Signal validation in electroencephalography research

Signal Validation in Electroencephalography Research

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. M. Rem, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen op
maandag 17 januari 2000 om 16.00 uur

door

Maarten van de Velde

geboren te Purmerend

Dit proefschrift is goedgekeurd door de promotoren

prof.dr.ir. J.E.W. Beneken

prof.dr. C.H.M. Brunia

copromotor: dr.ir. P.J.M. Cluitmans

Drukwerk: Universiteitsdrukkerij TUE

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Velde, Maarten van de

Signal validation in electroencephalography research / by Maarten van de Velde. – Eindhoven : Technische Universiteit Eindhoven, 2000.

Proefschrift. – ISBN 90-386-1710-0

NUGI 743

Trefw.: biomedische signaalverwerking / patientbewaking / elektro-encefalografie.

Subject headings: medical signal processing / patient monitoring / electroencephalography.

Validation is the Appreciation of Quality —
Quality is Intrinsic, an Abstraction of Reality

~~~~~~~~~~ ~~~~~~~~~~

~~~~~~~~~~~~~~~~~~~~

~~~~~~~~~ ~~~~~~~~~
~~~~~~~~~

Table of Contents

Published as:
M. van de Velde, G. van ERP and P.J.M. Cluitmans. *Muscle artefact detection in the normal human awake EEG.* Electroencephalography and Clinical Neurophysiology, 1998, 107: 149-158.

Introduction

ELECTROENCEPHALOGRAPHY RESEARCH

New inventions and technological innovations have been a driving force in electroencephalography (EEG), from the early days of its discovery until today (Geddes, 1995). Where the early commercial machines were limited in application and rather bulky, today's EEG recorders can be pocketsize while incorporating high-quality amplifiers and many channels.

From a clinician's point of view, it can be surprising and maybe even startling to see the impact of our technology-driven society, regardless of the 'efficiency' reasons for installing new equipment. *Surprising*, because of the extent of procedural changes and possibilities of new, additional recording settings. However, such changes can also be *startling*, when the installed new technology appears to have an indirect, seemingly degrading effect on the quality of procedures. For instance, simple (and fast) browsing through EEG pages during computerised recordings is still a much-wanted feature. Up-to-date digital EEG technology can be astounding also because of *printing* of EEGs in the envisaged *paperless* EEG laboratory: printouts are often the only universal medium when going to a meeting. Another example: since no consensus implementation of the Rechtschaffen and Kales (1968) sleep staging procedures is available, the 'established' automatic assessment in one laboratory may lead to serious discussion in another.

What we have seen in the past 10 years is the gradual move towards all-digital EEG laboratories. An average laboratory can save thousands of Euro's per annum on paper costs alone, apart from the cost of storage space and inefficient access to patient data. Doctors and nursing staff are increasingly aware of computerisation, and are eager to use tools that can reduce their workload. This is now within reach because of the increasing power and availability of computers, which in the early digital years were busy doing the data acquisition part, and now have processor-time to spare to perform additional calculations.

However, before automatic interpretation can be used reliably in day-to-day clinical practice, the process of scrutinising the data for contaminations should be performed. Paper analysis does not allow for an easy separation of validation and interpretation. In

principle, this process *can* be done in digitised recordings. Validation is an often overlooked, seldom systematically addressed issue. In clinical practice, validation is performed inherently at the time of reviewing the data. In scientific research, data sets are often described as being pruned: by a neurologist at best, but still mostly based on empirical knowledge. This introduces a rather high degree of subjectivity in data review and subsequent processing. Repetition of experiments will then produce different results even in the same data set, which invalidates the eventual incorporation of the computerised methods.

EEG VALIDATION

The area in-between the actual recording and the interpretation (or advanced processing) of neurophysiological signals is the main topic of this thesis. *How well do we measure what we want to measure?*

Two different signals in EEG research are represented in the data sets that were used to test various issues related to measurement validation: the *electroencephalogram* itself, and *evoked potentials*.

Human experience of EEG analysis has taught us that it is difficult to separate validation aspects from interpretation/diagnosis. Still, we have used human experts as a reference for the performance of our automatic methods. Therefore, we have tried to design objective evaluation procedures for the human assessment. Special focus is on the aspects related to accuracy and signal context.

Evoked potentials (EPs) are deterministic signals that can be obtained from advanced processing of the EEG, often recorded during a repeated task. An evoked potential represents the electrophysiological behaviour of a specific neural pathway, as measured on the scalp. We took the validation of EPs one step further than validation of EEGs: we performed artefact detection, and especially focused on objective assessment of signal quality.

Thesis outline

This thesis is subdivided into two major parts, preceded by this brief introduction and followed by the final discussion.

Part I – Time-related aspects of EEG validation

Chapter I-1 is the introduction of Part I. This chapter provides an overview of the field of EEG, describing the background of EEG research in general. Different types of artefacts in the EEG are categorised by physiological source and external interference. A mostly

technical point of view is taken in describing the literature on signal processing; this chapter also describes some technical issues in relation to practical measurement procedures.

Chapter I-2 describes a study that was performed at the Department of Clinical Neurophysiology, Kempenhaeghe, Heeze, in twenty-one normal adult subjects. The study was designed to test the performance and accuracy of muscle artefact detection. Several (classical) detection parameters are compared in this chapter.

One of the optimal parameters in chapter I-2 is used again in **Chapter I-3**. Here, a large clinical data set was used from the international European project "IMPROVE", which was recorded in the intensive care unit of Kuopio University Hospital, Finland. This data set presents a very diverse set of patterns, including artefacts, comprising seven 24-hour recordings in severely ill patients. Artefact detection is performed here by combining two processing methods, based on statistical amplitude analysis and autoregressive modeling.

Part II – Quality estimation of evoked potentials

Chapter II-1 introduces Part II, describing the background of evoked potential measurements. Artefacts and procedures specific to EPs are summarised. A brief description of statistical methods for artefact detection is included from previous research at the Department of Medical Electrical Engineering, Eindhoven University of Technology (TUE/EME). The principles of quality assessment as used in the subsequent chapters are introduced.

Chapter II-2 focuses on signal quality and recording time of auditory evoked potentials as measured in a clinical study in forty-one patients undergoing cardiac surgery. This data set was obtained from a previous study by the TUE/EME group, which was performed at the Catharina Hospital Eindhoven. The EP waveforms were scored independently by several researchers, providing an interesting data set for quality assessment and EP validation. **Chapter II-3** applies the findings from chapter II-2 in event related (long-latency) evoked potentials, which have different characteristics and need additional procedures for artefact detection and quality assessment. The data were taken from an experiment in ten subjects at the Department of Physiological Psychology, Tilburg University. **Chapter II-4** concludes Part II, and constitutes an encore piece about the application and optimisation of an alternative technique (Cluitmans, 1990) to conventional evoked potential recording. This study was again performed at Kempenhaeghe: the EEG and auditory evoked potentials were measured in fourteen volunteers during sleep.

The overall discussion reiterates the major themes of this thesis and summarises the principal findings of both EEG and EP studies.

The projects described in this thesis were sponsored by grant no. 94CH from the Co-operation Centre Tilburg and Eindhoven Universities (Samenwerkings Orgaan Brabantse Universiteiten), grant no. EEL80.1538 from the Dutch Technology Foundation (Stichting Technische Wetenschappen) and by Dräger Medical Electronics, Best, The Netherlands.

Part I –

Time-related aspects of EEG validation

I-1. Neurophysiological Measurements: Overview and Methods

Dealing with artefacts in recordings of physiological signals is an integral part of clinical decision making. Of course, adequately trained staff must be involved in the acquisition and interpretation of such signals. Especially in neurophysiological measurements, the validation process is essential, as in general the recorded signals are of very low amplitude and are therefore easily disturbed by other physiological signals or external sources. Interpretation and analysis is extra complicated because some of the resulting artefacts in the electroencephalogram (EEG) can adversely mimic 'normal' patterns. Different clinicians and EEG researchers also show subjective differences in interpretation, and artefacts for one particular procedure can be the signal of interest in another investigation. Effective and *objective* (computerised) detection of artefacts is needed, not only because of the amounts of data that are acquired easily today, but also for the improvement of the *quality* of clinical procedures.

This general overview first describes the background of neurophysiological signals. Next, the (technical) requirements for recording of electroencephalographic data are described, focusing on practical issues.

Artefacts in the EEG can be produced by different sources, and are categorised accordingly. The various types of artefacts are described along with previous work in artefact processing. Subsequently the research in EEG validation and artefact detection is reviewed in a section about the different (existing) methods for signal processing, followed by a section that briefly introduces technical issues in clinical EEG processing. The discussion at the end of this chapter introduces the research paths taken in the next chapters of Part I of this thesis.

I-1.1. INTRODUCTION: EEG MEASUREMENTS

When Hans Berger conducted the first measurements of electrical activity from the human brain (Berger, 1929), various critics opposed that the recorded signal was due to other physiological or mechanical activity (e.g., blood pulsation, respiration, or skeletal muscle), hence, from *artefactual* origin. At that time, the recording equipment consisted of the string galvanometer as developed by Einthoven for recording of the electrocardiogram. Apart from lack of sensitivity in this apparatus, Berger also had to tackle the problems of electrodes, and, of course, of artefacts. Already then, he

investigated different materials, electrode positioning and fixation, and electrode impedance, providing guidelines that still relate to those of today. Quite accurately, he described the "Elektroenkephalogramm" as oscillations with average duration of 90ms (alpha waves) and 35ms (beta waves), and amplitudes of 70-150µV and 20-30µV respectively (Geddes, 1995).

Since the days of Berger and the verification of his recordings by Jasper and Carmichael (1935), electroencephalography has taken its place as a standard laboratory investigation in clinical neurophysiology and neurology. It is used in the diagnosis of brain pathology, e.g., epilepsy, sleep disorders, and disorders of the nervous system. EEG recording is also used extensively in psychophysiological research and in the testing of drugs (pharmacology) (Pryse-Phillips, 1997).

I-1.1.1. Neurophysiology

The origin of the EEG

The electrical activity elicited by single nerve cells stems from the electrochemical processes underlying the generation of 'action potentials', essential for information transfer between nerve cells. The neuron consists of a cell body, dendrites (receptor, or afferent pathway) and axons (efferent pathway). A resting potential exists across the cell membrane, at an intracellular level of approximately –70mV. Neurotransmitters can change the permeability of the membrane mainly for sodium (Na^+) and potassium ($K^{+)}$ ions. An increased ion influx causes depolarisation of the membrane, resulting in a further change in permeability (the *Hodgkin cycle*: Kandel *et al.,* 1991). This process can trigger an 'action potential' wave that travels along the axon towards other neurons. At the 'synapse', the contact point between axon and dendrite, the transmission of the nerve impulse occurs. Thus, processes in other cells are affected.

Postsynaptic potentials can be of excitatory or inhibitory nature, respectively causing a reduction (depolarisation) or increase (hyperpolarisation) of the membrane potential. These EPSPs and IPSPs are the primary origins for the EEG recorded from the scalp. Because EEG recording electrodes are relatively far from the source of these neuron potentials, the actual potentials of the EEG on the skull are approximately 100 to 1000 times smaller than intracellular levels. Moreover, the recorded activity on one electrode on the scalp represents the *averaged* behaviour of about one million neurons in the cortex (of the approximate 100 billion in the brain). Large amplitudes in the EEG therefore require *synchronous rhythmic activity* in such neuronal populations. The rhythmic activity (especially alpha rhythm) has its origin in the thalamus, a deep structure of the brain, and is modified by dynamic feedback loops of inhibition and excitation (Schmidt, 1985; Lopes da Silva, 1987a; Fischbach, 1992).

EEG signal characteristics

The dynamic behaviour, or rhythms, of electrical activity recorded from the brain can be classified by amplitude and frequency. Action potentials at the cellular level can be recorded at amplitudes up to 100mV. EEG phenomena recorded on the scalp range in amplitude between 1μV up to 200μV in features observed in sleep and epilepsy.

Table I-1.1 General characteristics of electrical brain activity: amplitude and frequency dynamics (after Cohen, 1995).

Classification		**amplitude**	**frequency**	specifics
Potentials at the cellular level:				
Action potential		100mV	100Hz-2kHz	*transmembrane potentials*
IPSP / EPSP		2-10mV	5-100Hz	
Electroencephalographic activity (on the scalp):				
EEG bands	delta (δ)	2-100μV	0.5-4Hz	*e.g., deep sleep*
	theta (θ)	80μV	4-8Hz	
	alpha (α)	50μV	8-13Hz	*awake, relaxed, closed eyes*
	beta (β)	20μV	13-30Hz	
Special phenomena:				
spindles		50-100μV	8-15Hz	*sinusoid waves during sleep*
K-complexes		100-200μV	3-14Hz	*bi- or triphasic sleep phenomenon*
seizures		100-200μV	2-50Hz	*'sharp wave' activity (epilepsy)*

The description of EEG activity in the frequency domain as above is sometimes considered inadequate when non-linear dynamics in the signal are investigated. Non-linear analysis techniques have been used for EEG simulation and for model-based EEG interpretation. These studies mostly focus on discrimination between different 'brain states', e.g., in sleep studies or drug therapy studies, or try to assess differences in cognitive activity (Gallez and Babloyantz, 1991; Pradhan and Narayana Dutt, 1993; Fell *et al.*, 1996). However, non-linear dynamic features of the EEG or specific phenomena should be related to certain events and/or compared to linear measures to assess their value. Non-linear methods need not yield results significantly different from linear processing methods (Blinowska and Malinowski, 1991), or may produce inconsistent results (Palus, 1996).

I-1.1.2. Technical aspects of EEG recording

Polygraphy

EEG registration generally consists of simultaneous measurement of multiple signals (sometimes more than 100 in EEG topography research), mostly recorded on the scalp. Electrocorticography (recorded directly from the exposed cortex) and in-depth intracranial recordings using needle electrodes (e.g., in neurosurgery, epilepsy) are not discussed in this overview.

Standardisation of all aspects of EEG recording is difficult because of the diversity in illnesses and monitoring applications. Different protocols and different equipment regulations exist between hospitals for similar investigations (also: different 'schools' of EEG training). Entirely different procedures may be needed in one single patient (e.g., ambulant recording, laboratory recording, monitoring during surgery).

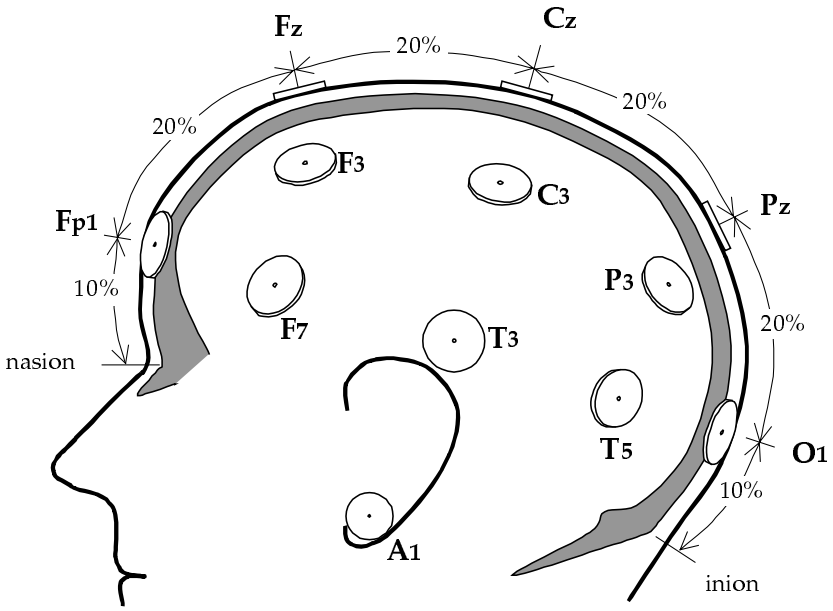

Figure I-1.1 The 10-20 electrode system: lateral view of left hemisphere. Electrodes on homologue positions over the right hemisphere are even-numbered: Fp2, F4, F8, C4, T4, P4, T6, O2, *and* A2 *on the earlobe (after Jasper, 1958).*

The only readily accepted instrumentation standard is the international 10-20 system for electrode positioning. The distance over the scalp at the midline between *nasion* and *inion* is used to position prefrontal electrodes (10%), frontal (30%), central (50%), parietal (70%) and occipital (90%) electrodes. Similar 10% and 20% inter-electrode distances are

used for positioning in the lateral direction. Additional positions and more closely spaced electrodes can be added if necessary (see Jasper, 1958; American Electroencephalographic Society, 1994b).

Apart from EEG channels, polygraphic measurements in neurophysiology may include other (electrical) signals. The EMG (*electromyography*, recording of muscle potentials) is recorded mostly on the chin (musculus mentalis) during sleep. EOG derivations (*electro-oculogram*, recording of eye movement potentials), horizontal and vertical channels, are used to detect REM sleep (§ I-1.4.1) and to identify eye movement artefact in EEG channels. Recording of the ECG (*electrocardiogram*, representing electrical activity of the heart) is standard. Monitoring clinical signs of other physiologic systems may be necessary during surgery, in intensive care, or for specific investigations.

Artefact prevention and event registration

An important issue for recording of high quality EEG signals is artefact prevention; standardised EEG recording techniques should be used as a guideline (e.g., American Electroencephalographic Society, 1986). Another obvious, but often overlooked approach to identify contaminations in the EEG is the direct monitoring of the source or external cause of the artefact. Separate recording devices may be needed to allow for retrospective identification of artefacts, e.g., the switching of electrocautery during surgery, or the sound levels of snoring during sleep. Proper annotation during the recording, or registration of such signals by means of event recording should be incorporated in the procedure (Lesser *et al.*, 1992; see the appendix of this thesis for a related paper on this subject).

I-1.1.3. Recording equipment

EEG electrodes

EEG recordings using scalp electrodes are most common. The skin below the electrode acts as impedance during the measurement, and is higher for low frequencies. Electrode impedance should be less than 5kΩ at 10Hz, and the EEG technician should strive towards equal impedance on all electrodes. Skin impedance can be reduced by proper preparation of the skin: scrubbing with a special paste or a blunt needle is common practice. Electrode leads should be as short as practically possible.

Classical types of EEG electrodes are metal cup electrodes, of silver/silver-chloride or gold, because of their favourable electrochemical properties. Newer self-adhesive electrodes or needle electrodes still have practical disadvantages, e.g., lasting visible marks on the skin, undesired sterilisation procedures (Neuman, 1995; Litscher *et al.*, 1996).

Electrode fixation. Because of drying of electrode paste and conductive gel electrode impedance will increase. Even worse, loose electrodes may result from patients moving

their heads (e.g., during sleep). Firm fixation of electrodes is achieved using glue (collodion) and adhesive pads. Especially in prolonged EEG monitoring, regular electrode (impedance) checks should be part of the recording protocol.

Amplification, filtering and digitisation

A concise, but detailed report of international recommendations for instrumentation standards can be found already in Barlow *et al.*, 1978. Today's standards mainly upgrade these specifications because of technical improvements. Some of the requirements are outlined below; more information is found in Spehlmann's EEG primer (Fisch, 1998).

The system must be equipped with a sufficient number of input channels (dependent on application, up to 64, or even 128 — Lesser *et al.*, 1992), with high input impedance (at least 10MΩ). *Calibration* pulses at microvoltage levels (e.g., 50μV) should be available to check the voltage scaling on screen or paper.

Electrode *impedance checks* must be available for all electrode channels. This facility must be highly accurate and allow for checks at different frequencies (0 to 1000Hz), at very low electrical currents through the electrodes during this measurement (IEC 601 regulation: below 10μA).

The *common-mode rejection* ratio of the amplifier must be sufficient to suppress noise and interfering signals synchronised at different electrode positions. At all possible input frequencies a rejection ratio of 1/10,000 or better is preferred. The common-mode rejection is altered when electrode impedances are not equal, which notably affects a *bipolar* recording. In this type of recording, the EEG is measured as the potential difference between the signal electrode and a 'reference' electrode with reference to a third electrode: the ground electrode (connected to the signal ground on the amplifier). In a true *unipolar recording*, the reference electrode and ground electrode are one. In practice, a virtual reference electrode may be obtained by deriving the 'common average' of all electrodes (e.g., Goldman, 1949). The difference in noise between signal electrode and reference electrode will be amplified in unipolar recordings, whereas 'common noise' at bipolar recording positions is not measured, provided that the amplification is equal at both signal and reference electrode.

Sample frequency. Several aspects are related to digitisation of the EEG. First of all, the sampling frequency (f_s) should be at least 100Hz for normal EEG recordings. An anti-aliasing low pass filter must be applied with a cut-off frequency well below the Nyquist frequency (half of the sample frequency), and a decay of 6 or 12dB/octave. A steeper decay can result in distortion of spikes and sharp waves near the cut-off frequency, and is therefore not advised (Spehlmann, 1981). Aliasing is the effect that occurs when bandlimiting was inadequate, e.g., a 50Hz noise component can corrupt the spectrum at a lower frequency of 30Hz for f_s = 80Hz (Mainardi *et al.*, 1995).

A low sampling rate of 100Hz may be chosen because of hardware constraints (e.g., Holter recording). In sleep recordings, and especially in epilepsy monitoring, the

investigated frequency range extends to 70Hz or higher, therefore requiring a sample frequency of at least 140Hz (but preferably higher, e.g., 200Hz). Very high sample frequencies of 5kHz up to 10kHz are required for instance in the recording of brainstem evoked potentials (see Part II).

AC-coupling. The characteristics for filtering of the lowest frequencies are usually given as the *time constant*, defining the time required for a return to 37% ($1/e$) of the DC baseline on square pulse input (DC/AC: direct/alternating current). The time constant TC corresponds to the low frequency cut-off point f by $TC = 1/(2\pi f)$. The most common time constants are between 0.05s and 1s; the effect of different settings is illustrated in Figure I-1.2, where low frequency artefact is completely abolished by TC=0.05s. When required, preservation of slow potentials can be controlled accordingly.

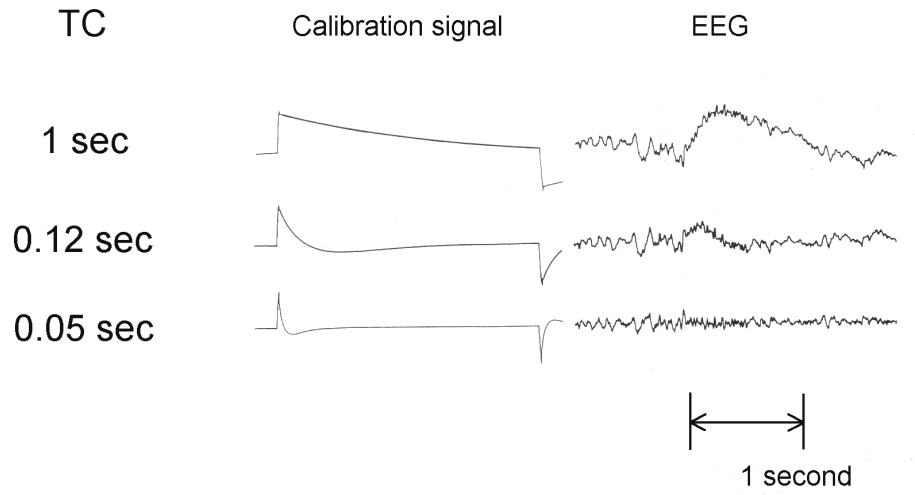

Figure I-1.2 Influence of filter settings: step-response and EEG example. Low frequency activity is suppressed by using shorter time constants (after Spehlmann, 1981).

The *input range* of the amplifier must be sufficient to record the normal amplitude range of the EEG. However, the amplifier's range should not be too small as it is often desirable to identify high amplitude artefact. In this respect, the *sensitivity* or signal resolution must be adequately chosen for the application. In the 1980's, 12-bit sample resolution was used in most research, whereas modern digital EEG equipment commonly uses 16-bit precision. For an input range of 2mV (i.e., –1000µV to +1000µV), 12-bit samples can represent the EEG at approximately 0.5µV resolution (72dB), versus 0.03µV resolution in 16-bits (96dB).

I-1.1.4. Practical issues in digital EEG

Standardisation. Digital EEGs are still relatively new for most neurophysiological laboratories. Standardisation is extremely important to exploit fully all advantages of the new technology. Standardisation of storage formats will greatly facilitate the exchange of EEG data (see e.g., Kemp *et al.*, 1992), which will advance the development and testing of (new) algorithms. An outline of practical design principles for digital EEG can be found in Lesser *et al.*, 1992 (but also in: Gorney, 1992; Burgess, 1993). Some of these issues will be highlighted below.

Archive issues

Apart from using a standard data format, the amount of data acquired during the recording must not be overlooked. A one-hour recording of 2 channels, 16-bit EEG at 100Hz sampling rate already results in 1.4 MB of data. A sleep investigation of 10 channels of 250Hz over a 10-hour period yields approximately 175MB. The archive systems (e.g., digital archive tape or a compact disk production unit) and infrastructure (computer network) must be dimensioned appropriately to handle these data streams.

Although the cost of archiving is ever decreasing per unit of storage space, data compression may be needed for economic use of available capacity in a specific environment. An excellent comparison of compression techniques by Antoniol and Tonella (1997), shows that in a 20 channel 8-bit 128Hz EEG recording a reduction of up to 58% is possible. This compression allowed them to send the data in real-time over a relatively low-performance telephone line, thus reducing cost of transmission time.

Visualisation of the EEG

The convenience of paper is often overlooked in the transition of EEG ink-writer recorders to a paperless, digital computer system. As obvious as it may seem, paper is one universal medium, allowing for high contrast, high resolution viewing. Display resolution on the computer screen should be as high as possible (e.g., at least 1280x1024 pixels for simultaneous display of ±10 channels) to approach paper resolution (Risk, 1993). Of course, paperless EEG has major advantages over paper because the user can zoom in on the data both in time and resolution. Display programs should at least allow for this horizontal and vertical zooming.

Reduced display resolution clearly affects the visibility of high frequencies in the EEG (Hirshkowitz and Moore, 1994). The display routine may need an anti-aliasing filter for proper data display; otherwise condensed display of data may result in spurious 'down-sampling' errors because of reduced pixel resolution.

Facilities such as automatic paging (forwards and backwards in time) are highly recommended for convenient browsing through the data (see e.g., Collura *et al.*, 1993; Thomsen *et al.*, 1997). Preferably, these facilities should be available in real time during

the recording. To overcome computational overload and to reduce the risk of data loss, most recording systems dedicate one computer to the on-line display of data while using another for signal manipulation (menu/algorithm selection; Lesser *et al.*, 1992). A useful feature of digital EEG is that it allows for redefinition of the electrode derivations. For instance, storing all channels with respect to a common average reference (discussed on p. 12) allows for recalculation of any derivation in retrospective analysis. This can be used to enhance certain characteristics in the EEG or to reduce the visual effects in the display of some artefacts, e.g., in topographic mapping of the EEG (Duffy *et al.*, 1994).

I-1.2. EEG ARTEFACTS

In this section, the different types of EEG artefacts are classified by their source of origin. Although in general we can recognise and distinguish artefacts of physiological origin, one should keep in mind the possibility of artefact mimicking EEG activity. For instance, part of the frequency characteristics of muscle artefact lie within normal EEG frequency range. Conversely, in electrocorticography, true brain activity on the cortex may seem artefactual (Barlow, 1986b).

I-1.2.1. Artefacts of non-physiological origin

Electrode artefacts

Electrode artefacts are more frequent when the electrode impedance increases during the recording of the EEG. Therefore, electrode impedance must be kept low throughout the measurement (p. 11). Electrode attachment, conductive gel and glue are always first to be checked when artefacts occur. But also the electrode head box (or jack box) and electrode leads are possible sources of artefact. Leads should not be curled and not be touched during the recording. Artefacts from poor connections are observed more frequently in ambulatory 'cassette' EEGs (Jayakar *et al.*, 1985).

Electrodes (leads) may pick up 50Hz (or 60Hz) sine wave patterns from the main power supply or other equipment, thus obscuring the EEG. Somewhat less severe mains interference is often caused by unequal electrode impedance at different positions, or by improper grounding of patients (Spehlmann, 1981). However, correct grounding may conflict with electrical safety regulations in some situations, e.g., in the operating room.

A characteristic artefact is the 'electrode pop', which is due to a sudden change in electrode contact resulting in a sharp spike in the recorded signal. A special filter can be developed to eliminate these artefacts, which resemble the calibration pulse of the EEG amplifier (Barlow, 1986a; see Figure I-1.3). This involves detection of the sharp leading edge, measurement of its amplitude, and subsequent subtraction of a generated

waveform of equal amplitude. However, the correction circuitry could trigger also at other spike-type EEG activity, introducing spurious new 'pops' in the signal (Barlow, 1986b).

DC drift, or base-line swaying artefact, is either related to changing electrode impedance or caused by movement of leads. Most often this causes very low frequency patterns in the recorded EEG, and therefore may be adjusted by (temporarily) decreasing the time constant of the recording equipment (when other adjustments failed). This is not an option of course, when focusing on slow potentials in the EEG.

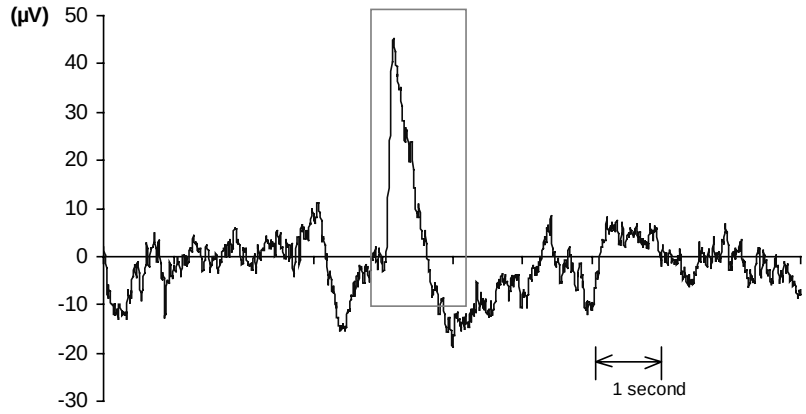

Figure I-1.3 Example of electrode 'pop'-artefact: a sudden sharp edge in the recorded signal, followed by an exponential decay, obscuring the EEG.

Equipment artefacts

A whole range of electrical apparatus can cause mains interference when electrical shielding or grounding is insufficient. Care has to be taken in the placement of patient and equipment, shielding for existing and possible electrical fields. Notorious are fluorescent lights, even worse when under variable intensity control. The electrical field around power cables, transformers, and antenna-equipped devices can possibly be picked up on electrode leads. During surgery, electrocautery by the surgeon generally causes high amplitude, high-frequency artefact in the EEG, rendering the signal useless for interpretation. Special cables, leads and filters can be used to reduce (50/60Hz) interference artefact, especially when wires must be placed in the vicinity of other equipment (Straw *et al.*, 1967; Van der Weide and Pronk, 1979; Ferdjallay and Barr, 1994; Stecker and Patterson, 1996).

Other machinery, e.g., respirators, perfusion pumps, and other (mechanical) actuators such as flush devices, cutting, drilling, suctioning, rubbing and washing can cause

(rhythmic) artefacts. Sims and colleagues (1973) already demonstrated that artefact from the respirator can be extremely variable in form, and can mimic for instance EEG 'burst' activity in a pattern called 'burst-suppression' (also see Klass, 1995).

In an intensive care environment EEG recordings can be troubled by the intravenous (i.v.) line, or by recurring measurements such as regular blood pressure checks. An intravenous infusion bottle can cause 'drip' artefact in the form of spikes in the EEG. This is most probably caused by a static charge on the i.v. fluid, resulting in a (small) electrical vibration upon each drop of the liquid. Grounding of the metal i.v. needle may diminish this artefact (Barlow, 1986b), however, as mentioned above, this may conflict with safety regulations.

Interference artefact can also be caused by mobile phones near the EEG recording equipment. Increased transmission power (e.g., phones of 5 Watt) necessitates a safe distance of at least 2 meters (Robinson *et al.*, 1997).

Spikes may be introduced by malfunctioning recording equipment, especially when high sample frequencies are used and data acquisition boards have to operate at performance limits. Usually such spikes are relatively easy to detect, and can be eliminated from the recording by interpolation of the signal values immediately preceding and following the spike (Cluitmans *et al.*, 1993).

I-1.2.2. Artefacts of physiological origin

As a general observation, physiological artefacts are greatly reduced in relaxed subjects. Proper patient information and a comfortable environment help to reduce muscle tension and anxiety.

Muscle artefact

Large signal disturbances can occur in the EEG from muscle activity, i.e., movement of the head, body and limbs, or from tension in the facial muscles, or from the tongue or jaw (e.g., clenched teeth, Keeney, 1981). While undesired movement of a subject in an experimental setting can be prevented mostly by clear instructions, involuntary movements or anxiety are often difficult to suppress (e.g., muscle tremor, shivering). Sustained muscle artefacts are caused by muscle tension in the face (e.g., frowning), neck, and also on the scalp (smaller in amplitude), or by repetitive actions such as talking, or chewing and swallowing while eating (large amplitudes).

Artefact caused by voluntary movement of the tongue (glossokinetic artefact) can generate a negative shift at the vertex of the scalp of 100μV. Brief muscle artefacts are caused by muscle twitches or brisk movements (Barlow, 1986b; Klass, 1995).

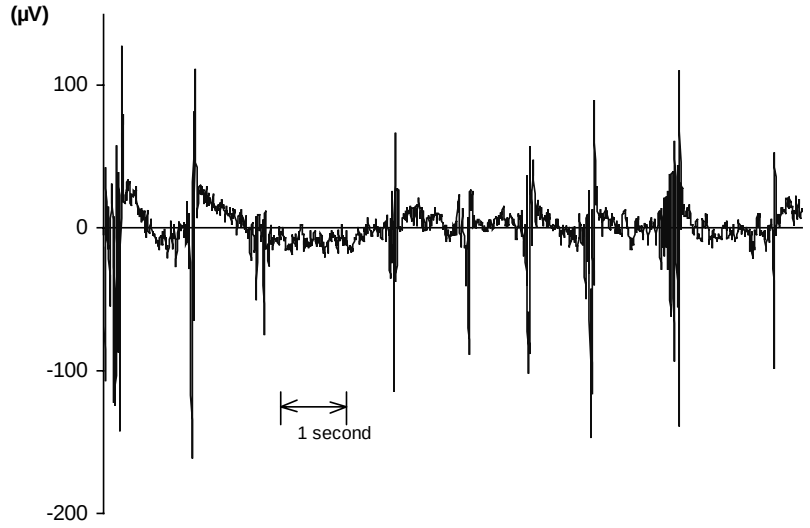

Figure I-1.4 Example of muscle artefact: chewing.

Eye movement

The eye acts like an electrical dipole in EEG recordings, being positive at the cornea, negative at the retina. Eye-artefacts are most prominent during REM sleep (§ I-1.4.1), but can also contaminate the EEG during drowsiness or light sleep. In awake state, blinks of the eye consist primarily of eyelid movement, which causes less significant artefact than movement of the eyeball. However, rhythmic activity (alpha range) can be observed in the EEG because of fluttering of the eyelid.

Eye movements cause a characteristic pattern in the EEG (see Figure I-1.5), consisting mostly of low frequency activity, and is distinctly observed in the EEG recorded from anterior (especially prefrontal) positions on the scalp. Recording of the EOG (electro-oculogram) using electrodes close to the eyes can be used as a reference, or, when not available, electrode positions F7, F8 may be used to ascertain detection of suspected eye artefact (Klass, 1995).

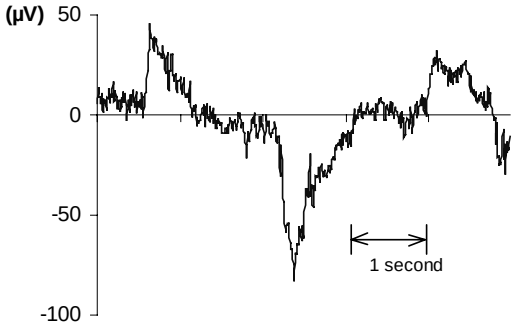

Figure I-1.5 Eye movement artefact in the EEG (position F7, common average reference).

Cardiovascular activity

Electrical activity from the heart can appear in the EEG, especially during recordings of very low voltage. Simultaneous recording of the electrocardiogram allows for the recognition of cardiac artefact in the EEG. Very sharp spikes can be observed from stimulation pulses generated by cardiac pacemakers (Brittenham, 1990).

Artefacts originating from the heart or circulatory system are sometimes caused by inadequate electrode placement. It is picked up more easily in the EEG recorded from wide inter-electrode distances, especially across the head to the left ear, and in subjects with short necks. Alternative reference electrode placement can be used to reduce this type of artefact (Barlow, 1986b; Spehlmann, 1981).

Respiration effects

The actions of the respiratory system can also cause some low frequency artefact in the EEG signals as recorded on the scalp because of the rhythmic movement of the chest, neck, and head. When using nasopharyngeal electrodes (inserted through the nasal cavities, for the localisation of an epileptic focus), rhythmic, fast activity artefact can be a major problem because of vibrations in the soft tissues near the electrode (Barlow, 1986b). The influence of respiration on the EEG is larger in small children. Excessive artefact from respiration effects may be observed during snoring.

An unusual, respiration related artefact was observed by Sahota *et al.* (1993) in the recording of a comatose child: semi-rhythmic activity of 5-7Hz occurred because of vibrations in the endotracheal tube of the ventilator system attached to the patient.

Skin impedance

Sweating is the main cause of the artefacts in the EEG originating from changes in skin impedance (electrodermal artefact). Sweat can also affect the conductive properties of electrodes through dissolving of the electrode gel. In the recorded signal on the scalp this will cause lasting effects when the electrode impedance eventually becomes too high (e.g., interference artefact, or slow rhythmic swaying of the base-line DC levels).

The effects of sweating can be reduced by adequate preparation of the skin before the procedure, and by recording in a comfortable position, in a cool environment (not too cool, because of risk of shivering). During sleep, 'base-line shift' artefact can also be associated with arousal (Barlow, 1986b).

I-1.3. EXISTING METHODS FOR ARTEFACT DETECTION

I-1.3.1. Time domain analysis of the EEG

Amplitude histograms of the EEG often show a symmetrical, essentially Gaussian distribution. The characteristics of Gaussian distribution can be summarised by calculating the amplitude mean and variance, and skewness (characterising symmetry) or curtosis ('flatness') measures, which can be used to detect different types of EEG or artefacts (Bronzino *et al.*, 1980). Rules for artefact detection are often based on amplitude thresholds that have been determined empirically. For instance, in the research by Flooh *et al.* (1982) an artefact was defined as the EEG amplitude exceeding a threshold six times the average amplitude of the preceding 10 seconds. This was combined with a minimal duration criterion for detected artefacts to ensure that an artefact was detected in its entirety. Arvidsson *et al.* (1977) reached 80% correct detection of artefacts based on 10 EEG features in the time domain. The detection rules were also determined empirically, and were mainly based on covariances between the calculated features and those obtained in a previous visual evaluation.

Although fixed amplitude thresholds may be used as a basic procedure, they can be very unspecific in the identification of certain artefacts (e.g., missed eye artefact at a 50μV detection threshold: Verleger, 1993). Still, max-min amplitude criteria are often selected because of simplicity (e.g., Kirkup *et al.*, 1997). Another very basic time domain procedure for *correction* of artefacts is subtraction of the average amplitude to correct for DC offset (at the output of the EEG amplifier) or base line swaying. This must be applied with care when investigating low frequency phenomena, e.g., in some evoked potential investigations (see part II of this thesis, § II-1.1.2).

Slope thresholds. Several successful studies were mentioned by Barlow (1979, 1986b) in which the first derivative or slope, as well as the second derivative was applied in the detection of fast activity. A slope (also called 'steepness') threshold was shown to be useful in the detection of (muscle) spike artefact (Scherg, 1982b; Cluitmans *et al.*, 1993). This has been implemented as a simple differentiator circuit (Barlow, 1983), where filtering of muscle spikes was achieved using a sample-and-hold circuit. This performed poorly on continuous muscle interference patterns, but its corrective properties could be improved by using an adaptive algorithm (Panych *et al.*, 1989). A danger of significant non-linear distortion exists in this method because of excessive 'hold' operations.

Other time domain parameters have been defined to capture frequency-related characteristics of the EEG. For instance, Hjorth's *normalised slope descriptor* 'mobility' is calculated as the standard deviation of the first derivative divided by the amplitude standard deviation. The first and second derivatives are used to calculate a 'complexity' measure, which can be interpreted as an estimate of signal bandwidth (Hjorth, 1970). The

descriptors are also related to the technique of computing and analysing the 'zero-crossing' frequency of the EEG (Saltzberg and Burch, 1971). In spite of their relative simplicity, Hjorth's and zero-crossing parameters have been useful in the exploration of prolonged EEG recordings in sleep, critical care and surgery, and can discriminate between normal and abnormal EEGs (see e.g., Pronk, 1982).

I-1.3.2. Frequency domain processing

Hans Berger already started description of EEG features in terms of rhythmic activity and frequency. The frequency spectrum of the EEG has been easy to compute since the publication of the Fast-Fourier Transformation (Cooley and Tukey, 1958; implementation issues described in Bracewell, 1986, after Hartley, 1942).

Frequency parameters were employed for rejection of artefacts in normal and abnormal EEGs by Gevins *et al.* in 1977. Individual algorithms (described in Gevins *et al.*, 1975) were used to detect head and body movements, large muscle potentials and eye-movement potentials. The system correctly detected only 65% of the artefacts identified by the consensus of three experts, versus 85% consensus between experts. The false positive rate for the computer system was rather high: 44%; however, no statistical difference was indicated. The investigators considered the algorithms inadequate for routine application in clinical EEG (Barlow, 1986b).

As shown by Levy (1980b) power spectrum information can be indispensable for the identification of artefacts: a rhythmic equipment artefact did not reveal itself clearly in unprocessed EEG traces, but did present a distinct regular pattern of higher harmonics in the spectrum.

Characteristic frequencies. A parameter that quantifies one aspect of frequency characteristics is the 'spectral edge frequency' or SEF (Rampil *et al.*, 1980), used in assessment of anaesthesia levels. SEF is defined as the highest frequency at which a significant amount of energy is present in the EEG (usually calculated at 90-95% of total power contents). Accordingly, the 'median peak frequency' (MPF) is located at the 50% energy level (see Figure I-1.6). However, the calculation of SEF is insensitive to changes in lower frequency bands and too sensitive to high-frequency spike activity; hence, it is not per se a reliable parameter. Both MPF and SEF can show large variances and inconsistent results (Thomsen *et al.*, 1991; Van de Velde and Cluitmans, 1991; De Beer *et al.*, 1992).

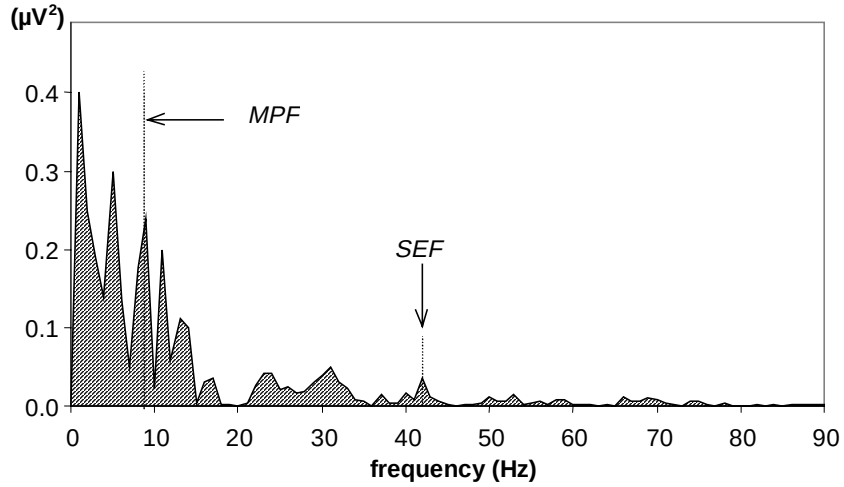

Figure I-1.6 EEG spectrum characteristics: median power frequency (MPF) and spectral edge frequency (SEF), at 50% and 95% of the power contents.

Short-time frequency analysis. The EEG power spectrum is usually calculated by applying frequency transformation to successive *epochs* of data. Longer epochs provide better frequency resolution, but then restrict time resolution; and too lengthy epochs may not be stationary (§ I-1.3.3). An improved time resolution can be obtained by using alternative *short-time* calculations, e.g. by shifting epochs forwards in time over short intervals (Kawabata, 1973: 0.5s). Very short-time shifts (e.g., 10ms) provide improved recognition in fast changing signals such as EMG (electromyogram) or speech signals (Hannaford and Lehman, 1986), or in epileptic seizure analysis (Williams *et al.*, 1995; Shamsollahi *et al.*, 1996).

Filtering

Filtering specific frequency bands from the EEG can be used to reduce muscle activity or mains interference. Muscle artefact is generally characterised as a (relatively) high-frequency phenomenon. However, heavy low pass filtering, using a cut-off frequency as low as 12.5Hz, is necessary to make sure that residual muscle activity can not resemble EEG beta activity. Obviously, this is not desirable in most recordings, because true beta activity and spike-type activity will be attenuated or even obscured. Modern adaptive algorithms can demonstrate very efficient artefact filtering while leaving intact important EEG features (Panych *et al.*, 1989; Neejärvi *et al.*, 1993; Roessgen *et al.*, 1993). Still, simultaneous display of unfiltered and filtered signal is always the safest solution for human interpretation (Barlow, 1986b; Klass, 1995).

Reduction of mains interference artefact (50/60Hz) by means of filtering should be used only if other corrective measures have failed. When the interference is persistent and significantly obtrusive, filtering must be applied. For this purpose, the use of an adaptive digital notch filter (rather than a fixed notch filter) can be advantageous (Ferdjallay and Barr, 1994).

Higher order spectra

Deviations from Gaussianity are characterised in the frequency domain by phase coupling of different frequencies, which can not be detected in the normal frequency transform (based on the autocorrelation, a second order function). Therefore, the calculation of higher order spectra is applied for investigation of non-linear behaviour in the EEG, of which especially the Fourier transform of the third-order correlation function has received much attention. This transformation is known as the 'bispectrum' (Bronzino, 1995).

Higher order spectra have been studied extensively for the identification of non-linearities in various signals (Cho *et al.*, 1992; Dalle Molle and Hinich, 1995), and can be useful for the detection of EEG of a transitional nature. For instance, a burst-suppression pattern in the EEG (see Figure I-1.11, p. 33) is characterised by non-linear phase coupling of frequencies (Muthuswamy *et al.*, 1999). Also for intraoperative monitoring of the level of anaesthetic suppression the bispectrum contains useful information. In a comparison study of a conventional spectrum and the bispectrum of the EEG recorded at three anaesthetic levels, the latter resulted in better detection of an intermediate level (Watt *et al.*, 1995: 83% versus 67% correct detection). However, a reliable measure of 'adequacy of anaesthesia' must include a combination of techniques (De Beer, 1996), and in this respect the bispectrum alone does not provide enough information. This is also illustrated in a measure known as the 'bispectral index', which uses the bispectrum, normal spectrum and time-domain EEG descriptors, and is correlated with data from a large database (Sigl and Chamoun, 1994; Todd, 1998).

I-1.3.3. Stationarity analysis

Stationarity

Statistics. Almost all EEG processing is based on parametric modeling of *periods* in the EEG time series. The investigated periods should more or less have constant statistical properties to validate such '*epoch based*' parameter extraction. We have already seen that derived features of EEG epochs are often hypothesised to follow a Gaussian, or normal distribution; this is usually taken as a requirement for stationarity. Different statistics have been developed to test the hypothesis theoretically, both in the frequency domain and in the time domain (Moulines *et al.*, 1993). The latter domain is preferred for

stationarity testing of short epochs, because of reduced frequency resolution in small sample sizes.

Testing can be performed by using the Chi-square test, or better, by using a Kolmogorov-Smirnov test (Lilliefors, 1967). The Shapiro-Wilk statistic (1965) provides even stricter testing of normality (Shapiro *et al.*, 1968; Gasser, 1975). Weiss (1986) again modifies and advocates a test procedure using Kolmogorov-Smirnov.

In signal theory, stationary stochastic processes are defined as processes whose statistics do not change in time (Cohen, 1995). Theoretically this can only be tested in an 'ensemble of (multiple) realisations' of the same process, and can therefore not be implemented in a real-life signal. In order to overcome the impossibility of ensemble testing, the signal is often assumed to be 'ergodic'. Ergodicity relates to processes in which every sequence or sizeable sample is equally representative of the whole. In practical terms, we can say that a stationary signal must have *several* time-invariant properties, requiring explicitly more than just one feature to describe the EEG signal (Salden, 1997). Apart from the standard statistical tests, the use of customised signal features provides a more pragmatic approach towards stationarity testing. We will further investigate this approach in chapter I-3.

Epoch length. An important issue is the choice of epoch length for appropriate processing of the EEG signal. The number of samples available in an epoch is of influence on the test for stationarity, because of varying interdependence between adjacent samples. Increased *in*dependence among cortical neural generators (less synchronous activity) may be related to increased Gaussian 'noise-like' EEG as recorded on the scalp (Elul, 1969).

Short epochs (1-2 seconds) are advised for EEG processing as best guarantee for 'wide-sense stationarity' (McEwen and Anderson, 1975), especially when using low sample frequencies (Persson, 1974, 1977; Fang *et al.*, 1987).

Autoregressive modeling

Autoregressive (AR) based modeling estimates the linear correlation to preceding samples of a discrete time series s_t which can be expressed as:

$$s_t = \left(a_0 + a_1 s_{t-1} + a_2 s_{t-2} + \ldots + a_p s_{t-p} \right) + e_t \qquad \text{(I-1.1)}$$

(a_0: DC-offset; a_1 .. a_p: AR coefficients; e_t: residual error at time t)

Zetterberg (1969) was the first to introduce AR models for EEG representation, and indicated that a model of order $p=5$ is already sufficient to give an accuracy of 5% in the AR parameters. Gersch (1970) showed that AR modeling can be used to obtain the frequency spectrum, providing improved resolution over other frequency transforms (Fenwick *et al.*, 1971), where higher orders ($p=10$) should be used for optimal accuracy (Jansen *et al.*, 1981a). Prefiltering of specific bands can further enhance peaks in the AR spectrum estimation (Narayana Dutt, 1994).

Normality of the EEG signal can be investigated from the residual errors already when using first or second order AR estimation (Pierce, 1985). Testing of residuals of a (very) high order AR model ($p=30$) was used in a more recent study to verify normality in EEGs obtained during anaesthesia (Bender *et al.*, 1992: using the Shapiro-Wilk statistic).

AR models describe the EEG at least as good as non-linear methods (Blinowska and Malinowski, 1991). Already a few spectral estimates or simple distance measures, calculated from successive AR estimations, can be sufficient for the detection of EEG changes (Goel *et al.*, 1994; Kong *et al.*, 1997, 1999). AR methods are very good at detection as well as simulation of non-stationarities (Vachon *et al.*, 1978; Kaipio and Karjalainen, 1997), and are therefore appealing for use in EEG validation studies.

Algorithms using autoregressive modeling can be implemented in numerous ways. This includes the generalised moving averaging (ARMA) model, where the error term e_t is also included in subsequent estimations:

$$s_t = (a_0 + a_1 s_{t-1} + \ldots + a_p s_{t-p}) + (b_1 e_{t-1} + \ldots + b_q e_{t-q} + e_t) \qquad \text{(I-1.2)}$$

Adaptive AR models such as the Kalman filter update the coefficients for every sample and are therefore also applicable to non-stationary data; this method is computationally more expensive and produces more data than consumed (Pardey *et al.*, 1996).

I-1.3.4. Other artefact detection techniques

EEG classification

An adaptive segmentation of the EEG into 'elementary patterns' can be achieved using a spectral error measure (SEM) derived from an AR model (Praetorius *et al.*, 1977). This was later extended into a complete clinical system that graphically summarised characteristics of EEG activity based on clustering of segments (Bodenstein *et al.*, 1985). The SEM method was not designed to indicate artefacts directly, but it can be used to detect for instance eye artefact (Dunstan and Marshall, 1991). A useful complementary screening program was developed by Krajca *et al.* (1991), which conveniently assists in the visual inspection of (over-)detected features.

Artificial intelligence (AI) and neural networks. The method used by Krajca included a fuzzy algorithm, allowing imprecise class descriptions, to obtain improved clustering of segments. An AI approach using fuzzy rules was already suggested by Jagannathan *et al.* (1982). Earlier, his group obtained an 80% agreement with visual classification in sleep EEG using AR modeling on 1s segments (Jansen *et al.*, 1981b). They added an artefact detection step (heuristic rules) that was quite successful in recognising 'low frequency' artefacts (Jansen *et al.*, 1982).

Another classification method is modeling through artificial neural networks (ANN). In ANNs, non-linear processing elements (PEs) are interconnected with weight factors in

different (hidden) layers. The application of ANNs is foremost in pattern recognition or prediction tasks, and in process control. When well managed, and by exploring the inner structures, they can also be applied to gain insight in multivariate data (Van Gils *et al.*, 1997a). ANN methods are often computation intense, and extensive training is needed in order to provide robust results. However, ANN structures can be used to simulate EEG waveforms, where up to thousands of PEs are needed to simulate epilepsy or awake EEGs (Mukesh and Nadkar, 1997). One particular system used ANNs for identification of artefacts (Wu *et al.*, 1997). However, this system was not tested on clinical EEG.

Wavelets. Wavelet analysis is successful where the time-frequency relationship is critical for optimal characterisation of waveforms. Where conventional Fourier transformation decomposes a signal in a sum of sine and cosine functions, a family of orthogonal Wavelet templates can focus on specific transient features (Thakor and Sherman, 1995; Bruce *et al.*, 1996). For example, Wavelet analysis has been successful in the detection of epileptic spikes (Clark *et al.*, 1995). Wavelet techniques are mostly used for detection of known waveforms in a noisy background signal.

Multichannel correlation analysis

Simple computerised correction of eye artefacts can be performed by subtraction of special EOG channels in other channels (see e.g., Barlow, 1986b). However, one must recognise the fact that the EOG reference signals can contain different types of eye artefacts (vertical, horizontal movement, blinks), or an EEG component that can be subtracted inadvertently from correlated EEG activity in other channels. Low pass pre-filtering of the EOG trace may enhance the robustness of these methods.

Successful correction of characteristic eye artefacts requires synchronous recording of multiple EOG channels, sometimes combined with a calibration phase (Van den Berg-Lenssen *et al.*, 1989; Brunia *et al.*, 1989). Other interesting techniques include ANN estimation of filter coefficients (Sadasivan and Narayana Dutt, 1994), eye source-waveform estimation from the EEG and subsequent subtraction (Berg and Scherg, 1994), or extraction of EEG components by independent component analysis (Vigário, 1997).

Multichannel signal processing has been found especially useful in epilepsy, for background EEG cancellation (James *et al.*, 1997), or to improve detection of characteristic spikes (see e.g., Glover *et al.*, 1989). This will be discussed in the next section, § I-1.4.2.

Reference signals

A subtraction method can be applied to reduce *ECG artefact*, similar to EOG subtraction as mentioned above. The relatively regular occurrence of the heart signal, as well as its more consistent waveform, makes it easier in principle to correct this artefact in the EEG. However, *artefacts in the ECG recording* can dramatically alter the EEG during 'correction'. Subtraction of a template ECG-complex (or average of preceding waves), using correlation analysis is preferable (Barlow, 1986b). This method was particularly

successful when processing all signals in digital instead of analogue form (Nakamura *et al.*, 1990).

Subject behaviour and environment. Other artefacts may be detected outside the realm of (EEG) signal processing. Practical solutions may be found from less than obvious sources.

In sleep recordings, the simplest form of artefact recognition (and 'elimination': Barlow, 1985) is the detection of excessive movement by means of an accelerometer. Such a device is also recommended for monitoring of body position in the investigation of snoring (Bloch, 1997), where sound analysis can identify the different types of snoring (Fiz *et al.*, 1996). When respiration monitoring is not possible or less reliable, a procedure using a microphone could possibly help to pick out these artefact-prone periods.

Diverse types of artefacts have been reported from diverse causes: e.g., sobbing, nystagmus (oscillation of the eyeballs) causing saw-tooth waves, etc., even lightning and earthquake artefacts (Keeney, 1981; Klass, 1995).

I-1.4. ISSUES IN CLINICAL EEG PROCESSING

The field of clinical EEG is too broad for a comprehensive review of all related data processing. Selected issues are briefly discussed below, focusing on general procedures and techniques that are relevant to signal validation or artefact detection.

I-1.4.1. Sleep research

Polysomnography is usually performed over the duration of an entire night, or at least 6.5 hours, in order to investigate normal and disturbed sleep or vigilance (Bloch, 1997). The most widely used standard for terminology and scoring of sleep stages is the manual by Rechtschaffen and Kales (1968), that facilitates the comparison of data among centres. A standard summary method is the hypnogram that graphically represents sleep stages in 20-30s epochs. Distinct signal characteristics define the different stages: awake, light to deep sleep (NREM1-4), and REM sleep ('rapid eye movements', associated with dreaming). Apart from the EEG channels, recording typically includes the EOG, EMG (chin), ECG, temperature, SpO_2 (oxygen saturation of the blood, recorded on the finger), respiration signals, as well as movement or body position.

Processing of sleep recordings requires elaborate training and is time consuming and expensive. No generally accepted standard exists for automatic sleep staging, but computerisation can improve efficiency and reduce cost when the entire laboratory is well organised and well geared towards the procedures involved (Doman *et al.*, 1995;

Kemp and De Weerd, 1996). International standards may be developed through intense collaboration between laboratories (Kemp, 1993; Dorffner, 1997).

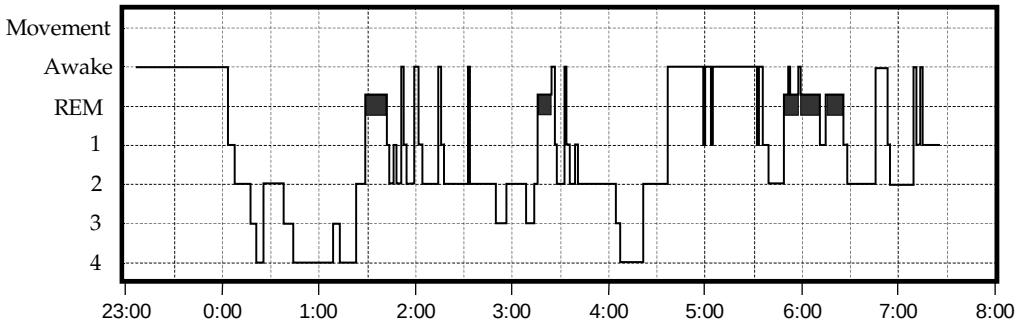

Figure I-1.7 Hypnogram: graphical display of sleep stages, based on manual EEG scoring in pages of 30 seconds (approximately 1,000 pages in this hypnogram).

The procedures and diagnosis involved in sleep disorders are very diverse. One of the most common and alarming disorders is obstructive sleep apnea (OSA), where breathing can stop for lengthy periods of time, and then recommences abruptly. Disturbances in the quality of the EEG channels can be observed in snoring, during periodic leg movements (restless leg syndrome), grinding of the teeth (bruxism), or caused by other physical spasms. However, as these are the phenomena under investigation in the first place, investigators speak of sleep 'events' rather than 'artefacts'. This indicates that a computerised validation method will probably also detect both artefacts and 'events', therefore requiring additional (human) evaluation to discriminate between the two.

Objective detection of subtle EEG changes caused by drowsiness is the subject of studies on vigilance levels. Several methods use human EEG inspection as a starting point to define wakefulness and drowsiness in terms of amplitude and frequency features. Such systems have reached adequate detection performance (90% correct classification of vigilance level), and have been suggested as computerised assistant tool (Värri *et al.*, 1992; Nakamura *et al.*, 1996). However, the detection of artefacts or a specific type of activity is still more difficult. For instance, the agreement between computer and human scores for prominent *alpha* activity reached only 79% (Hasan *et al.*, 1993).

Development of automated processing methods in sleep may be complicated because of non-linear characteristics of the EEG, as some researchers have indicated (Pradhan *et al.*, 1995). Non-linear methods are possible, and have physiological meaning, e.g., when studying the macroscopic aspects of sleep using Markov models (Kemp and Kamphuisen, 1986). Others have found non-linear effects to be significant only when investigating relatively long EEG segments (e.g., >20s, Fell *et al.*, 1996). Linear prediction schemes have

been used successfully in short epochs to investigate sleep EEG. For example, Aufrichtig and Pedersen (1992) used AR models for the characterisation of different types of sleep EEG, using an optimal model order *p* of *14* in 5s epochs. They also indicated that a lower order would be sufficient for shorter epoch lengths. In different states of sleep, optimal model orders have been found as *p=6* for wakefulness, *p=5* in REM sleep, *p=3* in deep sleep, using 1s epochs (Pardey *et al.*, 1996). So although EEG processing during sleep is difficult, the use of short epochs may reduce the need for all too complex models.

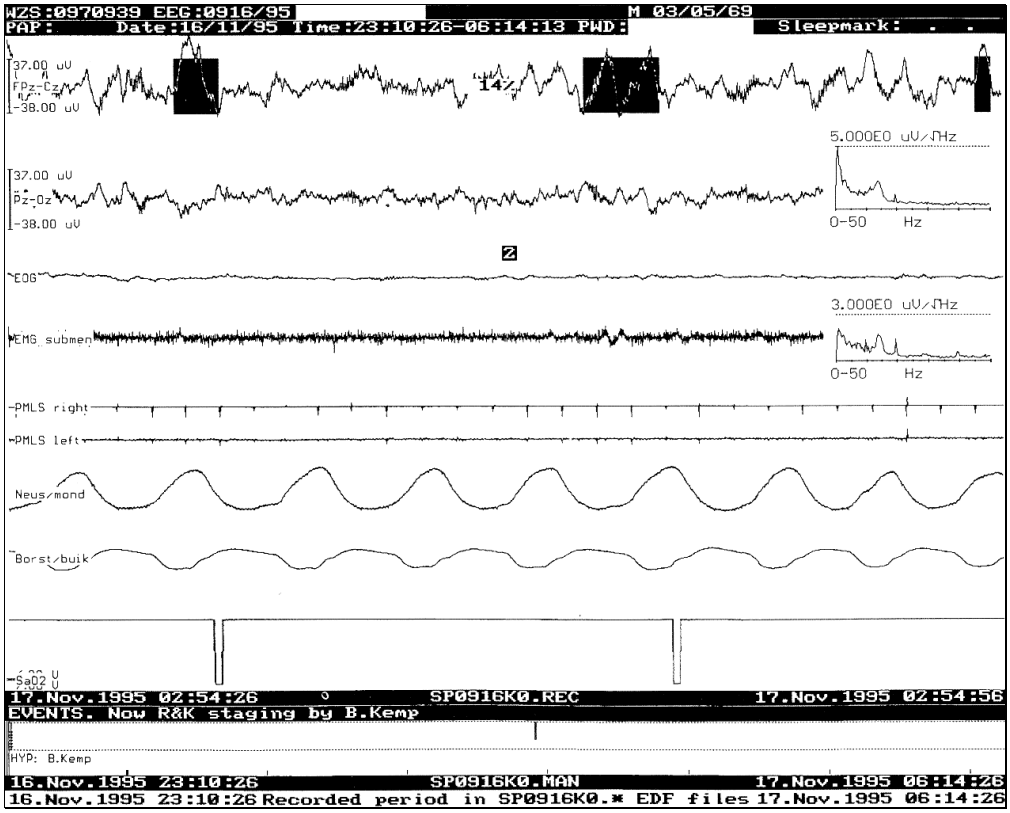

Figure I-1.8 Typical computerised processing of sleep signals consists of automatic detection of 'candidate' sleep events and artefacts (e.g., upper channel), requiring subsequent manual identification to improve the reliability of results (from Kemp and De Weerd, 1996).

I-1.4.2. Epilepsy: spike detection

Epilepsy is a neurological disorder in which *epileptic seizures* recur because of a lasting cerebral abnormality. During a seizure, motor control and mental capacity is disturbed. In patients suffering from epilepsy, diagnosis is based on analysis of epileptiform discharges (EDs) in the EEG. The type, location, and frequency of occurrence of ED waveforms are of primary concern in the clinical investigation.

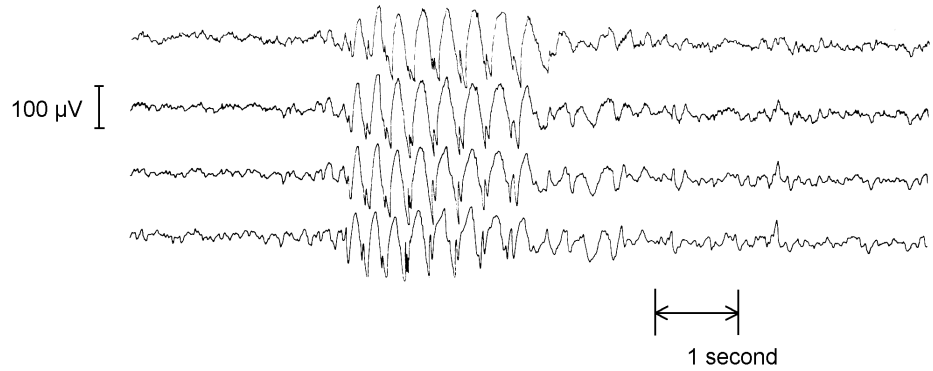

Figure I-1.9 Epileptic seizure activity: synchronous spike and wave complexes in the EEG (from Spehlmann, 1981).

Lengthy recordings and relatively few seizures are typical of epilepsy monitoring. As a consequent, a major interest in computerisation is data reduction and automatic detection of epileptiform discharges, in order to speed up the clinical process of data reviewing. The practical issues in the implementation of such visualisation tools are discussed in Collura *et al.* (1993), and also in Park *et al.* (1990), focusing especially on easy comparison of algorithms.

Time domain processing. Gotman (1982) was among the first to implement a system for the automatic recognition of spike-and-wave activity in the frequency range of 3-20/sec. Sixteen channels of EEG were cut into 2s-epochs, in which "half-waves" were identified from extrema in the signal in order to eliminate small amplitude superimposed on fast activity. Subsequent classification was based on average and variability of half-wave duration. The system showed a relatively large number of false detections, which was partly due to artefacts, but already detected an important proportion of the epileptic seizures.

Detection of epileptic activity can be improved by more detailed definition of waveform morphology in terms of amplitudes and durations for spike-and-sharp-wave complexes (SSWs). An example parameter-set is depicted in Figure I-1.10. A multichannel approach can combine spatial and temporal information and identify the most likely SSWs from multiple *candidate-SSW* detections in more than one channel. False detections

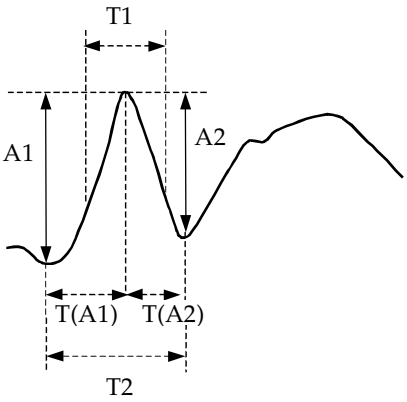

Figure I-1.10 Spike characterisation: amplitude and duration parameters

can then be recognised from a too dense occurrence of candidate SSWs (Glover *et al.*, 1986). Additional context information can be incorporated through detection of synchronous SSW, EOG or ECG activity, which further reduces the number of false detections (Glover *et al.*, 1989).

A computer system by Pietilä *et al.* (1994) classified segments of epilepsy-EEG based on a large set of amplitude and frequency features. This system showed a higher sensitivity than the original Gotman system: 31% versus 17% correct detection, but specificity was generally poorer. Here, the overall low performance (of both systems) was mainly due to artefacts and the occurrence of a lot of small amplitude spikes in some of the patients. Context based processing of spikes and artefacts, i.e., detection relative to the state of the EEG, has been incorporated in a more recent version of the Gotman system. For instance, the occurrence of eye blinks is typical for wakefulness (Gotman and Wang, 1991). The number of true detections was raised to 67% (was 41%), and the number of false detections was reduced to 27% (was 56%) (Gotman and Wang, 1992).

Successful incorporation of spatial and temporal information has been shown in an expert system approach by Dingle *et al.* (1993). The system's performance was evaluated versus only one human expert, but yielded 95% true detection at 59% false detection, showing a very low rate of 0.29 falsely detected EDs per hour (Jones *et al.*, 1996). Further improvements are investigated by using artificial neural networks (ANNs), for both filtering of EEG background (James *et al.*, 1997) and spike classification (James *et al.*, 1996). Enhancement of epileptiform activity from deep sources may be incorporated in the system through an advanced spatial filtering technique (Ward *et al.*, 1999).

Subjective clinical classification. A problem in the evaluation of different approaches is the inter-expert variability. Different styles for human scoring of EDs may be modeled in a computer program using ANNs (Webber *et al.*, 1994). However, this does not advance the objective detection of spikes, since human experts perform far from perfect. For instance, around 80% inter-expert correlation is prevalent in most studies (e.g., Wilson *et al.*, 1996); but an *intra*-expert reproducibility as low as 53% has been observed (one expert, two scoring sessions of identical data: Hostetler *et al.*, 1992). And in a study for classification of different types of seizures, three experts agreed on only 37% of cases, and disagreed on 17%. Processing of these seizures by an *expert system* resulted in correct classification of approximately 80% of consensus cases (Korpinen *et al.*, 1994; Benlamri *et al.*, 1997).

Alternative assessment techniques. Generalised investigation of seizures and outcome prediction of surgery may be performed by monitoring of intracranial EEG. Display of spectral parameters (including Hjorth's) during the surgical procedure can be effective, however, still requiring expert assessment (Alarcon *et al.*, 1995).

An interesting method for separating slow activity and spike-type activity is the use of Principal Components Analysis, where different 'PCA states' are obtained from the EEG signal (Vaz and Principe, 1995). Alternative frequency domain detection of spike activity can be performed by means of Wavelet analysis (Clark *et al.*, 1995). Other time-frequency techniques are discussed in Williams *et al.* (1995); some nice examples during seizures are also shown in Shamsollahi *et al.* (1996). However, analysis of the highly graphical time-frequency plots calls for new skills in display interpretation; studies are still mostly explorative.

I-1.4.3. EEG monitoring in critical care

Patient monitoring in the operating room (OR) or intensive care unit (ICU) involves many other clinical procedures, affecting the quality of EEG recording. Specialised topics like neurosurgery are beyond the scope of the current text.

Procedures for recording and processing

EEG monitoring during surgery or in ICU monitoring must follow regulations for electrical safety (discussed in Hull, 1994). Also from a recording perspective, demands on equipment and artefact processing are more strict in critical care environments. For instance, during electrocerebral inactivity, voltages can be as low as noise levels (Barlow, 1986b; Spencer, 1994).

Precautions for the prevention of electrode lead artefacts have been mentioned earlier (§ I-1.2.1), and are also important here in view of increased staff mobility near the patient. However, it is impossible to prevent all artefacts during the time of patient preparation, intubation or extubation. Electrode-to-scalp contact is affected when exposed to skin preparatives or other solutions used on the patient; using a plastic wrap over the electrodes can prevent this (Hanley and Charlton, 1982). Other effects are mostly due to specialised procedures or equipment interference.

Even more than in other applications of EEG monitoring, one must be aware of the fact that on-line monitoring of the patient's state in critical care environments cannot be performed using one single method or parameter. Because of interaction between different physiologic systems, methods, signals and other clinical data should be integrated (Van Gils *et al.*, 1997a, 1997b; Mainardi *et al.*, 1997).

The usefulness of EEG monitoring, in addition to routine monitoring of vital signs, has been shown during cardiac surgery, carotid endarterectomy, neurosurgery, spinal surgery and intensive care. When the number of EEG channels is restricted by practical

constraints, at least two channels at symmetrical positions near centre and middle of the scalp should be included. Apart from providing relevant EEG information, these positions show the least number of artefacts (e.g., bipolar channels C3-P3 and C4-P4, Schultz *et al.*, 1992). Evoked potentials (EPs) can be recorded in combination with EEG for monitoring of specific nerve pathways at risk, and are increasingly important for clinical patient monitoring (Pronk, 1986; De Beer, 1996). Evoked potentials will be discussed further in part II of this thesis.

Any (display) technique should allow for EEG characteristics to be compared over relatively long periods of time, e.g., 5 minutes, while keeping good time resolution (Levy, 1984). Constant monitoring of electrode impedance will help in the prevention and identification of artefacts (e.g., Levy *et al.*, 1980a).

Monitoring applications

Although anaesthetic and analgesic drugs are aimed at functional depression of the nervous system, patient monitoring in anaesthesia and critical care still does not routinely include the EEG! However, the use of different drugs is a major difficulty in the general applicability of EEG monitoring, since all drugs act differently on the nervous system and result in different changes in the EEG (Romer, 1986; Spencer, 1994). The quality of EEG monitoring and its interpretation heavily depends on the experience of dedicated staff members, who should be trained in both anaesthesia and clinical EEG (Williams *et al.*, 1985).

Unusual patterns in the EEG can be observed, often related to the ongoing clinical procedure. General slowing of EEG activity is observed during hypothermia below 35° centigrade, hypoxia (low O_2), or hypocapnia (low CO_2). EEG depression can also occur as a result of extreme hypercapnia, whereas just a slight CO_2 increase can excite the EEG.

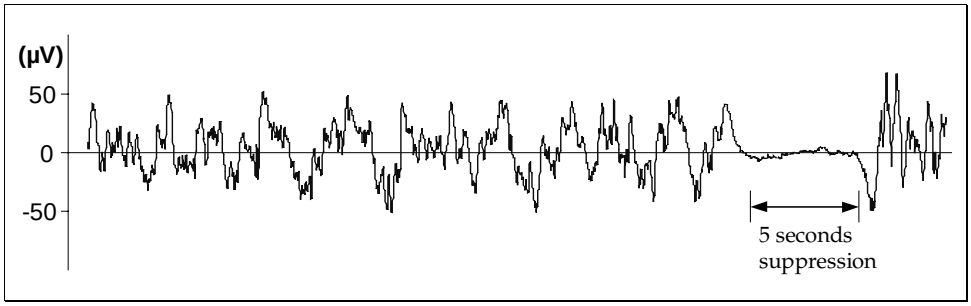

Figure I-1.11 Example of burst-suppression: alternating periods of activity and 'electrical silence' in the EEG. Either period can occur as a few seconds up to several minutes.

An alternating pattern of (relatively normal) EEG activity and 'flat' EEG is known as burst-suppression (Figure I-1.11). Burst-suppression is considered to represent reduced

cortical neuronal metabolic function, and has been observed in failure of oxygen supply to the brain and in (dose-related changes of) various anaesthetic drugs (Spencer, 1994; Prior, 1996). A burst-suppression pattern may be recognised easily from a (processed) EEG display, and can be detected for instance by tracking the inverse of signal variance (Van Gils *et al.*, 1997b). However, robustness in automatic detection is difficult to achieve because of the diversity of different manifestations of burst-suppression (Akrawi *et al.*, 1996; Litscher and Schwartz, 1997).

Spectral analysis. EEG monitoring of patient status in the OR or ICU has focused on the investigation of EEG spectral features. This can be accomplished using a cascaded ('waterfall') spectrum display called *compressed spectral array* (CSA, by Bickford, 1950; Bickford *et al.*, 1972), where spectra are plotted behind one another to allow for tracking of spectra in time. An alternative display was designed for better time resolution in the form of the *density spectral array* (DSA, by Fleming and Smith, 1979), but neither CSA or DSA is favourite (comparison in Levy *et al.*, 1980a). These techniques can be useful in the identification of major changes of the state of the patient (Young and Ornstein, 1985), but are less suited for the detection of important EEG features of short duration. Moreover, the CSA or DSA introduce an additional display that requires a great deal of attention in an already intense monitoring environment. In comparison, the *cerebral function analysing monitor* (CFAM) shows a processed trend display that compactly summarises the most important, clinically relevant information. The CFAM display is optimised for comparison of channels, amplitude and frequency information, and better suited for detection of periods of EEG suppression (Sebel *et al.*, 1983; Prior, 1996).

Autoregressive (AR) spectral measures have also been found useful, e.g., during neurosurgery, and should be considered as a potential application for long term monitoring to detect EEG changes, including transients and non-stationarity (Cerutti *et al.*, 1986). For instance, a fifth order AR model could reliably discriminate between levels of thiopental/enflurane anaesthesia in a study by Bender *et al.* (1991). EEG monitoring of more subtle changes in anaesthesia levels proved successful both in the CFAM system and in an AR-based hierarchical clustering system (Thomsen *et al.*, 1991; Thomsen and Prior, 1996). Further quantitative evaluation of these, and other studies of 'anaesthesia levels' is complicated because of the use of different drugs, and because of disagreement between neurologists even in the interpretation of 'stable/unstable' EEG during anaesthesia (Hinrichs *et al.*, 1996).

I-1.5. DISCUSSION

I-1.5.1. General considerations

This chapter described the many aspects of artefact prevention and detection in the EEG. Above all, artefacts should be *prevented* by high quality standards in both equipment and recording procedures. Removal, i.e., correction by filtering, of artefacts may be applied only for artefacts having frequency components outside the frequency range of the EEG. Specific artefacts and those that can be monitored on a separate channel (eye artefact, or electrical heart signals) may be removed from the EEG successfully. However, most artefacts present themselves within the same frequency range and some can even mimic EEG activity. Therefore, artefact removal should be considered only in situations where rejection is no option. Moreover, the occurrence of artefacts can reveal useful clinical information (Barlow, 1986b). For example, movement/muscle activity or eye artefacts in the EEG of a critically ill patient can actually indicate an improvement in the status of the patient.

I-1.5.2. Current approach

Proper comparison of different EEG validation methods is difficult because of several methodological problems. For a start, objective EEG evaluation is hindered by the subjective interpretation of human experts. Next, different studies use different definitions, and often consist of incongruent procedures resulting in incomparable results.

Furthermore, the variety of investigations, neurological phenomena, processing methods, analyses and diagnoses, hampers the design of a universal method for EEG validation.

From these observations, we cannot expect to reach 100% reliability in the detection of all EEG artefacts. However, we *can* investigate the use of several objective methods, and try to establish some common guidelines. We will strive for objectivity by using statistical methods for detection *and* evaluation, while focusing on the accuracy (time resolution) at which artefact detection should be performed.

Selected methods

Time domain methods seem to be most suitable for the detection of high amplitude artefact and for the detection of fast activity. Such processing is often based on simple amplitude thresholds, first and second derivatives. Nevertheless, time domain features are easy to implement and have been very successful in the detection of artefacts.

Another successful method is AR modeling, which has been shown useful for the detection of 'non-stationary' signals. The AR model parameters contain information about

the identity of the investigated EEG (or disturbance). Through 'within model' inspection of AR features, objective detection of deviating phenomena may be achieved. As a supporting fact, in the literature AR methods are often preferred over other (frequency domain) methods, being applied especially in classification of phenomena that lie within the frequency range of the EEG.

Focus: time related aspects of EEG validation

Epoch length has not been studied extensively in relation to detection performance of any method. We will investigate this aspect of signal validation for time domain and frequency domain detection of muscle artefact. This is one of the most frequent and obtrusive EEG contaminations. In this study (chapter I-2), we will objectively evaluate the performance of all processing methods.

Most automatic detection methods try to model the way human experts perform their visual EEG analysis, i.e., analysis of features/artefacts relative to the current state of the EEG, the *signal context*. The selected methods will be investigated in relation to signal context with special focus on the length of the EEG period that needs to be incorporated as context (chapter I-3).

Part II –

Quality of evoked potential measurements

II-1. Evoked Potential Measurements: Overview and Methods

Evoked potentials are signals that are derived from the EEG. Therefore, disturbances in the EEG (Part I) can affect the quality of the recorded evoked potential signal. Additionally, the detection and processing of artefacts in evoked potential measurements have some new aspects when compared to 'standard' EEG measurements. The measurement itself, e.g., the stimulation technique, may induce artefacts in the evoked potential, or the repetitive character of an otherwise minor EEG artefact may contaminate the evoked potential signal.

The different types of 'evoked potentials' and 'event related potentials' are described. Measurement requirements follow, and some clinical applications are presented. Artefacts and procedures specific to evoked potential recordings are then summarised, highlighting the intricacies of recording and processing. The overview continues with signal processing techniques, mostly related to evoked potential signal quality. The last section describes methods for correction of eye artefacts, feature recognition and non-linear analysis in evoked potential research.

II-1.1. EVOKED POTENTIAL MEASUREMENTS

The term 'evoked potential' (EP) is used as a general term for an electrical response of the nervous system that is associated with a sensory stimulus. The signal is usually recorded to quantify the response of a specific sensory pathway to one particular stimulus type. 'Event related potential' (ERP) is another general name, and is used here to indicate experiments of a psychophysiological character. 'Event' refers to any type of sensory stimulus, motor action, or cognitive task. Various stimulation schemes are used in ERP investigations of higher order processing in the brain. Such extended paradigms are used in fundamental research in cognitive psychology, whereas straightforward stimulation of one sensory modality is typical for clinical measurements.

II-1.1.1. Inventory of evoked potentials

EP investigations are used to assess conductive properties of sensory and motor pathways at various locations within the nervous system. In ERP investigations, sensory perception and cognitive processes are addressed. The investigations often focus on the measurement of (changes in) characteristic components; these are usually named after the polarity (Positive, Negative) and latency (ms) of the investigated peak (e.g., N100, P300).

Stimulation modalities

Below, the modalities of the main clinical application areas in EP research are introduced. Apart from these, other modalities can be stimulated; EPs based on olfactory (smell), gustatory (taste), and pain stimulation have been reported (Regan, 1989).

Auditory pathways. EPs from the auditory system can be recorded from the scalp by using clicks, tone pips, or bursts of mixed frequencies. Broadband clicks (wide frequency range) stimulate a large part of the cochlea, resulting in activation of a large number of nerve fibres in the acoustic nerve. The frequency characteristics of tone stimuli usually reflect the optimal hearing range of 300-4000Hz.

Auditory evoked potentials (AEPs) are classified by the investigated *post-stimulus* time window, distinguishing short latency (brainstem) components (BAEP: 0-10ms), middle latency components (MLAEP: 10-50ms) and long latency components (LLAEP: 100-500ms). BAEP and MLAEP measurements require very short stimuli (e.g., 100µs duration): when using monaural clicks, masking noise may be presented at the contralateral ear to reduce the influence of cross-stimulation by skull conduction of the ipsilateral stimuli. Condensation or rarefaction clicks (auditory stimulation starting respectively inwards or outwards the eardrum), or different types of earphones can slightly influence the brainstem response (Deltenre and Mansbach, 1993). LLAEP measurements are usually recorded using click or tone stimulation. Sound levels are set usually at approximately 70dB above hearing thresholds for short clicks (Stockard *et al.*, 1978; Regan, 1989, p. 150-151).

The middle latency components are generated in the medial geniculate body and primary auditory cortex. Later components are generated in brain areas related to (early) cognitive processing: the temporal and frontal cortex (Thornton and Sharpe, 1998).

Somatosensory pathways. Somatosensory evoked potentials (SEPs) are obtained by applying an electrical current to the arm or leg via skin electrodes or needle electrodes (but can be measured also through natural tactile stimulation). Electrical shocks have a typical duration between 100µs and 1ms, using lower current levels at longer durations, e.g., 25mA at 1ms. The response signal passes through the peripheral nerves and the spinal cord, propagating up to the sensory cortex (mid-central areas). The SEP components show longer delay times and altered characteristics when recorded farther from the stimulation site. Such changes can also be caused by nerve lesions. When

recorded on the scalp, the signal comprises early components from peripheral nerves and brainstem (<20ms), short- and medium-latency cortical components (20-100ms), the vertex potential (100-200ms), and an after-discharge. Most common SEP stimulation sites are the median and ulnar nerves in the arm, the peroneal nerve at the knee, or the tibial nerve at the ankle.

Early SEP components stem from the dorsal column and medial lemniscus deep in the brain. Medium-latency components are generated in the thalamus regions and primary somatosensory cortex; later components originate in the post-central gyrus and frontal cortex (Thornton and Sharpe, 1998).

Visual pathways. Visual evoked potentials (VEPs) are recorded during visual stimulation of a subject, using alternating checkerboard patterns, flashes or other light/intensity patterns. Stimulation is usually presented by means of a television screen, a panel of light emitting diodes or flash bulbs. Specification and standardisation of visual stimuli involves a great number of variables, e.g., colour, light-intensity, eye-fixation, pupil-diameter, distance of light-source, infrared filtering to prevent damage to the cornea (Regan, 1989).

Even more than in other EP measurements, a large number of different VEP types are known because of the diversity in stimulation modes. General observations of VEP signals when compared to other types of scalp evoked potentials are: higher amplitudes and relatively slow components. A common type of VEPs uses alternating checkerboard patterns, resulting in a characteristic positive peak at approximately 100ms after the stimulus (Spehlmann, 1985).

VEP measurements and related functional assessment can focus on different subsystems in visual processing, including luminance changes, colour vision, and spatial or motion perception. The corresponding perceptual pathways and neuronal areas have been studied in great detail. The principal sub-cortical region that processes visual information is the lateral geniculate nucleus (Kandel *et al.*, 1991). The primary visual cortex is located in the occipital area at the back of the head, but other regions in the brain are equally important in visual perception processes (Zeki, 1992).

Event related potentials

ERP measurements are used in the investigation of neurophysiologic correlates of cognition and attention. This vast topic is beyond the scope of the current introduction; only a few aspects of higher order processing are highlighted here (after Regan, 1989; McCallum, 1988, 1997).

ERP measurements are also used to investigate the presence or absence of specific components or patterns, and are characterised by amplitude and latency. For instance, a prominent negative peak at about 100ms (N100) is produced after presentation of a sound stimulus, even when repeated numerous times. Novel and deviant stimuli produce enlarged amplitudes. A related phenomenon, associated with attentional processes in the

brain, may be measured when the incoming stimulus and the expectation of the (preconditioned) subject are mismatched. The resulting ERP shows a more negative component around 200ms, hence the name *mismatch negativity* (MMN). Another *positive* peak at approximately 300ms can be found in an *auditory oddball* task, when a subject listens to a sequence of monotonous tones, randomly replaced by a deviant tone, which causes a pronounced 'P300' component. This can also be measured when the deviant stimulus is a memorised word in a sequence of words (Van Hooff *et al.*, 1996).

Slower reproducible potential changes can be measured in ERP paradigms designed for investigation of preparatory and anticipatory processes in the brain. For example, a slow negative going readiness potential (RP) can be measured prior to (self-paced) movements and the contingent negative variation (CNV) preceding a signalled movement. The latter ERP signal is measured in a warned reaction-time task, and may be related to priming of the brain (in the cortical areas) by the preparation between a warning stimulus and an imperative stimulus. We will revisit this last example from ERP research in chapter II-3.

Signal properties

Most of the activity recorded in a typical EP is generally of very small amplitude. However, EPs recorded within the brain near a neuronal generator site (involved in the early responses, <40ms) typically result in higher amplitudes (1-2mV) (Jellema, 1993; Rosenfalck, 1969). The table below does not list the properties of these 'in-depth' potentials, but mainly characterises non-invasive (scalp) EPs.

Table II-1.1 General characteristics of evoked potentials: amplitude and frequency dynamics (after Cohen, 1995).

Type	**amplitude**	**frequency**	specifics
Auditory evoked potentials	0.5-10μV	100Hz-3kHz	*recorded on vertex*
Somatosensory evoked potentials	1-10μV	2Hz-3kHz	*somatosensory cortex*
Visual evoked potential	1-20μV	1-300Hz	*occipital cortex*
Event related potential (general)	1-50μV	0.2-100Hz	*(e.g., P300, CNV*, in psychophysiology)*

* Contingent Negative Variation: see text above

II-1.1.2. General recording methods

In all EP measurements, the signal of interest is superimposed on the ongoing EEG activity recorded at the electrodes. Therefore, the EP is usually enhanced by averaging a

number of responses, thus cancelling the 'background' EEG that is not correlated to the stimulus presentation. Averaging and other EP signal processing will be discussed in detail in § II-1.3; recording procedures specific to EPs are summarised below (also see: American Electroencephalographic Society, 1984; Spehlmann, 1985; Regan, 1989; general EEG procedures have been discussed in chapter I-1).

Electrode positioning

The voltage and waveform of an EP depends on the locations of the electrodes, which are positioned according to the international 10-20 system, or closely related to these standard positions (example in chapter II-3, p. 118). A widely separated electrode pair is relatively insensitive to source location, whereas a closely spaced pair is sensitive to both location and orientation of nearby source(s), and can be used to isolate the contribution of a weak source. The location of the neuronal generators is also important for the choice of the reference electrode. A referential average over (all) electrodes can be used only if the EEG-activity is 'pseudo-random' (uncorrelated) over the included positions, which is often untrue during EP measurements (e.g., Tomberg *et al.*, 1990). When EP activity is widespread over the scalp, e.g., in short-latency EPs, a non-cephalic reference can be used. For sources in the outer cortex, bipolar recording from closely spaced electrodes (using a ground electrode at a distant, neutral position on the scalp) may be preferred.

Amplification, filtering and digitisation

Normal EEG recording uses time constants between 0.05s and 1s (i.e., filtering of DC components and low frequencies). In the recording of some slow changing potentials as the contingent negative variation (CNV) true DC recording has been advocated (Regan, 1989).

In equipment for general EP recording, the noise level of the amplifier must not exceed 3μV in the required frequency band of 0.1-5,000Hz. Sampling frequency and filter settings should take into account the frequency range of the EP under investigation. The dynamic ranges of Table II-1.1 are only a starting point; when focusing for instance on the *brainstem* part of an AEP, high-pass filtering is set at 100-500Hz, and low-pass filtering at 1,600-3,600Hz (Thornton, 1990), using a minimal sampling frequency of 10kHz (Grönfors and Juhola, 1995). Digital filters are preferred over analogue filters, because analogue filters can distort EP components (Spehlmann, 1985; Thornton, 1990; Grönfors and Juhola, 1993).

Equipment for on-line averaging and display should be able to show the ongoing background EEG together with the averaged waveforms. Where EP measurements in a difficult clinical setting have minimal requirements of only two channels, topographic ERP measurements may require up to 128 channels. Sufficient memory and processing resolution should be available to allow for the averaging of 4,000 trials in BAEP measurements (American Electroencephalographic Society, 1984). This number is often much lower in ERP research: e.g., approximately 60-100 trials constitute a CNV average.

Accurate event recording is important in EP measurements, especially for off-line processing. An example for encoding the stimulus and event information, correctly synchronised with the EEG measurement, can be found in the appendix of this thesis.

Stimulus presentation rate

Stimulation order and presentation rate in an ERP task is determined by the protocol of the experiment (e.g., tone, delay, flashes 2s later, etc.). In functional EP measurements stimulus presentation is applied at regular intervals long enough to record the response signal under investigation (e.g., 10/s for 100ms response). Faster stimulation will result in overlapping response intervals, introducing higher order effects through non-linear interaction of stimuli/responses (see also § II-1.4.3).

Steady-state potentials. Specific EP components — with sufficiently large amplitude — can be investigated in a *steady-state evoked potential* measurement. This is a repetitive EP waveform whose constituent discrete frequency components remain constant in amplitude and phase over a prolonged period. During the measurement, a selected peak is enhanced by accurately synchronising the periodic stimuli with the periodic interval of one harmonic frequency in the EP. Steady-state AEPs, SEPs, VEPs can be used in fundamental research of cognition and neuronal processing (Basar et al., 1987; Regan, 1989).

II-1.1.3. Clinical applications of evoked potentials

Guidelines for recording of EPs are very helpful in selection of equipment settings and techniques. Clinical EEG/EP recording requires the highest quality standards, especially in surgery and critical care (Levy *et al.*, 1984; American Electroencephalographic Society, 1994a; Prior, 1996). Equipment safety regulations must be checked at all times, in particular when nerve stimulation is used (Hull, 1994).

One of the main applications of EPs for clinical diagnosis is the objective evaluation of the functioning of a sensory pathway. A very useful application for instance in audiology studies is the assessment of hearing loss by examination of stimulus intensity versus observed peaks in the auditory brainstem response (e.g., Ödamar *et al.*, 1990). Lesions in a neural pathway can be detected from diminished peak amplitudes, increased latencies, or even absence of a normal EP component; monitoring for such changes is useful during surgery where a sensory pathway is at risk. For example, the recording of BAEPs during resection of an acoustic neuroma is crucial to the success of the operation, and is often combined with facial nerve monitoring (Schwartz *et al.*, 1985). BAEPs can also help in the diagnosis of vascular lesions, tumors or demyelination (Grundy *et al.*, 1982; Spehlmann, 1985). Recording of SEPs and subsequent evaluation of conduction velocities in the spinal cord is used during scoliosis correction, and SEPs are also an indicator of multiple sclerosis or spinal cord injury (Regan, 1989).

Monitoring during anaesthesia. Objective assessment of anaesthesia levels using EPs has been suggested by many authors (Raudzens, 1982; Grundy, 1985; Spencer, 1994). Changes in VEPs, SEPs and AEPs have been used to detect inadequate cerebral perfusion during cardiopulmonary bypass. A major problem in VEP measurements however, is the large inter- and intra-subject variability, which makes this modality less suited for intraoperative monitoring (Grundy, 1982; Levy *et al.*, 1984).

The relatively easy installation of AEP recording during surgery, and the sensitivity of its middle latency components to metabolic changes make this signal the first choice as an on-line monitor of anaesthesia levels (Thornton, 1991). A combination of EP and ERP techniques can be used to investigate the ability of the brain to process information during anaesthesia. For instance, in an auditory processing task, the P100-N100-P200 complex was found to be delayed and more positive going during anaesthesia when compared to the preoperative recordings (Van Hooff *et al.*, 1995, 1997). AEP monitoring during anaesthesia is promising, but careful evaluation of AEP features in relation to different aspects of anaesthesia is imperative. For instance, evidence suggests that AEP latencies are related to a hypnotic effect, while AEP amplitudes seem to correlate with analgesia (pain related effect) (De Beer, 1996; De Beer *et al.*, 1996).

II-1.2. SPECIFIC ARTEFACTS IN EP MEASUREMENTS

II-1.2.1. Muscle activity

As in EEG measurements, muscle artefact can be obtrusive in EP recording. Muscle artefacts notably affect the quality of later EP components starting at latency 10ms. This is mostly due to flexing of the muscles in the neck or jaw, and can be reduced by supporting the neck and head with a pillow, or by asking the subject to slightly drop the jaw.

Activity of muscles on the scalp can influence the signal, especially when related to the stimulation (§ II-1.2.2). For this reason, electrode positions on the mastoid bone just behind the ear are less suitable in AEP recording. Scalp muscle activity is not a serious problem in relaxed subjects, but is a possible source of artefact when recording from tense individuals (Regan, 1989).

The application of filters to eliminate muscle artefact is not a solution, because the frequency range of muscle artefact (20-300Hz) and the investigated EP waveforms are highly overlapping (see Table II-1.1). Contamination by muscle activity normally disappears during sleep, and is usually not a problem when muscle relaxants are used, e.g., in EP monitoring during surgery. However, drugs can also affect EPs.

II-1.2.2. Stimulus related artefacts

Stimulus-locked activity that originates in scalp muscles can contaminate SEPs recorded from scalp electrodes. These artefacts are called 'somatomotor' potentials (Regan, 1989, p. 292), sometimes having similar latencies and larger components than the neuronal SEP. Other stimulus-locked artefacts can be a major problem in ERP measurements where a subject has to make some kind of motor action (e.g., button-pressing) in response to stimuli. Even in preparing for this action, the subject's heightened alertness may contaminate the recording. Some stimulation can also cause an unwanted eye (blink) reflex, because of the startling effect of stimuli. Randomisation of stimuli is sometimes a solution, but may not completely abolish all stimulus related subject behaviour. Unconscious subject activity such as finger counting or (silent) 'talking' can cause unwanted readiness and motor potentials.

Stimulation artefact can also be caused by too loud auditory clicks or tones in AEP recording. Stimulus intensities over 70dB can induce muscle artefacts from the stapedius, which may obscure short-latency and early components in the AEP. Stimulus artefacts in BAEP recording may be observed in patients with hypoxic encephalopathy, and in (near) brain death subjects. Changing the polarity of the auditory stimuli, e.g., from rarefaction to condensation clicks, can diminish the artefact (Brittenham, 1990; Litscher *et al.*, 1995).

In clinical SEP measurements, a relatively large voltage is applied as stimulus. This typically results in a sharp and large artefact in the averaged SEP waveform, larger near the stimulus location. Adequate skin preparation and shielded leads are imperative to reduce electromagnetic coupling of stimulus and recording electrodes. The electrodes are best placed along equipotential lines (McLean *et al.*, 1996; Scott *et al.*, 1997). This artefact can also be minimised by post-processing, using special filters or neural network pattern correction (Grieve *et al.*, 1996).

II-1.2.3. Ocular artefacts

Artefacts caused by eye rotation are of serious concern in EP recording, because of the constant (DC) potential of several millivolts over the eyeball. A downward rotation of 10° already results in a negative shift of 50μV at the vertex on the scalp. This can be a major problem in ERP paradigms, where subject may systematically move or blink their eyes in relation to the stimulus sequence. For example, in CNV experiments, large involuntary eye movements are commonly synchronised with the preparatory interval, especially when the eyes are closed (Regan, 1989). In recordings of short-latency EPs, usually of very small amplitude, pre-filtering of the low frequency components (including the slow eye potentials) will effectively improve the EP quality. Fast eye activity may still be a problem during sleep recordings.

Special techniques to reduce the influence of eye artefacts in ERP measurements are discussed in a separate paragraph, § II-1.4.1.

II-1.2.4. Other artefacts

Mains interference artefact. Braiding of lead wires is advised to reduce 50/60Hz interference during EP monitoring (Stecker and Patterson, 1996). Phase synchronised triggering can effectively eliminate 50/60Hz artefact, and is recommended in recording settings such as the ICU or the operation theatre (Emerson and Sgro, 1985). Averaging in the alternating opposite phases of the sinusoidal signal will cancel out the interference. If this cannot be used, at least the inter-stimulus interval must be chosen *unsynchronised* to the mains period.

ECG artefact. Artefacts from the electrical activity of the heart can be minimised by choosing a symmetrical point for the ground electrode, e.g., on the forehead. When the ECG interference still presents problems, a procedure by Nakamura *et al.* (1990) can effectively eliminate the artefact before averaging the EP trials. They obtained an ECG average from a concurrent ECG recording, which after subtraction from the EEG still resulted in an R-top artefact in the EEG. With the exclusion of these remaining artefact periods, the authors demonstrated in an example that reliable short-latency SEPs could be obtained.

II-1.3. EP PROCESSING AND SIGNAL VALIDATION

II-1.3.1. Time domain processing

The averaging process

A computational procedure is needed to extract an evoked potential signal from the recorded EEG. The initial signal-to-noise power ratio (SNR), i.e., the ratio of EP (1-20μV, typically <5μV) versus background EEG (10-100μV) can be –20dB or even lower in single trials. By averaging the single trials (sweeps) in the EP measurement, the SNR is increased, under the *assumptions* that:

1) the response signal does not vary with time, i.e., the EP components have identical amplitude, latency, phase, and shape for all sweeps;
2) the noise (background EEG) is a random, zero-mean signal with constant stochastic properties during the measurement;
3) the noise/interference is not correlated to the stimuli (also see previous section).

When we define $s_i(t)$ as the EP response after the *i*th stimulus, and $n_i(t)$ as the EEG background signal that is uncorrelated to the stimulus, we can describe the recorded signal $x_i(t)$ as:

$$x_i(t) = s_i(t) + n_i(t) \tag{II-1.1}$$

This signal is averaged over N sweeps to obtain the evoked estimate $\hat{s}(t)$:

$$\hat{s}(t) = \bar{x}_N(t) = s(t) + \frac{1}{N}\sum_{i=1}^{N} n_i(t) \tag{II-1.2}$$

The noise term has an expected mean value of zero and variance σ_n^2, which is constant throughout the measurement. Now we find that the expected EP signal becomes

$$E\{\hat{s}(t)\} = s(t) + \frac{1}{N}\sum_{i=1}^{N} E\{n_i(t)\} = s(t) \tag{II-1.3}$$

(E: expectancy)

with an accompanying standard error (SE) of the average of

$$SE = \sqrt{E\{(\hat{s}(t) - s(t))^2\}} = \frac{\sigma_n}{\sqrt{N}} \tag{II-1.4}$$

The SE as given by this formula represents the residual noise amplitude that is proportional to the amplitude level in the background EEG. SE diminishes proportionally to the square root of the number of sweeps N included in the EP. The noise power (SE^2) diminishes proportionally to N. Theoretically, we can reduce this noise component to any desired level. However, in practical measurements the recording time is limited (e.g., constrained by protocol: limited attention span of subjects during ERP tasks, limited time during surgical procedures, etc.). Furthermore, a too lengthy measurement will result in invalid EP recording conditions (assumptions 1 and 2).

Amplitude analysis

Enhancement of the SNR can be achieved by exclusion of large EEG amplitudes. This can be performed by *clipping* of the EEG amplitude, where the analogue-to-digital conversion saturates at a pre-set level, or by *rejection* of individual trials (sweeps). Picton *et al.* (1984) concluded that clipping was more efficient than rejection, but, from a theoretical point of view this is not advised (2nd assumption above), because clipping alters the measured signal and significantly disturbs its stochastic properties.

In EP measurements with relatively large numbers of stimuli, rejection of sweeps can be implemented. Reduction of too large amplitudes in the background EEG is of major influence on the noise level in the EP waveform. Most commercial averagers have *standard* threshold borders for too high amplitudes to exclude the corresponding sweeps from the averaging process.

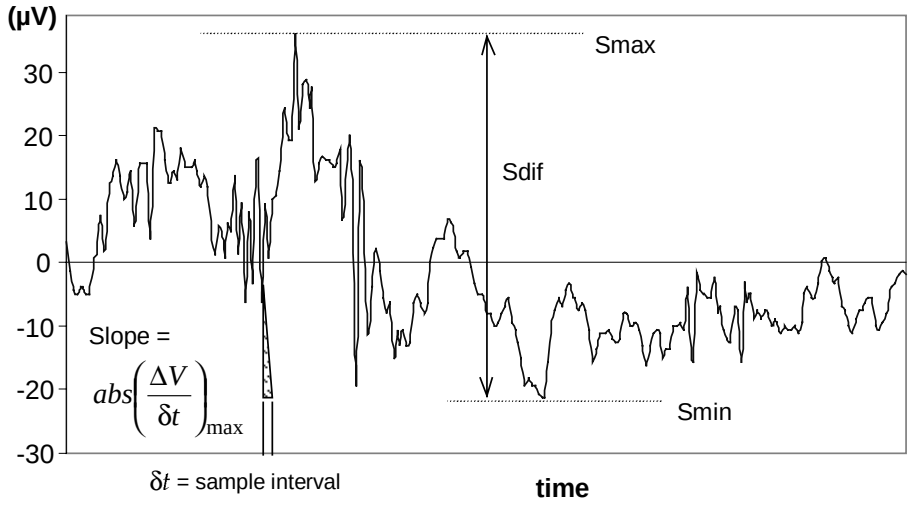

Figure II-1.1 <u>*Amplitude parameters*</u>*: maximum value* (Smax)*, minimum value* (Smin)*, difference in amplitude* (Sdif)*, the maximum of absolute slope values* (Slope) *in an epoch or sweep.*

Advanced amplitude detection and rejection of sweeps is beneficial for improved EP quality. Amplitudes can be investigated in the distributions of maximum, minimum, and their difference, as well as for the differential (slope) amplitude between two data samples (Figure II-1.1). The histogram's observations are taken from a large number of sweeps in a reference measurement, that were visually selected as artefact-free. Thresholds are calculated at $\mu \pm 3\sigma$ for the different signal parameters (μ, σ for mean, standard deviation; Cluitmans *et al.*, 1993; also see Oken, 1989). This results in the exclusion of only 1% of sweeps in artefact-free measurements (assuming a 'normal' reference distribution), while detecting artefactual sweeps when parameters fall outside of the histogram thresholds.

The effectiveness of artefact rejection by using this semi-automatic method was investigated in a study of AEP data obtained during surgery. The *effectiveness*, defined as the percentage of improved EP waveforms (visual assessment), was as high as 97%. Moreover, the effectiveness was substantially improved by the use of the *non-standard* borders for *low maximum* (LBSmax), and *high minimum* (UBSmin) amplitudes in sweeps. The slope threshold was not needed in this study, because muscle-relaxant drugs were administered (De Beer *et al.*, 1995); however, slope detection can be particularly useful for detection of fast activity such as muscle artefact (Scherg, 1982b; Cluitmans *et al.*, 1993).

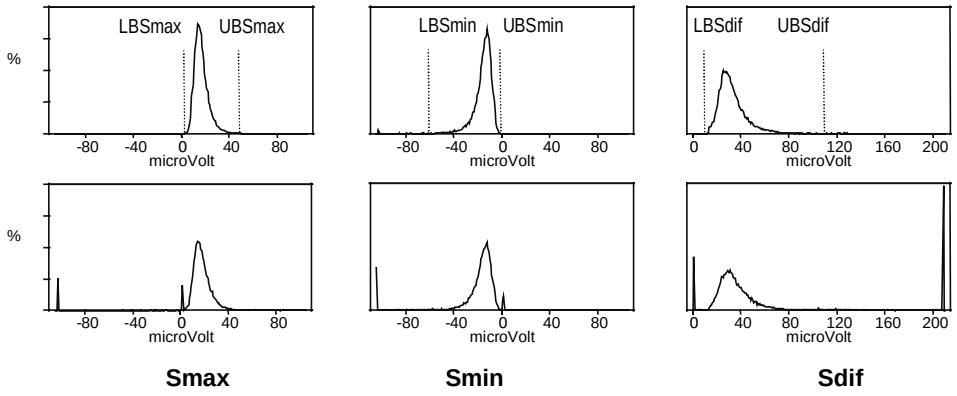

Figure II-1.2 Examples of amplitude histograms in auditory evoked potentials: (upper row) histograms for maximum, minimum and difference for artefact-free data, (bottom row) histograms during electrosurgery (from De Beer, Van de Velde and Cluitmans, 1995).

II-1.3.2. Frequency domain processing

Spectral analysis

Some evoked response signals in the EEG do not show clear phase- or time-locked features in relation to applied stimuli. A phenomenon called 'event related desynchronisation' (ERD) results in changes in the spectral intensity of different EEG bands, and is often studied for the alpha range (8-13Hz) in topographic distributions on the scalp. The ERD can be useful for identifying modality-specific (especially visual) attention processes (Bastiaansen *et al.*, 1999). These effects cannot be quantified through time averaging; the calculation is performed through spectrum-based statistical testing whether or not the changes are of a random nature.

Another signal that can be analysed favourably in the frequency domain is the steady-state evoked potential. This type of EP may be described by measuring its amplitude and phase as function of the intensity of the stimuli, or as a function of the stimulus rate (Lopes da Silva, 1987c). In addition, the variability of the frequency components can be investigated (Victor and Mast, 1991). Bispectrum analysis, i.e., detection of coupled harmonics, can be used for early detection of the response in steady-state VEP measurements (Husar and Henning, 1997).

Filtering

Specific peaks can be enhanced when the frequency range of the investigated EP is known and confined to a frequency band with limited activity of the background EEG.

Narrowband frequency filtering is commonly used in steady-state EPs to enhance the signal-to-noise ratio, because the recording focuses on known harmonics of the stimulation frequency.

Filtering of bands has been studied for optimal waveform morphology especially in the *auditory* evoked response because it includes a wide range of frequencies. Bandpass filtering of 200-1500Hz is suitable for the early components in brainstem AEPs (<10ms), and increases the reliability of latency assessment (Boston, 1981; Spivak and Malinoff, 1991). Digital filtering of frequencies below 15Hz effectively enhances middle latency components, as shown in MLAEPs in a study by Kraus *et al.* (1987). High-pass *analogue* filtering however, using steep filter settings (24-48dB/octave) may distort EP peaks (Scherg, 1982a), and is not advised (Sgro *et al.*, 1989b).

II-1.3.3. Waveform variability

We tend to interpret the averaged EP waveform as the representation of the neural pathway's typical response to one stimulus. However, single responses can be dissimilar: the first assumption in EP averaging (p. 83) may not hold. For instance, fatigue and habituation can produce large variability in some (e.g., olfactory) EPs. Variability is negligibly small in short-latency (brainstem) EP measurements, but changes can occur in specific procedures (see § II-1.1.3). For instance, changes in body temperature of 1°C result in 0.2ms latency shift in BAEP, or up to 2ms in MLAEP recording (Spehlmann, 1985; De Beer, 1996).

Single trial analysis is important in psychophysiological ERP research; variability in the response might be caused by diminished attention during the measurement. Trial-to-trial analysis can be used to investigate a possible temporal correlation between the 'background' EEG and the single response signal (Intriligator and Polich, 1995). For example, Gordon *et al.* (1994) indicated that only 40% of single-trial ERPs showed waveform morphology resembling the averaged ERP. Single trial latency variability in a P300 experiment for example, has been observed to vary between 200-600ms (Hansson *et al.*, 1996), clearly demonstrating the difference between the theoretical nomenclature 'P300' (at 300 ms) and actual observation.

Latency jitter can result in reduced EP amplitudes and unclear EP latencies because of waveform smoothing. The phenomenon may present some problems for instance in patients with multiple sclerosis (Regan, 1989). Prior knowledge can be used to filter latency jitter. For example, changes in brainstem EPs occur only over half a minute or longer (Paige *et al.*, 1996). Latency jitter can be corrected by using a correlation technique that aligns single sweeps (determination of optimal τ_i) to an EP template, also known as Woody filtering (Woody, 1967):

$$x_i(t) = s(t - \tau_i) + n_i(t) \qquad \text{(II-1.5)}$$

However, the underlying model still assumes some invariance of the EP properties, e.g., inter-peak latencies. To overcome this problem, the method can be extended by calculating cross-correlations for specific peaks, and aligning these peaks of the EP in each single sweep (McGillem and Aunon, 1977; Aunon, 1983). Still, it must be applied with care for the investigation of high-frequency phenomena embedded in low-frequency components (Challis and Kitney, 1990), and cross-correlation calculation is not particularly suitable for single sweeps with very low SNR, e.g., AEPs during anaesthesia.

Weighted averaging

Low SNR's (e.g., –20dB) are also a problem for the weighted-averaging method (Miskiel and Ödamar, 1987; Davila and Mobin, 1992). This method tries to optimise the contribution of single sweeps to the average EP; sweeps are weighted according to the presence of noise. A dynamic template tracking technique may optimise the correlation calculation (Jansen and Yeh, 1986; Picton *et al.*, 1988; Chan *et al.*, 1995; Gupta *et al.*, 1996). Implementation in hardware can overcome the extra computational demands of such adaptive techniques (MacLennan and Lovely, 1995). However, absolute amplitude information in the weighted-average can be maintained only when the background EEG is stationary (Hoke *et al.*, 1984), which is a rather strict requirement. Another practical disadvantage is encountered for instance in BAEP measurements, where a large set of reference templates may be needed (Ödamar *et al.*, 1990).

Autoregressive estimation

Cerutti *et al.* (1987, 1988) used an autoregressive (AR) modeling approach to obtain estimates for single trial EPs. The AR process characterises the background EEG and a 'template' of the expected EP waveform is used as the *exogenous* input process (ARX model). The template is obtained by the normal averaging method (as described in § II-1.3.1), but can consist of a limited amount of sweeps. In order to support single trial estimation under poor SNR conditions, filtering (Nishida *et al.*, 1993), pre-whitening of the template (Lange and Inbar, 1996), or continuous updating of the averaged template (Jensen *et al.*, 1996) can be used. Estimation of individual (statistical) components in the obtained EP can be performed by principal component analysis (Elkfafi *et al.*, 1997; Lange *et al.*, 1997).

Time-frequency analysis

Wavelet analysis (introduced in § I-1.3.4) can be used successfully for detection of SEP changes, where onset of hypoxia in brain injury was detected a few minutes earlier when compared to conventional averaging (Braun *et al.*, 1996). Other studies describe its use to record single trial P300 or CNV measurements (Geva *et al.*, 1997; Saatchi *et al.*, 1997), or use Wavelet decomposition to enhance late components (Hoppe *et al.*, 1996). The usefulness of time-frequency analysis depends heavily on the selected distribution

(Boudreaux-Bartels and Murray, 1995), and is reliable only in relatively favourable SNR conditions.

II-1.3.4. Quality estimation

Quantification of the signal-to-noise power ratio in evoked potentials is needed for objective assessment of the residual noise power in the (averaged) resulting waveform. We will now discuss quality estimation for the conventional averaging methods, as it is still the most important processing technique; even in alternative calculation methods as described in the previous paragraph, a conventional average is often used as a reference.

Quantitative SNR assessment

As we have seen, the noise amplitude (background EEG) is reduced proportional to the square root of the number of trials. Ergo, doubling of recording time increases the signal-to-noise power ratio only by 3dB. Let us consider an EP of intrinsic SNR of -20dB: using a stimulation rate of six per second (inter-trial-interval 167ms) we already need close to 3 minutes to arrive at an SNR of 10dB. Such long measurements are not desired, and may not even result in increased SNR because of habituation effects or adaptation of the sensory channel (Regan, 1989).

Two principal questions arise in practical averaging: (a) how to decide when a true EP signal is present; (b) how to compare averages measured at different times or from different sources. To deal with these questions, an early computer-program by Lowy and Weiss (1968) displayed two EP waveforms from different recording positions or from odd and even sweeps. However, they did not quantitatively estimate the signal-to-noise power.

The (±)-reference. The (±)-reference is the average signal that results from alternate addition and subtraction of sweeps, and division by *N* (the number of trials). By this process, the EP signal is eliminated, thus providing an estimate of the noise (Schimmel, 1967). When $\bar{x}_N(t)$ (formula II-1.2) represents the normal 'added' average after *N* trials, i.e., the summation of evoked response and residual noise, the (±)-reference $\bar{x}'_N(t)$ can be calculated for even numbers of trials *N*:

$$\bar{x}'_N(t) = \frac{1}{N}\sum_{i=1}^{N}(-1)^i \cdot \{s_i(t) + n_i(t)\} = \frac{1}{N}\sum_{i=1}^{N}(-1)^i \cdot n_i(t) = n'(t) \qquad \text{(II-1.6)}$$

$s_i(t) = s_j(t)$ for any *i*, *j*, therefore: $\sum_{i=1}^{N}(-1)^i \cdot s_i(t) = 0$

Schimmel *et al.* (1974) showed that the variance of $n'(t)$ can be expressed in terms of random noise: $\mathrm{var}\{n'(t)\} \approx \sigma_n^2 / N$. Now, we can calculate a *power ratio P*:

$$P = \frac{\mathrm{var}\{\bar{x}_N(t)\}}{\mathrm{var}\{\bar{x}'_N(t)\}} \tag{II-1.7}$$

$$= \frac{\mathrm{var}\{s(t)+n(t)\}}{\mathrm{var}\{n'(t)\}} \cong \frac{\mathrm{var}\{s(t)\}}{\mathrm{var}\{n'(t)\}} = \frac{\sigma_s^2}{\sigma_n^2/N} = N\left(\frac{\sigma_s^2}{\sigma_n^2}\right) \tag{II-1.8}$$

If *no* time-locked response is averaged, the variance ratio of II-1.7 will converge to 1, because both nominator and denominator are the result of an identical process. Therefore, formula II-1.8 only holds when a response signal is measured (Schimmel *et al.*, 1974). P factors larger than 1 (0 dB) represent a true signal-to-noise power ratio, increasing proportional to N.

An alternative implementation of the (±)-reference, inspired by the Lowy and Weiss approach, calculates the 'odd' and 'even' averages (see Figure II-1.3). Apart from SNR calculation, this procedure allows for visual comparison of sub-averages S1 and S2.

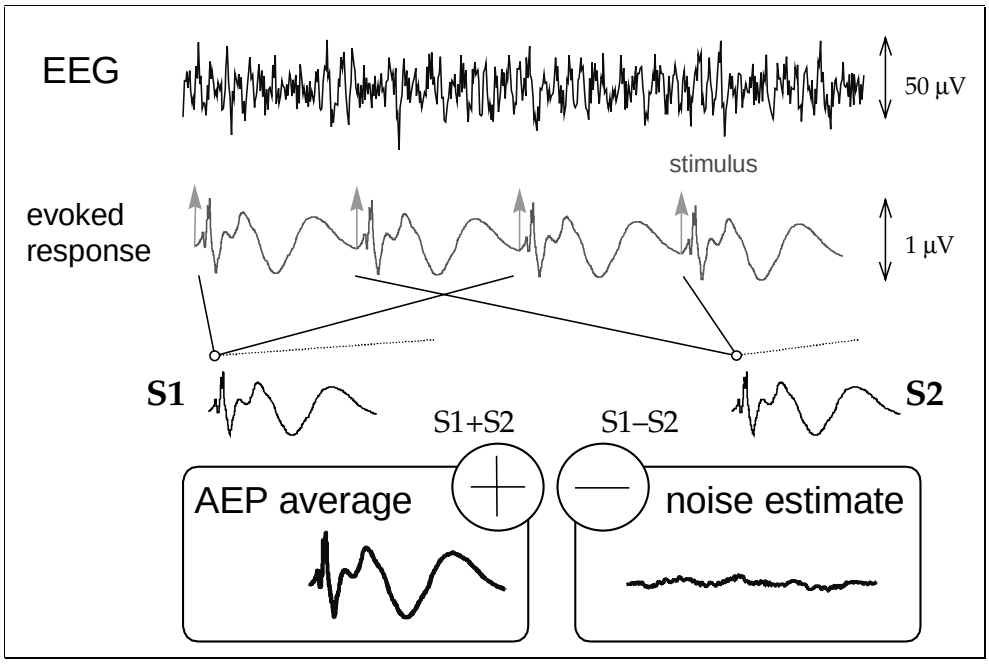

Figure II-1.3 Schematic diagram of a slightly modified implementation of the (±)-reference method. A noise estimate is obtained by subtraction of odd (S1) *and even* (S2) *averages. Addition of* S1 *and* S2 *produces the AEP average resulting from the measurement.*

Single point reference. A method derived from the previous procedure was developed by Elberling and Don (1984). They calculated a ratio of variances F_{sp} based on a single point reference in the evoked potential:

$$F_{sp} = \frac{\sigma_s^2}{\sigma_{sp}^2} \qquad \text{(II-1.9)}$$

where σ_s^2: variance of the *averaged response* $s(t)$, divided by

σ_{sp}^2: variance of the *background 'noise'*, estimated at $t = (sp)$

The F_{sp} statistic is also akin to signal-to-noise power ratio, and is a good alternative to the (±)-reference (Özdamar *et al.*, 1990). The time point *sp* can be arbitrarily chosen, and although it is thereby sensitive to time-locked artefacts (§ II-1.2.2), this allows for the investigation of the reproducibility of individual EP features.

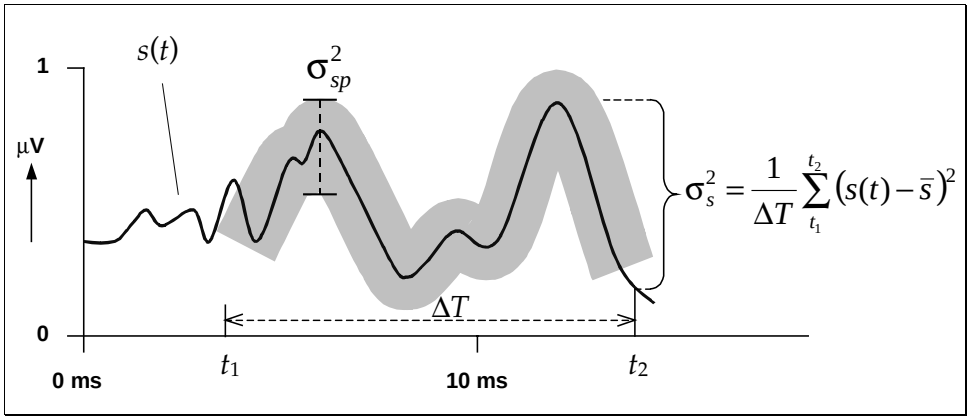

Figure II-1.4 Schematic explanation of the F_{sp} statistic calculation in a brainstem auditory evoked potential. The averaged signal converges to the indicated EP signal, and the calculated variance σ_s^2 converges accordingly. The variance σ_{sp}^2 will diminish during the recording, hence, the ratio σ_s^2/σ_{sp}^2 will increase.

Application in BAEP investigations. The mean power ratio P was investigated by Wong and Bickford (1980) in brainstem evoked potentials. They concluded that already for $P > 2$, there is a high likelihood that $\bar{x}(t)$ contains a non-random wave form, i.e., a physiological signal with a significant peak component. They also concluded that BAEP waveforms of $P > 30$ were associated with highly reproducible BAEPs, and low values ($P < 20$) were indicative of dissimilar BAEPs, or sources of noise.

In a BAEP study by Don *et al.* (1984), the F_{sp} statistic was concluded to have better precision than the (±)-reference for prediction of an adequate response. Here, the response signal was in the interval 4-14ms, and σ^2_{sp} was calculated at sp = 6ms, using $F_{sp} > 3.1$ as detection criterion (see Figure II-1.4). The statistic was recalculated for every 250 sweeps during the measurement. In another study, very similar to BAEP recording, F_{sp} values larger than 2.0 were found to reliably indicate the presence of an oto-acoustic ear response ('cochlear echo' — Lutman and Sheppard, 1990).

Related processing issues

Phase synchronised triggering. Careful implementation of a SNR calculation method is needed when EP triggering is synchronised with the mains signal. Alternated addition and subtraction (as in the (±)-reference) must not fall in opposite phases; alternating every two sweeps will solve this problem. The same applies to noise that is coherent at other frequencies (Emerson and Sgro, 1985; Sgro *et al.*, 1989b).

In *median averaging* of EP trials, the median value at each sample point over all trials is used instead of the mean value (Sgro *et al.*, 1989b). This is useful to filter outliers in the data, for example due to noise or other transient activity, and has been successful in the recording of motor unit action potentials (Nandekar and Sanders, 1989). A technical and practical disadvantage of this method is that it requires large amounts of computer memory, as all single sweeps must be stored during the measurement. The waveform itself may be reliable only when large numbers of trials are averaged. Statistical outlier detection may be preferred then (e.g., p. 84).

II-1.4. SPECIAL TOPICS IN EP PROCESSING

II-1.4.1. EOG detection and correction

As described before, eye movement and blinks cause low-frequency, high-amplitude artefact in the EEG, and can affect the potentials recorded anywhere on the scalp. Because of overlapping frequency ranges, this type of artefact is notably a problem in the recording of 'late' ERPs (P300, CNV). Although amplitude thresholds may be used for rejection of contaminated trials, this can be unspecific, especially when channels from anterior scalp positions are unavailable (Verleger, 1993). Rejection is undesirable, because it may result in a significant loss of data. Eye fixation methods, used to prevent eye artefacts, can introduce unwanted effects in cognitive ERP research (Weerts and Lang, 1973; Verleger, 1991). Therefore, *correction* of eye artefacts is advised. Various methods have been reviewed by Brunia *et al.* (1989).

A parametric method may be employed in ERP experiments, using three EOG derivations and incorporating the EEG activity through AR modeling. The estimated parameters can be used for proportional subtraction of the EOG influence in the target EEG/ERP channels. Parameters are preferably estimated from a calibration trial, but may be obtained directly from the measurement in off-line processing (Van den Berg-Lenssen *et al.*, 1989, 1994). Only one proportional coefficient per EEG channel is sufficient since different types of eye artefact propagate equally from the EOG to the EEG, as indicated again in a recent study by Croft and Barry (1998a,b).

Berg and Scherg (1994) used a model based on multiple source eye analysis. In their correction method, the artefactual EOG influence was calculated from the EEG, using the estimated spatial distribution of eye activity (dipole modeling). The method enhanced the precision of topographical EEG analysis. Calibration of the model was obtained from data containing systematic eye movements and blinks. In order to overcome the (theoretical) constraint of orthogonality of the model's source vectors, Vigário (1997) proposed a statistical method called *independent component analysis* as a tool to separate eye activity from EEG activity. Although its discriminative abilities were shown, a possible drawback of this method may be that it introduces a new set of traces for each electrode that need additional interpretation. In theory, the method could be used without any EOG channel, however, the use of simultaneous EOG recording was still advised for referential purposes.

II-1.4.2. Feature detection and classification

Clinical applications of EPs investigate the amplitude and location of characteristic peaks in the recorded waveforms. Techniques that use template matching may have no difficulty in tracking EP latencies (see § II-1.3.3). However, caution is advised when individual peaks are not always 'dominantly' present in the measurement. This was a problem for instance in an expert system by Boston (1989), which heavily relied on the consistency of amplitude parameters and correlation coefficients between intervals.

A rule-based expert system can favourably be extended by applying matched filtering pre-processing of the EP signal (Delgado, 1993). For instance, convolution of primary frequency components in BAEPs can 'roughly' estimate the important peak V (at 5-6ms). Followed by rule-based determination of peak and inter-peak latencies, over 80% correct localisation of peaks was reported. Performance was less in BAEP assessment of neurological abnormalities, and in noisy recordings (Delgado and Özdamar, 1994).

Artificial neural networks (ANN) have been shown as a useful means to identify peaks. Robust detection by ANNs may depend on adequately fitted bandpass filtering in the pre-processing step (Tian *et al.*, 1997), or can be accomplished by cascaded *classification* and subsequent *detection* of peaks (Van Gils, 1995). The methods employed by Van Gils reached expert performance for automatic recognition of BAEP peak V (around 90%

correct within 0.2ms), but assessment of middle-latency peaks (MLAEP) showed relatively large variations in latency (6-11ms). The data were obtained in noisy clinical conditions, and human consensus about latencies was low as well (5-9ms difference). Nevertheless, the system was very useful as an aid to reduce the time for human scoring.

For the detection of long-latency peaks (P300) low-pass filtering (<3.4Hz) was favoured over template matching (Smulders *et al.*, 1994). The latency was reliably calculated as the largest local maximum in the expected peak range (300-1000ms). The authors concluded that the method still had to be improved for automatic discrimination between presence and absence of the P300 component. This classification can be performed by using ANN-based pattern recognition, as shown in study by Gupta *et al.* (1995). Their method correctly classified around 80% of P300 ERPs in a patient suffering from head-injury.

II-1.4.3. Non-linear characteristics

The nervous system is known to exhibit non-linear behaviour. When a nerve cell is excited, an action potential can occur (see § I-1.1.1 — the mechanism leading to the action potential itself is non-linear). In a short period (a few milliseconds), just after its start, additional stimulation will not result in a new action potential. This period is called the refractory period.

From the above, we can expect changes in the character of the evoked electrical activity due to previously applied stimuli. This can be investigated by using random trains of stimuli where the inter-stimulus intervals are taken from a Poisson distribution (Krausz, 1975). The technique estimates a set of *kernels* h_1 .. h_N , where the first order kernel h_1 resembles the conventional averaged response (Cluitmans and Beneken, 1991). Higher order kernels describe the evoked response under the condition of prior stimuli. For example, the 2nd order $h_2(\delta)$ is related to the difference between the expected (linear) response and the true (non-linear) response of the system on a stimulus, as a function of the inter-stimulus interval δ. Because of the random inter-stimulus intervals, we can significantly increase the presentation rate. The interference effect of overlapping responses will average out, because of the lack of synchronicity between responses. The quality of EP signals may thus be enhanced when more responses are averaged in shorter recordings.

Sclabassi *et al.* (1977, 1982) have calculated the second order kernel h_2 in the somatosensory system of cats and humans, and clearly demonstrated a pattern that included inverted h_1 components. Further research by Cluitmans has indicated the usefulness of h_1 and h_2 of auditory evoked potentials obtained during anaesthesia. Non-linear effects were larger and better quantifiable in middle latency components (Cluitmans, 1990; Cerutti *et al.*, 1996).

II-1.5. DISCUSSION

Most of the aspects related to the recording and processing of evoked potentials were reviewed. When compared to the EEG, several things can be observed:

- Amplitudes in EPs are generally (very) low, but the frequency range extends to much higher frequencies (kHz) than in EEG.
- EPs are of deterministic nature. Waveforms are of expected morphology, and usually show relatively small variations over time.

Artefact detection. Amplitude criteria for *rejection* of trials during EP measurements are useful by virtue of the first observation above. The occurrence of high-amplitude artefacts is disastrous for EPs. Therefore, max-min controls are available in most commercial EP equipment. Further exploration of (statistical) amplitude thresholds is advantageous (De Beer *et al.*, 1995), and can be implemented as a well-structured and workable approach to artefact detection in the recording of EPs.

Validation by means of frequency analysis is difficult because of overlapping spectra of EEGs and EPs. Only at the ends of the frequency scale, high-pass filtering may be applied in Brainstem EPs (10ms), and low-pass filtering is useful in ERPs (1s). Other correction procedures are successful only a when a reference of the interfering artefact can be recorded and/or modeled. For instance, eye artefacts can be filtered reliably by proportional subtraction of pre-frontal EOG channels.

Quality assessment. The commonly applied procedure for recording EPs is averaging of single trials, thus cancelling non-stimulus related 'background EEG'. Although successful attempts have been reported for extraction of single trial EP signals, reference templates are still based on averaged responses. The averaging process requires a sufficient number of trials, but both from a theoretical point (response variability) as well as a practical point (limited time) the recording should be as short as possible. Surprisingly, the actual *quality* is seldom assessed nor discussed quantitatively.

Quantitative quality assessment is well possible by means of signal-to-noise ratio assessment. For this purpose, we need to calculate the (enhanced) power in the EP and estimate the (reduced) power in the background EEG. Several methods have been described in the literature, of which the (±)-reference and F_{sp} ratio have been most useful.

Promoting objective EP validation methods is one overall aim of Part II of this thesis. Therefore, the remaining chapters of this thesis will focus on quantitative quality assessment and improvement of EPs, in relation to clinical judgement, recording length and stimulation frequency.

II-3. Quality Assessment in Event Related Potential Measurements:

an explorative study of CNV lateralization

Abstract — Event related potentials are signals that can be evoked from the nervous system, and measured at the (human) scalp through methodical stimulation of one or more sensory modalities. The actual response signal is always deeply embedded in the background EEG activity, where signal-to-noise power ratio (SNR) for a single response can be as low as -20dB. Therefore, additional processing is needed to extract the evoked response signal from the measurement. The most common processing method is averaging of single trials (sweeps), by which the event related (time-locked to the stimulus) signal is enhanced and the unlocked signal is cancelled out. However, while this, or any other, processing technique is meant to improve the signal-to-noise ratio of the response signal, it is still uncommon to indicate the quality of the resulting waveform.

The current study explores a previously acquired data set, consisting of 'Contingent Negative Variation' (CNV) waveforms. This signal is measured in a task where a warning-stimulus and a subsequent response-stimulus are presented to a subject, which is to be followed by a prompt button-press action. This data set provides an interesting starting-point to investigate the usefulness of SNR assessment in event related potentials, because of the effect of response lateralization: the amplitudes of the (various parts of) waveforms are dependent on the limb used in the response action. This effect was found in the original study, and is indicated here also from the evaluation of the SNR. Several other statistical significant differences were found in the analyses, leading to the overall conclusion that the (±)-reference method provides useful information, corroborating the findings from other amplitude based analyses.

II-3.1. INTRODUCTION

The 'contingent negative variation' (CNV) can be measured at the scalp of a subject involved in a repeated task where first an auditory warning stimulus (WS) is presented, and after a short period a visual response-stimulus (RS) is presented, to be acted upon by immediate button-pressing. The interesting effect present in the CNV waveform, is the slowly increasing negative potential that appears before the response stimulus (see Figure II-3.1). The processes involved in between the WS and RS are related to anticipatory attention to the RS, and to motor preparation and execution of the response (Brunia and Damen, 1988; Brunia, 1993).

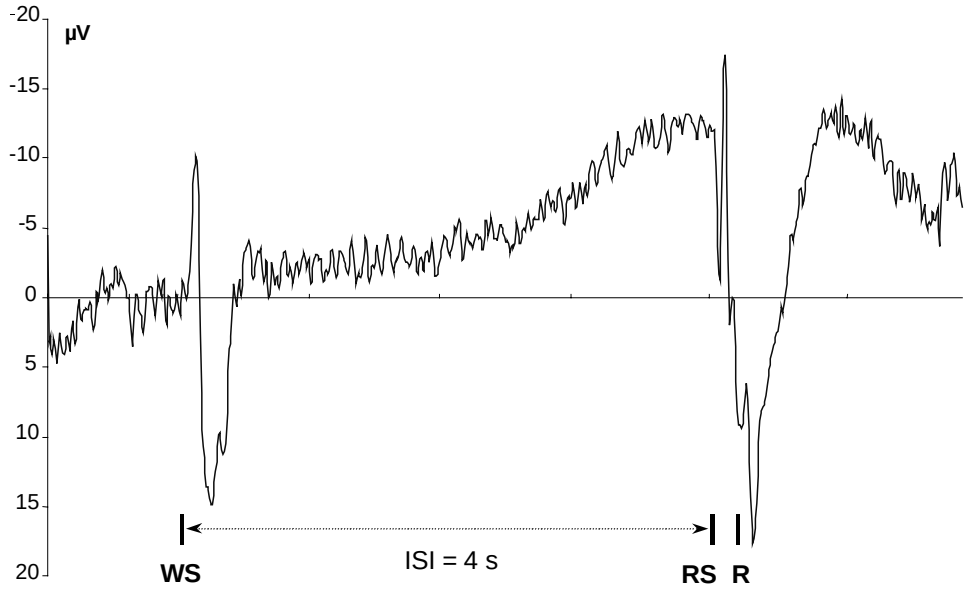

Figure II-3.1 The Contingent Negative Variation (CNV) occurs just before the response stimulus, after an initial auditory warning stimulus and a four second inter-stimulus-interval (ISI). Vertical markings indicate the timing of the warning stimulus (WS), response stimulus (RS), and average response (R) respectively. Note: negativity plotted upwards.

The data of the current study were obtained at the Department of Physiological Psychology, Tilburg University, where they have been used to investigate the described effects by dipole modeling from the recorded signals (Böker, 1994). The experiments involved the use of different ISI lengths (1 and 4 seconds), in order to try to separate the different psycho-physiological processes. Specifically, the CNVs of 4s ISI allowed for discriminative analyses on the parts of the CNV defined as the early wave and the late wave (in the interval between WS and RS). One of the main conclusions of this research

was the fact that no likely sources for stimulus anticipation were found for the paradigm used in this study (Böker, 1994).

In ERP research like this, inferences are made by comparing two or more signals, without having quantitative knowledge about the quality of the signals. There is a risk that the matched waveforms appear to have comparable quality when observed visually, and include approximately the same number of responses, but in fact have signal-to-noise ratios of different magnitude, therefore invalidating the conclusions. Still, even in recent publications focusing on measurement techniques, the SNR is not indicated as a standard measure of quality (e.g., Cohen and Polich, 1997; Saatchi *et al.*, 1997).

For the current study the CNV data of the study by Böker (1994) have been processed anew, but now including SNR calculation. This investigation serves two purposes. First, we will explore the quality of CNV averages, using a method for SNR assessment known as the (±)-reference (Schimmel, 1967). The use of this method has been restricted mainly to short-latency potentials, but it is equally suited for calculation of the quality of middle-latency and long-latency waveforms (chapter II-2). Second, we will show that this quantitative analysis can be used to investigate CNV characteristics. Specifically, the current study investigates the effect of lateralization: cerebral dominance caused by the response action from a left or right limb. The CNV late wave is more dominant on the contra-lateral hemisphere for finger responses; a counteracting paradoxical lateralization effect can be observed for foot responses.

We will test the hypothesis that signals obtained at the dominant side will also result in a higher signal-to-noise ratio. The electrodes over the non-dominant hemisphere will be farther away from the source of the response activity, hence will pick up more EEG background activity ('noise') that is unrelated to the stimuli.

II-3.2. METHODS

II-3.2.1. Calculation of signal-to-noise power ratio

The signal-to-noise power ratio (SNR) is calculated here through the (±)-reference method, which can be implemented by using two separate CNV averages during the averaging process. If all single responses are identical, the two averages will be alike, and differ only because of the non-normality of the background EEG. In the formulas below, x_1 and x_2 represent the summations of odd-numbered and even-numbered trials, consisting of respectively N_1 and N_2 trials. N_1 and N_2 will differ at most by one trial . The summation of x_1 and x_2 represents the end-result of the CNV average, the subtraction represents the (±)-reference. The variance of the summation, divided by the

variance of the difference signal results in a true signal-to-noise power ratio (Schimmel *et al.*, 1974) denoted by P, which we expressed in Decibels. In the original (±)-reference only even numbers of trials are allowed, resulting in $N_1 = N_2 = \frac{1}{2}N$ where N represents the total number of trials. However, in a typical CNV average N is relatively low (between 50 to 150), and in order to use all trials — including any incidental odd-numbered last trials — we used the following formula:

$$P_{dB} = 10\log\frac{\mathrm{var}\left\{\dfrac{x_1 + x_2}{N_1 + N_2}\right\}}{\mathrm{var}\left\{\dfrac{1}{2}\left(\dfrac{x_1}{N_1} - \dfrac{x_2}{N_2}\right)\right\}} \qquad \text{(II-3.1)}$$

The P_{dB}-factor was calculated in different post-stimulus intervals of the CNV signals, in order to investigate separately the SNR and lateralization effects of the stimulus, the CNV and the response components. A typical value for the *P*-factor in evoked potential waveforms of good quality is 13dB or higher (see Wong and Bickford, 1980; chapter II-2).

II-3.2.2. Subjects & recordings

The study included 10 right-handed subjects, 4 male and 6 female (age 20-31 years). The subjects were paid volunteers. During the experiments, the subjects were seated comfortably in a slightly reclining chair, which was placed in a sound attenuating, electrically shielded cubicle. A recording consisted of a series of trials, where each trial started with the presentation of an auditory warning-stimulus (WS) through a loudspeaker mounted on the wall 1m behind the subject (a 70 dB(A), 1000Hz tone of 45 ms duration). Either one or four seconds later a visual response-stimulus (RS) was presented by illumination (60ms duration) of four bright red LED's, placed 1.5m in front of the subject, in the centre of the visual field. The subjects were instructed to react as fast as possible to the presentation of the RS by flexing a pre-assigned finger or foot. They were also instructed to fix their gaze on the box containing the LED's (to preclude eye movements and blinks) and to prevent excessive body-movement, sneezing or swallowing except for a short period starting approximately 3s after a response.

Determination of reaction time. The subject rested his/her arms on the adjustable arms of the chair, holding a small cylinder (length 5.5cm) between thumb and index finger of each hand. The cylinders were mounted on top of the chair's arms. A switch at the end of each cylinder was operated by flexion of the two fingers, from which the reaction time was determined for a given trial. The feet rested on two separate foot-plates, elevated 30 degrees from the horizontal. The force needed to depress the plates until closure of the switch was adjusted by a spring to be subjectively equal on both sides. A small foot

flexion produced closure of the switch and stopped the reaction timer. Trials with reaction times outside the 100-500ms range were excluded from further processing.

A single registration consisted of 6 blocks of 25 trials, resulting in a CNV average of maximum 150 trials. Eight different CNV signals were obtained, for four limbs and for both CNV_1 and CNV_4, which were recorded on 2 separate days, about one week apart. Half the subjects reacted with finger responses on the first day and with foot responses on the second day; this was reversed in the other subjects. After each fourth block, the experiment changed from CNV_1 to CNV_4, or vice versa. This was counter-balanced between subjects as well as within subjects between both recording sessions. The response side was varied pseudo-randomly between blocks. A block, which took 3'45'' and 5'50'' for CNV_1 and CNV_4 respectively, was followed by a 1 to 2 minutes break. Individual trials were separated by intervals ranging from 6-10s, or 8-12s (steps of 0.5s) for CNV_1 and CNV_4 respectively.

After each fourth block, a longer break allowed for the recording of calibration pulses, and for the recording of eye movements for parameterisation of the EOG signals. The latter procedure was used in a method for off-line correction of eye movement artefacts in the EEG derivations. The horizontal EOG from the outer canthi and the vertical EOGs of both eyes were used for this procedure. This method is based on maximum likelihood parameter estimation of the propagation of the EOG into the different EEG channels. The background EEG activity in the EOG channels is modeled using an autoregressive function of order 3. Possible delays between EOG and EEG positions are also taken into account but are neglegibly small (Van den Berg-Lenssen *et al.*, 1989).

Additional off-line artefact rejection was based on the detection of high-amplitude peaks (>100μV), scanning the EEG for drift. Drift was detected on the basis of two criteria, one maximum threshold of 70μV for individual EEG samples, and another threshold (>30μV) for drift in 4 successive intervals after the baseline period of each trial (amplitudes with respect to the baseline level). These specifications describe the most liberal thresholds, and were manually optimised for individual measurements (Böker, 1994).

The EEG was recorded from 26 electrodes at scalp-positions (see Figure II-3.2), centred on vertex position Cz'. The recordings obtained at the electrode positions as indicated in the figure were referenced to the average of left and right mastoid electrodes, thus allowing for the investigation of lateralization effects. The EEG was digitised at 128 Hz sample frequence, using a 7^{th} order Butterworth low-pass filter at 30Hz (-42dB/octave) and a 30s time-constant. Recording of trials started one second before WS (baseline period) and lasted until 2.5 second after the response.

After the described pre-processing of the data the signals were converted from the local Tilburg file-format to the extended European Data Format (see appendix), which was used in the current study.

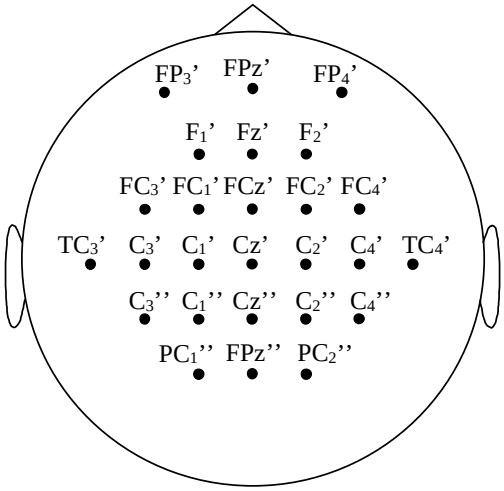

Figure II-3.2 Electrode scheme as used during the experiments: nomenclature is derived from the International 10-20 system. Primes (') denote positions one cm anterior and double primes ('') two cm posterior to the standard coordinates. Inter-electrode distances measure 10% of the nasion-inion distance.

II-3.3. RESULTS

On average, 102 trials were included in the intra-individual CNV waveforms (N=80, standard deviation 20, minimum 45, maximum 137). Apart from the different types of CNV (1s, 4s) and different limbs involved, the data set was processed twice: once averaged time-locked to the stimulus pattern (WS and RS, constant ISI within one type of CNV), and once averaged time-locked to each response (R) in individual trials.

Statistical significance between SNR factors between identical electrode positions in CNV_1 and CNV_4, and between opposite positions within one type of CNV was determined by the one-sided *t*-test for paired observations and complemented by calculating the non-parametric Wilcoxon ranked-sum statistic. These two complementary tests were used because of the relatively low number of observations (N=10), which prevents reliable testing of the form of the distributions. In a normal distribution, the Wilcoxon statistic is less powerful than the *t*-test, but it is more stable when the normality assumption is incorrect (Montgomery and Runger, 1994). Therefore, we used the significance of the Wilcoxon test to substantiate the significance of the *t*-test.

In general, the signal quality reached acceptable levels when calculating the SNR over the entire CNV signal, from WS to RS+1000ms. This resulted in an average SNR value of

approximately 15dB for CNV_1 signals. The SNR factors for the CNV_4 waves were approximately 4dB lower, which is mainly influenced by a relatively long interval of low amplitude signal in-between the WS and RS (see Figure II-3.1). Because of the different ISI lengths, these initial SNR factors were not compared statistically. We calculated the statistics of the SNR in different CNV components, using only equal intervals in CNV_1 and CNV_4 for the evaluations.

The period in between the WS and RS, the actual 'negative variation', showed very low SNR values (<5dB) using the current methods. As indicated before, the small amplitude fluctuations in this period result in an unpronounced waveform in x_1 and x_2, which invalidates the assumptions of the (±)-reference and produces unacceptably low *P* values (equation II-3.1, also see § II-1.3.4, p. 90). The effects of absolute amplitude measures in the ISI of the CNVs have been studied by Böker (1994), and are not part of the current investigation.

SNR evaluation of the WS component

We found that the SNR for *finger* responses is higher in CNV_1 than in CNV_4 for the WS component from t(WS) to t(WS+500ms) of the signal. A significant difference of approximately 2dB was found for the central positions in the left finger response (see Table II-3.1), and not in the right finger response. The average quality of the WS of the 'left finger' CNV_4 signals still reached an acceptable 15dB.

*Table II-3.1 CNV_1 (left finger response) from t(WS) .. t(WS+500ms): signal-to-noise power ratios, averages and standard devations over 10 subjects. The presented averages were approximately 2dB higher than at identical positions in CNV_4, significant in the central region (shaded): *) t-test (α=0.05), **) both t-test and Wilcoxon ranked-sum (α=0.05).*

	FP3' 11.5 (3.1)		FPz' 13.3 (2.9)		FP4' 11.5 (3.0)	
		F1' 14.9 (2.4)	Fz' 15.6 (2.3)	F2' 15.1 (2.2)		
	FC3' 15.9 (2.6)	**FC1' ** 16.6 (2.2)**	**FCz' ** 17.0 (2.0)**	**FC2' ** 16.8 (2.1)**	FC4' 16.1 (2.9)	
TC3' 15.4 (3.4)	**C3' ** 17.6 (2.9)**	**C1' ** 17.9 (2.4)**	**Cz' 17.6 (2.3)**	**C2' * 17.6 (2.4)**	**C4' * 17.5 (2.9)**	TC4' 15.8 (3.1)
	C3'' ** 17.7 (2.9)	C1'' * **17.7 (2.6)**	Cz'' 17.3 (2.5)	**C2'' * 17.4 (2.7)**	C4'' 17.0 (2.9)	
		PC1'' 16.1 (2.9)	PCz'' 16.0 (2.8)	PC2'' 15.7 (3.0)		

When comparing the WS component for the *foot* responses between CNV_1 and CNV_4 in the same interval as above, we found that the SNR was higher in the CNV_4. This

difference (approximately 0.5μV at the vertex, to 3μV at pre-frontal positions) was significant only for the pre-frontal and parietal positions in the 'left foot' CNV_4 (*t*-test, α=0.02). This points to a larger spread over the scalp of the WS component in this type of CNV.

Testing for SNR differences between hemispheres within CNV_1 or within CNV_4 did not reveal any lateralization in the WS component.

SNR evaluation of the (stimulus-locked) RS component

We found some interesting effects when focusing on the CNV activity around the RS, from 100ms before RS to the early part of the response t(RS+200ms). In the original study, this was the interval nearest to the response that still showed a significant amplitude effect for different ISIs (1s, 4s). In the current study, we found that the SNR factors for CNV_1 'left foot' (see Table II-3.2 and Figure II-3.3) and for CNV_4 'left finger' are significantly lateralized in the expected direction. In the CNV_4 finger responses, this effect was not clearly observed in the original study. A statistically insignificant, but similar lateralization pattern toward the right central and parietal positions was also observed in the other finger and foot responses for CNV_1 and CNV_4 .

*Table II-3.2 CNV_1 (left foot response) from t(RS-100ms) .. t(RS+200ms): signal-to-noise power ratios, averages and standard deviations over 10 subjects. Significant higher values (tested versus homologue location on opposite hemisphere) are indicated by shading: t-test *) α=0.05, **) α=0.01.*

	FP3' 6.3 (2.7)		*FPz'* 7.3 (3.1)		*FP4'* 7.0 (3.0)	
		F1' 7.2 (3.0)	*Fz'* 7.4 (2.9)	***F2'*** 7.6 (2.8)		
	FC3' 7.1 (2.6)	*FC1'* 7.3 (2.6)	*FCz'* 7.3 (2.1)	***FC2'*** 7.6 (2.2)	***FC4'*** 7.8 (2.8)	
TC3' 5.8 (3.4)	*C3'* 7.1 (2.7)	*C1'* 7.4 (2.1)	*Cz'* 7.6 (2.7)	***C2' **** **8.2 (2.4)**	***C4' ***** **8.6 (3.1)**	***TC4' ***** **8.3 (3.3)**
	C3'' 7.2 (3.6)	*C1''* 8.0 (2.7)	*Cz''* 8.6 (3.3)	***C2'' **** **9.0 (3.4)**	***C4'' ***** **9.1 (3.5)**	
		PC1'' 8.3 (4.4)	*PCz''* 8.8 (4.3)	***PC2'' **** **9.6 (4.9)**		

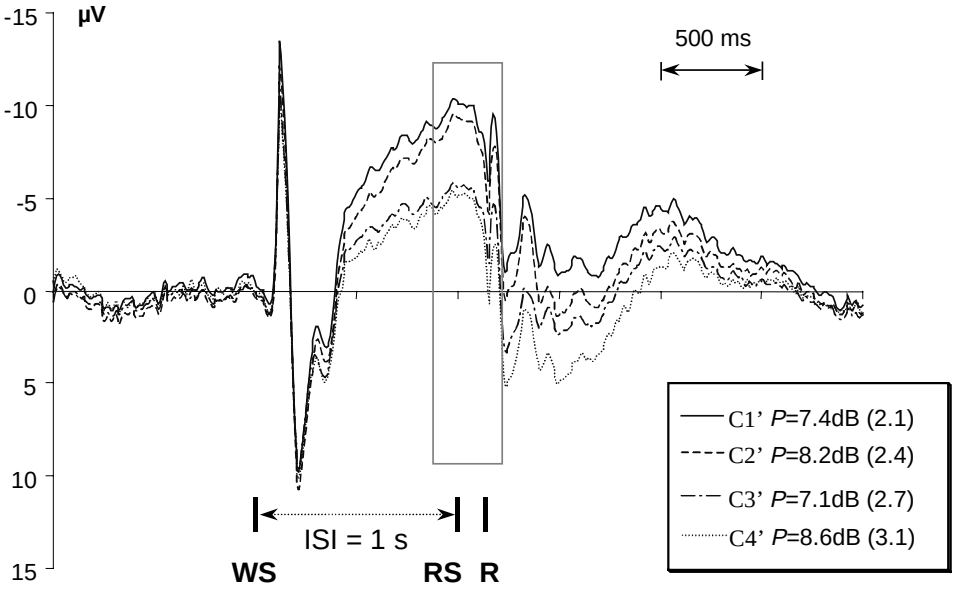

Figure II-3.3 Grand averages of the CNV_1 'left foot' waveforms: stimulus-locked average waveforms of 10 subjects. The signal-to-noise power ratios P were calculated in the interval t(RS-100ms) .. t(RS+200ms) (indicated in legend: averages and standard deviations over subjects). The differences were significant between electrodes positions C1', C2', and between C3', C4'. Negativity plotted upwards.

SNR evaluation of the (response-locked) R component

The largest effect of SNR and dominance occurs within the 'left finger' CNVs that were averaged time-locked to the responses. Here, we focused on the response component only, for the interval t(R) to t(R+500ms), which presents a motor component at the contralateral side that does not occur ipsilaterally. This effect is significant for CNV_1 , and highly significant for CNV_4 signals (see Table II-3.3).

In both CNV_1 and CNV_4 , a small *paradoxical* lateralization effect for foot responses was found (higher SNR over the ipsilateral hemisphere). This effect was most clearly observed in the SNR values of the CNV_1 'right foot' responses (Table II-3.4). No clear statistical pattern was found. Only when comparing the *left* foot response between CNV_1 and CNV_4 , we found that the SNR was significantly higher in the CNV_1 for the ipsilateral positions FC1p, C3p, (*t*-test, α=0.05), and FC3p, TC3p (α=0.02). These SNR values in CNV_1 showed a remarkable symmetrical pattern, whereas the CNV_4 SNR values

Table II-3.3 *CNV_4 (left finger response) from t(R) .. t(R+500ms): signal-to-noise power ratios, averages and standard deviations over 10 subjects. Significant higher values (tested versus homologue location on opposite hemisphere) are indicated by shading: *) t-test (α=0.02), **) both t-test and Wilcoxon ranked-sum (α=0.05), and ***) Wilcoxon (α=0.01).*

	FP3' 9.7 (3.4)		FPz' 11.0 (2.9)		FP4' 10.1 (2.1)	
		F1' 10.3 (2.5)	Fz' 11.0 (2.6)	**F2' * 11.2 (2.2)**		
	FC3' 8.6 (2.3)	FC1' 9.8 (2.6)	FCz' 11.8 (2.6)	**FC2' ** 12.5 (2.0)**	**FC4' *** 11.3 (1.6)**	
TC3' 8.7 (2.6)	C3' 9.3 (2.5)	C1' 10.6 (2.8)	Cz' 13.2 (3.0)	**C2' ** 13.8 (2.2)**	**C4' *** 13.1 (2.4)**	**TC4' *** 12.0 (2.1)**
	C3'' 10.1 (3.1)	C1'' 10.6 (3.5)	Cz'' 12.2 (3.3)	**C2'' * 12.8 (2.8)**	**C4'' ** 13.1 (2.1)**	
		PC1'' 9.7 (3.6)	PCz'' 10.0 (3.6)	**PC2'' * 11.0 (2.9)**		

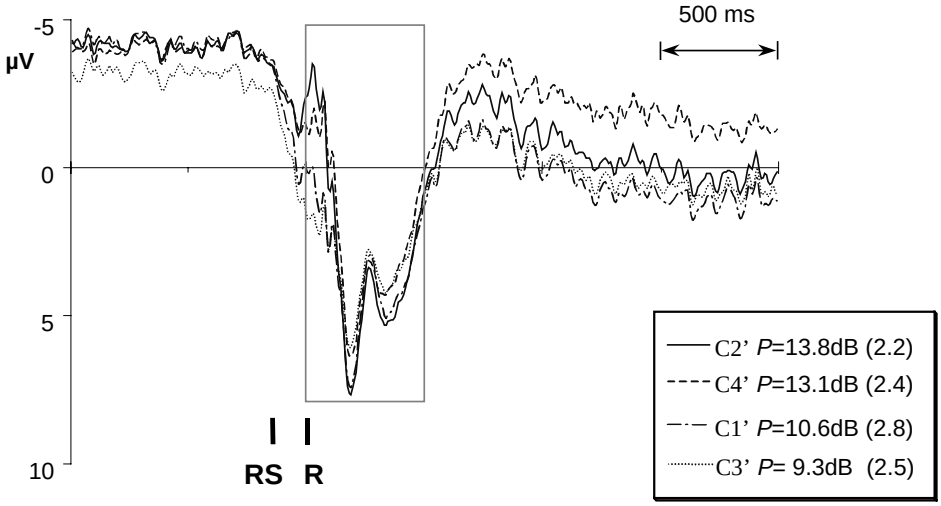

Figure II-3.4 Grand averages of the response component of the CNV_4 'left finger' waveforms: response-locked average waveforms of 10 subjects. The marker 'R' indicates the average response time, following directly after the RS (approximate position indicated), 4 seconds after the initial warning stimulus. The signal-to-noise power ratios P were calculated in the interval t(R) .. t(R+500ms) (indicated in legend: averages and standard deviations over subjects). The P values on right-side positions C2' and C4' were significantly higher than corresponding electrodes C1' and C3' on the left hemisphere. Negativity plotted upwards.

were lower on the left hemisphere. Therefore, we can only conclude that a small paradoxical lateralization exists in the CNV_1 foot responses, observed most clearly for right foot movements.

*Table II-3.4 CNV$_1$ (right foot response) from t(R) .. t(R+500ms): signal-to-noise power ratios, averages and standard deviations over 10 subjects. This response-locked average shows a tendency (bold typeface) toward paradoxical lateralization in the right frontal and central regions (** only significant for FP4' versus FP3', t-test, α=0.01).*

	FP3' 8.5 (4.0)		FPz' 10.5 (4.4)		**FP4' ** 10.3 (3.8)**	
		F1' 10.9 (4.1)	Fz' 11.4 (3.0)	F2' **11.2 (3.6)**		
	FC3' 10.3 (3.8)	FC1' 11.9 (3.7)	FCz' 13.1 (3.3)	FC2' **12.5 (3.3)**	FC4' **10.8 (3.8)**	
TC3' 10.5 (3.2)	C3' 12.0 (3.0)	C1' 13.3 (3.9)	Cz' 15.3 (4.1)	C2' **14.0 (3.2)**	C4' **12.3 (2.9)**	TC4' 10.1 (3.8)
	C3" 12.1 (3.3)	C1" 13.6 (3.8)	Cz" 15.4 (4.3)	C2" 13.6 (3.7)	C4" 12.2 (3.4)	
		PC1" 12.3 (3.5)	PCz" 12.6 (4.1)	PC2" 11.9 (4.1)		

II-3.4. CONCLUSIONS

In general, the signal-to-noise ratio is reasonably high: dependent on the investigated component, the SNR is approximately 15dB, but reaches up to 20dB for individual measurements. The highest SNRs were mostly located at or near the central electrode positions.

The significant difference between the stimulus components of CNV_1 and CNV_4 is probably related to the effect of overlapping 'early' and 'late' wave within the 1-second ISI of the CNV_1 response. The early wave of CNV_4 is much lower in amplitude, whereas the same interval after the WS in CNV_1 contains more reproducible signal components. This results in a higher SNR. However, for right side response signals, this effect was not significant. The conclusions regarding lateralization of the response component complement some of the findings from the original study by Böker (1994):

1) We found a small paradoxical lateralization effect for SNR of the reponse component related to foot responses, which is slightly more pronounced for right responses in CNV_1.
2) In the current study a large significant lateralization effect for SNR values occurs in the interval 100ms prior to the RS, to RS+200ms: for CNV_1 left foot responses, and for CNV_4 left finger responses (contra-lateral dominance in the central, parietal and temporal regions).

The paradoxical lateralization effect relates to the original observations in the CNV late wave. The source activity for finger responses can be explained by a deep dipole below the vertex, and a symmetrical pair of dipoles at lateral positions. For foot responses source activity is significant only in the central region, and shows a tendency toward largest lateralization at medial positions for left foot responses, and lateral positions for right foot responses (Böker, 1994).

Most of the significant effects seem to be related to the 'left response' signals. Because all subjects were right-handed, probably the effort for initiating and physically executing the response was easier for the right-sided response trials. The data on reaction times versus response-side did not reveal a significant effect (Böker, 1994), indicating an insignificant physical effect. If we conclude from this that the tasks must have presented a more 'difficult' cognitive task for left responses, the significance found is probably mostly related to the motor preparations of the left responses. Of course, this effect would have to be investigated using a dedicated paradigm where the anticipatory and preparatory processes are manipulated.

The current method for SNR assessment is less suited for evaluation of the interval between warning stimulus and response stimulus in the CNVs. This slow negative slope has been investigated by comparing the average amplitude values (with respect to the baseline level) in the 1000ms interval prior to the RS, in intervals of 100ms (see Böker, 1994). Still, for convincing inferences about the amplitude changes in these intervals the validity of the waveforms should be assessed. The current study has produced this reassuring evidence: the SNR of the entire CNV_4 wave was approximately 11dB, despite the long low amplitude ISI, and SNR was approximately 15dB for the entire CNV_1 wave.

Overall, we come to the conclusion that SNR assessment using the (±)-reference method provides useful measures that corroborate the findings from other amplitude based analyses, and complement these results by quantifying the signal quality. For this reason, and to facilitate better comparison among experiments, we must strongly advocate the use of these objective, quantitative signal descriptors.

Discussion & Conclusions

The studies presented in this thesis focus on generally applicable methods for the validation of electroencephalographic measurements: the electroencephalogram itself (EEG), and evoked potential (EP) measurements. Validation is the process of scrutinising the signals for undesired disturbances. Especially when subsequent analyses are going to be performed automatically, reliable results are only obtained when the validation step is integrated in the overall processing scheme.

EEG VALIDATION

Context related 'event' detection

Most of the tools for (semi-automated) EEG analysis described in the literature are aimed at detection of specific EEG phenomena (e.g., Nakamura *et al.*, 1996; Pietilä *et al.*, 1994). No software tool specifically targets validation of the EEG; pruning of EEG data is often implied, or based on heuristics. An example of a (detailed) description of an artefact detection step, using empirical rules for amplitude and frequency parameters, is found in Jansen *et al.* (1982).

Because of the diversity of EEG investigations we chose to restrict our research to the detection of muscle artefact and to the detection of 'events' in prolonged EEG recordings in ICU patients. Muscle artefact presents one of the most frequent and obtrusive contaminations of the EEG, and the ICU study consisted of a relatively large data set incorporating diverse patterns of EEG. Therefore, the investigated methods are potentially useful in other EEG investigations.

Current approach: time-related aspects of EEG validation

In the study of muscle artefact detection (chapter I-2) we did not use a segmentation procedure to obtain stationary periods of EEG (e.g., as performed by Krajca *et al.*, 1991). Rather, we processed the EEG based on fixed epoch lengths of short duration, which may be considered stationary (Jansen, 1991). The expert's scoring was not restricted to

fixed epochs. In each epoch, we calculated several amplitude and frequency domain parameters. Per subject we used constant thresholds and statistical thresholds based on the parameter histograms of an artefact-free reference measurement. The statistical approach also required the use of fixed epoch lengths. Not studied before, was the relation between epoch length and detection performance. Based on the criteria of 'scattering' (short periods useless for subsequent analysis) and overall detection performance we found that the optimal epoch length is approximately 1 second, for detection of muscle artefact in the normal human awake EEG.

In the muscle artefact study, we observed that constant thresholds were better than statistical thresholds obtained from the reference measurements; performance was optimal using a Slope or 'high beta' parameter. Of course this approach was not suited for application in prolonged recordings containing different EEG patterns, e.g., sleep EEG. The Slope detection algorithm was therefore extended into a more adaptive version (chapter I-3), which successfully detected transient (high-frequency) 'events' like muscle artefacts as statistical outliers based on the context of the ongoing EEG signal. Context related evaluation of autoregressive (AR) parameters was useful for detection of high-amplitude, low frequency artefacts. Using both the AR error-term as well as the AR model parameters is not common, but highly contributed to the specificity of the AR-based detection process. The new approach of context incorporation also allowed us to investigate the optimal length of the context period: a relatively short period of EEG, limited to 20-40 seconds, was optimal and resulted in the detection of 90% of the artefacts in the ICU data. The accompanying positive prediction was not so high: 53%.

Recommendations for enhanced context related 'event' marking

The artefact detection process based on the shifting context-window as described in chapter I-3, can not be used for instance to mark sustained periods of muscle artefact or longer low-frequency artefacts. This drawback can be overcome by a more advanced implementation as illustrated in Figure 1. The context model is now extended to a sequence of 'context', 'test', and 'prescan' windows, which relates to the way clinicians browse back and forth pages of an EEG recording. In each of these windows the Slope distributions and AR parameters are calculated as described in chapter I-3. Statistical differences between the prescan-window and the test-window can be tested analogous to the procedures as described for context-window and test-window, but are used to detect the *end* of a significant change.

Additional improvements to the detection algorithm can be made. For instance, shifting the detection windows forward in time per epoch, instead of per 'page' (e.g., 10 seconds) will result in higher specificity, because artefact marking can be performed more accurately. Small epoch lengths (e.g., 0.5 s) can be used to make the method more sensitive to very short phenomena. Also, different window lengths and different thresholds (not necessarily based on the *F*-distribution) can be used. Using an example

threshold of $\lambda_2^2/\lambda_1^2 > 2$, and further settings as illustrated in Figure 1, resulted in the detection of the artefacts as depicted in Figure 2. The settings were not specifically tuned for this particular artefact, but we must admit that it is one of the 'typical best' examples.

These improvements have to be implemented carefully, since the new settings can result in a highly specific detection process, thereby missing a lot of artefacts. One additional enhancement that should be investigated however, is the use of *adaptive* context and test windows. When a significant difference is detected in the test window, the context window should not shift forwards in time, until the ongoing signal in the test window is within a 'valid' range, based on λ_2^2/λ_1^2. This approach also enables accurate marking of artefacts, without using a 'prescan' window.

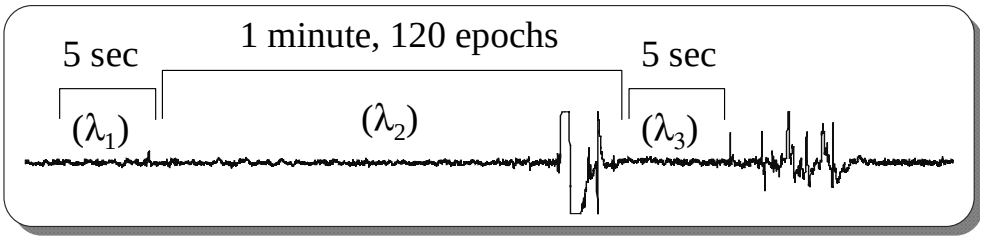

Figure 1 Alternative implementation of context related EEG validation, tuned for high specificity, using 'context' and 'prescan' windows of 5 seconds, and a 'test' window of 1 minute. In the current example, epoch length is 0.5 seconds.

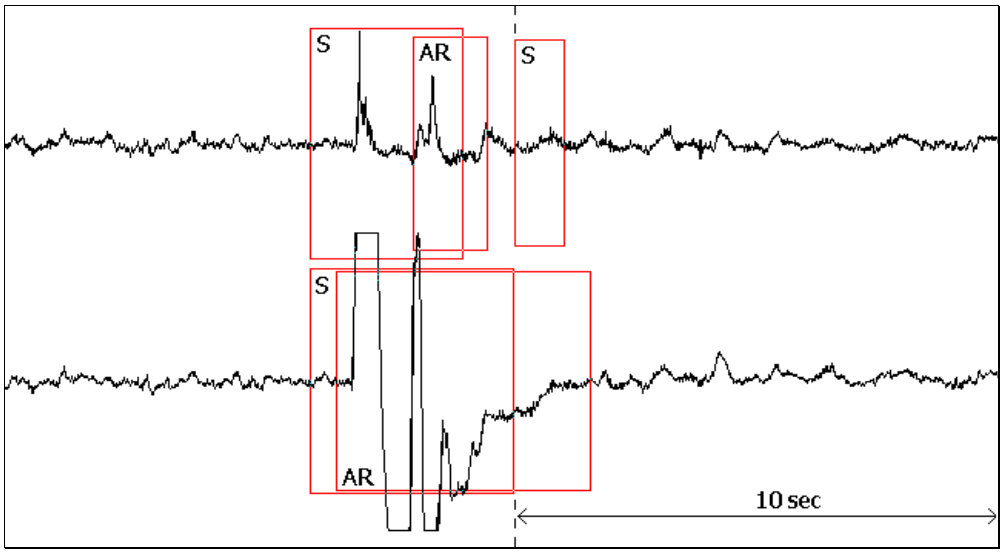

Figure 2 Examples of artefact detection using the alternative context detection. Indicated are the markings by the autoregressive model (AR) and Slope (S) detection.

EEG validation: integration in the evaluation process

The context related detection procedure as presented in this thesis is not aimed at detection of specific clinical EEG phenomena. The algorithms merely indicate the statistical differences between periods or epochs of EEG data. As already indicated by the relatively low positive prediction of our method (chapter I-3), we have observed that such statistical 'changes' in the EEG are often not artefactual. However, most of these non-artefacts still presented clinically relevant EEG 'events', and were characterised by a relatively low variability score (based on AR model). All high variability scores (e.g., $\lambda_2^2/\lambda_1^2 > 20$) indicated true artefacts (marked as such by both observers).

Therefore, the EEG validation procedure should be combined with a feature detection program, which can process the detected 'events' of lower variability. A general EEG processing scheme of measurement, validation, analysis and clinical review is depicted in Figure 3. The feature detection module can of course consist of several specialised programs; this advanced (automated) analysis step focuses on specific EEG phenomena (e.g., arousals, spindles, and spikes). This design will result in a higher true positive rate for detected features from the valid periods than processing of the raw EEG signal. Moreover, 'events' can be dealt with separately, possibly using more robust settings or special pre-filtering in the feature detector. When using such a combination of methods, an optional artefact classification module can also benefit from such an approach (see Figure 3). Thus conceived, the system allows for optimal tuning of both the validation process as well as the advanced analysis process.

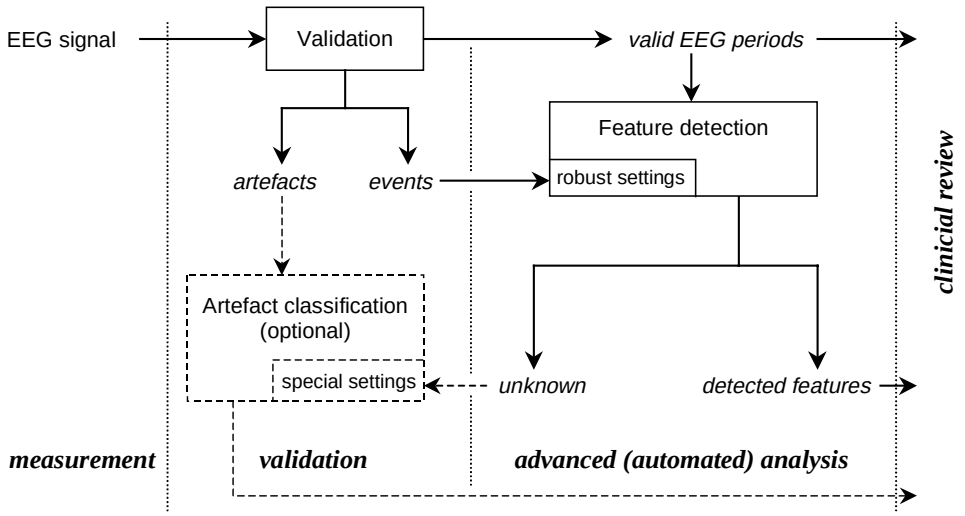

Figure 3 Schematic diagram of an EEG evaluation process, incorporating validation and feature detection. Dashed lines indicate optional processing steps.

Objective EEG validation

As observed in many EEG studies — the studies in this thesis are no exception — differences between human observers are unavoidable in visual EEG evaluations (e.g., Wilson *et al.*, 1996, 1999). However, we can try to model the human process of validating EEG signals by detailed analysis of the correlation between 1) each individual observer's scoring set, and 2) all parameters used in the automatic evaluation. This perusal will result in a different set of parameter rules for each observer. Next, the consensus of scores from several observers can be used to split the rules into generic, and observer-specific rules.

Instead of calculating the consensus set from the different scores, we can also have the observers meet in an consensus session, in order to agree on some of their differences in a joint effort (Pietilä *et al.*, 1994). Ideally, different clinical backgrounds among experts should be avoided (Williams *et al.*, 1985). From the 'joint effort consensus scores' we can gain insight in several processes. The difference between the initial scoring of an expert and his/her adjusted scores in the consensus session relates to the level of individual 'uncertainty' of scores, much like the intra-expert 'reproducibility' comparison as performed in chapter I-2. Comparison of the joint effort scores versus the calculated consensus set provides an indication for the acceptable difference between consensus scores and computer scores.

When applying these principles we may be able to tailor the computerised procedure to the need of a particular clinical expert or laboratory, using a custom selection of observer-specific rules, leaving the generic rules intact. In general, this is a proper procedure when the validation tool is only used as an advisory system for the neurophysiological expert. Requirements of sensitivity and specificity will be different for (semi-) automated analysis.

An early advisory/classification system was described by Bodenstein *et al.* (1985). Their program provided a summary of a prolonged EEG recording, where different types of EEG were clustered into groups of 'alike' periods based on a spectral distance criterion. However, a major drawback of this approach is the rigid reduction of a prolonged recording to only a few raw traces of EEG (representative samples of 5 seconds for each 'type' of EEG) on the final one-page 'summary sheet'. A similar classification method was described recently by Agarwal *et al.* (1998), who improved, and slightly enlarged the summary sheet to two pages, displaying more raw EEG traces. Still, immediate access to all original EEG traces by 'clicking' on a salient detail in the compressed display should be a standard feature, even when it requires a more powerful computer or increased network communication.

The absence of a universal *analytical* model of the EEG hampers the scientific research on EEG validation and feature detection. Therefore, human observers are involved in the evaluations, and subsequently all investigations will be afflicted with subjective

interpretation of the EEG signal. The use of empirical rules and heuristics based on the stochastic EEG process is inevitable, and the factor of human error must be taken into account. Although explicitly marked artefacts will probably indicate true artefacts, we cannot rule out the possibility of artefacts being *missed*, even by the expert's eye. The amount of missed artefacts may even be significant, especially in a consensus set, resulting in an exaggerated false detection rate (i.e., a low positive prediction) of any detection method that is compared to the human scores.

EVOKED POTENTIALS

Quality of EP measurements

Quantitative EP assessment

A major difference between EEG and EP measurements is the fact that the EEG is a stochastic signal, whereas EPs are discrete signals, having known properties. Although some types of evoked signals can change during the recording (e.g., long latency EPs, evoked K-complexes during sleep), we have good reasons to believe that waveform shape changes relatively little during the recording of most *sensory* EPs. These EPs can be acquired in relatively short measurements: between 1 minute (middle latency EPs) and 10 minutes (long latency EPs).

Based on these observations, two existing methods for objective signal-to-noise ratio (SNR) estimation were compared in a large data set of middle latency auditory EPs (MLAEPs, chapter II-2). Because of the difficult recording conditions, EPs of different quality were present in the data set, which allowed for a thorough investigation of SNR related to visual, *clinical* judgement of EP quality. We found that the (±)-reference method, originally designed by Schimmel (1967), resulted in the best classification of clinically 'acceptable' and 'unacceptable' waveforms.

The EP study in chapter II-2, as well as the subsequent explorative study in long-latency CNV signals (Contingent Negative Variation, see chapter II-3) indicates that calculation of the SNR can be performed reliably in the post-stimulus interval around relatively stable EP components. Since the MLAEPs were obtained during anaesthesia, the middle latency components were relatively variable. It appeared that the brainstem interval of 5-15ms is the best predictor for signal quality of the entire MLAEP waveform, when using the (±)-reference method.

The quality of the MLAEP signal can also be assessed quantitatively by estimating the noise contribution as the differential signal between successive intermediate averages (chapter II-4). This new 'convergence' factor *Qc* and the (±)-reference *P* factor were used

to find the optimal stimulation rate in the high-frequency random stimulation method (*h1*) developed by Cluitmans (1990). The results of this study indicated that the acquisition time required for MLAEP waveforms of acceptable quality can be reduced to one minute or shorter, using a mean click rate of 80/second or 90/second.

Artefact detection

Artefact detection using amplitude parameters has been very successful for the improvement of evoked potentials (EPs) in clinically difficult recording conditions (see chapter II-1 p. 85, chapter II-2 p. 99, and in the original study of De Beer *et al.*, 1995). This semi-automatic rejection method uses statistical thresholds based on just a few artefact-free EP measurements (visual assessment) from the investigated set of EPs, all recorded under the same conditions.

In the MLAEP study described in chapter II-4, high-pass filtering of the EEG was used to reduce the adverse influence of slow-wave sleep. This pre-processing did not filter all artefacts. However, detailed analysis of the SNR quality factors revealed that the evolution curves of *P* (±-reference), and *Qc* (convergence) of individual measurements should be analysed to detect loss of the EP response caused by artefacts (or other adverse effects). In an on-line implementation, the expected logarithmic increase of the SNR curve (power) can be tracked in both *P* and *Qc* quality factors based on both brainstem and middle latency intervals; 'no-response' sweeps should be recognised easily. In this respect, the applied quality factors *P* and *Qc* are complementary.

Application in clinical monitoring

Indication of the SNR is still not a standard feature on most commercial EP averaging equipment. The studies in Part II of this thesis have shown that several methods are readily available; incorporation of an objective indication of SNR improves the comparison between studies, and clinical applications will benefit from an increased awareness about the quality of measurements. A signal-to-noise estimate can also be used to decide when the evoked potential average has sufficient quality. Currently, the recording length of EP measurements is based on empirical rules, e.g., 4000 sweeps are used in a BAEP recording.

In a clinical application, a frequent status update of the investigated EP is desired. This can for instance be implemented by constant updating of the average: adding new responses, and deleting the 'oldest' sweeps. A practical and reliable implementation of such a 'moving average' EP can be achieved by using an SNR evaluation module, which controls the adding and deleting of sweeps based on a SNR threshold.

Monitoring of MLAEPs can be improved through the use of high-frequency random stimulation, according to a Poisson distribution. As stated in the discussion in chapter II-2, the observed average recording time for clinically acceptable MLAEPs was 100

seconds; however, this was biased towards higher *N* (number of sweeps), because of the procedures used. In chapter II-4 the recording time for adequate MLAEPs was investigated in detail: the results show that 60 seconds is feasible. When combining the *h1* random stimulation method with other detection techniques, e.g., wavelets (Geva, 1998), ARX (Jensen *et al.*, 1996), or neural networks (Fung *et al.,* 1999), the recording time may be reduced even more.

OVERALL CONCLUSIONS

Signal validation of the EEG

The application of relatively simple statistical principles was very successful for the detection of signal disturbances in the investigated EEG data sets.

In a study for the detection of muscle artefact in the EEG, the use of short epochs of EEG data (0.5 to 2 seconds) yielded the best results. The analysis using a Slope threshold (first derivative) or absolute 'high beta' power showed the best results, matching the expert's performance.

A model for EEG signal context, consisting of autoregressive parameters and Slope parameters in successive epochs, was very successful in the detection of non-stationarities and artefacts in a large dataset as measured in critically ill patients during their stay in the intensive care unit. This EEG study pointed out that objective computerised validation may be applied generally when the signal context is taken into account in the model(s) or parameters used. An adequate signal period providing enough context information can be limited to 20 to 40 seconds.

Evoked potential validation

Validation through the use of amplitude parameters is very successful for quality improvement of EPs. Quantification of the quality of EP measurements can be implemented by using the (±)-reference method, or by the 'convergence' method.

The high-frequency random stimulation method, using a mean click rate of 80/second or 90/second, yields higher quality MLAEP signals, which can be acquired faster than conventional, regular stimulation methods.

EEG validation: future directions

For a successful incorporation of automated validation algorithms, clinical experts must be open to computerised detection of artefacts. The use of objective, statistical

evaluation methods can help to raise general acceptance. Although computer methods are not perfect, the results are highly reproducible.

Validation cannot be separated from the (clinical) application area: an additional *feature recognition* module may be used for instance for detection of a particular EEG phenomenon, using pre-processed EEG obtained through a tailored *validation* step. Both procedures can benefit from this integration: the overlap between artefact ('event') detection and feature detection can be evaluated in several steps. This approach is illustrated in Figure 3. The positive prediction and specificity of this integrated validation/detection process will be higher than in the individual methods.

A better understanding of the way (groups of) clinicians perform EEG evaluation is essential for the development of automated digital EEG analysis. This is probably more important than technological progress, e.g., improved amplifiers, electrodes, or 'ideal' (50Hz) filters cannot prevent subjective interpretation. In conclusion, validation of EEGs must be taken seriously; the confirmation that authentic EEG data is used, is a prerequisite for reliable incorporation of any automated digital EEG analysis in the future.

Notwithstanding the fact that advanced (automated) analyses and clinical review focus on specific EEG phenomena, one must not forget that artefacts can also reveal useful clinical information (Barlow, 1986b). For example: muscle activity or eye movements in the EEG of an ICU patient can actually indicate a status improvement. Therefore: 1) EEG signals should *always* be available in the original format, preferably accompanied by detailed annotations, and 2) end-conclusions about physiological changes should be made by an expert clinician.

REFERENCES

Agarwal, R., Gotman, J., Flanagan, D. and Rosenblatt, B. *Automatic EEG analysis during long-term monitoring in the ICU.* Electroenceph. Clin. Neurophysiol., 1998, 107: 44-58.

Akaike, H. *A new look at the statistical model identification.* IEEE Trans. Autom. Control, 1974, 19, 6: 716-723.

Akrawi, W.P., Drummond, J.C. Kalkman, C.J. and Patel, P.M. *A comparison of the electrophysiologic characteristics of EEG burst-suppression as produced by isoflurane, thiopental, etomidate, and propofol.* J. Neurosurg. Anesthesiol., 1996, 8: 40-46.

Alarcon, G., Binnie, C.D., Elwes, R.D.C. and Polkey, C.E. *Power spectrum and intracranial EEG patterns at seizure onset in partial epilepsy.* Electroenceph. Clin. Neurophysiol., 1995, 94, 5: 326-337.

American Electroencephalographic Society. *Guidelines for Clinical Evoked Potential Studies.* J. Clin. Neurophysiol., 1984, 1: 3-53.

American Electroencephalographic Society. *Guidelines in EEG.* J. Clin. Neurophysiol., 1986, 3: 131-168.

American Electroencephalographic Society. (1994a) *Guideline eleven: guidelines for intraoperative monitoring of sensory evoked potentials.* J. Clin. Neurophysiol., 1994, 11: 77-87.

American Electroencephalographic Society. (1994b) *Guideline thirteen: Guidelines for standard electrode position nomenclature.* J. Clin. Neurophysiol., 1994, 11: 111-113.

Anderson, C.W., Stolz, E.A. and Shamsunder, S. *Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks.* IEEE Trans. Biomed. Eng., 1998, 45, 3: 277-286.

Antoniol, G. and Tonella, P. *EEG data compression techniques.* IEEE Trans. Biomed. Eng., 1997, 44, 2: 105-114.

Arvidsson, A., Friberg, S. and Matoušek, M. *Computer controlled classification of EEG activity.* Electroenceph. Clin. Neurophysiol., 1977, 43: 443.

ASTM. Specification for Transferring Digital Neurophysiological Data Between Independent Computer Systems (ASTM E1467-94). American Society for Testing and Materials, Washington, 1994.

Aufrichtig, R. and Pedersen, S.B. *Order estimation and model verification in autoregressive modeling of sleep recordings.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1992. Paper number: 2653.

Aunon, J.I. *Digital signal processing of evoked potentials.* IEEE frontiers of Engineering and Computing in Health Care, 1983; 656-657.

Barlow, J.S., Kamp, A., Morton, H.B., Repoche, A., Shipton, H. and Tchavdarov, D.B. *EEG instrumentation standards (revised 1977).* Report of the Committee on EEG Instrumentation Standards, International Federation of Societies for Electroencephalography and Clinical Neurophysiology. Electroenceph. Clin. Neurophysiol., 1978, 45: 144-150.

Barlow, J.S. *Computerized clinical electroencephalography in perspective.* IEEE Trans. Biomed. Eng., 1979, 26, 7: 377-391.

Barlow, J.S. *Muscle spike artifact minimization in EEGs by time-domain filtering.* Electroenceph. Clin. Neurophysiol., 1983, 55: 487-491.

Barlow, J.S. *A general-purpose automatic multichannel electronic switch for EEG artifact elimination.* Electroenceph. Clin. Neurophysiol., 1985, 60: 174-176.

Barlow, J.S. (1986a) *Automatic elimination of electrode-pop artifacts in EEG's.* IEEE Trans. Biomed. Eng., 1986, 33, 5: 517-521.

Barlow, J.S. (1986b) *Artifact processing (rejection and minimization) in EEG data processing.* In: Handbook of Electroencephalography and Clinical Neurophysiology. Revised edition. Vol. 3B: Applications of Analytical Techniques. F.H. Lopes da Silva and W.H. Storm van Leeuwen (eds). Amsterdam: Elsevier, 1986; 15-62.

Basar, E., Rosen, B., Basar-Eroglu, C. and Greitschus, F. *The associations between 40 Hz-EEG and the middle latency response of the auditory evoked potential.* Int. J. Neuroscience, 1987, 33: 103-117.

Bastiaansen, M.C.M., Böker, K.B.E., Cluitmans, P.J.M. and Brunia, C.H.M. *Event-related desynchronization related to the anticipation of a stimulus providing knowledge of results.* Clinical Neurophysiology, 1999, 110: 250-260.

Bender, R., Schultz, B., Schultz, A. and Pichlmayr, I. *Identification of EEG patterns occuring in anesthesia by means of autoregressive parameters.* Biomed. Technik, 1991, 36, 10: 236-240.

Bender, R., Schultz, B., Schultz, A. and Pichlmayr, I. *Testing the gaussianity of the human EEG during anaesthesia.* Meth. Inf. Med., 1992, 31: 56-59.

Benlamri, R., Batouche, M., Rami, S. and Bouanaka, C. *An automated-system for analysis and interpretation of epileptiform activity in the EEG.* Comp. Biol. Med., 1997, 27, 2: 129-139.

Berg, P. and Scherg, M. *A multiple source approach to the correction of eye artifacts.* Electroenceph. Clin. Neurophysiol., 1994, 90, 3: 229-241.

Berger, H. *Über das Elektroenkephalogramm des Menschen* (About the human electroencephalogram. In German). Arch. Psychiatr. Nervenkr., 1929, 87: 527.

Bickford, R.G. *Automatic electroencephalographic control of anaesthesia.* Electroenceph. Clin. Neurophysiol., 1950, 2: 93-96.

Bickford, R.G., Billinger, T.W., Flemming, N.I. and Stewart L. *The compressed spectral array (CSA): a pictorial EEG.* Proc. San Diego Biomedical Symposium, 1972, 11: 365-370.

Blinowska, K.J. and Malinowski, M. *Non-linear forecasting of the EEG time series.* Biol. Cybern., 1991, 66: 159-165.

Bloch, K.E. *Polysomnography: a systematic review.* Technology and Health Care, 1997, 5: 285-305.

Böker, K. *Spatiotemporal dipole models of slow cortical potentials.* Ph.D. Thesis. Tilburg: Tilburg University, 1994.

Bodenstein, G. and Praetorius, H.M. *Feature extraction from the electroencephalogram by adaptive segmentation.* Proc. IEEE, 1977, 65, 5: 642-652.

Bodenstein, G., Schneider, W. and Malsburg, C.V.D. *Computerized EEG pattern classification by adaptive segmentation and probability-density-function classification. Description of the method.* Comput. Biol. Med., 1985, 15, 5: 297-313.

Boston, J.R. *Spectra of auditory brainstem responses and spontaneous EEG.* IEEE Trans. Biomed. Eng., 1981, 28, 4: 334-341.

Boston, J.R. *Automated interpretation of brainstem auditory evoked potentials: a prototype system.* IEEE Trans. Biomed. Eng., 1989, 36, 5: 528-532.

Boudreaux-Bartels, G.F. and Murray, R. *Time-frequency signal representations for biomedical signals.* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 866-885. ISBN 0-8493-8346-3.

Box, G.E.P. and Jenkins, G.M. *Time series analysis, forecasting and control.* Revised edition. London: Holden-Day, 1976. ISBN 0-8162-1104-3.

Bracewell, R.N. *The Hartley transform.* New York: Oxford University Press, 1986.

Braun, J.C., Hanley, D.F. and Thakor, N.V. *Detection of neurological injury using time-frequency analysis of the somatosensory-evoked potential.* Evoked Potentials - Electroenceph. Clin. Neurophysiol., 1996, 100, 4: 310-318.

Brittenham, D. *Artifacts. Activities not arising from the brain.* In: Current Practice of Clinical Electroencephalography. 2nd edition. D.D. Daly and T.A. Pedley (eds). New York: Raven Press, 1990; 85-105.

Bronzino, J.D., Forbes, W. and Morgane, P.J. *Quantitative indices of the EEG amplitude histograph.* IEEE Front. Eng. Health Care, 1980: 186-189.

Bronzino, J.D. *Principles of electroencephalography*. In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 201-212. ISBN 0-8493-8346-3.

Bruce, A., Donoho, D. and Gao, H. *Wavelet Analysis.* IEEE Spectrum, 1996, 33, 10: 26-35.

Brunia, C.H.M. and Damen, E.J.P. *Distribution of slow potentials related to motor preparation and stimulus anticipation in a time estimation task.* Electroenceph. Clin. Neurophysiol., 1988, 69: 234-243.

Brunia, C.H.M., Möks, J., Van den Berg-Lenssen, M.M.C. and Coelho, M. *Correcting ocular artifacts in the EEG, a comparison of several methods.* J. Psychophysiol., 1989, 3: 1-50.

Brunia, C.H.M. *Stimulus preceding negativity: arguments in favour of non motoric slow waves.* In: Slow Potential changes in the Human Brain. W.C. McCallum and S.H. Curry (eds). New York: Plenum Press, 1993: 147-161.

Brunner, D.P., Vasko, R.C., Detka, C.S., Monahan, J.P., Reynolds III, C.F. and Kupfer, D.J. *Muscle artifacts in the sleep EEG: automated detection and effect on all-night EEG power spectra.* J. Sleep Res., 1996, 5: 155-164.

Burgess, R.C. *Digital electroencephalographs.* J. Clin. Neurophysiol., 1993, 10, 3: 378-392.

Carson, E. and Saranummi, N. *Improving control of patient status in critical care: the IMPROVE project.* Comp. Meth. Progr. Biomed., 1996, 51: 1-130.

CEN (TC251/WG5/PT5-021). Vital Signs Information Representation, Interim Report, revision 1.2, Annex B, Evaluation of present biosignal archive interchange formats. European Committee for Standardisation, Brussels, 1995.

CEN (TC251/WG5/PT5-021). Vital Signs Information Representation, Annex A (Normative). The Medical Data Information Base (MDIB). Nomenclature, Data Dictionary and Codes. Version 1.78 (PT5-021 Pre-Final Version). October 1997.

Cerutti, S., Liberati, D. and Mascellani, P. *Parameter extraction in EEG processing during riskful neurosurgical operations.* Signal Proc., 1985, 9: 25-35.

Cerutti, S., Liberati, D., Avanzini, G., Franceschetti, S. and Panzica, F. *Classification of the EEG during neurosurgery. Parametric identification and Kalman filtering compared.* J. Biomed. Eng., 1986, 8: 244-254.

Cerutti, S., Gaselli, G., Liberati, D. and Pavesi, G. *Single sweep analysis of visual evoked potentials through a model of parametric identification.* Biol. Cybern., 1987, 56: 111-120.

Cerutti, S., Chiarenza, G., Liberati, D. Mascellani, P. and Pavesi, G. *A parametric method of identification of single-trial event-related potentials in the brain.* IEEE Trans. Biomed. Eng., 1988, 35, 9: 701-711.

Cerutti, S., Carrault, G., Cluitmans, P.J.M., Kinie, A., Lipping, T., Nikolaidis, N., Pitas, I., Signorini, M.G. *Non-linear algorithms for processing biological signals.* Comp. Meth. Prog. Biomed., 1996, 51: 51-73.

Challis, R.E. and Kitney, R.I. *Biomedical signal processing (in four parts), Part 1, Time-domain methods.* Med. Biol. Eng. Comput., 1990, 28: 509-524.

Chan, F.H.Y., Lam, F.K., Poon, P.W.F. and Qiu, W. *Detection of brain-stem auditory-evoked potential by adaptive filtering.* Med. Biol. Eng. Comput., 1995, 33, 1: 69-75.

Cho, Y.S., Kim, S.B., Hixson, E.L. and Powers, E.J. *A digital technique to estimate second-order distortion using higher order coherence spectra.* IEEE Trans. Signal Processing, 1992, 40: 1029-1040.

Clark, I., Biscay, R., Echeverría, M. and Virués, T. *Multiresolution decomposition of non-stationary EEG signals: a preliminary study.* Comput. Biol. Med., 1995, 25, 4: 373-382.

Cluitmans, P.J.M. *Neurophysiological monitoring of anesthetic depth.* Ph.D. Thesis, Eindhoven: Eindhoven University of Technology, 1990. ISBN 90-9003453-6.

Cluitmans, P.J.M. and Beneken, J.E.W. *Non-linear analyis of sensory evoked potentials.* Proc. 13th IEEE Eng. Med. Biol. Soc. Conf., 1991, 1: 419-420.

Cluitmans, P.J.M., Jansen, J.W. and Beneken, J.E.W. *Artefact detection and removal during auditory evoked potential monitoring.* J. Clin. Mon., 1993, 9, 2:112-120.

Cohen, A. *Biomedical signals: origin and dynamic characteristics; frequency domain analysis.* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 805-827. ISBN 0-8493-8346-3.

Cohen, J. and Polich, J. *On the number of trials needed for P300.* Int. J. Psychophysiol., 1997, 25, 3: 249-255.

Collura, T.F., Jacobs, E.C., Braun, D.S. and Burgess, R.C. *EView – a workstation-based viewer for intensive clinical electroencephalography.* IEEE Trans. Biomed. Eng., 1993, 40: 736-744.

Cooley, J.W. and Tukey, J.W. *An algorithm for the machine calculation of complex Fourier series.* Math. Comput., 1965, 19: 297-301.

Coppola, R. Tabor, R. and Buchsbaum, M.S. *Signal to noise ratio and response variability measurements in single trial evoked potentials.* Electroenceph. Clin. Neurophysiol., 1978, 44: 214-222.

Croft, R.J. and Barry, R.J. (1998a) *EOG correction: a new perspective.* Electroenceph. Clin. Neurophysiol., 1998, 107: 387-394.

Croft, R.J. and Barry, R.J. (1998b) *EOG correction: a new aligned-artifact average solution.* Electroenceph. Clin. Neurophysiol., 1998, 107: 395-401.

Cunha, M.B., Cunha J.P. and Oliveira e Silva, T. *SIGIF: A digital signal interchange format with application in neurophysiology.* IEEE Trans. Biomed. Eng., 1997, 44, 5: 413-418.

Dalle Molle, J.W. and Hinich, M.J. *Trispectral analysis of stationary random time series.* J. Acoust. Soc. Am., 1995, 97, 5: 2963-2978.

Davila, C.E. and Mobin, M.S. *Weighted averaging of evoked potentials.* IEEE Trans. Biomed. Eng., 1992, 39, 4: 338-345.

De Beer, N.A.M. Cluitmans, P.J.M. and Grundy, B.L. *The use of the median frequency and spectral edge frequency to estimate depth of anesthesia.* Proc. 10th World Congress of Anaesthesiologists, The Hague, 1992; A51.

De Beer, N.A.M., Van de Velde, M. and Cluitmans, P.J.M. *Clinical evaluation of a method for automatic detection and removal of artefacts in auditory evoked potential monitoring.* J. Clin. Mon., 1995, 11: 381-391.

De Beer, N.A.M. *Monitoring adequacy of anesthesia using spontaneous and evoked electroencephalographic activity.* Ph.D. Thesis, Eindhoven: Eindhoven University of Technology, 1996. ISBN 90-386-0317-7.

De Beer, N.A.M., Van Hooff, J.C., Cluitmans, P.J.M., Korsten, H.H.M. and Grouls, R.J.E. *Hemodynamic responses to incision and sternotomy in relation to the auditory evoked potential and spontaneous EEG.* Br. J. Anaesth., 1996, 76, 5: 685-693.

Delgado, R.E. *Automated identification and interpretation of auditory brainstem responses.* Ph.D. Thesis. Miami: University of Miami, 1993. Univ. Microfilms N° ADG93-31516.

Delgado, R.E. and Özdamar, Ö *Automated auditory brainstem response interpretation.* IEEE Eng. Med. Biol., 1994, 13, 2: 227-237.

Deltenre, P. and Mansbach, A.L. *A new descriptor of the dual character of the input-output behavior of the cochlea, with implications for signal-to-noise ratio estimation of brain-stem auditory potentials evoked by alternating polarity clicks.* Electroenceph. Clin. Neurophysiol., 1993, 88, 5: 377-388.

Dingle, A.A., Jones, R.D., Carroll, G.J. and Fright, W.R. *A multistage system to detect epileptiform activity in the EEG.* IEEE Trans. Biomed. Eng., 1993, 40: 1260-1268.

Doman, J., Detka, C., Hoffman, T., Kesicki, D., Monahan, J.P., Buysse, D.J., Reynolds III, C.F., Coble, P.A., Matzzie, J. and Kupfer, D.J. *Automating the sleep laboratory: implementation and validation of digital recording and analysis.* Int. J. Biomed. Comput., 1995, 38: 277-290.

Don, M., Elberling, C. and Waring, M. *Objective detection of averaged auditory brainstem responses.* Scand. Audiol., 1984, 13: 219-228.

Dorffner, G. *Toward a new standard of modeling sleep based on polysomnograms - the SIESTA project.* Proc. 4th European Conf. Eng. Med. (ESEM 1997).

Duffy, F.H., Hughes, J.R., Miranda, F., Bernad, P. and Cook, P. *Status of quantitative EEG (QEEG) in clinical practice, 1994.* Clin. Electroenceph., 1994, 25, 4: R6-R22.

Dunstan, F.D.J. and Marshall, R.W. *The detection of artefacts in EEG series.* Stat. Med., 1991, 10: 1719-1731.

Elberling, C. and Don, M. *Quality estimation of averaged auditory brainstem responses.* Scand. Audiol., 1984, 13: 187-197.

Elkfafi, M., Shieh, J.S., Linkens, D.A. and Peacock, J.E. *Intelligent signal-processing of evoked-potentials for anesthesia monitoring and control.* IEE Proc. Control Theory Appl., 1997, 144, 4: 354-360.

Elul, R. *Gaussian behavior of the electroencephalogram changes during performance of mental task.* Science, 1969, 164: 328-331.

Emerson, R.G. and Sgro, J.S. *Phase synchronized triggering: a method for coherent noise elimination in evoked potential recording.* Electroenceph. Clin. Neurophysiol., 1985, 61: 17P.

Erwin, R.J. and Buchwald, J.S. *Midlatency auditory evoked responses: differential effects of sleep in the human.* Electroenceph. clin. Neuroph., 1986, 65: 383-392.

Fang, P., Finette, S., Jasaitis, D.K. and Klein, S.L. *Stationarity and Gaussianity analysis of EEG under different sampling rates and signal lengths.* Proc. IEEE 9th Ann. Conf. Eng. Med. Biol. Soc., 1987: 1256-1257.

Fell, J., Röschke, J. and Schäffner, C. *Surrogate data analysis of sleep electro-encephalograms reveals evidence for nonlinearity.* Biol. Cybern., 1996, 75: 85-92.

Fenwick, P.B.C., Michie, P., Dollimore, J. and Fenton, G.W. *Mathematical simulation of the electroencephalogram.* Biomed. Comput., 1971, 2: 281-307.

Ferdjallay, M. and Barr, R.E. *Adaptive digital notch filter design on the unit circle for the removal of powerline noise from biomedical signals.* IEEE Trans. Biomed. Eng., 1994, 41: 529-536.

Fisch, B.J. *Spehlmann's EEG Primer.* 2nd revised and enlarged edition. Amsterdam: Elsevier, 1998. ISBN 0-444-81420-5.

Fischbach, G.D. *Mind and brain.* Scientific American, 1992, 267, 3: 24-33.

Fiz, J.A., Abad, J., Jane, R., Riera, M., Mananas, M.A., Caminal, P., Rodenstein, D. and Morera, J. *Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep-apnea.* European Respiratory J., 1996, 9, 11: 2365-2370.

Fleming, R.A. and Smith, N.T. *Density modulation: a technique for the display of three-variable data in patient monitoring.* Anesthesiology, 1979, 50: 543-546.

Flooh, E., Köner, E., Ladurner, G. and Lechner, H. *EEG-Nachtschlafableitungen: Auswertung mittels automatischer Datenanalyse* (EEG-night-sleep-recordings: automatic analysis. In German). Z. EEG-EMG, 1982, 13: 157-160.

Fung, K.S.M., Chan, F.H.Y., Lam. F.K. and Poon, P.W.F. *A tracing evoked potential estimator.* Med. Biol. Eng. Comput., 1999, 37: 218-227.

Gallez, D. and Babloyantz, A. *Predictability of human EEG: a dynamical approach.* Biol. Cybern., 1991, 64: 381-391.

Gasser, T. *Goodness-of-fit tests for correlated data.* Biometrika, 1975, 62: 563-570.

Gasser, T., Bäher, P. and Möks, J. *Transformations towards the normal distribution of broad band spectral parameters of the EEG.* Electroenceph. Clin. Neurophysiol., 1982, 53: 119-124.

Gasser, T., Bäher, P. and Steinberg. H. *Test-retest reliability of spectral parameters of the EEG.* Electroenceph. Clin. Neurophysiol., 1985, 60: 312-319.

Geddes, L.A. *Electroencephalography (Historical Perspectives 5).* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 2774-2783. ISBN 0-8493-8346-3.

Gersch, W. *Spectral analysis of EEG's by autoregressive decomposition of time series.* Math. Biosciences, 1970, 7: 191-204.

Geva, A.B., Pratt, H. and Zeevi, Y.Y. *Multichannel wavelet-type decomposition of evoked potentials: model-based recognition of generator activity.* Med. Biol. Eng. Comput., 1997, 35: 40-46.

Gevins, A.S., Yeager, C.L., Diamond, S.L., Spire, J.P., Zeitlin, G.M. and Gevins, A.H. *Automated analysis of the electrical activity of the human brain (EEG). A progress report*. IEEE Proc., 1975, 63: 1382-1399.

Gevins, A.S., Yeager, C.L., Zeitlin, G.M., Ancoli, S. and Dedon, M.F. *On-line computer rejection of EEG artifact.* Electroenceph. Clin. Neurophysiol., 1977, 42: 267-274.

Ghosh, I.R., Prior, P.F., White, S.R., Gade, J., Jensen, K., Langford, R.M., Rosenfalck, A. and Thomsen, C.E. *Artefact assessment in prolonged EEG-polygraphic recordings in intensive care.* Electroenceph. Clin. Neurophysiol. 1997, (abstract in press).

Glover, J.R., Ktonas, P.Y. Raghavan, N. Uruñela, J.M., Velamuri, S.S. and Reilly, E.L. *A multichannel signal processor for the detection of epileptogenic sharp transients in the EEG.* IEEE Trans. Biomed. Eng., 1986, 33, 12: 1121-1128.

Glover, J.R., Raghavan,N., Ktonas, P.Y. and Frost, J.D. *Context-based automated detection of epileptogenic sharp transients in the EEG: elimination of false positives.* IEEE Trans. Biomed. Eng., 1989, 36, 5: 519-527.

Goel, V., Brambrink, A.M., Baykal, A. and Thakor, N.V. *Auto-regressive analysis of EEG reveals brain's response to injury.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1994. Paper number: 239.

Goldman, D. *The use of a new type 'indifferent' electrode.* Electroenceph. Clin. Neurophysiol., 1949, 1: 523.

Gordon, E., Haig, A., Rogers, G., Rennie, C., Anderson, J., Barry, R., Hook, S. and Meares, R. *Beyond averaging - classification of single-trial ERP subtypes using globally optimal vector quantization (abstract).* Int. J. Psychophysiol., 1994, 18, 2: 107-108.

Gorney, D.S. *The practical guide to digital EEG.* Am. J. EEG Technol., 1992, 32: 260-289.

Gotman, J., Ives, J.R. and Gloor, P. *Frequency content of EEG and EMG at seizure onset: possibility of removal of EMG artefact by digital filtering.* Electroenceph. Clin. Neurophysiol., 1981, 52: 626-639.

Gotman, J. *Automatic recognition of epileptic seizures in the EEG.* Electroenceph. Clin. Neurophysiol., 1982, 54: 530-540.

Gotman, J. *The use of computer in analysis and display of EEG and evoked potentials.* In: Current Practice of Clinical Electroencephalography. 2nd edition. D.D. Daly and T.A. Pedley (eds). New York: Raven Press, 1990; 51-83.

Gotman, J. and Wang, L.-Y. *State-dependent spike detection: concepts and preliminary results.* Electroenceph. Clin. Neurophysiol., 1991, 79: 11-19.

Gotman, J. and Wang, L.-Y. *State-dependent spike detection: validation.* Electroenceph. Clin. Neurophysiol., 1992, 83: 12-18.

Grieve, R.C.W., Parker, P.A. and Hudgins, B. *Training neural networks for stimulus artifact reduction in somatosensory evoked potential measurements.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 915.

Grönfors, T. and Juhola, M. *On digital filtering of auditory brainstem responses.* Medical Progress through Technology, 1993, 19: 145-151.

Grönfors, T. and Juhola, M. *Effect of sampling frequencies and averaging resolution on medical parameters of auditory brainstem responses.* Comput. Biol. Med., 1995, 25, 5: 447-454.

Grundy, B.L. Monitoring of sensory evoked potentials during neurosurgical operations: methods and applications. Neurosurgery, 1982: 11, 4: 556-575.

Grundy, B.L., Jannetta, P.J., Procopio, P.T., Lina, A., Boston, J.R. and Doyle, E. *Intraoperative monitoring of brain-stem auditory evoked potentials.* J. Neurosurg., 1982, 57: 674-681.

Grundy, B.L. *Evoked potential monitoring.* In: Monitoring in Anesthesia and Critical Care Medicine. C.D. Blitt (ed). New York: Churchill Livingstone, 1985; 345-411.

Gupta, L., Molfese, D.L. and Tammana, R. *An artificial neural-network approach to ERP classification.* Brain & Cognition, 1995, 27, 3: 311-330.

Gupta, L., Molfese, D.L., Tammana, R. and Simos, P.G. *Nonlinear alignment and averaging for estimating the evoked-potential.* IEEE Trans. Biomed. Eng., 1996, 43, 4: 348-356.

Hanley, J.A. and Charlton, M.H. *EEG in the operation room: artifacts and unusual waveforms.* Am. J. EEG Technol., 1982, 22: 135-141.

Hannaford, B. and Lehman, S. *Short time Fourier analysis of the electromyogram: fast movements and constant contraction.* IEEE Trans. Biomed. Eng., 1986: 1173-1181.

Hansson, M., Gansler, T. and Salomonsson, G. *Estimation of single event-related potentials utilizing the Prony method.* IEEE Trans. Biomed. Eng., 1996, 43, 10: 973-981.

Hartley, R.V.L. *A more symmetrical Fourier analysis applied to transmission problems.* Proc. IRE, 1942, 30: 144-50.

Hasan, J., Hirvonen, K., Värri, A., Häkkinen, V. and Loula, P. *Validation of computer analysed polygraphic patterns during drowsiness and sleep onset.* Electroenceph. Clin. Neurophysiol., 1993, 87: 117-127.

Hellmann, G., Kuhn, M., Prosch, M. and Spreng, M. *Extensible biosignal (EBS) file format: simple method for EEG data exchange.* Electroenceph. Clin. Neurophysiol., 1996, 99: 426-431.

Hinrichs, H., Heinze, H.J. and Gaab, M.R. *Neurophysiologisches Monitoring bei neurochirurgischen Gefäßoperationen: Spezifische technische Anforderungen und deren Umsetzung* (Neurophysiological monitoring of neurosurgical vessel-operations: technical specification and implementation. In German). Z. EEG-EMG, 1992, 23: 195-202.

Hinrichs, H., Feistner, H. and Heinze, H.J. *A trend-detection algorithm for intraoperative EEG monitoring.* Med. Eng. Physics, 1996, 18, 8: 626-631.

Hirshkowitz, M. and Moore, C. *Issues in computerized polysomnography sleep.* Sleep, 1994, 17, 2: 105-112.

Hjorth, B. *EEG analysis based on time domain properties.* Electroenceph. Clin. Neurophysiol., 1970, 29: 306-310.

Hoke, M., Ross, B., Wickesberg, R.E. and Lütkenhöner, B. *Weighted averaging – theory and application to electric response audiometry.* Electroenceph. Clin. Neurophysiol., 1984, 57: 484-489.

Hoppe, U., Eysholdt, U. and Weiß, S. *A sequential detection method for late auditory evoked potentials.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 217.

Hostetler, W.E., Doller, H.J. and Homan, R.W. *Assessment of a computer program to detect epileptiform spikes.* Electroenceph. Clin. Neurophysiol., 1992, 83: 1-11.

Hull, C.J. *Electrical hazards of patient monitoring.* In: Monitoring in Anaesthesia and Intensive Care, P. Hutton and C. Prys-Roberts (eds). London: W.B. Saunders Company, 1994; 56-77. ISBN 0-7020-1407-9.

Husar, P. and Henning, G. *Bispectrum analysis of visually evoked potentials.* IEEE Trans. Med. Biol., 1997, 16, 1: 57-63.

Intriligator, J. and Polich, J. *On the relationship between EEG and ERP variability.* Int. J. Psychophysiol., 1995, 20, 1: 59-74.

Jagannathan, V., Bourne, J.R., Jansen, B.H. and Ward, J.W. *Artificial intelligence methods in quantitative electroencephalogram analysis.* Comput. Prog. Biomed., 1982, 15: 249-258.

James, C.J., Jones, R.D., Bones, P.J. and Carroll, G.J. *The self-organising feature map in the detection of epileptiform transients in the EEG.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 269.

James, C.J., Hagan, M.T., Jones, R.D., Bones, P.J. and Carroll, G.J. *Multireference adaptive noise canceling applied to the EEG.* IEEE Trans. Biomed. Eng., 1997, 44, 8: 775-779.

Jansen, B.H., Bourne, J.R. and Ward, J.W. (1981a) *Autoregressive estimation of short segment spectra for computerized EEG analysis.* IEEE Trans. Biomed. Eng., 1981, 28, 9: 630-638.

Jansen, B.H., Hasman, A. and Lenten, R. (1981b) *Piecewise analysis of EEGs using AR-modeling and clustering.* Comput. Biomed. Res., 1981, 14: 168-178.

Jansen, B.H., Hasman, A. and Lenten, R. (1981c) *Piecewise EEG analysis: an objective evaluation.* Int. J. Biomed. Comput., 1981, 12: 17-27.

Jansen, B.H., Bourne, J.R. and Ward, J.W. *Identification and labelling of EEG graphic elements using autoregressive spectral estimates.* Comput. Biol. Med., 1982, 12: 97-106.

Jansen, B.H. and Yeh, Y.-S. *Single trial evoked potential analysis by means of crosscorrelation and dynamic time-warping.* Signal Processing, 1986, 11: 179-186.

Jansen, B.H. and Dawant, B.M. *Knowledge-based approach to sleep EEG analysis – A feasibility study.* IEEE Trans. Biomed. Eng., 1989, 36: 510-518.

Jansen, BH. *Quantitative analysis of electroencephalograms: is there chaos in the future?* Int. J. Biomed. Comput., 1991, 27: 95-123.

Jasper, H.H. and Carmichael, L. *Special article: Electrical potentials from the intact human brain.* Science, 1935, 81: 51.

Jasper, H.H. *The ten-twenty electrode system of the International Federation.* Electroenceph. Clin. Neurophysiol., 1958, 10: 370-375.

Jayakar, P. B., Patrick, J.P., Sill, J., Shwedyk, E. and Seshia, S.S. *Artifacts in ambulatory cassette electroencephalograms.* Electroenceph. Clin. Neurophysiol., 1985, 61: 440-443.

Jellema, T. *Cortical generators of somatosensory evoked potentials : a current source density analysis.* Ph.D. Thesis. Tilburg: Tilburg University, 1993.

Jensen, E.W., Lindholm, P. and Henneberg, S.W. *Autoregressive modeling with exogenous input of middle-latency auditory-evoked potentials to measure rapid changes in-depth of anesthesia.* Meth. Inf. Med., 1996, 35, 3: 256-260.

Jones, R.D., Dingle, A.A., Caroll, G.J., Green, R.D., Black, M., Donaldson, I.M., Parkin, P.J., Bones, P.J. and Burgess, K.L. *A system for detection epileptiform discharges in the EEG: real-time operation and clinical trial.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 854.

Kandel, E.R., Schwartz, J.H. and Jessell, T.M. *Principles of neural science.* 3rd edition. London: Prentice Hall, 1991. ISBN 0-8385-8068-8.

Kaipio, J.P. and Karjalainen, P.A. (1997) *Simulation of nonstationary EEG.* Biol. Cybern., 1997, 76: 349-356.

Kawabata, N. *A nonstationary analysis of the electroencephalogram.* IEEE Trans. Biomed. Eng., 1973, 20, 6: 444-452.

Keeney, S. *Artifact: Sources and solutions.* Am. J. EEG Technol., 1981, 21: 147-158.

Kemp, B. and Kamphuisen, H. *Simulation of human hypnograms using a Markov chain model.* Sleep, 1986, 9: 405-414.

Kemp, B., Värri, A., Rosa, A.C., Nielsen, K.D. and Gade, J. *A simple format for exchange of digitized polygraphic recordings.* Electroenc. clin. Neurophysiol., 1992, 82: 391-393.

Kemp, B. *A proposal for computer-based sleep/wake analysis.* J. Sleep Res. 2, 1993: 179-185.

Kemp, B. and De Weerd, A.W. *Een slaapcentrum. Technische opzet, diagnostische en therapeutische mogelijkheden* (A sleep-center. Technical plan, diagnostic and therapeutic possibilities. In Dutch). Klinische Fysica, 1996, 2: 13-16.

Kirkup, L. Searle, A., Craig, A., McIsaac, P. and Moses, P. *EEG-based system for rapid on-off switching without prior learning.* Med. Biol. Eng. Comput., 1997, 35: 504-509.

Klass, D.W. *The continuing challenge of artifacts in the EEG.* Am. J. EEG Technol., 1995, 35: 239-269.

Kong, X., Luo, X. and Thakor, N.V. *Detection of EEG changes via a generalized Itakura distance.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1997; 1540-1542.

Kong, X., Brambrink, A., Hanley, D.F. and Thakor, N.V. *Quantification of injury-related EEG signal changes using distance measures.* IEEE Trans. Biomed. Eng., 1999, 46: 899-901.

Korhonen, I., Ojaniemi, J., Nieminen, K., Van Gils, M., Heikelä A. and Kari, A. *Building the IMPROVE Data Library.* IEEE Eng. Med. Biol., 1997, 16, 6: 25-32.

Korpinen, L., Pietilä T. Peltola, J. Nissilä M., Keränen, T., Tuovinen, T., Falck, B., Petránek, E.S. and Frey, H. *Evaluation of Epilepsy Expert – a decision support system.* Comp. Meth. Prog. Biomed., 1994, 45: 223-231.

Krajca, V., Petranek, S., Patakova, I. and Värri, A. *Automatic identification of significant graphoelements in multichannel EEG recordings by adaptive segmentation and fuzzy clustering.* Int. J. Biomed. Comput., 1991, 28: 71-89.

Kraus, N., Reed, N., Smith, D.I., Stein, L. and Cartee, C. *High-pass filter settings affect the detectability of MLRs in humans.* Electroenceph. Clin. Neurophysiol., 1987, 68: 234-236.

Krausz, H.I. *Identification of nonlinear systems using random impulse train inputs.* Biol. Cybern., 1975, 19: 217-230.

Lange, D.H. and Inbar, G.F. *A robust parametric estimator for single-trial movement related brain potentials.* IEEE Trans. Biomed. Eng., 1996, 43, 4: 341-347.

Lange, D.H., Pratt, H. and Inbar, G.F. *Modeling and estimation of single evoked brain potential components.* IEEE Trans. Biomed. Eng., 1997, 44, 9: 791-799.

Lesser, R.P., Webber, W.R.S. and Fischer, R.S. *Design principles for computerised EEG monitoring.* Electroenceph. Clin. Neurophysiol., 1992, 82: 239-247.

Levy, W.J., Shapiro, H.M., Marushak, G. and Meathe, E. (1980a) *Automated EEG processing for intraoperative monitoring: a comparison of techniques.* Anesthesiology, 1980, 53: 223-236.

Levy, W.J., Shapiro, H.M. and Meathe, E. (1980b) *The identification of rhythmic EEG artifacts by power-spectrum analysis.* Anesthesiology, 1980, 53: 505-507.

Levy, W.J. *Intraoperative EEG patterns: implications for EEG monitoring.* Anesthesiology, 1984, 60: 430-434.

Levy, W.J., Grundy, B.L. and Smith, N.T. *Monitoring the electroencephalogram and evoked potentials during anesthesia.* In: Monitoring in Anesthesia. 2nd edition. L.J. Saidman and N.T. Smith (eds). Boston: Butterworth, 1984; 227-267. ISBN 0-409-95072-6.

Levy, W.J. *Effect of epoch length on power spectrum analysis of the EEG.* Anesthesiology, 1987, 66: 489-495.

Lilliefors, H.W. *On the Kolmogorov-Smirnov test for normality with mean and variance unknown.* J. Amer. Stat. Ass., 1967, 62: 399-402.

Litscher, G., Schwarz, G. and Kleinert, R. *Brain-stem auditory-evoked potential monitoring - variations of stimulus artifact in brain-death.* Evoked Potentials – Electroenceph. Clin. Neurophysiol., 1995, 96, 5: 413-419.

Litscher, G., Keyl, G., Schwarz, G. and Soyer, H.P. *Neurophysiologische Signalableitung. Neue technische und praxisbezogene Aspekte zu EEG-Ableiteelektroden* (Neurophysiological signal recording. EEG electrodes: new technical and practical aspects. In German). Biomed. Technik, 1996, 41: 106-110.

Litscher, G. and Schwartz, G. *Burst-Suppression-Erkennung beim pEEG* (Detection of burst suppresion in the pEEG. In German). Biomed. Technik, 1997, 42: 12-15.

Lopes da Silva, F. (1987a) *Dynamics of EEGs as signals of neuronal populations: models and theoretical considerations.* In: Electroencephalography: Basic Principles, Clinical Applications and Related Fields. 2nd edition. E. Niedermeyer and F. Lopes da Silva (eds). Baltimore-Munich: Urban & Schwarzenberg, 1987; 15-28.

Lopes da Silva, F. (1987b) *EEG-analysis: theory and practice.* In: Electroencephalography: Basic Principles, Clinical Applications and Related Fields. 2nd edition. E. Niedermeyer and F. Lopes da Silva (eds). Baltimore-Munich: Urban & Schwarzenberg, 1987; 685-711.

Lopes da Silva, F. (1987c) *Event-related potentials: methodology and quantification.* In: Electroencephalography: Basic Principles, Clinical Applications and Related Fields. 2nd edition. E. Niedermeyer and F. Lopes da Silva (eds). Baltimore-Munich: Urban & Schwarzenberg, 1987; 763-772.

Lowy, K. and Weiss, B. *Assessing the significance of averaged evoked potentials with an on-line computer: the split-sweep method.* Electroenceph. Clin. Neurophysiol., 1968, 25: 177-180.

Lutman, M.E. and Sheppard, S. *Quality estimation of click-evoked otoacoustic emissions.* Scan. Aud., 1990, 19, 1: 3-7.

MacLennan, A.R. and Lovely, D.F. *Reduction of evoked-potential measurement time by a TMS320 based adaptive matched-filter.* Med. Eng. Physics, 1995, 17, 4: 248-256.

Madler, C., Keller, I., Schwender, D. and Pöppel, E. *Sensory information processing during general anaesthesia: effect of isoflurane on auditory evoked neuronal oscillations.* Br. J. Anaesth., 1991, 66: 81-87.

Mainardi, L.T., Bianchi, A.M. and Cerutti, S. *Digital biomedical signal acquisition and processing.* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 828-852. ISBN 0-8493-8346-3.

Mainardi, L.T., Yli-Hankala, A., Korhonen, I., Signorini, M.G., Bianchi, A.M., Takala, J., Nieminen, K. and Cerutti, S. *Monitoring the autonomic nervous system in the ICU through cardiovascular variability signals.* IEEE Eng. Med. Biol., 1997, 16, 6: 64-75.

Makhoul, J. *Linear prediction: a tutorial review.* Proc. IEEE, 1975, 63, 4: 561-580.

Mason, S.M., Su, A.P. and Hayes, R.A. *Simple online detector of auditory evoked cortical potentials.* Med. Biol. Eng. Comput., 1977, 15: 641-647.

McCallum, W.C. *Potentials related to expectancy, preparation and motor activity.* In: T.W. Picton (Ed.). Human event-related potentials. EEG Handbook (revised series), 1988, 3: 427-534.

McCallum, W.C. *Attention.* In: Encyclopædia Britannica, 15th edition. Chicago: Encyclopædia Britannica, 1997, 14; 391-396. ISBN 0-85229-633-9.

McEwen, J.A. and Anderson, G.B. *Modeling the stationarity and Gaussianity of spontaneous electroencephalographic activity.* IEEE Trans. Biomed. Eng., 1975, 22, 5: 361-369.

McFarland, W.H., Vivion, M.C., Wolf, K.E. and Goldstein, R. *Reexamination of effects of stimulus rate and number on the middle components of the averaged electroencephalic response.* Audiology, 1975, 14: 456-465.

McGillem, C.D. and Aunon, J.I. *Measurements of signal components in single visually evoked potentials.* IEEE Trans. Biomed. Eng., 1977, 24: 232-241.

McLean, L., Scott, R.N. and Parker, P.A. *Stimulus artifact reduction in evoked-potential measurements.* Archives Phys. Med. Rehabil., 1996, 77, 12: 1286-1292.

Mendel, M.I. and Goldstein, R. *Early components of the averaged electroencephalic response to constant level clicks during all-night sleep.* J. Speech Hearing Res., 1971, 14: 829-840.

Miskiel, E. and Özdamar, Ö *Real-time weighted averaging system for evoked potentials.* Proc. IEEE Eng. Med. Biol. Soc. Conf., 1987; 9: 589-590.

Montgomery, D.C. and Runger, G.C. *Applied statistics and probability for engineers.* New York: John Wiley & Sons, 1994.

Moulines, E., Dalle Molle, J.W., Choukri, K. and Charbit, M. *Testing that a stationary time-series is Gaussian: time-domain vs. frequency-domain approaches.* IEEE Signal Processing Workshop on Higher-Order Statistics, 1993, p. 336-340.

Mukesh, D. and Nadkar, R.Y. *Neural-network modeling of human electro-encephalogram patterns.* Current Science, 1997, 72, 4: 261-265.

Muthuswamy, J., Sherman, D.L., and Thakor, N.V. *Higher-Order Spectral Analysis of Burst Patterns in EEG.* IEEE Trans. Biomed. Eng., 1999, 46, 1: 92-99.

Nakamura, M., Shibasaki, H. and Nishida, S. *Method for recording short latency evoked-potentials using an ECG artifact elimination procedure.* J. Biomed. Eng., 1990, 12: 51-56.

Nakamura, M., Sugi, T., Ikeda, A., Kagigi, R. and Shibasaki, H. *Clinical application of automatic integrative interpretation of awake background EEG: quantitative interpretation, report making, and detection of artifacts and reduced vigilance level.* Electroenceph. Clin. Neurophysiol. 1996, 98: 103-112.

Nandekar, S.D. and Sanders, D.B. *Median averaging of electromyographic motor unit action potentials: comparison with other techniques.* Med. Biol. Eng. Comput., 1989, 27: 566-571.

Narayana Dutt, D. *Spectral estimation of EEG signals using cascaded inverse filters.* Int J. Biom. Comput., 1994, 36, 4: 251-256.

Neejärvi, J., Väri, A., Fotopoulos, S. and Neuvo, Y. *Weighted FMH filters.* Signal Proc., 1993, 31, 2: 181-190.

Neuman, M.R. *Biopotential electrodes.* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 745-757. ISBN 0-8493-8346-3.

Nielsen-Bohlman, L. Knight, R.T., Woods, D.L. and Woodward, K. *Differential auditory processing continues during sleep.* Electroenceph. Clin. Neurophysiol., 1991, 79: 281-290.

Nieminen, K., Langford, R.M., Morgan, C.J., Takala, J. and Kari, A. *A clinical description of the IMPROVE data library.* IEEE Eng. Med. Biol., 1997, 16, 6: 21-24.

Nishida, S., Nakamura, M. and Shibasaki, H. *Method for single-trial recording of somatosensory evoked potentials.* J. Biomed. Eng., 1993, 15: 257-262.

Oken, B. *Filtering and aliasing of muscle activity in EEG frequency analysis.* Electroenceph. Clin. Neurophysiol., 1986, 64: 77-80.

Oken, B.S. and Chiappa, K.H. *Short term variability in EEG frequency analysis.* Electroenceph. Clin. Neurophysiol., 1988, 69, 3: 191-198.

Oken, B.S. *Statistics for evoked potentials.* In: Evoked Potentials in Clinical Medicine. 2nd edition. K.H. Chiappa (ed). New York: Raven Press, 1989; 593-608. ISBN 0-88167-569-5.

Özdamar, Ö., Delgado, R.E., Eilers, R.E. and Widen, J.E. *Computer methods for online hearing testing with auditory brain-stem responses.* Ear and Hearing, 1990, 11, 6: 417-429.

Paige, A.L., Özdamar, Ö. and Delgado, R.E. *Two-dimensional spectral processing of sequential evoked potentials.* Med. Biol. Eng. Comput., 1996, 34: 239-243.

Palus, M. *Nonlinearity in normal human EEG - cycles, temporal asymmetry, nonstationarity and randomness, not chaos.* Biol. Cybern., 1996, 75, 5: 389-396.

Panych, L.P., Wada, J.A. and Beddoes, M.P. *Practical digital filters for reducing EMG artefact in EEG seizure recordings.* Electroenceph. Clin. Neurophysiol. 1989, 72: 268-276.

Pardey, J., Roberts, S. and Tarassenko, L. *A review of parametric modeling techniques for EEG-analysis.* Med. Eng. Physics, 1996, 18, 1: 2-11.

Park, S., Principe, J.C., Smith, J.R. and Reid, S.A. *TDAT – Time domain analysis tool for EEG analysis.* IEEE Trans. Biomed. Eng., 1990, 37, 8: 803-811.

Persson, J. *Comments on estimations and tests of EEG amplitude distributions.* Electroenceph. Clin. Neurophysiol., 1974, 37: 309-313.

Persson, J. *Comments on "Modeling the stationarity and gaussianity of spontaneous electroencephalographic activity" (McEwen and Anderson, 1975).* IEEE Trans. Biomed. Eng., 1977, 24: 302.

Picton, T.W., Hink, R.F., Perez-Abalo, M., Linden, R.D. and Wiens, A.A. *Evoked potentials. How now?* J. Electrophysiol. Tech., 1984, 10: 177-221.

Picton, T., Hunt, M., Mowrey, R., Rodriguez, R. and Maru, J. *Evaluation of brain-stem auditory evoked potentials using dynamic time warping.* Electroenceph. Clin. Neurophysiol., 1988, 71: 212-225.

Pierce, D.A. *Testing normality in autoregressive models.* Biometrika, 1985, 72: 293-297.

Pietilä T. Vapaakoski, S. Nousiainen, U., Väri, A., Frey, H., Häkkinen, V. and Neuvo, Y. *Evaluation of a computerized system for recognition of epileptic activity during long-term EEG recording.* Electroenceph. Clin. Neurophysiol., 1994, 90: 438-443.

Pradhan, N. and Narayana Dutt, D. *A nonlinear perspective in understanding the neurodynamics of EEG.* Comput. Biol. Med., 1993, 23, 6: 425-442.

Pradhan, N., Sadasivan, P.K., Chatterji, S. and Narayana Dutt, D. *Patterns of attractor dimensions of sleep EEG.* Comput. Biol. Med., 1995, 25, 4: 455-462.

Praetorius, H.M., Bodenstein, G. and Creutzfeldt, O.D. *Adaptive segmentation of EEG records: a new approach to automatic EEG analysis.* Electroenceph. Clin. Neurophysiol., 1977, 42: 84-94.

Prior, P. *The rationale and utility of neurophysiological investigations in clinical monitoring for brain and spinal cord ischaemia during surgery and intensive care.* Comp. Meth. Prog. Biomed., 1996, 51: 13-27.

Pronk, R.A.F. *EEG processing during cardiac surgery.* Ph.D. Thesis. Amsterdam: Vrije Universiteit Amsterdam, 1982.

Pronk, R.A.F. *Peri-operative monitoring.* In: Handbook of Electroencephalography and Clinical Neurophysiology. Revised series. Vol. 2: Clinical Applications of Computer Analysis of EEG and other Neurophysiological Signals. F.H. Lopes da Silva, W. Storm van Leeuwen and A. Rémond (eds). Amsterdam: Elsevier, 1986; 93-130.

Pryse-Phillips, W.E.M. *Nervous system diseases and disorders.* In: Nerves and nervous systems. Encyclopædia Britannica, 15th edition. Chicago: Encyclopædia Britannica, 1997, 24; 839-861. ISBN 0-85229-633-9.

Rampil, I.J., Sasse, F.J., Smith, N.T., Hoff, B.H. and Flemming, D.C. *Spectral edge frequency – a new correlate of anesthetic depth.* Anesthesiology, 1980, 53: S12.

Raudzens, P.A. *Intraoperative monitoring of evoked potentials.* Annals New York Academy of Sciences, 1982, 388: 308-326.

Rechtschaffen, A. and Kales, A. (eds) and Berger, R.J., Dement, W.C., Jacobson, A., Johnson, L.C., Jouvet, M., Monroe, L.J., Oswald, I., Roffwarg, H.P., Roth, B., Walter, R.D. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.* Public Health Service, U.S. Government Printing Office, Washington, D.C., 1968.

Regan, D. *Human Brain Electrophysiology – Evoked potentials and evoked magnetic fields in science and medicine.* Amsterdam: Elsevier, 1989. ISBN 0-444-01324-5.

Risk, W.S. *Viewing speed and frequency resolution in digital EEG.* Electroenceph. Clin. Neurophysiol., 1993, 87: 347-353.

Robinson, M.P., Flintoft, I.D. and Marvin, A.C. *Interference to medical equipment from mobile phones.* J. Med. Eng. Techn., 1997, 21: 141-146.

Roessgen, M., Boashash, B. and Deriche, J. *Preprocessing noisy EEG data using time-frequency peak filtering.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1993. Paper number: 808.

Romer, J. *Neuroanesthesia: nursing implications.* In: Acute Neuroscience Nursing. Concepts & Care. J. Lundgren. Boston: Jones & Bartlett Publ., 1986; 67-77. ISBN 0-86720-355-2.

Rosenfalck, P. *Intra- and extracellular potential fields of active nerve and muscle fibers - a physico-mathematical analysis of different models.* Akademisk Forlag, Kopenhagen, 1969.

Royston, P. *Shapiro-Wilk W test and its significance level.* Algorithm AS R94. Applied Statistics, 1995, 44, 4.

Saatchi, M.R., Gibson, C., Rowe, J.W.K. and Allen, E.M. *Adaptive multiresolution analysis based evoked-potential filtering.* IEE Proc.-Sci. Meas. Technol., 1997, 144, 4: 149-155.

Sadasivan, P.K. and Narayana Dutt, D. *Minimization of EOG artefacts from corrupted EEG signals using a neural network approach.* Comput. Bio. Med., 1994, 24, 6: 441-449.

Sahota, P.K., Bray, F. and Garcia-Mendez, L. *Respiration-related artifact presenting as intermittent rhythmic theta activity.* Am. J. EEG Technol., 1993, 33: 121-128.

Salden, J.H.L. *EEG stationarity detection using classical and auto-regressive methods.* M.Sc. Thesis. Eindhoven: Eindhoven University of Technology, 1997.

Saltzberg, B. and Burch, N.R. *Period analytic estimates of moments of the power spectrum: a simplified EEG time domain procedure.* Electroenceph. Clin. Neurophysiol., 1971, 30: 568-570.

Scherg, M. (1982a) *Distortion of the middle latency auditory response produced by analog filtering.* Scand. Audiol., 1982, 11: 57-60.

Scherg, M. (1982b) *Simultaneous recording and separation of early and middle latency auditory evoked potentials.* Electroenceph. Clin. Neurophysiol., 1982, 54: 339-341.

Schimmel, H. *The (±) reference: accuracy of estimated mean components in average response studies.* Science, 1967, 157: 92-94.

Schimmel, H., Rapin, I. and Cohen, M.M. Improving evoked response audiometry with special reference to the use of machine scoring. Audiology, 1974, 13: 33-65.

Schimmel, H., Rapin, I., Cohen, M.M. Improving evoked response audiometry. Results of normative studies for machine scoring. Audiology, 1975, 14: 466-479.

Schmidt, R.F. *Integrative functions of the central nervous system.* In: Fundamentals of Neurophysiology. 3rd edition. R.F. Schmidt (ed). New York: Springer-Verlag, 1985; 270-316. ISBN 0-387-96147-X.

Schultz, B., Bender, R., Schultz, A. and Pichlmayr, I. *Reduktion der Anzahl von EEG-Ableitungen für ein routinemäßiges Monitoring auf der Intensivstation* (Electroencephalographic monitoring in the ICU – Reduction of the number of recorded channels. In German). Biomed. Technik, 1992, 37, 9: 194-199.

Schwartz, D.M., Bloom, M.J. and Dennis, J.M. *Perioperative monitoring of auditory brainstem responses.* Hearing J., November 1985: 9-13.

Sclabassi, R.J., Hinman, C.L., Kroin, J.S. and Risch, H. *The modulatory effect of prior input upon afferent signals in the somatosensory system.* Proc. IEEE Joint Autom. Control Conf. 1977: 787-795.

Sclabassi, R.J., Vries, J.K. and Bursick, D.M. *Somatosensory evoked potentials to random stimulus trains.* Annals New York Academy of Sciences, 1982: 695-701.

Scott, R.N., McLean, L. and Parker, P.A. *Stimulus artefact in somatosensory evoked potential measurement.* Med. Biol. Eng. Comput., 1997, 35: 211-215.

Sebel, P.S., Maynard, D.E., Mauor, E. and Frank, M. *The cerebral function analyzing monitor (CFAM).* Br. J. Anaesth., 1983, 55: 1265-1270.

Sgro, J.A., Emerson, R.G. and Pedley, T.A. (1989a) *Methods for steadily updating the averaged responses during neuromonitoring.* In: Neuromonitoring in Surgery. J.E. Desmedt (ed). Amsterdam: Elsevier, 1989; 49-60.

Sgro, J.A., Emerson, R.G. and Stanton, P.C. (1989b) *Advanced techniques of evoked potential acquisition and processing.* In: Evoked Potentials in Clinical Medicine, 2nd edition. K.H. Chiappa (ed). New York: Raven Press, 1989; 609-629. ISBN 0-88167-569-5.

Shamsollahi, M.B., Coatrieux, J.L., Senhadji, L. Wendling, F. and Badier, J.M. *On some time-requency signatures in stereo-electroencephalography (SEEG).* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 919.

Shapiro, S.S. and Wilk, M.B. *An analysis of variance test for normality (complete samples).* Biometrika, 1965, 52: 591-611.

Shapiro, S.S., Wilk, M.B. and Chen, H.J. *A comparitive study of various tests for normality.* Am. Stat. Ass. J., Dec. 1968, 63: 1343-1372.

Sigl, J.C. and Chamoun, N.G. *An introduction to bispectral analysis for the electroencephalogram.* J. Clin. Mon., 1994, 10: 392-404.

Sims, J.K., Aung, M.H., Bickford, R.G., Billinger, T.W. and Shattuck, C.M. *Respirator artifact mimicking burst-suppression during electrocerebral silence.* Am. J. EEG Technol., 1973, 13: 81-87.

Smulders, F.T.Y., Kenemans, J.L. and Kok, A. *A comparison of different methods for estimating single-trial P300 latencies.* Electroenceph. Clin. Neurophysiol., 1994, 92: 107-114.

Spehlmann, R. *EEG Primer.* Amsterdam: Elsevier, 1981. ISBN 0-444-80299-1.

Spehlmann, R. *Evoked Potential Primer: Part C. Visual, auditory, and somatosensory evoked potentials in clinical diagnosis.* Boston: Butterworth, 1985; 190-277.

Spencer, E.M. *Monitoring of the central nervous system and the effects of anaesthesia.* In: Monitoring in Anaesthesia and Intensive Care, P. Hutton and C. Prys-Roberts (eds). London: W.B. Saunders Company, 1994; 256-283. ISBN 0-7020-1407-9.

Spivak, L.G. and Malinoff, R.L. *Effect of digital filtering of ABR on scorer variability.* Am. J. Otology, 1991, 12, 1: 20-24.

Stecker, M.M. and Patterson, T. *Strategics for minimizing 60 Hz pickup during evoked-potential recording.* Evoked Potentials - Electroenceph. Clin. Neurophysiol., 1996, 100, 4: 370-373.

Stockard, J.J., Stockard, J.E. and Sharbrough, F.W. *Nonpathologic factors influencing brainstem auditory evoked potentials.* Am. J. EEG Technol., 1978, 18: 177-209.

Straw, R.N., McAdam, D. Berry, C.A. and Mitchell, C.L. *A simple cable for reduction of movement artifact in electroencephalographic recordings.* Electroenceph. Clin. Neurophysiol., 1967, 22: 90-92.

Strejc, V. *Least squares parameter estimation.* Automatica, 1980, 16: 535-550.

Swets, J.A. *Measuring the accuracy of diagnostic systems.* Science, 1988, 240: 1285-1293.

Thakor, N.V. and Sherman, D. *Wavelet (time-scale) analysis in biomedical signal processing.* In: The Biomedical Engineering Handbook. J.D. Bronzino (ed). Boca Raton: CRC Press, 1995; 886-906. ISBN 0-8493-8346-3.

The IBIS project (EU BIOMED2, BMH4-97-2570). Improved Monitoring for Brain Dysfunction in Intensive Care and Surgery, 1997-2000.

The SIESTA project (EU BIOMED2, BMH4-CT97-2040). A New Standard for Integrating Polygraphic Sleep Recordings into a Comprehensive Model of Human Sleep and Its Validation in Sleep Disorders, 1997-2000.

Thomsen, C.E., Rosenfalck, A. and Nørregaard-Christensen, K. *Assessment of anaesthetic depth by clustering analysis and autoregressive modelling of electroencephalograms.* Comp. Meth. Prog. Biomed., 1991, 34: 125-138.

Thomsen, C.E. and Prior, P.F. *Quantitative EEG in assessment of anaesthetic depth — a comparative study of methodology.* Br. J. Anaesth., 1996, 77: 172-178.

Thomsen, C.E., Gade, J., Nieminen, K., Langford, R.M., Ghosh, I.R., Jensen, K., Van Gils, M., Rosenfalck, A., Prior, P.F. and White, S. *Collecting EEG signals in the IMPROVE data library.* IEEE Eng. Med. Biol., 1997, 16, 6: 33-40.

Thornton, C. *Assessment of graded changes in the Central Nervous System, during general anaesthesia and surgery in man, using the auditory evoked response.* Ph.D. Thesis. University of London, 1990.

Thornton, C. *Evoked potentials in anaesthesia.* European J. Anaesthesiology, 1991, 8: 89-107.

Thornton, C. and Sharpe, R.M. *Evoked responses in anaesthesia.* Br. J. Anaesth. 1998, 81: 771-781.

Tian, J.L., Juhola, M. and Grönfors, T. *Latency estimation of auditory brain-stem response by neural networks.* Artificial Intell. Med., 1997, 10, 2: 115-128.

Todd, M.M. *EEGs, EEG processing, and the Bispectral index.* Anesthesiology, 1998, 89, 4: 815-817.

Tomberg, C., Nolè, P., Ozaki, I. and Desmedt, J.E. *Inadequacy of the average reference for the topographic mapping of focal enhancements of brain potentials.* Electroenceph. Clin. Neurophysiol., 1990, 77: 259-265.

Vachon, B., Dubuisson, B. et Samson-Dollfus, D. *Etude automatique de l'EEG: une méthode de détection des non stationnarites* (Automatic EEG processing: a method for detection of non-stationarities. In French). Int. J. Biom. Comput., 1978, 9: 147-162.

Van de Velde, M. and Cluitmans, P.J.M. *EEG analysis for monitoring of anesthetic depth.* M.Sc. Thesis. Eindhoven: Eindhoven University of Technology, 1991. EUT report 91-E-254. ISBN 90-6144-254-0.

Van de Velde, M. Cluitmans, P.J.M. and Declerk, A.C. (1993a) *Optimization of stimulation frequency in middle latency auditory evoked potentials in humans.* Proc. 2nd European Conf. Eng. Med. (ESEM 1993). Technology and Health Care, 1993, 1: 340-341.

Van de Velde, M., Cluitmans, P.J.M., Declerck, A.C., Griep, P., Beecher, L., Verbeek, H.M.J.C. and Voets, E. (1993b) *Assessment of methods for measuring middle latency auditory evoked potentials in humans during sleep.* In: Sleep-Wake Research in The Netherlands. Leiden: Dutch Society for Sleep-Wake Research, 1993, 4; 195-199. ISBN 90-73675-04-9.

Van de Velde, M., De Beer, N.A.M. and Cluitmans, P.J.M. *Evaluation of automatic quality assessment of auditory evoked potentials in clinical data.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1996. Paper number: 493.

Van den Berg-Lenssen, M.M.C., Brunia, C.H.M. and Blom, J.A. *Correction of ocular artifacts in EEGs using an autoregressive model to describe the EEG; a pilot study.* Electroenceph. Clin. Neurophysiol., 1989, 73: 72-83.

Van den Berg-Lenssen, M.M.C., Van Gisbergen, J.A.M. and Jervis, B.W. *Comparison of two methods for correcting ocular artefacts in EEGs.* Med. Biol. Eng. Comput., 1994, 32: 501-511.

Van der Weide, H. and Pronk, R.A.F. *Interference suppression for EEG recording during open heart surgery.* Electroenceph. Clin. Neurophysiol., 1979, 46, 5: 609-612.

Van Gils, M. *Peak identification in auditory evoked potentials using artifical neural networks.* Ph.D. Thesis. Eindhoven: Eindhoven University of Technology, 1995. ISBN 90-386-0036-4.

Van Gils, M., Jansen, H., Nieminen, K., Summers, R. and Weller, P.R. (1997a) *Using artificial neural networks for classifying ICU patient states.* IEEE Eng. Med. Biol., 1997, 16, 6: 41-47.

Van Gils, M., Rosenfalck, A., White, S., Prior, P., Gade, J., Senhadji, L., Thomsen, C.E., Ghosh, I.R., Langford, R.M. and Jensen, K. (1997b) *Signal processing in prolonged EEG recordings during intensive care.* IEEE Eng. Med. Biol., 1997, 16, 6: 56-63.

Van Hooff, J.C., De Beer, N.A.M., Brunia, C.H.M., Cluitmans, P.J.M., Korsten, H.H.M., Tavilla, G. and Grouls, R. *Information-processing during cardiac-surgery - an event-related potential study.* Evoked Potentials - Electroenceph. Clin. Neurophysiol., 1995, 96, 5: 433-452.

Van Hooff, J.C., Brunia, C.H.M, and Allen, J.J.B. *Event-related potentials as indirect measures of recognition memory.* Int. J. Psychophysiol., 1996, 21: 15-31.

Van Hooff, J.C., De Beer, N.A.M., Brunia, C.H.M., Cluitmans, P.J.M. and Korsten, H.H.M. *Event-related potential measures of information-processing during general-anesthesia.* Electroenceph. Clin. Neurophysiol., 1997, 103, 2: 268-281.

Väri, A., Nielsen, K., Penzel, T., Hellmann, G., Macerata, A. and Gottlieb, L. *File Exchange Format for Vital Signs.* Preliminary description prepared for CEN/TC251/WG5 by Project Team PT5-021, rev. 0.60, September 28, 1997.

Väri, A., Hirvonen, K., Hasan, J., Loula, P. and Häkkinen, V. *A computerized analysis system for vigilance studies.* Comp. Meth. Progr. Biomed., 1992, 39: 113-124.

Vaz, F. and Principe. J.C. *Neural networks for EEG signal decomposition and classification.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1995. Paper number: 182.

Verleger, R. *The instruction to refrain from blinking affects auditory P3 and N1 amplitudes.* Electroenceph. Clin. Neurophysiol., 1991, 78: 240-251.

Verleger, R. *Valid identification of blink artefacts: are they larger than 50 μV in EEG records?* Electroenceph. Clin. Neurophysiol., 1993, 87, 6: 354-363.

Victor, J.D. and Mast, J. *A New Statistic for Steady-State Evoked-Potentials.* Electroenceph. Clin. Neurophysiol., 1991, 78, 5: 378-388.

Vigário, R.N. *Extraction of Ocular Artifacts from EEG Using Independent Component Analysis.* Electroenceph. Clin. Neurophysiol., 1997, 103, 3: 395-404.

Ward, D.-M., Jones, R.D., Bones, P.J. and Carrol, G.J. *Enhancement of deep epileptiform activity in the EEG via 3-D adaptive spatial filtering.* IEEE Trans. Biomed. Eng., 1999, 46, 6: 707-716.

Watt, R.C., Sisemore, C., Kanemoto, A. Dakwar, P. and Mylrea, K. *Artificial neural networks used with bispectral analysis for intra-operative EEG monitoring.* Proc. IEEE Eng. Med. Biol. Soc. Conf. 1995. Paper number: 40.

Webber, W.R.S., Litt, B. Wilson, K. and Lesser, R.P. *Practical detection of epileptiform discharges (EDs) in the EEG using an artifical neural network: a comparison of raw and parameterized EEG data.* Electroenceph. Clin. Neurophysiol., 1994, 91: 194-204.

Weerts, T.C. and Lang, P.J. *The effects of eye fixation and stimulus and response location on the Contingent Negative Variation.* Biol. Psychol., 1973, 1: 1-19.

Weiss, M.S. *Testing correlated "EEG-like" data for normality using a modified Kolmogorow-Smirnow statistic.* IEEE Trans. Biomed. Eng., 1986, 33: 1114-1120.

Williams G.W, Lüders, H.O., Brickner, A., Goormastic, M. and Klass, D.W. *Interobserver variability in EEG interpretation.* Neurology, 1985, 35: 1714-1719.

Williams, W.J., Zaveri, H.P. and Sackellares, J.C. *Time-frequency analysis of electrophysiology signals in epilepsy.* IEEE Eng. Med. Biol. , 1995, 14, 2: 133-143.

Wilson, S.B., Harner, R.N., Duffy, F.H., Tharp, B.R., Nuwer, M.R. and Sperling, M.R. *Spike detection. I. Correlation and reliability of human experts.* Electroenceph. Clin. Neurophysiol., 1996, 98: 186-198.

Wilson, S.B., Turner, C.A., Emerson, R.G. and Scheuer, M.L. *Spike detection II: automatic, perception-based detection and clustering.* Clinical Neurophysiology, 1999, 110: 404-411.

Wong, P.K.H. and Bickford, R.G. *Brain stem auditory evoked potentials: the use of noise estimate.* Electroenceph. Clin. Neurophysiol., 1980, 50: 25-34.

Woody, C.J. *Characterisation of an adaptive filter for the analysis of variable latency neuroelectric signals.* Med. Biol. Eng., 1967, 5: 539-554.

Wu, J., Ifeachor, E.C., Allen, E.M., Wimalaratna, W.K. and Hudson, N.R. *Intelligent artefact identification in electroencephalography signal processing.* IEE Proc.-Sci. Meas. Technol., 1997, 144, 5: 193-201.

Young,W.L. and Ornstein, E. *Compressed spectral array during during cardiac arrest and resuscitation.* Anesthesiology, 1985, 62: 535-538.

Zeki, S. *The visual image in mind and brain.* Scientific American, 1992, 267, 3: 42-50.

Zetterberg, L.H. *Estimation of parameters for a linear difference equation with application to EEG analysis.* Math. Biosciences, 1969, 5: 205-226.

Summary

This thesis describes the research on validation of electroencephalographic measurements: the electroencephalogram (EEG, in Part I), and evoked potentials (EPs, in Part II). These neurophysiological measurements are generally of very low amplitude and are easily disturbed by other physiological signals or external sources. The studies on EEG validation focused on time-related aspects of validation; the research on EP validation focused on objective assessment of EP signal quality. The work is positioned in-between the actual recording and the (human or automated) clinical review process.

The general overview in chapter I-1 describes the background of neurophysiological signals and the (technical) requirements for recording of electroencephalographic data, focusing on practical issues. The various types of artefacts are described along with previous work in artefact processing. Subsequently the research in EEG validation and artefact detection is reviewed, consisting of a section about existing methods for signal processing and of a section that briefly introduces technical issues in clinical EEG processing.

The study presented in chapter I-2 deals with automatic detection of muscle artefact in 21 normal human subjects, using (classical) time domain and frequency domain methods. Distributions as calculated from a reference period in each subject were used to investigate the statistics of the parameter ranges. Performance of the automatic detection was compared to human (visual) assessment, using per-subject thresholds, and constant (empirical) thresholds for the entire data set. The results indicate that a 1-second epoch length was optimal for detection of muscle artefact. The analysis using a Slope (first derivative) threshold or absolute 'high beta' power (>25Hz) showed the best results in sensitivity (80%) and specificity (90%), matching the expert's performance. Constant threshold settings performed better than statistical thresholds per subject.

Chapter I-3 tackles the problem of artefact detection in seven 24-hour EEG recordings in the intensive care unit (ICU). ICU recordings have received less attention than e.g., epilepsy monitoring, although recordings in this environment present an interesting application area. The investigated artefact detection methods were based on statistical differences between signal parameters, using time-varying autoregressive (AR) modeling and Slope detection. The study focused on the optimal settings for context incorporation by testing the algorithms for different time windows and epoch lengths against the artefact markings made by two human observers.

The results of the ICU study indicated that a relatively short period (20-40 seconds) provided sufficient context information. The combined AR and Slope detection parameters yielded good performance, detecting approximately 90% of the artefacts as indicated by the consensus score of the human observers. However, the accompanying positive prediction was rather low: only 53%. This rather low positive prediction was possibly adversely influenced by the use of consensus scores; the consensus may also have excluded some true artefacts. In addition, it would seem sensible to err towards high sensitivity (at a cost of lesser positive prediction); this would allow observers to visually analyse events detected by automation, and categorise them as artefact/non-artefact.

Part II describes the research on objective assessment of evoked potential (EP) quality. An EP represents the electrophysiological behaviour of a specific neural pathway, as measured on the scalp. EPs are signals that can be obtained from advanced processing of the EEG, often recorded during a repeated task. The different types of EPs are described in the introduction to Part II, in chapter II-1, which also focuses on artefacts and procedures specific to EP recordings, and concludes by describing the investigated methods for quality assessment.

Chapter II-2 describes a study where the usefulness of the ()-reference method and the 'single-point' method for residual noise estimation was investigated in middle latency auditory evoked potentials (MLAEPs) obtained during cardiac surgery. The use of auditory evoked potentials in clinical monitoring is still mainly restricted to research, but its potential as an indicator for levels of anaesthesia is well established. However, the auditory evoked potential is a very small signal (investigated post-stimulus interval 0-200ms, amplitudes <10μV) that can be disturbed easily, which is especially true in the operating room, where sub-optimal measurement conditions are encountered. Therefore, automatic monitoring of the quality of the recording is desired.

The visual screening of the signals by human observers was used as a reference for the performance evaluation of the methods. The best results were obtained when the computation was restricted to the brainstem part of the MLAEPs. High performance for detection of evoked potential waveforms 'acceptable for clinical use' was obtained by the ()-reference method, by combining the quality factor P with the number of sweeps N contributing to the average. For a selected combination of $P > 8\text{dB}$ and $N > 2400$ (corresponding to a minimal recording time of 30 seconds), the sensitivity for detecting 'acceptable' MLAEP waveforms was 87%, with 82% specificity.

The study in chapter II-3 explores a previously acquired data set, consisting of 'Contingent Negative Variation' (CNV) waveforms. This signal is measured in a task where a warning-stimulus and a subsequent response-stimulus are presented to a subject, which is to be followed by a prompt button-press action. The data set provided an interesting starting-point to investigate the usefulness of SNR assessment in event related potentials, which present slower components and larger amplitudes when compared to

the MLAEP waveforms of chapter II-2. In particular the effect of response lateralization was examined: the amplitudes of (various parts of) these CNV waveforms were dependent on the limb used in the response action. This effect was found in the original study and was indicated here also from the evaluation of the SNR. Several other statistical significant differences were found in the analyses, leading to the overall conclusion that the (±)-reference method provides useful information, corroborating the findings from other amplitude based analyses.

Minimisation of the data-acquisition time required for middle latency auditory evoked potentials (MLAEPs) is studied in chapter II-4. Shorter measurements can reduce the influence of non-stationary EEG and artefacts and provide a more frequent status update, which is crucial for a practical application of these measurements in clinical monitoring.

The study investigated the amplitude characteristics and quality of MLAEPs obtained using both conventional stimulation and random high-frequency stimulation during stages 'Wake' and stationary NREM2 sleep in 14 human subjects. The evaluation of signal quality focused on the estimation of signal-to-noise ratios, based on the (±)-reference and a new 'convergence' calculation method. The overall results show the high-frequency random stimulation method, using a mean click rate of 80/second or 90/second, yields higher quality MLAEP signals, which can be acquired faster than conventional, regular stimulation methods.

In conclusion, the research described in this thesis has focused on general applicable validation methods and the evaluation of the methods in several clinical data sets. The EEG was modeled through parameterisation of amplitude and/or frequency characteristics in short signal periods (epochs). In a study for the detection of muscle artefact in the EEG, the use of short epochs of EEG data (0.5 to 2 seconds) yielded the best results. The analysis using a Slope threshold (first derivative) or absolute 'high beta' power showed the best results, matching the expert's performance. A model for EEG signal context, consisting of autoregressive parameters and Slope parameters in successive epochs, was very successful in the detection of non-stationarities and artefacts in a large dataset as measured in critically ill patients during their stay in the intensive care unit. This EEG study pointed out that objective computerised validation may be applied generally when the signal context is taken into account in the model(s) or parameters used. A signal period providing adequate context information can be limited to 20 to 40 seconds. Further research into the performance of different context models and different parameter settings is necessary for accurate marking of artefacts/events.

Validation through the use of amplitude parameters was very successful for quality improvement of EPs. Because EPs are discrete signals, having known properties, and EP waveform shape changes relatively little during the recording, signal context (the ongoing measurement) can be used to quantify EP quality. This can be implemented objectively by using the (±)-reference method, or by using the 'convergence' method,

which have been evaluated in Part II of this thesis. The correlation between the calculated signal-to-noise power ratios and clinical (visually assessed) quality has been demonstrated clearly. Furthermore, the quality factors were useful for the evaluation of the high-frequency random stimulation method for middle latency evoked potentials.

Samenvatting

Dit proefschrift beschrijft het onderzoek naar validatie van elektro-encefalografische metingen, met betrekking tot het elektro-encefalogram (EEG, in Deel I) en 'evoked potential' metingen (EPs, in Deel II). De gemeten voltages bij dit soort neurofysiologische metingen zijn in het algemeen erg klein en kunnen makkelijk worden verstoord door andere fysiologische signalen of externe bronnen. In de EEG studies zijn vooral de tijdsaspecten van validatie onderzocht; het onderzoek naar EP validatie was gericht op het objectief bepalen van signaal kwaliteit. Het onderzoek is gepositioneerd tussen de eigenlijke EEG/EP metingen en het (menselijke of geautomatiseerde) klinische beoordelingsproces.

In de algemene inleiding van Deel I (hoofdstuk I-1) wordt de achtergrond van neurofysiologische signalen beschreven in relatie tot de (technische) eisen voor het meten van het EEG, met speciale aandacht voor een aantal praktische zaken. De verschillende types van artefacten (verstoringen) worden beschreven aan de hand van de bestaande literatuur. Hierna wordt het onderzoek naar EEG validatie en artefact detectie beschreven in een sectie over bestaande methoden voor signaalverwerking, en in een sectie waarin een korte introductie wordt gegeven over signaalverwerking in een drietal hoofd-onderwerpen van de klinische neurofysiologie.

De studie in hoofdstuk I-2 behandelt automatische detectie van spierartefacten in 21 normale menselijke proefpersonen, gebruik makend van (klassieke) tijd-domein en frequentie-domein methoden. Op basis van de statistische verdelingen van de berekende parameters, is onderzocht hoe goed de automatische detectie presteerde in relatie tot menselijke (visuele) beoordeling. De prestaties zijn bepaald voor het gebruik van parameter-grenzen per individuele proefpersoon (statistisch), en voor het gebruik van constante parameter-grenzen (empirisch) over de hele data-set. De resultaten geven aan dat een tijdsduur van 1 seconde voor een 'epoch' (stukje signaal) optimaal was voor de detectie van spierartefacten. De artefact detectie met behulp van parameter-grenzen gebaseerd op de 1e afgeleide ('Slope') en gebaseerd op de frequentie-band boven 25Hz ('high beta') gaven de beste resultaten: een sensitiviteit van 80% en een specificiteit van 90% werden hiermee behaald. Dit benaderde de prestaties van de menselijke expert. Ook bleek dat constante parameter-grenzen betere prestaties tot gevolg hadden dan individuele instellingen per proefpersoon.

Hoofdstuk I-3 gaat in op het probleem van artefacten in zeven 24-uurs registraties van patiënten tijdens hun verblijf in de intensive care unit (ICU). Patiëntbewaking in de ICU door middel van het EEG krijgt minder aandacht dan bijvoorbeeld metingen ten behoeve van epilepsie, maar registraties in deze omgeving vormen een interessant toepassingsgebied. De onderzochte methoden voor artefact detectie zijn gebaseerd op statistische verschillen tussen signaal parameters, gebruik makend van in de tijd variërende autoregressieve (AR) modellen en Slope detectie. De studie ging nader in op het bepalen van een optimale context door de algoritmes te testen voor verschillende tijdvensters en epoch lengtes ten opzichte van de artefact markeringen van twee menselijke beoordelaars.

De resultaten van de ICU studie geven aan dat een relatief korte periode (20-40 seconden) voldoende context informatie bevat. De gecombineerde detectie met AR en Slope parameters detecteerde ongeveer 90% van alle artefacten, zoals gemarkeerd door beide menselijke beoordelaars. De hier bij behorende positieve predictie was echter enigszins laag: slechts 53%. Deze lage positieve predictie is mogelijk beïnvloed door het gebruik van consensus markeringen; hierdoor zijn waarschijnlijk ook echte artefacten buiten beschouwing gelaten. Het is overigens zinvol in dit soort studies om een hogere sensitiviteit te verkiezen boven een hoge positieve predictie, om zo min mogelijk verstoringen te missen. Na een initiële automatische markering, kunnen gemarkeerde 'events' op basis van een nadere visuele beoordeling van de signalen relatief snel worden ingedeeld als zijnde artefact, of niet-artefact.

Deel II beschrijft het onderzoek naar het objectief bepalen van de kwaliteit van evoked potential (EP) metingen. Een EP representeert het elektrofysiologische gedrag van een specifieke zenuwbaan, en wordt meestal gemeten op de schedel door het EEG signaal verder te bewerken. EPs worden in het algemeen gemeten door een proefpersoon een bepaalde stimulus herhaald aan te bieden, of deze een bepaalde taak te laten herhalen. De verschillende types van EPs worden beschreven in de inleiding van Deel II, in hoofdstuk II-1, waarin ook nader wordt ingegaan op artefacten en meetprocedures specifiek voor EP metingen. Aan het eind van deze inleiding worden de onderzochte methoden beschreven, met betrekking tot de rest van Deel II.

Hoofdstuk II-2 beschrijft een studie waarin het gebruik van de ()-reference methode en de 'single-point' methode voor het schatten van de ruis is geëvalueerd in middle latency auditieve evoked potentials (MLAEPs) zoals gemeten tijdens hartchirurgie. Hoewel auditieve evoked potentials ten behoeve van klinische patiëntbewaking nog voornamelijk alleen in onderzoek worden gebruikt, zijn deze signalen mogelijk geschikt om het niveau van anesthesie te bepalen. MLAEPs zijn echter hele kleine signalen (het onderzochte post-stimulus interval beslaat 0-200ms, met amplitudes <10μV) welke makkelijk kunnen worden verstoord. Dit is vooral een probleem in een operatiekamer, waar de meetomstandigheden verre van optimaal zijn. Daarom is het automatisch bepalen van de signaalkwaliteit gewenst.

De visuele beoordeling van de MLAEPs door menselijke beoordelaars is gebruikt als referentie voor het bepalen van de prestaties van de gebruikte methoden. De beste resultaten werden behaald bij het gebruik van slechts het brainstem gedeelte (post-stimulus interval 5-15ms) van de gemeten signalen. Vooral de (±)-reference methode presteerde goed in het classificeren van golfvormen, welke als 'acceptabel voor klinisch gebruik' waren ingedeeld. De combinatie van de berekende kwaliteitsfactor P en een minimaal aantal sweeps N, bijvoorbeeld voor $P > 8$dB en $N > 2400$ (een minimale registratieduur van 30 seconden), leverde een sensitiviteit en specificiteit op van respectievelijk 87% en 82%.

De studie van hoofdstuk II-3 beschrijft een eerste evaluatie van de signaalkwaliteit in een data-set, bestaande uit eerder geregistreerde 'Contingent Negative Variation' (CNV) signalen. Het CNV is een signaal dat wordt gemeten gedurende een taak waarin een twee waarschuwings-stimuli elkaar opvolgen, waarbij een proefpersoon onmiddellijk na het tweede waarschuwings-signaal op een knop dient te drukken. Deze data-set vormde een interessant beginpunt om het automatisch bepalen van signaalkwaliteit te onderzoeken in event related potentials, waar langzamere componenten en hogere amplitudes worden gemeten dan in de MLAEP golfvormen van hoofdstuk II-2. Hier is vooral het effect van response lateralisatie onderzocht: de amplitudes in (verschillende delen van) de CNV zijn afhankelijk van het lichaamsdeel dat is gebruikt bij het indrukken van de knop. Dit effect was al aangetoond in de originele studie, en is hier ook gevonden door evaluatie van de signaal-ruis verhouding. De analyses lieten meerdere interessante significante verschillen zien, met als eindconclusie dat de (±)-reference methode bruikbare informatie oplevert, te gebruiken voor het staven van andere op signaal-amplitude gebaseerde analyses.

Minimalisatie van de registratie tijd van MLAEPs is onderzocht in hoofdstuk II-4. Korte metingen kunnen de invloed van niet-stationariteiten en artefacten in het EEG verminderen. Bovendien kan het MLAEP signaal dan vaker worden bepaald, hetgeen van cruciaal belang is voor een praktische toepassing van deze metingen in de klinische patiëntbewaking.

De studie onderzocht de amplitude karakteristieken en de kwaliteit van de MLAEP signalen, zoals gemeten met conventionele stimulatie en gerandomiseerde hoog-frequent stimulatie gedurende de slaapstadia 'waak' en NREM2 in 14 proefpersonen. Voor de evaluatie van de signaalkwaliteit (signaal-ruis verhuiding) is gebruik gemaakt van de (±)-reference, en van een nieuwe methode gebaseerd op het berekenen van de 'convergentie' van het signaal. De resultaten laten zien dat de gerandomiseerde hoog-frequent stimulatie, bij een gemiddelde klik-snelheid van 80/seconde of 90/seconde, een hogere kwaliteit MLAEP signalen oplevert, welke sneller kunnen worden gemeten dan met de conventionele, regelmatige stimulatie methode.

Dit promotiewerk heeft zich gericht op algemeen toepasbare validatie methoden, en de evaluatie in een aantal experimentele en klinische studies. Het EEG is hierbij gemodelleerd door parametrisatie van amplitude- en/of frequentie-eigenschappen in korte tijdsintervallen (epochs). In een studie naar detectie van spierartefacten in het EEG is aangetoond dat zeer korte epochs (0.5 tot 2 seconden) de beste resultaten opleveren, waarbij zowel de 1e-afgeleide (amplitude) als een frequentie-band de prestaties van een menselijke expert benaderen. Een model voor EEG signaalcontext, bestaande uit autoregressieve parameters van meerdere korte epochs, was succesvol voor het detecteren van niet-stationairiteiten en artefacten in een grote klinische dataset, bestaande uit metingen van patiënten gedurende hun verblijf in de intensive care unit. Uit deze EEG studie bleek dat objectieve validatie door een computer programma mogelijk is wanneer de signaalcontext meegenomen wordt in de gebruikte modellen en parameters. Verder onderzoek naar verschillende modellen voor signaalcontext is nodig om te bepalen welke parameter instellingen optimaal zijn voor het nauwkeurig markeren van artefacten en events.

Validatie middels amplitude parameters is zeer succesvol gebleken in het verbeteren van de kwaliteit van evoked potentials. Omdat het EP signaal redelijk kan worden bepaald, is de signaalcontext (de gemeten responsen) hier te gebruiken voor het kwantificeren van de EP kwaliteit. Gangbare methoden modelleren dit met behulp van de amplitude variabiliteit in EP en EEG; dit is hier verder uitgewerkt. Een objectieve bepaling van de signaalkwaliteit kan worden geïmplementeerd met gebruikmaking van de (±)-reference methode, en de 'convergentie' methode. De hiermee berekende kwaliteit is duidelijk gerelateerd aan klinische (visueel bepaalde) kwaliteit. Ook waren deze methoden bruikbaar in de evaluatie van een alternatieve, hoog-frequent, random stimulatie meetmethode voor middle latency auditieve evoked potentials.

Nawoord

Het onderzoekswerk beschreven in dit proefschrift is, zoals ook al blijkt uit de auteurs bij verschillende hoofdstukken, natuurlijk niet door één persoon verricht. De inhoud van dit boekje had nooit tot stand kunnen komen zonder de sturing, de inzet, het meeleven, en bijdragen van een groot aantal mensen.

Eerst en vooral wil ik Pierre Cluitmans bedanken. Pierre, je was al vanaf mijn afstudeeronderzoek een inspiratiebron voor het ontdekken van de wereld der neurofysiologie, en de (on)mogelijkheden van signaalverwerking in dit vakgebied. Bedankt voor alle tips, aangewezen richtingen, en het samen opzetten van onderzoek. Je kritische leeswerk en opmerkingen waren vanaf het eerste begin van dit proefschrift tot aan de laatste veranderingen bijzonder waardevol. Het was een groot genoegen om jou als begeleider en collega te hebben.

Professor Beneken wil ik vooral danken voor de inspiratie om niet alleen op een wetenschappelijke manier onderzoek te verrichten, maar het daarna ook goed op te schrijven. Nadat ik enigszins had geworsteld met de grote lijn van het proefschrift, kwam het overzicht in het schrijfwerk pas na de suggestie om het geheel in twee duidelijk herkenbare delen op te splitsen. Ook professor Brunia wil ik hartelijk danken voor de interesse in mijn (grotendeels niet-psychofysiologische) onderzoekswerk, het secure leeswerk, en het commentaar op de eerdere versies van dit proefschrift. Alle opmerkingen waren even waardevol als leerzaam, waardoor het vooral een beter afgerond geheel vormt.

Op deze plaats ook een belangrijk woord van dank richting het Epilepsiecentrum (en Slaap/Waakcentrum) Kempenhaeghe. De samenwerking met de afdeling Klinische Neurofysiologie was uitstekend, en een uitkomst qua locatie. Alle hulde voor de fijne contacten, de praktische en kritische houding, en het uitgevoerde werk: Johan Arends, Laurel Beecher, Guus Declerck, Gerard van Erp, Paul Griep, de technische staf, en de zeer kundige EEG-laboranten (met name Erna en Joke bij het onderzoek in II-4). Mede door de tijd op Kempenhaeghe is mijn belangstelling voor signaalverwerking en automatisering in de klinische neurofysiologie gegroeid. Zonder jullie waren grote delen van dit proefschrift niet tot stand gekomen. Op deze plaats wil ik ook alle proefpersonen bedanken voor hun wezenlijke bijdrage; zonder meetgegevens geen onderzoek!

Een andere bijzondere samenwerking vond plaats in het kader van de Europese projecten IMPROVE en IBIS. De meetings waren warm en inspirerend; nieuwe ideeën

ontstonden vooral dankzij de informele sfeer. In relatie tot dit proefschrift wil ik speciaal noemen: Bob Ghosh, Mark van Gils, Ilkka Korhonen, Pamela Prior. Thanks for the cooperation, your interest in my work, thank you for all the shared ideas.

Koen Böker wil ik bedanken voor het behulpzaam zijn bij het 'uitlenen' van zijn data, alsmede voor zijn commentaar op de eerste versie van het hoofdstuk II-3. Ook de andere Tilburgse collega's hartelijk dank voor het samen discussieren, het begeleiden van afstudeerders: Geert van Boxtel en Greet van den Berg-Lenssen.

Terug naar de 'eigen' universiteit. Stagiairs, afstudeerders (met name Steven van Dijk en Marco Salden), en alle collega's van de groep Medische Elektrotechniek heel erg hartelijk bedankt voor de plezierige samenwerking, de leuke activiteiten, de interessante uitjes, en de gezellige koffie-uurtjes. Weg van de vakgroep, waar ik ruim zeven jaar heb gewerkt, is toch anders. Geert, Hans, Harrie, Herman, Sjef, Sjoerd, Wim, Yvonne, en Andriana, Bart, Bert, Eelco, Erik, Gulian, Harald, Johan, Luc, Marcel, Mark, Nicole, Pieter, Piet-Hein, Raymond, Rob, Wendela, Wim: bedankt!

Een van de belangrijkste bijdragen kwam van Margit Horsthuis, partner in 'food'-crime en andere belangrijke dingen. De inspiratie uit bijv. chocolade, toetjes, en taartjes niet alleen te hoeven ophalen was heerlijk; allerbelangrijkst was het dat je mij door de moeilijke perioden van het schrijven heen hebt 'gecoached'.

Ook de belangstelling van mijn ouders, broers, en andere familie, heeft, net als alle contacten in muzikale sfeer, bijgedragen aan het afronden van dit werk.

Curriculum Vitae

Maarten van de Velde was born in The Netherlands, December 12, 1965. He graduated at the Medical Electrical Engineering (EME) group at Eindhoven University of Technology (TUE), in 1991. His Master's project focused on the analysis of spectral features of electroencephalographic signals during anaesthesia. Subsequently, he has worked at TUE/EME on signal-processing software for the group's research on evoked potentials, and was part of the development team for standardised data acquisition in a large international clinical study. From 1994 to 1998 he conducted his Ph.D. project at the TUE, collaborating with the Dept. of Clinical Neurophysiology at Kempenhaeghe (Heeze, The Netherlands) and the European IMPROVE/IBIS projects (patient monitoring in ICU). This work focused on signal validation, artefact detection and quality estimation of spontaneous and evoked electroencephalographic brain activity, which resulted in the current thesis.

From 1998 to 1999 he was employed in the areas of functional design and object oriented software development at Origin IT-Services, The Netherlands. Currently, he is pursuing a career that combines professional software development and scientific research. Apart from biomedical engineering and neuroscience in general, his research interests include signal processing and modeling/simulation of (bio-)medical processes.

In his spare time, he enthusiastically plays the clarinet in a chamber orchestra and in the Edison Quintet 큐익, for which he also arranges new repertoire. Previously, he has played the clarinet and saxophone in numerous ensembles, ranging from small solo-performances to big-band and symphonic orchestras.

"What does that mean, Expiremental Proseedcake?" said Pooh.
"For I am only a Bear with Very Little Brain, and long words Bother me."

(after A.A. Milne)

Statements pertaining to the thesis "Signal Validation in Electroencephalography Research" by Maarten van de Velde

1. Artefacts in the EEG can be recognised from a context period of 1 minute.
   This thesis

2. a) Quality assessment of sensory evoked potentials by a computerised method performs at least as good as visual inspection.
   b) Quantitative assessment of the signal-to-noise ratio of evoked potentials adds significantly to further signal analysis.
   This thesis

3. The complexity of algorithm design for any analysis method increases proportional to the number of domain experts involved.

4. Finding an artefact in EEG research resembles finding an artefact in archaeology: the beginning marks the start of a systematic search for a complete model.
   (based on personal communication with P.J.M. Cluitmans)

6. The lack of a generally accepted format for clinical data recording hinders the development of wider applications for patient monitoring using the EEG.
   This thesis

7. The AWK scripting language is more useful for data processing than a spreadsheet program.

8. Data-fusion resolves diffusion and confusion.

9. 'Knowledge is power', but true learning is more than cyclic in(ter)ference.
   ('Imagination is more important than knowledge' – A. Einstein)

10. Often statistics are used like a drunken man uses lamp posts — for support rather than illumination.
    (Fortune cookie)

11. We are only consciously aware of changing sensory input or active mental thought. Therefore, the Zen-aspired 'timeless' higher level of consciousness cannot reach individual consciousness.

12. The musical importance of *embouchure* for playing a wind instrument is underscored by its relation to the physical harmony of smiling.

Eindhoven, 17 January 2000