with the high number of sensors in the current MEG and EEG devices, a practical problem of the analysis is the sheer

amount of data. One way to cope with this, employed in Study IV, is volume visualization with thresholding, which simultaneously reveals the time, frequency, and spatial distribution of power changes (increases and decreases viewed and thresholded separately).

Event-related power changes: Time–frequency transforms with both wavelets showed pronounced power increases in the post-stimulus signal. The DOG2 wavelet provides a time resolution high enough for simple comparison of time–frequency events and the time-domain averaged ERRs. The power increases coincided accurately with the time-domain P1m, N1m, and P2m responses. Furthermore, each power increase had a single maximum, as is expected from the Fourier representation of a monophasic waveform. Thus, all power increases of notable magnitude were accounted for by the phase-locked ERRs. Here, this result is highlighted in Figure 5 showing DOG2 transforms of time-domain averaged ERRs (Fig. 5a) and of the same but unaveraged data (Fig. 5b). The qualitative appearance of the transforms is the same, although the maxima of the transient peaks occur at higher frequencies in the transform obtained from the single-trial data. This is because the high frequency part of the transient peaks has a smaller time span than the low frequency part and is more easily cancelled in the phase-sensitive time-domain averaging processes, as some latency variation is always present.

Because the Morlet-wavelet comprises several deflections, it is sensitive to the P50m-N100m and N100m-P200m complexes rather than to single peaks. The interval between peaks provides the frequency where the CWT maxima were found. In addition to these maxima, harmonics were observed: When peaks of two non-sinusoidal ERRs coincide with deflections of a wavelet with certain frequency (scale), they will also coincide when the wavelet is three times that frequency. Further, the Morlet-wavelet spread the evoked activity to prestimulus side to an extent that hindered the analysis of low frequency data.

With both wavelets, small magnitude power decreases ($< 3\%$) were observed with the auditory stimulation. These were located to the same brain areas as the auditory evoked activity but also to the parietal areas. The decreases were most pronounced at around 17 Hz and occurred after the auditory ERRs. However, the actual onset times of the power reductions may be concealed by the power increases of the ERRs.

2.4 Exploiting *a priori* information of signal structure in single-trial analysis (Study V)

Background: In Study I, a pronounced latency jitter was observed in the active condition using selective averaging of single trials according to behavioral reaction times. In Studies I and V, the averaged ERRs were of higher amplitude in the active than in the passive condition. With there being no behavioral responses in the passive condition, selective averaging could not be used to determine whether the differences in the averaged ERRs were due to differences in the amplitudes of the single trial ERRs or with the apparent increase in attention concentrating the single-trial responses more accurately in the active condition. In Study V, it was also observed that the amplitude of the averaged ERRs decreased as a function of decreasing sound slope. Here, the goal was to obtain single-trial-level explanations to the amplitude effects of Study V.

Several methods based on different *a priori* assumptions have been developed for single-trial analysis. Methods that are based on the spatial distribution of ERRs include the use of PCA (Lugger *et al.*, 1998), ICA (Tang *et al.*, 2002), spatial filtering (Hillebrand *et al.*, 2005; and also RSW of Study II), and linear integration (Parra *et al.*, 2002).[8] Of the one-dimensional single-trial analysis techniques, template matching is the best known (Woody, 1967). The development of one-

---

[8] These multidimensional signal-analysis methods cannot typically be applied to clinical EEG measurements because they rely on large sensor arrays.

dimensional single-trial analysis methods, however, may have been hampered by the lack of knowledge of the generation mechanism of ERRs. For example, if the phase resetting of ongoing oscillations produced the ERRs, a good strategy would be to perform a signal decomposition on unaveraged data and then reconstruct the single-trial ERR from the coefficients that correlate best with the averaged ERR (Quian, Quiroga and Garcia, 2003). However, as the auditory ERRs were shown in Studies II and III to be additive to ongoing oscillations, a different strategy, presented below, may be more effective.
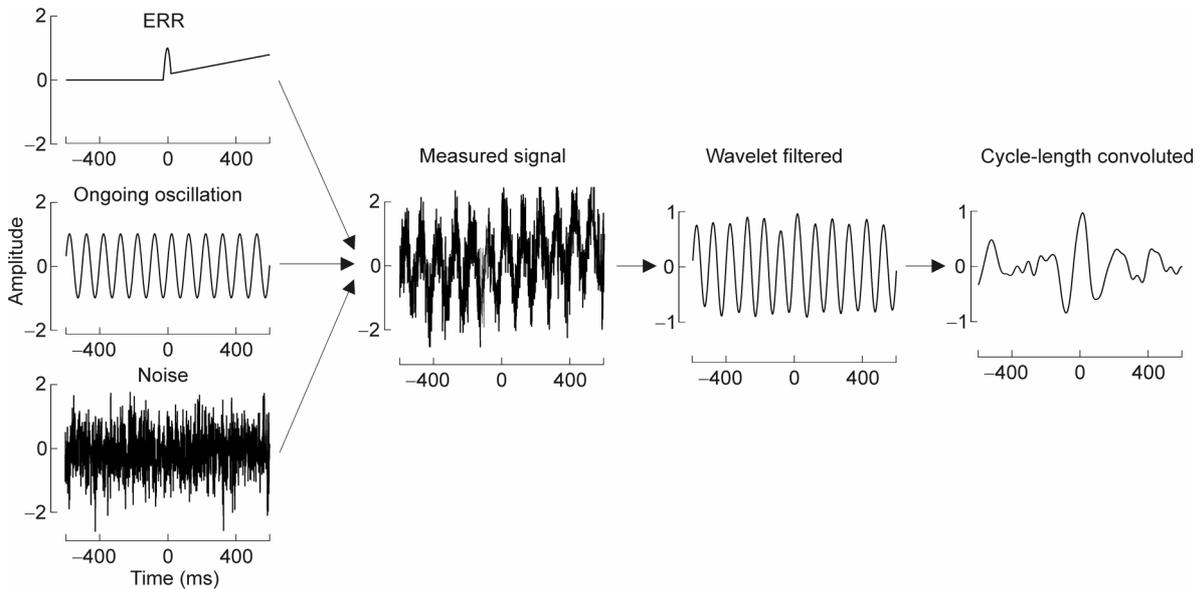
Fig. 6: Demonstration (simulated data) of the two first steps of CCWW. The measured signal has three components: ERR whose peak latency is to be resolved, ongoing oscillation, and noise. The cycle length of the ERR and of the oscillation are the same. Wavelet filtering at the frequency of the ongoing oscillation essentially removes the noise but reveals no discernible peak. The peak of the ERR is made visible by convoluting the wavelet-filtered signal with a square wave whose length is one cycle of the center pass-band frequency of the wavelet.

Cycle-length convoluted weighted-wavelet (CCWW) method: CCWW is based on the properties of CWT, the cancellation of the effect of the incidental phase of the background oscillations, and the weighting of the wavelet coefficients to emphasize the evoked activity over ongoing activity. These three points are elaborated below.

- The properties of CWT are suitable for accurate latency examination (see Section 1.5), and the wavelet can be selected so that it efficiently captures the signal features of interest. As the interest here is in short-duration transient events, the wavelet should be of a high time resolution. The already available data and the demonstrated capability of the DOG2 to unambiguously describe ERRs (see Fig. 5) favor its selection as the basis of the CWT that is the first step of CCWW (DOG2 has recently been used for similar purposes by Bostanov and Kotchoubey, 2004).

- When an ERR occurs in the presence of ongoing oscillations, the wavelet-filtered signal is composed of both the oscillation and of the ERR, and thus may not show a discernible peak at time of the ERR (see also Section 2.2). Here, the solution is to describe the ERR as a local disturbance in the oscillation. Consider first that the mean of a wavelet-filtered stationary oscillation over a (moving) cycle-length of the center pass-band frequency of the wavelet is zero. Now, when an oscillation is perturbed by an additive component, the mean of the wavelet-filtered signal over the cycle-length is not zero but reflects the perturbation. This is the result of the span of the wavelet being more than one cycle of its

center pass-band frequency: The increase in local mean is surrounded by opposing effects that bring the total mean back to zero. The disturbance in the local mean can be made visible by convoluting the wavelet-filtered signal with an appropriate waveform, the simplest of which is a square wave whose length is one cycle of the corresponding frequency of the wavelet. The effect of the wavelet filtering and cycle-length convoluting is demonstrated with simulated data in Figure 6. Filtering of wavelet coefficients using either hard or soft thresholding is a common technique (e.g. Addison, 2002) but the current technique of cycle-length-specific convoluting appears to be novel.

- The final stage of CCWW is the weighting of wavelet coefficients to emphasize ERR activity over ongoing activity. All wavelets are first scaled to be of equal magnitude, after which they are given an appropriate weight according to the frequency-specific SNR determined in Study IV (see also Fig. 5). The coefficients are acquired from the peak latency of the N1m, which is taken as a model for the onset of the ERR (see also Section 3.1). The CCWW signal is obtained by wavelet filtering, cycle-length convoluting, and averaging over the weighted coefficients.

CCWW applied: Here, the goal was to examine whether the amplitude effects in the averaged ERRs of Study V were attributable to changes in single-trial amplitudes or to latency jitter. To this end, the single-trial peak latencies were obtained with CCWW, after which the amplitudes were collected at the determined peak latencies from the band-pass-filtered and baseline-corrected unaveraged data. In this examination, the task of CCWW was to separate the onset ERR peaks from the high-magnitude noise, sustained brain activity, and background oscillations.

The amplitude decrease in the averaged responses with longer sound slopes was not reflected in single-trial amplitudes and is therefore likely to result from increased latency jitter. This increase in jitter appears inevitable, because with longer sound durations, the sound detection is susceptible over a longer period to stochastic events in the neural processes and auditory environment. The amplitude increase in the averaged ERRs as a function of task (active vs. passive), however, was accompanied by single-trial amplitude increases, indicating that the attention directed to the task increases the actual strength of brain activation. To demonstrate that the amplitudes obtained from the CCWW-determined latencies reflect the actual ERR amplitudes and not random noise values, the single-trial amplitude estimates were correlated with the averaged ERR amplitudes of the subjects; the CCWW results accounted for the inter-subject variation with notable accuracy ($r = 0.96$).

2.5 Accentuating oscillatory processes from frequency data (Study VI)

Background: In the examinations so far, ongoing brain activity has mainly played the role of noise, but the ongoing activity is of considerable interest to neuroscience, and furthermore is an important source of information regarding the status of the brain. Spectral estimation is the obvious way to examine the ongoing processes, but this is complicated by the $1/f$ slope (see e.g. Freeman *et al.*, 2000, and Fig. 2) which, rather than the oscillatory processes, is the main determinant of the spectral values. Another complication to spectral estimation is that the spectral peaks signifying the presence of oscillatory processes are often broad and ill-defined. Moreover, with the spectral peaks not being accurately aligned in, for example, subject population, averaging over data sets additionally spreads the spectral peaks along the $1/f$ slope. Below, we present a solution to the above problems.

Partition-referenced spectral estimation (PRSE): As the $1/f$ slope is an inherent part of the spectrum, the strategy chosen for removing it is not the development of spectral estimation *per se*,

but the exploitation of mutual information in spectral estimates obtained with different window lengths from the same data. Here, PRSE (not related to fractional Fourier transform, for which see e.g. Ozaktas *et al.*, 1994) employs the periodogram, but it may also be implemented, with additional considerations, using other spectral estimation methods.

The frequency response of the rectangular time-domain window of periodogram is a sinc function and the width of its main lobe is inversely related to the time window length[9]. The periodgram can be considered to operate by filtering the examined signal with the sinc-function shaped band-pass filters and thereby acquiring the magnitudes of the frequency components. If the signal contains a stationary single-frequency oscillation, then the magnitude of the spectral peak of the oscillation increases with the narrowing of the main lobe of the sinc function (for more details, see e.g. Hayes, 1996). Then again, if a process has a continuous, smooth spectrum such as the ideal $1/f$ noise, the width of the main lobe has little effect on the obtained spectral estimate. Now, by calculating the spectrum from a time window (full-length spectrum) and then partitioning the window into parts (partition to two parts used here) and taking the mean of the spectra calculated from these parts, one obtains a reference spectrum (of the same data) for the full-length spectrum. The difference between the reference and full-length spectrum is that the spectral peaks of approximately stationary oscillations are sharper and of higher amplitude in the full-length spectrum. In addition, the lowest frequencies with a wavelength of the approximate length of the full-length data window are not represented in the reference spectrum and hence discarded from PRSE. Dividing the full-length spectrum with the reference spectrum removes the noise slope and yields a level estimate (mean = 1), where the peaks of the oscillations remain.

Oscillatory processes have a characteristic shape in the divided spectrum: The spectral peaks are broader in the reference spectrum; therefore, in the divided spectrum, the peaks are surrounded by downward deflections. This has the practical consequence that when one averages over the divided spectra, peaks not accurately aligned effectively cancel each other out. Thus, prior to averaging over inhomogeneous data sets, one should use a nonlinear transform that enhances the peaks more than the downward deflections. As the mean of the divided spectrum is 1, a calculation of moments (10th moment used here) transforms the data appropriately. The prerequisite for employing moments is an SNR that guarantees that the peaks of the oscillations are the prominent components of the divided spectrum.

In PRSE, the visibility of an oscillation depends on its level of stationarity. One can attain a measure of this level by using several window lengths in the estimation (25 used in Study VI, logarithmically spaced between 500–4000 ms). Here, data obtained above the auditory brain areas during auditory stimulation were examined with half-overlapping data segmentation (similar to the Welch method).

PRSE applied: PRSE expectedly removed the $1/f$ slope. At low frequencies and long window lengths, the approximately $1/f$-shaped (original) spectrum has large magnitude variations, the main lobes of the periodogram are narrow, and there are fewer estimates available, all of which makes this part of the data noisy. Nevertheless, most of the obtained data was highly usable. The grand-averaged PRSE calculated over the shortest half of the window lengths displays a double peak at around 10 Hz (Fig. 7a), which is the result of four subjects having a peak at 9–10 and two at 11–12 Hz. The 9–12 Hz oscillation is in line with previous findings of tau rhythm and its inter-subject variation (Lehtelä *et al.*, 1997; Bastiaansen *et al.*, 2001). PRSE thus appears effective in removing the $1/f$ slope and accentuating the peaks of the oscillatory processes.

A mapping of the same PRSEs, but focused on the frequency range of interest, allows us to explore the dynamics of the oscillation (Fig. 7b). The 10 Hz activity is clearly seen with short windows and exhibits a maximum at around 750 ms but is poorly visible when the window length is over 1 second. The maximum indicates a duration over which the process is meaningfully describable as a near-stationary oscillation.

---

[9] The width of the main lobe of the sinc function at half height is $2\Delta t/N$, where Δt is the sampling interval and N is the number of samples in the rectangular time window.
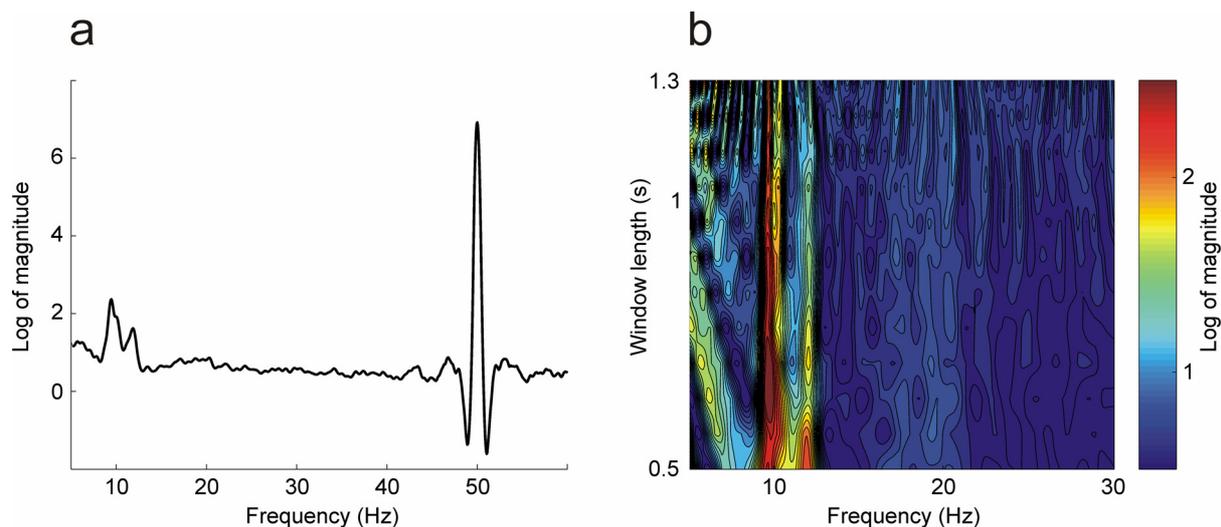
Fig. 7: The grand-averaged PRSE of MEG data obtained from auditory brain areas and averaged over window lengths of 0.5–1.3 s is displayed in (**a**). In (**b**), the oscillatory activity around 10 Hz is highlighted by plotting PRSEs according to window length.

2.6 Mapping the envelope modulation structure of noise-buried processes (Study VI)

Background: The methodological starting point of this work was described at the end of Section 1.5: the development of compact two-dimensional data representation where the envelope modulation structure of each frequency (as defined by pass-bands of wavelets) would be directly comparable with that of the other frequencies. Another goal of the current examination is related in Study IV, where event-related power decreases around 17 Hz were observed in the auditory brain areas. At that time, examinations with spectral estimates did not reveal the presence of an ongoing oscillation around this frequency (hence these attempts were not reported), and here it was hoped that the current methodological developments could shed new light on the situation.

Fractally scaled envelope modulation (FSEM): Here we approach FSEM as an extension to time–frequency methods. These methods enable the detection of even minute event-related power changes because of two factors: a high SNR (obtained with stimulus time-locked averaging) and a specific, highly useful baseline (provided by the prestimulus power level at each frequency band, see Section 2.3). Neither of these two is directly available for examining the ongoing processes but, through a byway, FSEM can be considered to employ both.

The starting point here is CWT of data. The SNR of CWT obtained from raw MEG or EEG data is not usually adequate for meaningful examination and, additionally, the amount of data tends to be impractically large. Consequently, the aim is to gather information over time in order to improve the SNR in a manner that does not lead to the cancellation of a relevant signal and yet retains some information of the temporal evolution of the processes[10]. Here, this is achieved using envelopes of each wavelet scale (modulus of complex CWT or Hilbert transform of real CWT). With spectral estimation of envelopes, one can characterize the typical temporal patterns that a frequency component exhibits (for application of these techniques to MEG data, see Linkenkaer-Hansen *et al.*, 2001, 2004). By examining the envelope spectrum of relatively short

---

[10] One such technique was used in the first published MEG recording, where SNR was improved by averaging according to phase of alpha rhythm obtained from simultaneous EEG recording (Cohen, 1968).

data segments, one can improve the SNR with averaging but without the time-domain cancellation. The time evolution and magnitudes, however, are so different in small and large scales (high and low frequencies) that there would not seem to be much of a point in obtaining the envelope spectrum from each scale and comparing them. Furthermore, the envelope spectra have the same unfavorable properties as the normal power spectrum (e.g. $1/f$ shape).

The solution to these problems is, simultaneously, a solution to the goal of acquiring a useful baseline. The envelopes of each wavelet scale are resampled so that an equal number of samples represents one cycle of the center frequency of each scale. That is, the sampling rate is adaptively changed and not related to resampling as "wavestrapping" (for wavestrapping see Bullmore *et al.*, 2004, where the relationship between DWTs and fractals is also elaborated). The spectra obtained from the resampled envelopes with equal sample-size windows are then directly comparable according to their frequency content. The spectra are further scaled to be of equal mean magnitude. These steps constitute fractal scaling. Its rationale is that if any scale contains any processes other than noise, then the distribution of the envelope spectral power differs from that of the neighboring scales. The scales without distinct processes provide the baseline. With each frequency (wavelet scale) mapped to its own array, the resolution of the examination is literally raised to a new dimension when compared to spectral estimation. The relevant signal components are made palpable by normalizing the data over scales with respect to each envelope modulation frequency. As no scale-dependent trends were observed when tested with simulated data in Study VI, it appears that all deviations from the neighboring, normalized values must be due to relevant processes being present in the data.

Finally, fractals are objects that display self-similarity over a range of scales, and it is often customary to define the fractal property of a phenomenon via the Hurst exponent (Addison, 2002). Here the Hurst exponent was not used, but the data were scaled so that the (self-)similarity of the envelope spectra could be examined over scales, which is enabled by the fractal property of the wavelet transform (shape of a wavelet is the same regardless of the scale), warranting the term fractal scaling.

FSEM applied: Here the examined data are the same as with PRSE (Section 2.5). The CWTs that underlie the FSEMs were performed with an emphasis on time, frequency, or compromise between the two by employing DOG2, Morlet-12, and Morlet-6 wavelets (see Section 2.3), respectively. FSEM expectedly provided a description of the two prominent entities known to be present in the data: the auditory ERRs and the 50 Hz mains component. The examination was not performed in a stimulus time-locked manner, but the used constant stimulus repetition interval provided a strong signal modulation that was readily observed with FSEM. As ERRs are localized in time rather than in frequency, they were best observed with DOG2-based FSEM. The 50 Hz mains noise is a prominent, frequency-localized component and hence clearly visible with FSEMs obtained with wavelets emphasizing frequency resolution (Fig. 8).

The advantages of FSEM, however, do not lie in the analysis of these prominent entities but in the analysis of irregular noise-buried signals and more generally in the examination of modulations. Here, the 9–12 Hz tau rhythm, observed also with spectral estimates, was evident, but, importantly, FSEM revealed an oscillation in the 15–18 Hz range that could not be detected by spectral estimation. The frequency range of the latter oscillation coincides exactly with the frequency range of the power reductions observed in Study IV. FSEM further showed that the oscillations had maxima in the envelope spectra, which indicate that the oscillations exhibited modulations in the 3–7 second range. The time course of the oscillations was thus largely irrespective of auditory stimulation, although a local maximum can be observed in Figure 8 for the 15–18 Hz oscillation that corresponds with the stimulation interval, in accordance with the results of Study IV.
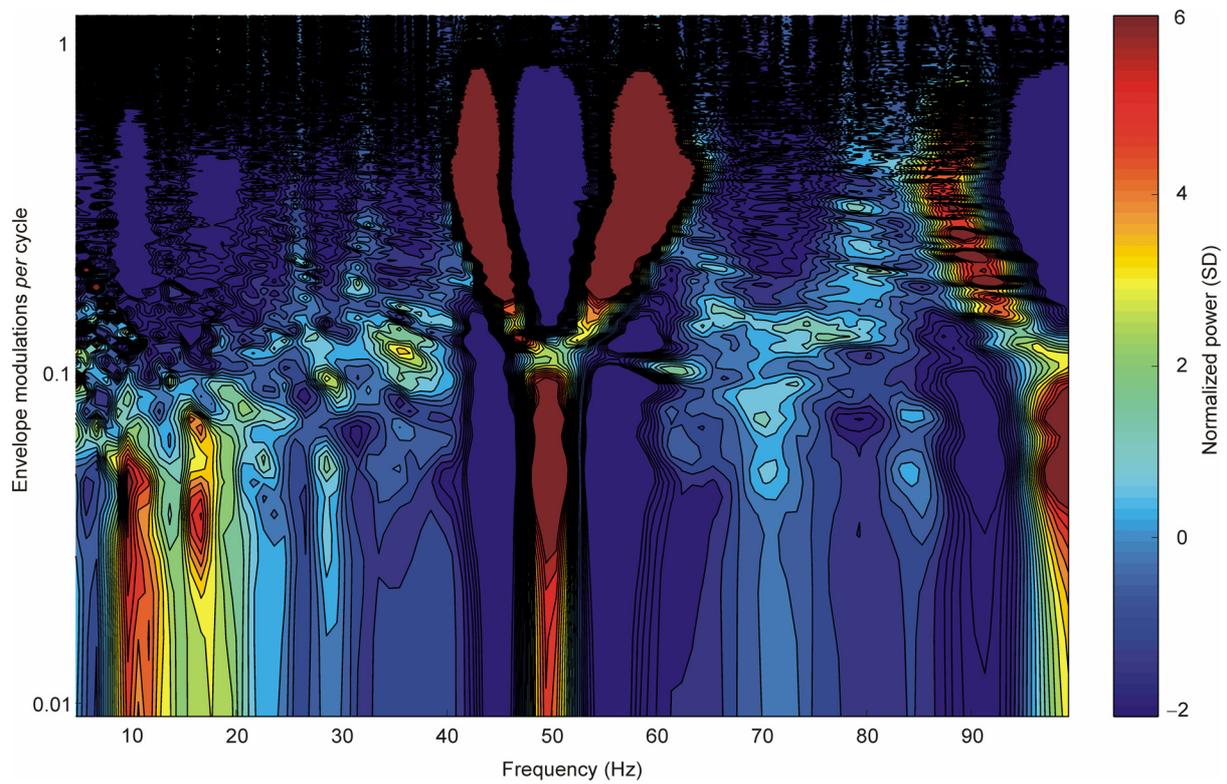
Fig. 8: Grand-averaged FSEM representation (Morlet-6, envelope spectra acquired using data windows whose length was 100 cycles) of an MEG signal obtained from the auditory brain reveals, for example, ongoing oscillations at around 10 and 17 Hz as well as their modulations (envelope modulation axis is logarithmic).

3. DISCUSSION

3.1 Meeting the thesis objectives

This thesis had six objectives (Section 1.6) that were all concerned with the examination of time, frequency, and time–frequency structure of MEG and EEG data. How these objectives were met is discussed below. In addition to these objectives, this thesis may provide insights into to two more general topics: the analysis and processing of elusive signals buried in noise (Section 3.2) and the role of transient and ongoing activity in neural information processing (Section 3.3).

Objective 1: In Studies I and V, it was found that tones slowly increasing in intensity from inaudible to audible give rise to prominent ERRs. The abrupt onset peak of these ERRs contrasts with the smooth and gentle intensity increase of the stimuli; instead, it reflects the discrete change in perception with the inaudible sound becoming audible. This ERR provides the sought linkage between perception and MEG responses by being closely correlated with the behaviorally measured sound detection. Previous studies have indicated a link between brain events and behavioral sound detection using near-threshold stimuli and *a posteriori* classification of ERRs based on behavioral decision criteria (Hillyard *et al.*, 1971; Parasuraman *et al.*, 1980, 1982; Paul and Sutton, 1972). The current results complement these observations in providing a predictive temporal linkage.

The onset of the ERR of Studies I and V was generated in the same areas as the N1m but can be seen to differ from the N1(m) in that it does not reflect the sound onset typically accorded to the N1(m) (Näätänen and Picton, 1987). In defining the N1(m) to be a stimulus-onset response, it is not considered that only the part of the stimulus that is at least nearly audible is likely to contribute to the generation of the N1(m). By further noting the continuum of response morphologies in Study V from the N1(m)-like ERR with short sound slopes to the step-like entity with long sound slopes, it appears that the N1(m) and the ERR of Studies I and V are generated by the same neural processes. With the used slope stimuli, this ERR appears to be a sound onset response—not that of the physical onset, but that of the perceptual onset.

The relationship between the N1(m) and perception is, then again, controversial (Section 1.3), and Studies I and V can only do little to clarify this controversy (Section 3.3). Nevertheless, by exploiting the high time resolution of MEG, Studies I and V succeed in describing the stimulation conditions where MEG indexes behavioral sound detection. These results thus pave the way for utilizing non-invasive brain measures in the objective evaluation of the functionality and development of the human auditory system. However, the conclusions of several recent studies have cast a shadow on the notion that ERRs directly index neural information processing.

Objective 2: The objective was to resolve whether ERRs are generated by stimulus-induced phase reorganization of ongoing oscillations or by processes additive to ongoing activity. The popular idea of organized oscillations (e.g. Başar, 1980; Gruber *et al.*, 2005; Hamada 2005; Jansen *et al.*, 2003; Karakaş *et al.*, 2000; Klimesch *et al.*, 2004; Makeig *et al.*, 2002; Sayers *et al.*, 1974) would appear to indicate that ERRs do not directly reflect the processing of sensory information in the human brain. That is, by its very nature, ongoing activity will stochastically, and without stimuli, produce all the states that can be produced with phase resetting. This might suggest that ERRs are more of an epiphenomenon than measurable projections of the neural information processing.

In Studies II and III, the evidence held in favor of the organized oscillations model was found to be inconclusive. In particular, an inter-trial phase-coherent state observed at the time of ERRs, which has been held to mark the phase coherence of the ongoing oscillations, was found to be explainable equally well by additive ERR influencing the phase distribution of the signal. This ambiguity will follow all methods that examine the phase of a band-limited signal. The inter-trial phase coherence, however, would manifest itself also as inter-trial amplitude coherence, which can be examined from a broadband signal. In Studies II and III, this broadband examination provided unambiguous evidence that auditory stimulation did not result in phase reorganization of ongoing oscillations, and, in addition, the presence of additive components that accounted for the ERRs could be established.

Studies supporting the conclusions of this thesis have begun to emerge. In fact, in 1983, Jervis *et al.* had already found evidence that Fourier components at the poststimulus stage had greater power than at the prestimulus, and they concluded that the additive model is more plausible. However, the study by Jervis *et al.* has been suggested to provide inconclusive results (Jansen *et al.*, 2003) or even to support the organized oscillations model (Hamada, 2005). More recently, Shah *et al.* (2004) have taken a clear stance against the organized oscillations model on the basis that, in their intracortical measurements from monkeys, visual responses were found to emerge from a baseline where the amplitude of the ongoing activity was only a fraction of the amplitude of the evoked response. Kirchfield (2005) reanalyzed Brandt's (1997) data and observed that superposition of visual ERRs to ongoing oscillations without resetting their phase accurately accounted for the data. Studies II and III are most closely paralleled by Yeung *et al.* (2004), which meticulously shows, mainly based on simulations, that none of the evidence provided by Makeig *et al.* (2002) distinguishes between the reorganization of ongoing oscillations and the additive model. Moreover, Yeung *et al.* reach the conclusion that this question would need to be examined using broadband signals and preferably with spatial weighting; both techniques were realized and used in Study II (inter-trial amplitude variance and RSW).

In conclusion, the reviewed evidence suggests that the additive model is a generally valid generation mechanism for the ERRs. Studies II and III, in parallel with Yeung *et al.*, provide an explanation for all common results that have been interpreted as supporting the organized oscillations model in terms of the additive model. In addition, the inter-trial amplitude variance introduced in Study II is, thus far, the only proposed method that can distinguish between an inter-trial phase-coherent state of ongoing oscillations and the emergence of phase-coherent additive ERRs. However, one might entertain a possibility of various phase-locked and non-phase-locked induced or dampened oscillatory processes that could account for any observed inter-trial amplitude variance result. In order to exclude this possibility one should examine, using time–frequency methods, what event-related processes actually are present in the measured signal.

Objective 3: In Study IV, a conclusive mapping of auditory event-related processes was sought along with prerequisites for the effective use of CWT in this task. It was found that wavelets of high time resolutions could be readily used for analysis of auditory ERRs. This is because they are sensitive to individual ERR peaks, whereas wavelets of a higher frequency resolution are sensitive to wave complexes or oscillations. A central point for the effective application of CWT was found to be the frequency (wavelet scale) specific baseline. The most notable problem with the baseline is that the wavelets may spread the power of ERRs to an extent that a true baseline level may be difficult to acquire.

Another objective was to clarify the relationship of the time-domain averaged ERRs and the processes observed in the time–frequency plane: All the power increases could be accounted for by the ERRs, which further supports the conclusions of Studies II and III. Study IV also showed that power decreases following the ERRs, while small in magnitude, are an integral part of the auditory event-related processes. The effect of auditory stimulation on ongoing oscillations thus appears to be a slight intermittent dampening. Study IV, however, did not address the relationship between the event-related power decreases and composition of ongoing processes, whose efficient examination required further methodological developments (Study VI).

Objective 4: Here the objective was to exploit the information obtained in Studies II–IV to uncover the single-trial level explanations for the amplitude effects of averaged ERRs in Study V. The cycle-length convoluted weighted-wavelet (CCWW) method was devised for this purpose. CCWW performed agreeably by showing that the amplitude increase in averaged ERRs as a function of task (in the active, or response, condition, the ERRs were of higher amplitude than in the passive) was accompanied by increases in single-trial ERR amplitudes. The amplitude decrease of averaged ERRs, as a function of decreasing sound slope, then again, could be attributed to increased latency jitter of the single-trial ERRs.

CCWW employs a novel technique, cycle-length convolution of wavelet-filtered signal, to eliminate the effect of the incidental phase of background oscillations. This convolution technique may be useful also in other applications, whereas in its current form the full CCWW method should only be used for examining single, relatively isolated transient deflections. The results of Studies II–IV are likely to be more important for single-trial analysis than the solutions of Study V.

Objective 5: Partition-referenced spectral estimation (PRSE), introduced in Study VI, achieves the objective of providing an unambiguous representation of the oscillatory processes of MEG data. The $1/f$ slope was removed, and the peaks of oscillations, when SNR was sufficient, could be accentuated to an arbitrary degree. PRSE is based on simple manipulations of the spectra, but I have not found a record of any closely similar technique. Unlike some methods used for removing the $1/f$ slope (Buzsáki and Draguhn, 2004; Dumermuth and Molinari, 1987), PRSE is insensitive to unexpected or non-monotonic shapes of the noise profile and avoids possible over- and under-fitting problems as well as the requirement of searching parameters for fitting. Because of its simplicity and shown functionality, PRSE may be generally useful in the analysis of noise-buried oscillations. It should, for example, provide a suitable input, with the artifactual contribution of background activity eliminated, for localization algorithms when the goal is to determine the sources oscillatory processes (e.g. Jensen and Vanni, 2002).

Due to removing the noise slope, an inherent part of the spectrum, a PRSE signal representation may not be considered a true spectrum. Instead, PRSE describes what is normally considered the relevant contents of the spectra. PRSE also provides a measure of stationarity, which indicates the optimal window length for spectral estimation in the sense that shorter window lengths will yield a blunter spectral peak, while increasing the window length will not sharpen the peak but will reduce the SNR. However, when the goal is to obtain a measure of the duration of the oscillatory states or to reveal oscillatory processes with a low level of stationarity, fractally scaled envelope modulation (FSEM) estimation instead of PRSE may be recommended.

Objective 6: Fractally scaled envelope modulation (FSEM), introduced in Study VI, meets the objectives of the desired two-dimensional time-series data representation very well. FSEM describes the statistical temporal behavior of each frequency (as defined by pass-bands of wavelets) via its envelope spectrum and employs a fractal scaling that makes the envelope spectra of different frequencies directly comparable with one another, despite their inherent magnitude and time-scale differences. By providing a statistical description of the temporal behavior of the frequency components, FSEM achieves an SNR that is beyond that of the non-averaged time–frequency estimates. Furthermore, by mapping each frequency component to an array and with the fractal scaling providing a functional baseline, it achieves a resolution beyond that of the spectral estimation.

Studies I–V and PRSE have all given their separate contributions for elucidating the time (ERRs), frequency (ongoing oscillations), and time–frequency (ERD) structure of MEG signals, but FSEM manages to describe all these processes simultaneously. It is, however, obvious that FSEM is not an ideal method for examining all processes that it manages to reveal, but that a preferable alternative may be found from the wide range of current signal processing methods (see e.g. Cohen, 1995; Mainardi *et al.*, 2002; Thakor and Tong, 2004). FSEM may, however, be recommended for examining signals when SNR is low but plenty of data are available. Its fractal

scaling also ensures that it is a generally effective method for examining modulation structure of a signal, and for the elemental task of detecting weak, unstable, oscillatory processes, it may be the most effective of all currently available methods.

3.2 On the analysis of elusive signals in noise

The level of stationarity and the analysis methods of this thesis: A signal is usually very well described if one has an accurate model of the processes generating it and information of the conditions that influence the processes. Modeling of the system is, however, often not possible because of the complexity of the processes or a lack of knowledge; instead, signal transforms and representations are used to uncover the constituents of the signal. In MEG and EEG research, there have recently been suggestions that the signal is nonlinear and therefore unsuitable for analysis with linear methods such as the Fourier transform (e.g. Freeman *et al.*, 2003). However, the Fourier transform and DWT provide frequency and time–frequency representations of a signal, respectively, which allow perfect reconstruction of the time-domain signal; hence, they certainly are valid signal representations. In addition, the division between linear and non-linear methods seems misplaced. That is, while the Fourier transform, CWT, and DWT are linear transforms, the derived amplitude and power spectra or the type of time–frequency transforms presented in this thesis are all nonlinear data descriptions, yet are often classified as linear (e.g. Thakor and Tong, 2004). Therefore, rather than describing signals in terms of linearity and non-linearity, this thesis work suggests that the stability of processes, or the level of stationarity, is a useful concept in evaluating signals.

Traditional signal analysis has mainly operated under the assumption that the examined signal is generated by processes exhibiting a high level of stationarity, whereby spectral estimation is a suitable analysis technique (e.g. Hayes, 1996), but the methods introduced in this thesis also work with such signals. If the examined process is buried in some other type of noise than white, then PRSE, which specifically exploits the stability of oscillatory processes, can be recommended for removing the noise slope. An oscillation with a stable modulation structure usually constitutes a relatively stationary process. Such processes are well-suited for analysis with FSEM, whose fractal scaling allows detecting even small and diffused peaks in the modulation spectrum that would easily be missed with conventional envelope analysis techniques.

When considering pronouncedly nonstationary processes, one may differentiate between processes that are relatively stationary over some periods and those that are in a continuous mode of change. Speech is an example of a signal that, while definitely nonstationary, contains segments that are approximately stationary and are generated by mechanisms that are relatively simple (at least when compared to the brain), and thus suitable for linear modeling (e.g. Varho, 2001). In this thesis, no strict assumptions of the level of stationarity of MEG signals have been made; instead, both PRSE and FSEM provide ways to measure the level of stationarity. The measure gained with PRSE is derived from the width and morphology of spectral peak, and while it is related to the level of stationarity of the oscillation, the relationship may not be a straightforward one. FSEM, then again, measures the intervals of the modulations of the oscillation, which coincides with the intuitive notion of the duration of the oscillatory state. FSEM results were also robust in the sense that the same modulation durations were obtained in Study VI with FSEMs employing different wavelets and envelope spectrum estimation window lengths. Thus, FSEM may be a useful analysis tool in tasks ranging from those where the goal is to detect the presence weak signal sources, such as exposing submarines from the $1/f$-distributed ocean noise, to those where the goal is to identify the unknown structure of signal, such as in radio astronomy.

The above considers the examination of the whole signal, whereas event-related processes of MEG and EEG contribute to the measured signal only intermittently. Approximately time-

and phase-locked event-related processes are routinely examined with event time-locked averaging, but an approach where they are considered as transient additions to more stationary background activity appears useful for obtaining single-trial level descriptions (e.g. Lange *et al.*, 1997). The cycle-length convoluting of the wavelet-filtered signals, introduced in Study V, is an example of this approach; even when an ERR and an ongoing oscillation have opposing phases, the cycle-length convolution unambiguously reveals the existence of the additive ERR and facilitates the detection of peak latency. In other signal-analysis fields, this approach is often applied in an inverse manner; for example, Hadjileontiadis and Panas (1998) used it to separate transient heartbeat artifacts from more stationary lung sounds.

All approximately event-time-locked processes can be examined with time–frequency methods, but Study IV highlights some difficulties that arise when a technique (wavelet) suited to examining oscillatory processes is used for examining transient events. Without contrary information, the event-related processes should be considered to be devoid of stable oscillatory structure to an extent that they are best examined using time–frequency methods of the highest time resolutions. Indeed, it is a peculiarity of MEG and EEG research that frequency-domain methods are commonly used to represent signals not meaningfully described by the concept of frequency (a cautious reader may find insights on the above notion from Kramarenko and Tan, 2002).

On the pitfalls and safe passages of MEG and EEG analysis: One purpose of this thesis text has been to provide guidelines for obtaining methodologically solid results. Here, some of that material is compressed to ten educational observations.

1. Obtaining an undistorted baseline is as important as obtaining the actual signal. A short stimulus presentation interval is particularly problematic in this respect.

2. Negative observations are not a good basis for inferences. Not observing stimulus-related power changes with single-trial spectral estimation is a specifically questionable negative result.

3. Amplitude changes in averaged ERRs can be accounted for by a number of factors other than changes in the single-trial ERR amplitudes (e.g. changes in single-trial latency alignment or changes in source visibility with changing source location or orientation).

4. With averaged ERRs, narrowing of the examined frequency band should only be used for emphasizing events already visible in a broadband, preferably unfiltered signal. A related point is that the removal of transient events (e.g. ERRs) to uncover hypothesized underlying oscillations should not be attempted with frequency-domain methods.

5. Observing events on a frequency band does not imply that the observed processes are oscillatory; frequency and time–frequency representation can be difficult interpret mainly because of how they link events separated in time.

6. Defining brain processes into narrow frequency bands (e.g. < 2 Hz) is likely to yield erroneous results.

7. Inter-trial phase coherence on a frequency band of oscillation does not guarantee that it is the oscillation that is in the phase-coherent state.

8. A power increase in an averaged time–frequency estimate does not readily imply changes in the ongoing brain processes whereas a power decrease does (if the baseline is undistorted).

9. Observing that distinct frequency bands have similar time courses can mean that distinct processes are communicating but may also mean that a single process is represented at distinct frequencies.

10. If the possible motor artifacts can be controlled, using same stimulation in two or more recording conditions where subjects have different instructions or tasks is one of the most reliable ways of obtaining results that actually reflect information processing in the human brain.

3.3 Transient and ongoing processes in neural information processing

Setting the stage: At an elementary level, the discussion of neural information processing focuses on the question of whether the firing rate of neurons or the precise temporal pattern of the firings encodes information (e.g. Borst and Theunissen, 1999; deCharms and Zador, 2000). The conjecture that the temporal patterns are important for information processing is associated with oscillations, because information could be encoded into phase relations between oscillations of different neural populations without concomitant change in firing rate. Indeed, there is a plethora of articles emphasizing the role of synchronized oscillations in neural information processing (for reviews see e.g. Buzsáki and Draguhn, 2004; Schnitzler and Gross, 2005; Varela *et al.*, 2001). Nevertheless, the role or even the abundant existence of synchronized oscillations is not conclusively established (e.g. Pareti and Palma, 2004), and alternative explanations for the results favoring temporal coding with oscillations have emerged. For example, Harris (2005) describes how the internal dynamics of the cortical circuit can result in a temporal structure beyond what is expected from firing rate coding, despite the rate being the primary means of encoding information. Recent evidence (Johansson and Birznieks, 2004) has revived the hypothesis that neural coding is performed with first spike timing (Thorpe, 1990). This form of coding uses transient spike trains or even single spikes with respect to some available time reference to encode information (VanRullen *et al.*, 2005). Shadlen and Movshon (1999) have gathered a load of criticism to slow down the bandwagon of synchronized oscillations, but their potent argumentation appears to have been ignored rather than defeated.

In this thesis work, no evidence supporting the role of synchronized oscillations in neural information processing was encountered, whereas various examples were met where the effects of the transient (non-oscillatory) ERRs could be misinterpreted to reflect oscillatory activity (see also Yeung *et al.*, 2004). Moreover, the auditory ERRs were shown to provide for unambiguous information representation in not reflecting phase-organized states of ongoing oscillations, which entail the problem that they occur stochastically and thus without stimulation. The current results highlight the role of transient, non-oscillatory activity in neural information processing.

The possible roles of transient and ongoing processes: The ERR observed in Studies I and V is a step-like sign that the subject is able to, within a couple of hundred milliseconds, behaviorally indicate that a sound has been detected. The non-proportionality of the onset of the ERR compared to the smooth sound slope is in line with the idea that increasing sound intensity is an ecologically important signal associated with approaching sound source (Neuhoff, 1998). Recurrent excitation of cortical circuits (Douglas *et al.*, 1995), which provides rapid amplification of the input signal, has been considered as a neural level mechanism to underlie the generation of the N1(m) (May, 1999) and could also account for the ERR of Studies I and V. With such a mechanism in place in the cerebral cortex and with the shown correlation between behavioral and brain responses, one might suggest that the role of this abrupt cortical activation is to provide a representation of the sensory stimuli with its basic features encoded in the location according to the topographical organization. This suggestion is supported, for example, by the study of Lu *et al.* (2001), which indicates that in the auditory system of monkeys, a locally increased firing rate encodes most auditory stimulus properties and that only slowly varying stimulus features are encoded with the modulation of the firing rate. It should, however, be noted that although the activity is elicited by the sensory stimuli, the evolution of activity is likely to be determined by the organization of the neural circuit, which has been molded by the past life of the animal, rather than directly reflecting the dynamics of the stimuli (Fiser *et al.*, 2004; Harris, 2005).

The neural representation of the stimulus refers here to the part that is necessary and central (but not sufficient alone) among the ensemble of neural events that underlie the perception of the stimulus. The complexity of the human brain, however, discourages

generalizations. That is, the abrupt amplified neural activation is suitable for providing a high SNR representation of weak auditory stimuli. However, when sounds are clearly audible, the brain may employ a different processing strategy, wherein some neural processes preceding those manifested in the N1(m) provide the information of the existence of the stimuli. This would explain why the N1(m) type of activity is sometimes found to be linked to perceptual sound detection (Parasuraman *et al.*, 1980, 1982; Studies I and V) and sometimes not (e.g. Jaskowski *et al.*, 1994). With strong signals, the main role of the neural processes underlying the N1(m) may be to mediate information of the more elaborate features of the stimuli (Roberts *et al.*, 2000).

On the basis of the current data, it is hardly possible to suggest or support any role that induced oscillations might have; no induced additive oscillations were observed. It may, however, be noted that the coding scheme with transient activity patterns is highly suitable, specifically, for the auditory modality. In other sensory modalities, the stimulus features are often less dynamic and therefore more suitable to be encoded with more stable activity patterns. For example, it has been suggested that a steady-state condition in the somatosensory system is maintained with (standing wave) oscillations (Schoffelen *et al.*, 2005). From the visual system, there is evidence that stable images produce oscillations, whereas more dynamic stimuli result in transient, non-oscillatory activity (Kruse and Eckhorn, 1996).

A small event-related power reduction, or ERD, was observed in Study IV (for possible functional roles of ERD, see Neuper and Pfurtscheller, 2001), but, overall, the ongoing oscillations appeared highly independent of the ERRs. This, however, does not imply that the ongoing activity is irrelevant to neural information processing. On the contrary, the results of this thesis are in line with those obtained invasively with voltage sensitive dyes (for a review of the method, see Grinvald and Hildesheim, 2004). These have indicated that evoked responses in the mammalian visual cortex are accounted for by linear summation of evoked responses and ongoing activity (Arieli *et al.*, 1996). This indicates that the ongoing activity affects the evoked responses (e.g. their amplitude) in a straightforward manner and, subsequently, neural information processing. Several studies have discussed and provided evidence for the possible beneficial role of noise in neural information processing (for review, see Moss *et al.*, 2004). These studies have mostly referred to the role of noise in enabling a weak stimulus to exceed a threshold in a nonlinear mechanism. The current situation, however, appears different; the neural responses to stimuli of sufficient intensity are unlikely to require contribution from noise for their elicitation. Instead, the role of background oscillations might be to provide variability to the neural responses to stimuli. In simple serial computer systems, added variability in the input is likely to lower the performance of the system, and may also do so in complex and adaptive animal brains in a short time span. However, for long-time survivability of the brain, the variable input may be essential in facilitating learning, for example, by preventing us from being stuck on a local maximum or minimum while seeking an optimal solution to a problem.

3.4   Concluding remarks

In this thesis, the time–frequency information structure of electromagnetic brain signals has been examined via MEG measurements and methodological developments; some elemental questions of MEG and EEG research have been tackled along with the topic of processing of noise-buried signals. As a scientific endeavor, this thesis work shares features common to contemporary research. The work provides for journal articles, but is not an extensive, orderly examination of a single question covered by a single scientific discipline. Hence, it is in order to consider the consistency and validity of the thesis work.

Studies I and V examined whether and how the MEG signals reflect human cognitive operations. It was found that smooth sound slopes with intensity increasing from undetectable to audible elicited event-related responses (ERRs) whose step-like morphology reflected perceptual

sound detection rather than the acoustical structure of the stimuli. Moreover, the latencies of the ERRs accurately predicted the behavioral reaction times to sound detection. These results should contribute to the evidence that provides credibility for MEG and EEG as objective measures of human cognition. As the ERR of Study I and its correlation with behavioral measures could be replicated in Study V, at least some level of consistency was achieved. In addition, the behavior of this ERR was examined using a fair range of sound slopes. Nevertheless, these results provide only the starting point for understanding the dynamics of this ERR and for its possible use in hearing assessment.

The examination of the generation mechanism of ERRs can be considered as the most thorough part of the thesis work. In Studies II and III, the popular suggestion that ERRs are the result of phase resetting and consequent reorganization of ongoing brain oscillations was found to originate from inconclusive evidence. Subsequently, in Studies II and III, the presence of additive components could be established, and a novel analysis technique showed that auditory stimulation does not give rise to a phase-coherent state of ongoing oscillations. In Study IV, a detailed time–frequency mapping of auditory event-related processes was performed. An ongoing oscillation was found to be slightly and intermittently dampened by stimulation, but more importantly, and in line with Studies II and III, the averaged ERRs were found to accurately account for the power increases observed in the time–frequency plane. In addition to the results of Studies II–IV, the unipolar ERR shapes in Studies I and V as well as the results of Study VI, indicating that the modulations of ongoing oscillations were mainly irrespective of stimulation, are all evidence that supports the view that the ERRs are additive to ongoing brain oscillations.

The methodological work of this thesis was aimed at providing solutions to problems that emerged from the analysis of MEG and EEG signals, but this work may also be viewed from a broader perspective. Study V introduced the cycle-length convoluted weighted-wavelet (CCWW) method that employs *a priori* information of ERRs to single-trial analysis. The method itself is unlikely to find an application outside its intended purpose, but the cycle-length specific convolution, which allows effective isolation of perturbations from oscillations, is a persuasive, more widely applicable technique.

Simplicity is usually one of the most important assets of a method when considering its overall applicability. Partition-referenced spectral estimation (PRSE), introduced in Study VI, scores high in this respect and yet provides a data representation where oscillations are accentuated and the noise slope of the spectral estimate is removed. PRSE was not tested very extensively in this thesis, but its simplicity should ensure that no serious complications arise if it is applied more generally to data analysis. The main methodological development of the thesis, fractally scaled envelope modulation (FSEM), introduced in Study VI, requires a number of signal processing steps. The added complexity may be justified by the information obtained of the modulations of oscillatory processes. Moreover, FSEM revealed oscillations that could not be detected with spectral estimation; FSEM may be the most effective method yet for detecting unstable oscillatory processes buried in prominent noise. In addition, the idea of a data decomposition where all components are made directly comparable despite their inherent scale differences, the fractal scaling, is certainly applicable elsewhere.

To conclude, this thesis is far from being the most focused and concentrated examination possible. Nevertheless, Studies I–VI follow in succession from the developments of each previous study and offer mutually supportive results that have, I hope, succeeded in providing genuinely new insights into various disciplines. Specifically, the examinations of this thesis come together for two general propositions: (1) the structures and their duration in signal both determine and facilitate the most effective analysis method and (2) transient brain activity, reflected in ERRs, is a feasible means of representing information in the human brain.

4. REFERENCES

Addison PS, 2002. The Illustrated Wavelet Transform Handbook. IOP press, Bristol and Philadelphia.

Adrian ED, Matthews BHC, 1934. The Berger rhythm; potential changes from the occipital lobes in man. Brain 57, 355–384.

Angelidou A, Strintzis MG, Panas S, Anogianakis G, 1992. On AR modelling for MEG spectral estimation, data compression and classification. Comput Biol Med 22, 379–387.

Anurova I, 2005. Processing of Spatial and Nonspatial Auditory Information in the Human Brain. PhD thesis, University of Helsinki.

Arieli A, Sterkin A, Grinvald A, Aertsen A, 1996. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868–1871.

Baillet S, Mosher JC, Leahy RM, 2001. Electromagnetic brain mapping. IEEE Signal Process Mag 18, 14–30.

Başar E, 1980. EEG-Brain Dynamics: Relation Between EEG and Brain Evoked Potentials. Elsevier, New York.

Bastiaansen MC, Böcker KB, Brunia CH, de Munck JC, Spekreijse H, 2001. Event-related desynchronization during anticipatory attention for an upcoming stimulus: a comparative EEG/MEG study. Clin Neurophysiol 112, 393–403.

Berger H, 1929. Über das Elektroenkephalogram des Menschen. Arch f Psychiat 87, 527–570.

Borst A, Theunissen FE, 1999. Information theory and neural coding. Nat Neurosci 2, 947–957.

Bostanov V, Kotchoubey B, 2004. Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. Psychophysiology 41, 259–268.

Bracewell R, 1999. The Fourier Transform and Its Applications (3rd ed). McGraw-Hill, New York.

Bradley AP, Wilson W, 2004. On wavelet analysis of auditory evoked potentials. Clin Neurophysiol 115, 1114–1128.

Brandt ME, 1997. Visual and auditory evoked phase resetting of the alpha EEG. Int J Psychophysiol 26, 285–298.

Budinger TF, 1998. PET instrumentation: what are the limits? Semin Nucl Med 28, 247–267.

Bullmore E, Fadili J, Maxim V, Sendur L, Whitcher B, Suckling J, Brammer M, Breakspear M, 2004. Wavelets and functional magnetic resonance imaging of the human brain. Neuroimage 23, Suppl 1, S234–S249.

Buzsáki G, Draguhn A, 2004. Neuronal oscillations in cortical networks. Science 304, 1926–1929.

Clynes M, 1969. The dynamics of R–M function. In: Average Evoked Potentials: Methods, Results and Evaluations. Donchin E, Lindsley DB (Eds). US Government Printing Office, Washington, 363–374.

Cohen D, 1968. Magnetoencephalography: Evidence of magnetic fields produced by alpha rhythm currents. Science 161, 784–786.

Cohen L, 1995. Time-frequency Analysis. Prentice Hall, New Jersey.

Coifman R, Wickerhauser M, 1992. Entropy-based algorithms for best basis selection. IEEE Trans Inform Theory 38, 713–718.

Crowley KE, Colrain IM, 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin Neurophysiol 115, 732–744.

Darvas F, Pantazis D, Kucukaltun-Yildirim E, Leahy RM, 2004. Mapping human brain function with MEG and EEG: methods and validation. Neuroimage 23, Suppl 1, S289–S299.

Dawson GD, 1951. A summation technique for detecting small signals in a large irregular background. J Physiol (Lond) 115, 2P–3P.

deCharms RC, Zador A, 2000. Neural representation and the cortical code. Annu Rev Neurosci 23, 613–647.

Donchin E, Ritter W, McCallum WC, 1978. Cognitive psychophysiology: The endogenous components of ERP. In: Event-Related Brain Potentials in Man. Callaway E, Tueting P, Koslow S (Eds). Academic Press, New York, 349–441.

Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH, 1995. Recurrent excitation in neocortical circuits. Science 269, 981–985.

Dumermuth G, Molinari L, 1987. Spectral analysis of the EEG. Neuropsychobiology 17, 85–99.

Durka PJ, 2003. From wavelets to adaptive approximations: time-frequency parametrization of EEG. Biomed Eng Online 2, 1.

Eddins AC, Peterson JR, 1999. Time-intensity trading in the late auditory evoked potential. J Speech Lang Hear Res 42, 516–525.

Fiser J, Chiu C, Weliky M, 2004. Small modulation of ongoing cortical dynamics by sensory input during natural vision. Nature 431, 573–578.

Freeman WJ, Holmes MD, Burke BC, Vanhatalo S, 2003. Spatial spectra of scalp EEG and EMG from awake humans. Clin Neurophysiol 114, 1053–1068.

Freeman WJ, Rogers LJ, Holmes MD, Silbergeld DL, 2000. Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. J Neurosci Methods 95, 111–121.

Gabor D, 1946. Theory of communication. J Inst Elec Eng 93, 429–457.

Geddes LA, Baker LE, 1989. Principles of Applied Biomedical Instrumentation. Jonh Wiley & Sons, New York.

Godey B, Schwartz D, de Graaf JB, Chauvel P, Liegeois-Chauvel C, 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin Neurophysiol 112, 1850–1859.

Gonzalez RC, Woods RE, 2002. Digital Image Processing. Prentice Hall, New Jersey.

Griffiths TD, Warren JD, 2004. What is an auditory object? Nat Rev Neurosci 5, 887–892.

Grinvald A, Hildesheim R, 2004. VSDI: a new era in functional imaging of cortical dynamics. Nat Rev Neurosci 5, 874–885.

Gröchenig K, 2001. Foundations of Time-Frequency Analysis. Birkhäuser, Boston.

Gruber WR, Klimesch W, Sauseng P, Doppelmayr M, 2005. Alpha phase synchronization predicts P1 and N1 latency and amplitude size. Cereb Cortex 15, 371–377.

Hadjileontiadis LJ, Panas SM, 1998. A wavelet-based reduction of heart sound noise from lung sounds. Int J Med Inform 52, 183–190

Hamada T, 2005. A neuromagnetic analysis of the mechanism for generating auditory evoked fields. Int J Psychophysiol 56, 93–104.

Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV, 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys 65, 413–497.

Hanes D, Schall J, 1996. Neural control of voluntary movement initiation. Science 274, 427–430.

Hari R, Salmelin R, 1997. Human cortical oscillations: a neuromagnetic view through the skull. Trends Neurosci 20, 44–49.

Harris FJ, 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. Proc IEEE 66, 51–83.

Harris KD, 2005. Neural signatures of cell assembly organization. Nat Rev Neurosci 6, 399–407.

Hayes MH, 1996. Statistical Digital Signal Processing and Modeling. John Wiley & Sons, New York.

Haykin S, 1999. Neural Networks: A Comprehensive Foundation. Prentice Hall, New Jersey.

Heath RG, 1972. Pleasure and brain activity in man. Deep and surface electroencephalograms during orgasm. J Nerv Ment Dis 154, 3–18.

Helmholtz HLF, 1853. Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche, Poeggendorffische. Ann Phys Chem 89, 211–233 and 354–377.

Hillebrand A, Barnes GR, 2002. A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. Neuroimage 16, 638–650.

Hillebrand A, Singh KD, Holliday IE, Furlong PL, Barnes GR, 2005. A new approach to neuroimaging with magnetoencephalography. Hum Brain Mapp 25, 199–211.

Hillyard SA, Squires KC, Bauer JW, Lindsay PH, 1971. Evoked potential correlates of auditory signal detection. Science 172, 1357–1360.

Huffman DA, 1952. A method for the construction of minimum-redundancy codes. Proc Inst Radio Eng 40, 1098–1101.

Ikeda H, Wang Y, Okada YC, 2005. Origins of the somatic N20 and high-frequency oscillations evoked by trigeminal stimulation in the piglets. Clin Neurophysiol 116, 827–841.

Jansen BH, Agarwal G, Hegde A, Boutros NN, 2003. Phase synchronization of the ongoing EEG and auditory EP generation. Clin Neurophysiol 114, 79–85.

Jansen BH, Bourne JR, Ward JW, 1981. Autoregressive estimation of short segment spectra for computerized EEG analysis. IEEE Trans Biomed Eng 28, 630–638.

Jaskowski P, Rybarczyk K, Jaroszyk F, 1994. The relationship between latency of auditory evoked potentials, simple reaction time, and stimulus intensity. Psychol Res 56, 59–65.

Jensen O, Vanni S, 2002. A new method to identify multiple sources of oscillatory activity from magnetoencephalographic data. Neuroimage 15, 568– 574.

Jervis BW, Nichols MJ, Johnson TE, Allen E, Hudson NR, 1983. A fundamental investigation of the composition of auditory evoked potentials. IEEE Trans Biomed Eng 30, 43–50.

Johansson RS, Birznieks I, 2004. First spikes in ensembles of human tactile afferents code complex spatial fingertip events. Nat Neurosci 7, 170–177.

Kähkönen S, Mäkinen V, Jääskelainen IP, Pennanen S, Liesivuori J, Ahveninen J, 2005. Serotonergic modulation of mismatch negativity. Psychiatry Res 138, 61–74.

Kaiser J, Birbaumer N, Lutzenberger W, 2002. Magnetic oscillatory responses to lateralization changes of natural and artificial sounds in humans. Europ J Neurosci 15, 345–354.

Kandel E, Schwartz J, Jessell T, (Eds) 2000. Principles of Neural Science (4th ed). Elsevier, New York.

Karakaş S, Erzengin OU, Başar E, 2000. The genesis of human event-related responses explained through the theory of oscillatory neural assemblies. Neurosci Lett 285, 45–48.

Kirschfeld K, 2005. The physical basis of alpha waves in the electroencephalogram and the origin of the "Berger effect". Biol Cybern 92, 177–185.

Klimesch W, Schack B, Schabus M, Doppelmayr M, Gruber W, Sauseng P, 2004. Phase-locked alpha and theta oscillations generate the P1-N1 complex and are related to memory performance. Brain Res Cogn Brain Res 19, 302–316.

Knuutila JET, Ahonen A, Hämäläinen M, Kajola M, Laine PP, Lounasmaa OV, Parkkonen LT, Simola JTA, Tesche CD, 1993. A 122-channel whole-cortex SQUID system for measuring the brain's magnetic fields. IEEE Trans Magn 29, 3315–3320.

Kramarenko AV, Tan U, 2002. Validity of spectral analysis of evoked potentials in brain research. Int J Neurosci 112, 489–499.

Krause CM, 1999. Event-related EEG changes during auditory processing. In: Event-Related Desynchronization. Pfurtscheller G, Lopes da Silva FH (Eds). Elsevier, Amsterdam, 195–201.

Kruse W, Eckhorn R, 1996. Inhibition of sustained gamma oscillations (35–80 Hz) by fast transient responses in cat visual cortex. Proc Natl Acad Sci USA 93, 6112–6117.

Kuhlman WN, 1978. Functional topography of the human mu rhythm. Electroencephalogr Clin Neurophysiol 44, 83–93.

Lange DH, Pratt H, Inbar GF, 1997. Modeling and estimation of single evoked brain potential components. IEEE Trans Biomed Eng 44, 791–799.

Lehtelä L, Salmelin R, Hari R, 1997. Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. Neurosci Lett 222, 111–114.

Lewicki MS, 2002. Efficient coding of natural sounds. Nat Neurosci 5, 356–363.

Linkenkaer-Hansen K, Nikouline VV, Palva J, Ilmoniemi RJ, 2001. Long-range temporal correlations and scaling behavior in human brain oscillations. J Neurosci 21, 1370–1377.

Linkenkaer-Hansen K, Nikouline VV, Palva J, Kaila K, Ilmoniemi RJ, 2004. Stimulus-induced change in long-range temporal correlations and scaling behavior of sensorimotor oscillations. Eur J Neurosci 19, 203–211.

Loveless NE, Brunia CMH, 1990. Effects of rise-time on late components of the auditory evoked potential. J Psychophysiol 4, 369–380.

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

Lugger K, Flotzinger D, Schlogl A, Pregenzer M, Pfurtscheller G, 1998. Feature extraction for on-line EEG classification using principal components and linear discriminants. Med Biol Eng Comput 36, 309–314.

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

Lyons RG, 2001. Understanding Digital Signal Processing. Prentice Hall, New Jersey.

Mainardi LT, Bianchi AM, Cerutti S, 2002. Time-frequency and time-varying analysis for assessing the dynamic responses of cardiovascular control. Crit Rev Biomed Eng 30, 175–217.

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ, 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Mäkelä AM, Alku P, Mäkinen V, Tiitinen H, 2004. Glides in speech fundamental frequency are reflected in the auditory Nlm response. Neuroreport 15, 1205–1208.

Mäkelä AM, Alku P, Mäkinen V, Valtonen J, May P, Tiitinen H, 2002. Human cortical dynamics determined by speech fundamental frequency. Neuroimage 17, 1300–1305.

Makhoul J, 1975. Linear prediction: A tutorial review. Proc IEEE 63, 561–580.

Mäkinen V, 2002. Methods and Measurement for Relating MEG Responses to Human Hearing. MSc thesis, Helsinki University of Technology.

Mallat SA, 1998. Wavelet Tour of Signal Processing. Academic Press, San Diego.

Malmivuo J, Plonsey R, 1995. Bioelectromagnetism – Principles and Applications of Bioelectric and Biomagnetic Sields. Oxford University Press, New York.

Marple SL, 1987. Digital Spectral Analysis. Prentice Hall, New Jersey.

May PJC, 1999. Memory Traces in Human Auditory Cortex. PhD thesis, King's College, London.

May P, Tiitinen H, 2001A. Human cortical processing of auditory events over time. Neuroreport 12, 573–577.

May P, Tiitinen H, 2001B. Adaptation and excitation effects underlying human auditory sensory memory. Proceedings of the 12th International Conference on Biomagnetism, Helsinki, 74–78.

May PJ, Tiitinen H, 2004. The MMN is a derivative of the auditory N100 response. Neurol Clin Neurophysiol.

Milotti E, 2002. 1/f noise: a pedagogical review. arXiv: physics/0204033.

Mitra PP, Pesaran B, 1999. Analysis of dynamic brain imaging data. Biophys J 76, 691–708.

Moore BCJ, (Ed) 1995. Hearing. Academic Press, San Diego.

Moss F, Ward LM, Sannita WG, 2004. Stochastic resonance and sensory information processing: a tutorial and review of application. Clin Neurophysiol 115, 267–281.

Murakami S, Hirose A, Okada YC, 2003. Contribution of ionic currents to magnetoencephalography (MEG) and electroencephalography (EEG) signals generated by guinea-pig CA3 slices. J Physiol 553, 975–985.

Murakami S, Zhang T, Hirose A, Okada YC, 2002. Physiological origins of evoked magnetic fields and extracellular field potentials produced by guinea-pig CA3 hippocampal slices. J Physiol 544, 237–251.

Näätänen R, 1992. Attention and Brain Function. Erlbaum, New Jersey.

Näätänen R, Picton TW, 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.

Neuhoff JG, 1998. Perceptual bias for rising tones. Nature 395, 123–124.

Neuper C, Pfurtscheller G, 2001. Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. Int J Psychophysiol 43, 41–58.

Niedermeyer E, LopesdaSilva F, (Eds) 2004. Electroencephalography: Basic Principles, Clinical Applications and Related Fields (5th ed). Lippincott Williams & Wilkins, Baltimore.

Nuttall AH, 1981. Some windows with very good sidelobe behavior. IEEE Trans Acoust Speech Signal Process 29, 84–91.

Nuwer MR, 1998. Fundamentals of evoked potentials and common clinical applications today. Electroencephalogr Clin Neurophysiol 106, 142–148.

Nyquist H, 1928. Certain topics in telegraph transmission theory. Trans Am Inst Elec Eng 47, 617–644.

Ozaktas HM, Barshan B, Mendlovic D, Onural L, 1994. Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms. J Opti Soc Am 11, 547–559.

Palomäki K, 2005. Studies on Auditory Processing of Spatial Sound and Speech by Neuromagnetic Measurements and Computational Modeling. Dr Tech thesis, Helsinki University of Technology.

Parasuraman R, Beatty J, 1980. Brain events underlying detection and recognition of weak sensory signals. Science 210, 80–83.

Parasuraman R, Richer F, Beatty J, 1982. Detection and recognition: Concurrent processes in perception. Percept Psychophys 31, 1–12.

Pardey J, Roberts S, Tarassenko L, 1996. A review of parametric modelling techniques for EEG analysis. Med Eng Phys 18, 2–11.

Pareti G, De Palma A, 2004. Does the brain oscillate? The dispute on neuronal synchronization. Neurol Sci 25, 41–47.

Parkkonen L, Mäkelä J, 2002. MEG sees deep sources: Measuring and modelling brainstem auditory evoked fields. Proceedings of the 13th International Conference on Biomagnetism, Jena, 107–109.

Parra L, Alvino C, Tang A, Pearlmutter B, Yeung N, Osman A, Sajdal P, 2002. Linear spatial integration for single-trial detection in encephalography. Neuroimage 17, 223–230.

Paul DD, Sutton S, 1972. Evoked potential correlates of response criterion in auditory signal detection. Science 177, 362–364.

Penny WD, Kiebel SJ, Kilner JM, Rugg MD, 2002. Event-related brain dynamics. Trends Neurosci 25, 387–389.

Percival DB, Walden AT, 1993. Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. Cambridge University Press, Cambridge.

Pfurtscheller G, Lopes da Silva FH, (Eds) 1999A. Event-Related Desynchronization. Elsevier, Amsterdam.

Pfurtscheller G, Lopes da Silva FH, 1999B. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110, 1842–1857.

Phillips DP, Hall SE, Boehnke SE, 2002. Central auditory onset responses, and temporal asymmetries in auditory perception. Hear Res 167, 192–205.

Picton TW, Woods DL, Proulx GB, 1978A. Human auditory sustained potentials I. The nature of the response. Electroencephalogr Clin Neurophysiol 45, 186–197.

Picton TW, Woods DL, Proulx GB, 1978B. Human auditory sustained potentials II. Stimulus relationships. Electroencephalogr Clin Neurophysiol 45, 198–210.

Popivanov D, Mineva A, 1999. Testing procedures for non-stationarity and non-linearity in physiological signals. Math Biosci 157, 303–320.

Qiu J, Paw UKT, Shaw RH, 1995. Pseudo-wavelet analysis of turbulence patterns in three vegetation layers. Boundary-layer Meteor 72, 177–204.

Quian Quiroga R, Garcia H, 2003. Single-trial event-related potentials with wavelet denoising. Clin Neurophysiol 114, 376–390.

Regan D, 1989. Human Brain Electrophysiology: Evoked potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, Amsterdam.

Renvall H, 2003. Temporal Processing of Sensory Information in Developmental Dyslexia: Neuromagnetic and Psychophysical Studies. PhD thesis, University of Helsinki.

Roberts TP, Ferrari P, Stufflebeam SM, Poeppel D, 2000. Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. J Clin Neurophysiol 17, 114–129.

Rossi S, Rossini PM, 2004. TMS in cognitive plasticity and the potential for rehabilitation. Trends Cogn Sci 8, 273–279.

Ryhänen T, Seppä H, Ilmoniemi RJ, Knuutila J, 1989. SQUID magnetometers for low-frequency applications. J Low Temp Phys 76, 287–386.

Sams M, Paavilainen P, Alho K, Näätänen R, 1985. Auditory frequency discrimination and event-related potentials. Electroencephalogr Clin Neurophysiol 62, 437–448.

Sanders LD, Newport EL, Neville HJ, 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat Neurosci 5, 700–703.

Sayers BM, Beagley HA, Henshall WR, 1974. The mechanism of auditory evoked EEG responses. Nature 247, 481–483.

Schaul N, 1998. The fundamental neural mechanisms of electroencephalography. Electroencephalogr Clin Neurophysiol 106, 101–107.

Schnitzler A, Gross J, 2005. Normal and pathological oscillatory communication in the brain. Nat Rev Neurosci 6, 285–296.

Schoffelen JM, Oostenveld R, Fries P, 2005. Neuronal coherence as a mechanism of effective corticospinal interaction. Science 308, 111–113.

Selig KK, 2002. Uncertainty principle revisited. Electron Trans Numer Anal 14, 165–177.

Shadlen MN, Movshon JA, 1999. Synchrony unbound: a critical evaluation of the temporal binding hypothesis. Neuron 24, 67–77.

Shah AS, Bressler SL, Knuth KH, Ding M, Mehta AD, Ulbert I, Schroeder CE, 2004. Neural dynamics and the fundamental mechanisms of event-related brain potentials. Cereb Cortex 14, 476–483.

Shannon CE, 1948. A mathematical theory of communication. Bell System Tech J 27, 379–423 and 623–656.

Shannon CE, 1949. Communication in the presence of noise. Proc Inst Radio Eng 37, 10–21.

Smith D, 2001. Digital Signal Processing Technology: Essentials of the Communications Revolution. American Radio Relay League, USA.

Sutoh T, Yabe H, Sato Y, Hiruma T, Kaneko S, 2000. Event-related desynchronization during an auditory oddball task. Clin Neurophysiol 111, 858–862.

Tang AC, Pearlmutter BA, Malaszenko NA, Phung DB, 2002. Independent components of magnetoencephalography: single-trial response onset times. Neuroimage 17, 1773–1789.

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O, 1995. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr Clin Neurophysiol 95, 189–200

Thakor NV, Tong S, 2004. Advances in quantitative electroencephalogram analysis methods. Annu Rev Biomed Eng 6, 453–495.

Thomson DJ, 1982. Spectrum estimation and harmonic analysis. Proc IEEE 70, 1055–1096.

Thorpe SJ, 1990. Spike arrival times: A highly efficient coding scheme for neural networks. In Parallel Processing in Neural Systems and Computers. Eckmiller R, Hartmann G, Hauske G (Eds). Elsevier, North-Holland, 91–94.

Tiihonen J, Hari R, Kajola M, Karhu J, Ahlfors S, Tissari S, 1991. Magnetoencephalographic 10-Hz rhythm from the human auditory cortex. Neurosci Lett 129, 303–305.

Tiitinen H, May P, Reinikainen K, Näätänen R, 1994. Attentive novelty detection in humans is governed by pre-attentive sensory memory. Nature 372, 90–92.

Torrence C, Compo GP, 1998. A practical guide to wavelet analysis. B Am Meteor Soc 79, 61–78.

Tseng SY, Chen RC, Chong FC, Kuo TS, 1995. Evaluation of parametric methods in EEG signal analysis. Med Eng Phys 17, 71–78.

Turner R, Howseman A, Rees GE, Josephs O, Friston K, 1998. Functional magnetic resonance imaging of the human brain: data acquisition and analysis. Exp Brain Res 123, 5–12.

Uusitalo MA, Ilmoniemi RJ, 1997. Signal-space projection method for separating MEG or EEG into components. Med Biol Eng Comput 35, 135–140.

Uutela K, Hämäläinen M, Somersalo E, 1999. Visualization of magnetoencephalographic data using minimum current estimates. Neuroimage 10, 173–180.

VanRullen R, Guyonneau R, Thorpe SJ, 2005. Spike times make sense. Trends Neurosci 28, 1–4.

Varela, FJ, Lachaux JP, Rodriguez E, Martinerie J, 2001. The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2, 229–239.

Varho S, 2001. New Linear Predictive Methods for Digital Speech Processing. Dr Tech thesis, Helsinki University of Technology.

Virtanen J, 1998. EEG Combined with MEG and TMS in Studies of Human Brain Function. Dr Tech thesis, Helsinki University of Technology.

Vrba J, 2002. Magnetoencephalography: the art of finding a needle in a haystack. Physica C 368, 1–9.

Wallace BE, Wagner AK, Wagner EP, McDeavitt JT, 2001. A history and review of quantitative electroencephalography in traumatic brain injury. J Head Trauma Rehabil 16, 165–190.

Welch PD, 1967. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust 15, 70–73.

Woody CD, 1967. Characterization of an adaptive filter for the analyses of variable latency neuroelectric signals. Med Biol Eng 5, 539–553.

Yeung N, Bogacz R, Holroyd CB, Cohen JD, 2004. Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. Psychophysiology 41, 822–832.