

Event-Related Brain Potentials in Emotion Perception Research, Individual Cognitive Assessment, and Brain-Computer Interfaces

Dissertation

der Fakultät für Informations- und Kognitionswissenschaften
der Eberhard-Karls-Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Phys. Vladimir Bostanov
aus Sofia, Bulgarien

Tübingen
2003

Tag der mündlichen Qualifikation: 16. Juni 2004

Dekan: Prof. Dr. Ulrich Güntzer

1. Berichterstatter: Priv.-Doz. Dr. Boris Kotchoubey

2. Berichterstatter: Prof. Dr. Rolf Ulrich

For my father, Prof. Dr. Vesselin Bostanov, who taught me the principles of thermodynamics:

- 1. There is no free lunch – you always pay in some way (conservation of energy).*
- 2. You cannot be perfect – some part of your directed actions always dissipates in pure heat (increase of entropy).*

Contents

Introduction	9
1 Recognition of Affective Prosody: ERPs to Emotional Exclamations	15
1.1 Introduction	15
1.2 Experiments	18
1.2.1 Oddball	20
1.2.2 Priming	22
1.3 Methods	22
1.3.1 Participants	22
1.3.2 Stimuli	23
1.3.3 Procedure	23
1.3.4 Data Analysis	24
1.4 Results	25
1.4.1 Oddball	25

1.4.2	Priming	29
1.5	Discussion	31
2	Continuous Wavelet Measures of ERP Components	37
2.1	Introduction	37
2.2	Continuous Wavelet Transform	38
2.3	The total-average-CWT method	39
2.4	Results	41
2.4.1	Oddball	42
2.4.2	Priming	43
2.5	Discussion	44
3	ERP Assessment with the τ-CWT Method	47
3.1	Introduction	47
3.2	The τ -CWT Method	54
3.2.1	CWT	54
3.2.2	τ -CWT	55
3.2.3	Computational Algorithm	57
3.3	Experiments	59
3.4	Methods	61
3.4.1	Participants	61

<i>Table of Contents</i>	7
3.4.2 Stimuli	61
3.4.3 Procedure	61
3.4.4 Data Analysis	62
3.5 Results	67
3.5.1 Oddball	67
3.5.2 Priming	74
3.6 Discussion	80
4 The τ-CWT as a Pattern Recognition Method in BCIs	85
4.1 Introduction	85
4.2 BCI Study 1: P300 Speller	87
4.2.1 Experimental Procedure	88
4.2.2 Data Analysis and Results	91
4.3 BCI Study 2: Self-Regulation of SCPs	99
4.3.1 Experimental Procedure	99
4.3.2 Data Analysis and Results	101
4.4 Discussion	101
Summary and Conclusions	103
Bibliography	106

Introduction

PSYCHOPHYSIOLOGY has been growing increasingly important as an experimental science that reveals the intricate relationships between psychological and physiological processes (Birbaumer and Schmidt 2003). The peripheral psychophysiological responses – skin conductance, heart rate, the various electromyographic responses (EMG), etc. – constitute one of the three major components (physiological, behavioral, cognitive) of any modern theoretical model of emotion (Öhman and Birbaumer 1993). The investigation of the electrophysiological responses of the central nervous system (CNS) plays an ever more important role in the study of cognitive processes (Birbaumer and Schmidt 2003, pp. 492–503). The modern fundamental psychological research relies heavily on psychophysiological experimental work.

Apart from its importance for testing the theoretical models of psychology, psychophysiology has found various applications in psychological diagnostics, psychotherapy, forensic psychology. Biofeedback training (Birbaumer and Rief 2000) and the notorious lie detector (Lykken 1991) are only two of a plethora of examples of how psychophysiological methods can be adopted for practical purposes.

The most popular applications of psychophysiology have used peripheral measures. However, CNS electrophysiological responses and the methods of their measurement and assessment as applied diagnostic methods of cognitive functioning have been gaining increasing attention by neuroscientists and clinicians in the last decade. Event-related brain potentials (ERPs) are the most popular kind of brain responses

used in both fundamental and applied research. ERPs are extracted from the electroencephalogram (EEG) and comprise various positive and/or negative waves (components) of different duration that reflect the responses of the brain to changes (events) in the external or internal environment of the organism (Picton and Hilliard 1988). ERPs are time-locked with certain latency to these events, which in the most experimental paradigms are external stimuli of various complexity that are presented visually or acoustically to the participants.

For about 15 years ERPs have been used for assessment of cortical information processing in patients with severe disorders of consciousness, such as coma and vegetative state (Kotchoubey, Lang, Bostanov, and Birbaumer 2002). This technique possesses two advantages that make it particularly important for the diagnostics and prognosis of such patients. First, it has the highest possible temporal resolution and follows cortical information processing in real time. Second, the technique is very easy, so that ERPs can be recorded directly at a patient's bedside practically any time when necessary; the patient need not be especially prepared or moved somewhere.

To date, a lot of data has been accumulated that many coma patients continue to process information at different levels, as indicated by the ERP components P300 and the mismatch negativity (MMN) (Gott, Rabinovicz, and DiGiorgio 1991; Guerit, Verougstraete, Tourtchaninoff, Debatisse, and Witdoeck 1999; Kane, Moss, Curry, and Butler 1998; Mutschler, Chaumeil, Marcoux, Wioland, Tempe, and Kurtz 1996; Signorino, D'Acunto, Cercaci, Pietropaoli, and Angeleri 1997). The presence of these components in some 20–50% of coma patients, as well as the prognostic value of the ERP technique, is not questioned any longer (Fischer, Morlet, Bouchet, Luaute, Jourdan, Salord, et al. 1999; Fischer, Morlet, and Giard 2000; Kane, Curry, Butler, and Cummings 1993; Kane, Curry, Rowlands, Minara, Lewis, Moss, Cummins, and Butler 1996; Kane, Butler, and Simpson 2000). The same diagnostic procedures have been tested with completely paralyzed patients whose cognitive abilities cannot be expressed in their behavior. The results show that these cognitive abilities, ranging

from simple sine tone discrimination to speech perception, are grossly underestimated by the initial neurological diagnosis and that the ERP technique (together with other functional imaging techniques such as positron emission tomography (PET)) can provide a better picture of the patient's cognitive status (Kotchoubey, Lang, Bostanov, and Birbaumer 2002; Kotchoubey, Lang, Baales, Herb, Maurer, Mezger, Schmalohr, Bostanov, and Birbaumer 2001; Kotchoubey, Lang, Bostanov, and Birbaumer 2003; Laureys, Faymonville, Degueldre, Fiore, Damas, Lambermont, Janssens, Aerts, Franck, Luxen, Moonen, Lamy, and Maquet 2000).

All of the experimental and theoretical work presented in this dissertation has been inspired by the general idea of applying the ERP technique for practical purposes: cognitive diagnostics (Kotchoubey et al. 2002) and Brain-Computer Interfaces (BCI; Farwell and Donchin 1988; Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey, Kübler, Perelmouter, Taub, and Flor 1999) for paralyzed people. As it often happens, this markedly applied research has also produced some valuable general results, whose importance goes beyond the particular applications. Most of the results presented here, are described and discussed from the wider perspective of fundamental research rather than with regard to the clinical applications. The reason for this approach is that thus far, the developed methods have only been demonstrated with healthy persons and have not yet been tested enough (or at all) with patients.

Organization and Overview of the Dissertation

In Chapter 1, two new ERP paradigms are introduced, which were developed for the diagnostics of a particular cognitive function – the recognition of affective prosody, especially emotional voice quality. A new ERP component, the N300, is found to reflect this recognition process. The N300 is interpreted as analogical to the well known N400 response to semantically inappropriate words.

Chapters 2 and 3 address the important issue of ERP component detection and quantification mostly from the perspective of individual data assessment, which is crucial for any reliable ERP diagnostics. Chapter 2 shows how the Continuous Wavelet Transform (CWT) can be used in ERP data analyses. A novel assessment method, the total-average-CWT, is introduced and demonstrated on the ERP data acquired in the emotional prosody experiments presented in Chapter 1. The results show a clear superiority of the CWT method to the standard assessment methods.

In Chapter 3 a new ERP assessment method is introduced – the \mathfrak{t} -CWT, which is based on the CWT and Student's \mathfrak{t} -statistic. It is a largely improved version of the total-average-CWT method presented in Chapter 2. The \mathfrak{t} -CWT is demonstrated on two prototypical ERP paradigms – oddball and semantic priming – and its power is compared to the power of various other assessment procedures – area measurement and peak picking, Principal Component Analysis (PCA), Discrete Wavelet Transform (DWT).

In Chapter 4, the detection and quantification method introduced in Chapter 3, the \mathfrak{t} -CWT, is applied in the classification of ERP trials for the purposes of BCI. In this application, the \mathfrak{t} -CWT is used as a general feature extraction method, which provides the optimal variables describing the pattern that best discriminates between the ERPs reflecting different cognitive processes. The method has been validated in the International BCI Classification Contest 2003, where it was a winner (provided best classification) on two ERP data sets acquired in two different BCI paradigms – P300 speller and Slow Cortical Potential (SCP) feedback. These results are presented in Chapter 4.

Acknowledgement

The author heartily thanks Dr. Boris Kotchoubey for the wise guidance and for the patience. He greatly appreciates the *freedom* he was granted to follow his intu-

itions and ideas, the encouragement for his endeavors and the understanding for his mistakes.

A Note Added after Submission of the Dissertation

The results of the studies presented in this dissertation were reported in three journal articles one of which is still in preparation and two were published after the submission of the dissertation. The studies on affective prosody described in Chapter 1 and the total-average CWT method described in Chapter 2 were reported in an article published in March 2004 in *Psychophysiology* (Bostanov and Kotchoubey 2004). The BCI applications of the τ -CWT method described in Chapter 4 were reported in a paper for the Special Issue on Brain-Machine Interfaces of *IEEE Transactions on Biomedical Engineering* from June 2004, where the results of the BCI Competition 2003 were published (Bostanov 2004). The applications of the τ -CWT in the ERP assessment and diagnostics described in Chapter 3 were reported in an article currently still in preparation, which will be submitted to *Clinical Neurophysiology*.

Chapter 1

Recognition of Affective Prosody: ERPs to Emotional Exclamations

The affective state of a speaker can be identified from the prosody of her/his speech. Voice quality is the most important prosodic cue for emotion recognition from short verbal utterances and nonverbal exclamations, the latter conveying pure emotion, void of all semantic meaning. In the study presented in this chapter, two context violation paradigms, passive oddball and priming, were adopted for the study the ERPs reflecting this recognition process. A new negative wave, the N300, was found in the ERPs to contextually incongruous exclamations. This component was interpreted as analogical to the well known N400 response to semantically inappropriate words. The N300 appears to be a real-time psychophysiological measure of spontaneous emotion recognition from vocal cues, which could prove a useful tool for the examination of affective-prosody comprehension for the purposes of both fundamental psychological research and applied clinical diagnostics.

1.1 Introduction

IN THE BEGINNING was the Scream and the Scream became a Word.¹ This may well be the evolution theorist's version of the story about the genesis of language. All birds and mammals have a large repertoire of vocal and bodily signals (displays) to express their motives and emotions (Darwin 1872; Scherer and Kappas 1988).

¹In analogy with: "In the beginning was the Word, and the Word was with God, and the Word was God." John 1.1. *The New Oxford Annotated Bible*. New Revised Standard Version, Oxford University Press, New York (1991).

Primates show some elements of semantic communication, uttering for instance, different cries of fear in the face of different kinds of danger (Seyfarth, Cheney, and Marler 1980). Chimpanzees are even capable of learning some elementary propositional language and rudimentary abstract thinking (Premack 1971). Yet humans are the only speaking creatures. Even they, however, use nonverbal utterances rather than words, to express intense or sudden emotions (Goffman 1978). Humans have inherited such screams, squeaks, moans, growls, etc., termed "sounds of nature" (Naturlaute), "interjections", "emotional vocalizations", "response cries" or "affect bursts", from their ancestors in evolution and these sounds, together with facial expressions and bodily gestures, comprise the most natural and ancient language of emotion communication whose expressiveness no words can ever achieve (Scherer 1994; Scherer 1985). These behavioural patterns seem to be quite similar in humans and lower primates – naive human listeners can identify correctly some basic emotions from monkey vocalizations (Leinonen, Linnankoski, Laakso, and Aulanko 1991; Linnankoski, Laakso, Aulanko, and Leinonen 1994).

Although nonverbal affective exclamations are used relatively rarely in human everyday social communication, spoken language still possesses a variety of nonverbal means to express emotions (Banse and Scherer 1996). These paralinguistic features make up the *affective*, or *emotional prosody* of speech². The ability to identify the speaker's emotions from prosodic cues is a distinct cognitive function, which is often impaired in neurological patients with right brain damage (RBD) (Ross, Thompson, and Yenkosky 1997). Neuropsychological research has shown that certain right hemispheric lesions may even cause complete loss of this ability (*sensory aprosodia*) without any significant impairment of other cognitive functions (Ross 1981). A double dissociation has been found between comprehension of emotional prosody and identification of affective facial expressions (Adolphs and Tranel 1999; Anderson and Phelps 1998). The former function is associated with the right posterior sylvian cortices (Ross, Orbelo, Burgard, and Hansel 1998), and the latter – with the amygdala

²Affective prosody stands in opposition to *linguistic prosody*, which serves semantic purposes, e.g. discriminating an assertion from a question.

(Adolphs, Tranel, Damasio, and Damasio 1994). Emotional-prosody comprehension is a cognitive function of high clinical importance. Besides in RBD, it has been found impaired in various neurological and psychic disorders: Parkinson's disease (Breitenstein, Daum, and Ackermann 1998), Huntington's disease (Speedie, Brake, Folstein, Bowers, and Heilman 1990), schizophrenia (Ross, Orbelo, Cartwright, Hansel, Burgard, Testa, and Buck 2001), alcoholism (Monnot, Nixon, Lovallo, and Ross 2001), depression (Emerson, Harrison, and Everhart 1999), attention-deficit and hyperactivity disorder (Manassis, Tannock, and Barbosa 2000), alexithymia (Lane, Sechrest, Riedel, Weldon, Kaszniak, and Schwartz 1996).

Unfortunately, the term "prosody" is often understood in a narrower sense, referring to the *suprasegmental* features of speech defined only on a large time scale, that is, pitch and loudness contours and speech and articulation rates (Crystal 1969; Breitenstein, Daum, and Ackermann 1998). In this sense it does *not* include *voice quality*, which is defined on a small time scale and is acoustically described by the spectral characteristics (timbre) of the vocal sound signal, as well as by rapid, possibly aperiodical changes in frequency which are not perceived as inflections, but rather as specific vibrations or harshness of the voice (Laver 1980). Clearly, the shorter an emotional utterance (verbal or nonverbal) is, the less prosodic features are present in the narrow sense of the term, and the more important the voice quality is for affect recognition. Thus, it is reasonable to hypothesize that short exclamations convey emotion predominantly by voice quality. Furthermore, Scherer (1986) suggested that voice quality is the key to the differentiation of emotions in affective speech.

On the other hand, the term "sensory aprosodia", as introduced by Ross (1981), denotes impaired comprehension of *all* affective components of language, including emotional voice quality. RBD patients were found equally impaired in affect recognition from sentences and single-segment nonverbal utterances ("aaaahhhhhh") as tested by the Aprosodia Battery (Ross, Thompson, and Yenkosky 1997). This finding suggests that they missed all affective prosodic cues, including voice quality,

which is probably the main medium of emotion in short unisegmental vocalizations. There is apparently some inconsistency in definitions, which is perhaps one reason why voice quality has received so little attention from neuroscientists.

Phonetics defines voice quality in terms of *settings*. One possible definition of a setting is: a long-term average configuration of the vocal organs, biasing all speaker's phonation and articulation for a certain time period (Laver 1980). Thus settings cause audible variations of those acoustical parameters, which constitute the physical description of voice quality. Settings can be controlled either deliberately by the speaker, or reflexively by physiological changes (muscle tension, etc.) triggered by his or her affective state (Laver 1980; Scherer 1986; Scherer 1994). In the latter case, we can hypothesize that the same settings are at work in nonverbal and verbal utterances and that emotional vocalizations and emotional speech share the same or similar voice quality and, possibly, other prosodic features as well.

In conclusion we can summarize the three most important features of nonverbal emotional exclamations. First, they convey the speaker's affective state much better than words do. Second, they most probably abstract some key prosodic features, primarily the voice quality, of (verbal) emotional speech. And third, the underlying physiological mechanisms are similar in all primates.

1.2 Experiments

Event-related brain potentials (ERP) have proved extremely useful in studying normal (Kutas and Hillyard 1980) and impaired (Hagoort, Brown, and Swaab 1996) semantic speech comprehension, as well as the processing of linguistic prosody (Steinhauer, Alter, and Friederici 1999). ERP components of latency from about 200ms to 500ms are perfectly suited for testing affect recognition from short, one-syllable verbal or nonverbal utterances, because the voice quality, which can be regarded as the main medium of emotion in this case, is present and perceivable from the

very onset of a vocalization. This allows for good time locking and, consequently, small variance in ERP component latencies. However, only few studies have employed ERPs to investigate recognition of emotional voice quality in particular and affective prosody in general. Twist, Squires, Spielholz, and Silverglide (1991) investigated ERPs to emotional prosodic stimuli in an oddball task. The participants listened to one-syllable words spoken with neutral (frequent stimuli) and surprised (rare, target stimuli) intonation. They were instructed to press a button on the occurrence of the rare, emotional stimuli. The P300 ERP component to the targets had a diminished amplitude and delayed latency in RBD patients compared to left BD patients and healthy controls. Erwin, Van Lancker, Guthrie, Schwafel, Tanguay, and Buchwald (1991) tested emotion recognition in autistic participants and found a surprisingly normal P300 response to the rare targets in an oddball task with happy and angry prosodic stimuli. However, since only two different stimuli were presented in the latter experiment, participants might have discriminated them merely by their physical differences and occurrence frequencies (see below).

Other researchers addressed lateralization issues in affective prosody processing. Erhan, Borod, Tenke, and Bruder (1998) investigated ERPs to emotionally spoken nonsense syllables in a dichotic listening task. Pihan, Altenmuller, and Ackermann (1997) and Pihan, Altenmuller, Hertrich, and Ackermann (2000) studied direct-current (DC) components of the electroencephalogram (EEG) acquired during listening to sentences with different emotional prosody.

The main purpose of the study presented in this chapter was to find a reliable, real-time psychophysiological measure of *immediate* and *spontaneous* recognition of affective prosody, specifically emotional voice quality, presumably captured in its purest and prototypical form, as extant in exclamations. Using nonverbal vocalizations as stimulus material should also prevent the production of any linguistic prosody on the part of the speaker, which may confound the subsequent acoustical analysis of affective prosody (Leinonen et al. 1997). On the part of the listener, it forestalls any semantic processing, which possibly interferes with emotion recogni-

tion. Moreover, as already mentioned, emotional exclamations are a language *per se* and hence deserve special interest in their own right.

An active oddball task (with an instruction defining the targets) has two important disadvantages: first, it tests discrimination rather than recognition, and second, discrimination could be achieved mainly on the basis of physical differences between, and occurrence frequencies of, the two acoustic signals without making much use of their emotional connotation. (However, the poorer performance of the RBD participants reported by Twist et al. (1991) suggests at least some affect discrimination along with physical discrimination in the active prosodic oddball task.) That is why, in this study, two other paradigms, based on *context violation* rather than on stimulus infrequency or novelty, were adopted for the investigation of emotion recognition. An important feature of meaningful stimuli is that they can build a context, which induces some expectations about the next-coming stimuli. For instance, when one hears the sentence: "He spread the warm bread with...", one expects the usual ending, "butter", and when the unexpected word "socks" comes instead, the cognitive processes related to the context violation are reflected by the N400 ERP component in one's EEG (Kutas and Hillyard 1980). This effect has been replicated many times with various kinds of visual and acoustic stimuli: written and spoken words and sentences (Bentin, McCarthy, and Wood 1985; Bentin, Kutas, and Hillyard 1993; Kutas and Hillyard 1980; McCallum, Farmer, and Pocock 1984), pictures of objects (McPherson and Holcomb 1999) and human faces (Jemel, George, Olivares, Fiori, and Renault 1999), environmental sounds (Van Petten and Rieffers 1995), etc..

1.2.1 Oddball

If in a *passive* oddball paradigm (without any specific task) two categories of stimuli are presented instead of two single stimuli, the rare stimuli may elicit an N400 rather than a P300 (Schlaghecken 1998). The stimuli must have some meaning, so that the

frequent stimuli generate a context and provoke some expectations, which are then broken by the occurrences of the deviant stimuli. By the same mechanism, an N400 to rare, incongruous, *non-target* stimuli can be found in an active oddball (Bentin, Mouchetant-Rostaing, Giard, Echallier, and Pernier 1999).

The first experiment of the present study was a passive oddball with emotional vocalizations. The category of the frequent stimuli comprised four exclamations of joy: "Yeeh!", "Heey!", "Wowh!" and "Oooh!". A single exclamation of woe: "Oooh!", served as a deviant stimulus. All five exclamations had the same occurrence frequency of 20%. The same vowel "o" was purposefully chosen for both the sad exclamation and one of the joyful exclamations – without an active instruction to assign one of the vocalizations as a target, there was no clearly defined rare stimulus. Thus the expressed emotion was the only feature distinguishing "Oooh!" (woe) as deviant. The context violation effect was readily discernable: Listening to the (randomized) stimulus sequence gave the impression of a man who was expressing his joy, but now and then unexpectedly uttered a sound of despair and deep sorrow. (Note the difference between this design and the standard passive oddball paradigm with only two meaningless sine tones presented in Chapter 3.)

The experiment still shared some of the shortcomings of a standard oddball paradigm. First, emotional differences were inevitably expressed by some physical parameters which left at least the theoretical possibility of "mechanical", rather than, or along with, emotion discrimination. Second, only two mutually converse emotions were included, which raised the question, whether woe was really spontaneously recognized, or it was merely inferred by discrimination, possibly based more on different arousal level than on valence (Banse and Scherer 1996). And third, since joy is a positive emotion and woe is a negative one, the question arose whether the ERP to the rare stimulus was not in fact valence-specific, and reflecting motivational evaluation rather than, or along with, recognition processes (Ito, Larsen, Smith, and Cacioppo 1998).

1.2.2 Priming

The second experiment was aimed to replicate the results of the first experiment and to answer the three open questions that were formulated above. A paradigm from the study of Van Petten and Riefelder (1995), which demonstrated the conceptual relationships between spoken words and environmental sounds, was adopted. In that study, an N400 was found to sounds preceded by inconsistent words, e.g. a sound of helicopter rotor (target) preceded by the word "dog" instead of "helicopter" (prime). In the priming experiment described in this chapter, 9 different spoken emotion names were presented as primes and 9 corresponding emotional vocalizations were presented as targets. The inconsistent combinations were: "joy"-[grief], "pleasure"-[rage], "surprise"-[disappointment], "disappointment"-[surprise], "grief"-[joy], "disgust"-[terror], "rage"-[fright], "fright"-[pleasure], and "terror"-[disgust]. (Square brackets denote an exclamation expressing the respective emotion.) By taking words rather than vocalizations as primes, any possible priming by physical features was forestalled. Including a number of different emotions was thought to engender more recognition and less simple discrimination. Finally, by constructing three types of inconsistent pairs: positive-[negative], negative-[positive], and negative-[negative], it is aimed to ensure that the elicited ERP response was not valence-specific.

1.3 Methods

1.3.1 Participants

There were 19 participants (10 male, 9 female, mean age \approx 26 years) in the oddball experiment and 29 participants (16 male, 13 female, mean age \approx 23 years) in the priming experiment. According to self report, all participants were right-handed, their mother tongue was German and none of them had a history of mental or neuro-

logical diseases. Most of them were students and all were paid for their participation (15DM/hour). None of the participants in the first experiment took part in the second one.

1.3.2 Stimuli

All stimuli were recorded digitally at 22.05kHz/16bit sampling rate. All exclamations were uttered by the first author – male, non-native German speaker (Bulgarian) – and the emotion names were spoken by a female, native German speaker. Neither of the speakers was a professional actor. Stimulus duration varied from 750ms to 870ms in the oddball and from 630ms to 980ms in the priming experiment. In the oddball, "Oooh!" (woe) and "Oooh!" (joy) had approximately the same duration \approx 840ms. In the priming, the longest emotion name, "disappointment" (German: "Enttäuschung"), lasted approximately 790ms.

1.3.3 Procedure

In the oddball experiment, each of the 5 exclamations occurred 60 times in a randomized sequence of 300 trials, which were presented at a constant rate of 1 stimulus per 1.1s. The word-exclamation pairs in the priming experiment were presented in a randomized sequence of 108 pairs, where each of the nine exclamations was repeated 12 times – 6 times preceded by the correct emotion name and 6 times preceded by a wrong name. The stimulus onset asynchrony between a word and an exclamation was 1s and the presentation rate was 1 pair per 3s. Digitized EEG (time resolution: 2ms/step (500Hz), voltage resolution: $0.1678\mu V$ /step) was continuously recorded from 9 scalp positions according to the 10-20 system: Fz, Cz, Pz, F3, F4, C3, C4, P3, P4. In the oddball, all electrodes were referenced to the linked mastoids; in the priming the nose was used as reference and subsequent off-line rereferencing to the mastoids did not yield any considerable change in results. Electrical eye activity

was recorded by bipolar acquisition from the following sites: FP2 and a site below the right eye – for vertical eye movements and eye blinks; F7 and F8 – for horizontal eye movements.

In both experiments, the only instruction given to the participants was to listen attentively. At the end of each experiment the participants were asked to describe in detail what they had heard (verbal reports).

1.3.4 Data Analysis

The acoustic features of the oddball stimuli were analyzed with the short-time Fourier Transform (STFT). The EEG was filtered on-line (bandpass: 0.1Hz-70Hz, notch: 50Hz). EEG epochs were created off-line. Eye-blink and eye-movement artefacts, both vertical and horizontal, were corrected off-line by a computerized procedure (Miller, Gratton, and Yee 1988; Gratton, Coles, and Donchin 1983). Trials with voltage exceeding $90\mu V$ in any EEG channel were considered artefacts and were excluded from further analysis. Averages for each participant as well as grand averages over all participants were calculated for each experimental condition using a 100ms prestimulus baseline.

All ERP components were detected by visual inspection of the grand average waveforms and quantified with the classical *area* and *peak* measures obtained from the individual-participant averages. The area measure was computed as the average voltage in a time window determined by visual inspection of the grand average waveforms. The peak measure was defined as the maximum voltage in the same time window for positive ERP components and as the minimum value for negative waves.

The ERP components were also quantified by *wavelet* measures. The corresponding theory and the obtained results from participant-average and *single-trial* analyses are presented in Chapter 2.

All ERP measures were tested for statistical significance by three-way analyses of variance (ANOVA) with factors: condition(2) \times FCP(3) \times LMR(3), where FCP abbreviates: frontal/central/parietal, and LMR abbreviates: left/mid/right. The experimental conditions were "Oooh" (joy)/"Oooh!" (woe) in the oddball (the other joyful exclamations were not considered) and consistent/inconsistent in the priming. FCP and LMR were taken as within-participant or within-trial factors, whereas the experimental condition was taken as a within-participant factor for the participant-average ERP measures. Left/right and frontal/parietal asymmetries were studied by linear contrasts; predominance at the vertex Cz was studied by quadratic contrasts. (With threefold topography factors, a main effect or an interaction with the condition factor may reflect either an asymmetry or a central deviation, or both; contrasts specify which of the three is true.)

The area and peak computations were performed with MATLAB 6.0, and the ANOVAs – with SPSS 11.0.

1.4 Results

For the sake of brevity and clarity, only significant ANOVA probability values (p -values) and no effect sizes are reported.

1.4.1 Oddball

Verbal reports

All participants identified correctly the emotions expressed by the presented exclamations. "Yeeh!", "Heey!", "Wowh!" and "Oooh!" (joy) were described as "cheerful", "joyful", "merry", expressing "pleasure", "positive surprise", "admiration". "Oooh" (woe) was described as "mourning", "suffering", expressing "grief", "des-

peration”, ”pain”.

Acoustical analysis

Figure 1.1 shows that emotional voice quality, and more precisely emotional timbre, was the only prosodic feature which distinguished ”Oooh!” (woe) from the joyful exclamations.

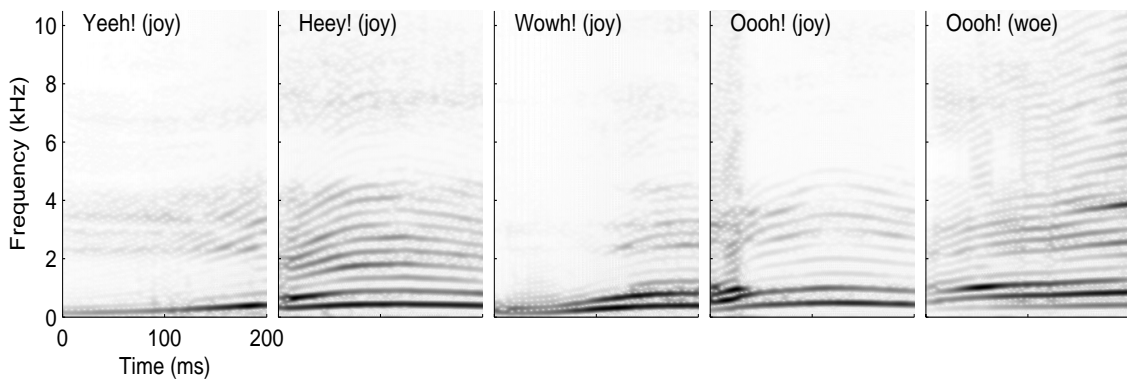


Figure 1.1: STFT spectrograms of the first 200ms of each of the exclamations in the oddball experiment. Darker shades indicate higher intensities. The first dark stripe from the bottom represents the fundamental frequency f_0 , the second represents the first harmonic f_1 , etc. Two principal spectral parameters discriminating between the emotions are readily discernable: first, the energy distribution between f_0 and f_1 is in favour of f_0 for joy and clearly in favour of f_1 for woe, and second, the proportion of energy contained in frequencies above f_{11} ($\approx 4.5\text{kHz}$) is also fairly larger for woe than for joy. Other acoustic parameters – f_0 -mean (pitch), f_0 -contour (intonation), amplitude envelope (accentuation), – vary among all exclamations and have approximately mean values for ”Oooh!” (woe).

ERP measures

The ERP components found in the waveforms can be seen in Figure 1.3. The direct comparison between ”Oooh!” (woe) and ”Oooh!” (joy) at Fz is shown in Figure 1.2.

The N300 amplitude was significantly larger for ”Oooh!” (woe) than for ”Oooh!” (joy) by the main effect of condition with the area measure measure ($p = 0.03$) but not with the peak measure. No significant topographic effects (as analysed by

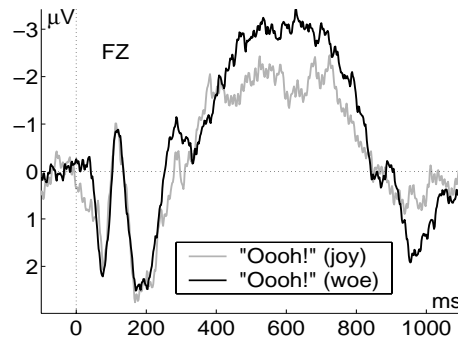


Figure 1.2: Grand average waveforms at Fz for "Oooh!" (woe) and "Oooh!" (joy) in the oddball experiment. The ERP to "Oooh!" (joy) is almost the same as the average of all 4 joyful exclamations (see Figure 1.3).

quadratic contrasts) were found for the N300 component with either the area or the peak measure (but see Chapter 2).

The N600 amplitude was significantly larger for "Oooh!" (woe) than for "Oooh!" (joy) at frontal sites by the condition \times FCP interaction with both participant-average measures: area ($p = 0.03$) and peak ($p = 0.04$). No significant topographic effects (as analysed by quadratic contrasts) were found for the N600 component with either the area or the peak measure (but see Chapter 2).

The P1000 amplitude was significantly larger for "Oooh!" (woe) than for "Oooh!" (joy) at frontal sites by the main effect of condition with the peak measure ($p = 0.04$), but not with the area measure, as well as by the condition \times FCP interaction with area measure ($p = 0.01$), but not with the peak measure.

No significant lateralization (linear-contrast effects of LMR or LMR \times condition) was found for any of the components (even with the single-trial wavelet measures, see Chapter 2).

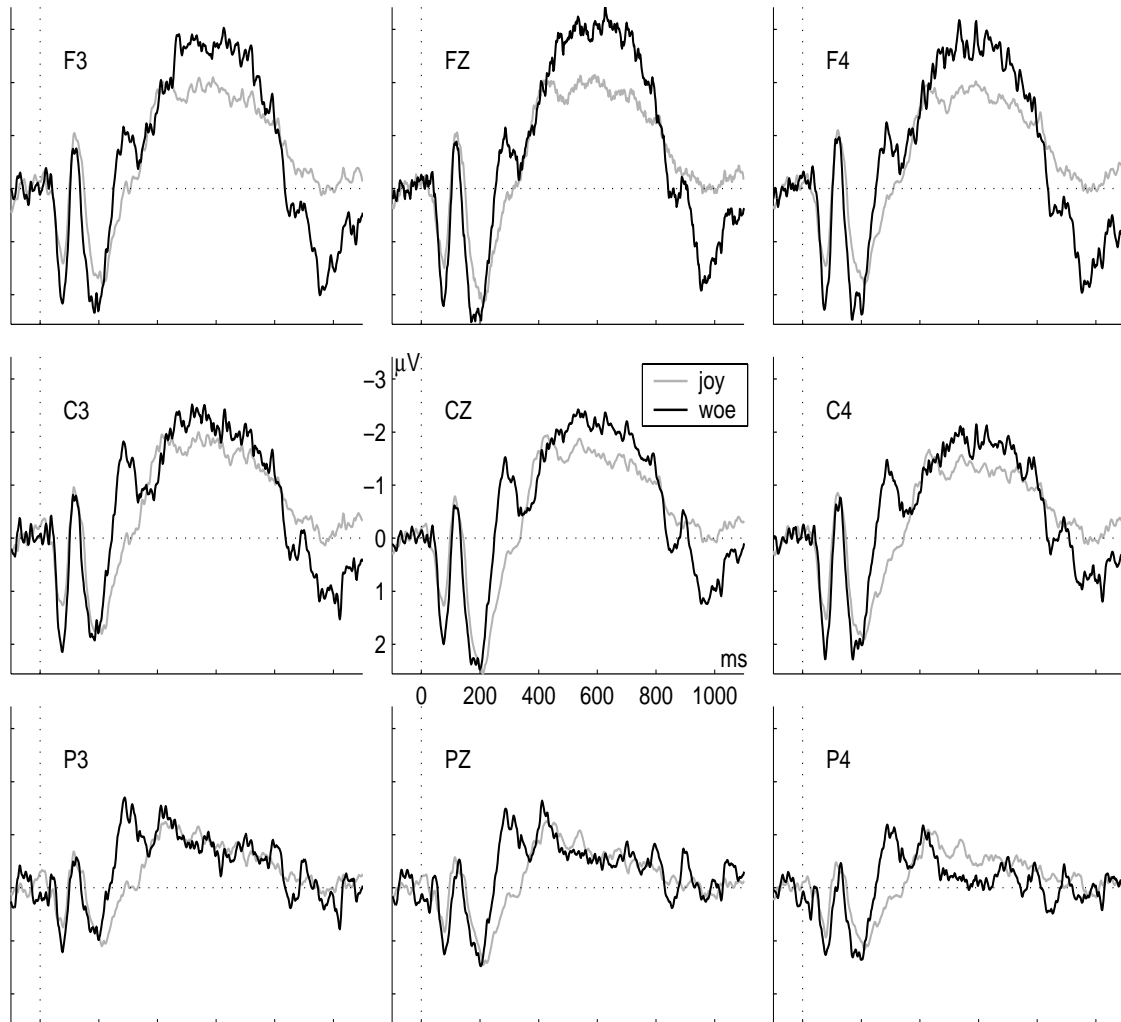


Figure 1.3: Grand average waveforms for the 4 joyful exclamations: "Yeeh!", "Heey!", "Wowh!", "Oooh!" (all 4 averaged together), and for the doleful one: "Oooh!", in the oddball experiment. Besides the P50/N100/P200 complex, three further prominent ERP components can be distinguished: N300 – an early negative wave with peak latency ≈ 300 ms, duration ≈ 100 ms, peak-to-peak amplitude $\approx 1.5\mu V$, and broad scalp distribution; N600 – a later, slower negative wave with latency ≈ 600 ms, duration ≈ 400 ms, and centro-frontal distribution; P1000 – a late positive wave with latency ≈ 1000 ms, duration ≈ 150 ms, and centro-frontal distribution.

1.4.2 Priming

Verbal reports

All participants in the priming experiment identified consistent and inconsistent word-exclamation pairs and reported that each of the 9 emotional words was followed some times by the corresponding vocalization and some times by (an) inappropriate one(s).

ERP measures

The ERP components found in the waveforms can be seen in Figure 1.4.

The N300 amplitude was significantly larger for the inconsistent vocalizations by the main effect of condition with the peak measure ($p = 0.03$), but not with the area measure. No significant topographic effects (as analysed by quadratic contrasts) were found for the N300 component with either the area or the peak measure (but see Chapter 2).

No significant effects were found for the P500 component with either the area or the peak measure (but see Chapter 2).

No significant lateralization (linear-contrast effects of LMR or LMR \times condition) was found for any of the components (even with the single-trial wavelet measures, see Chapter 2).

The significance of the N300 effect was also tested separately for each of the three kinds of word-exclamation pairs: positive-[negative] ("joy"-[grief], "pleasure"-[rage], "surprise"-[disappointment]), negative-[positive] ("grief"-[joy], "fright"-[pleasure], "disappointment"-[surprise]), and negative-[negative] ("disgust"-[terror], "rage"--[fright], "terror"-[disgust]). No significant effects were found with either the area or the peak measure (but see Chapter 2).

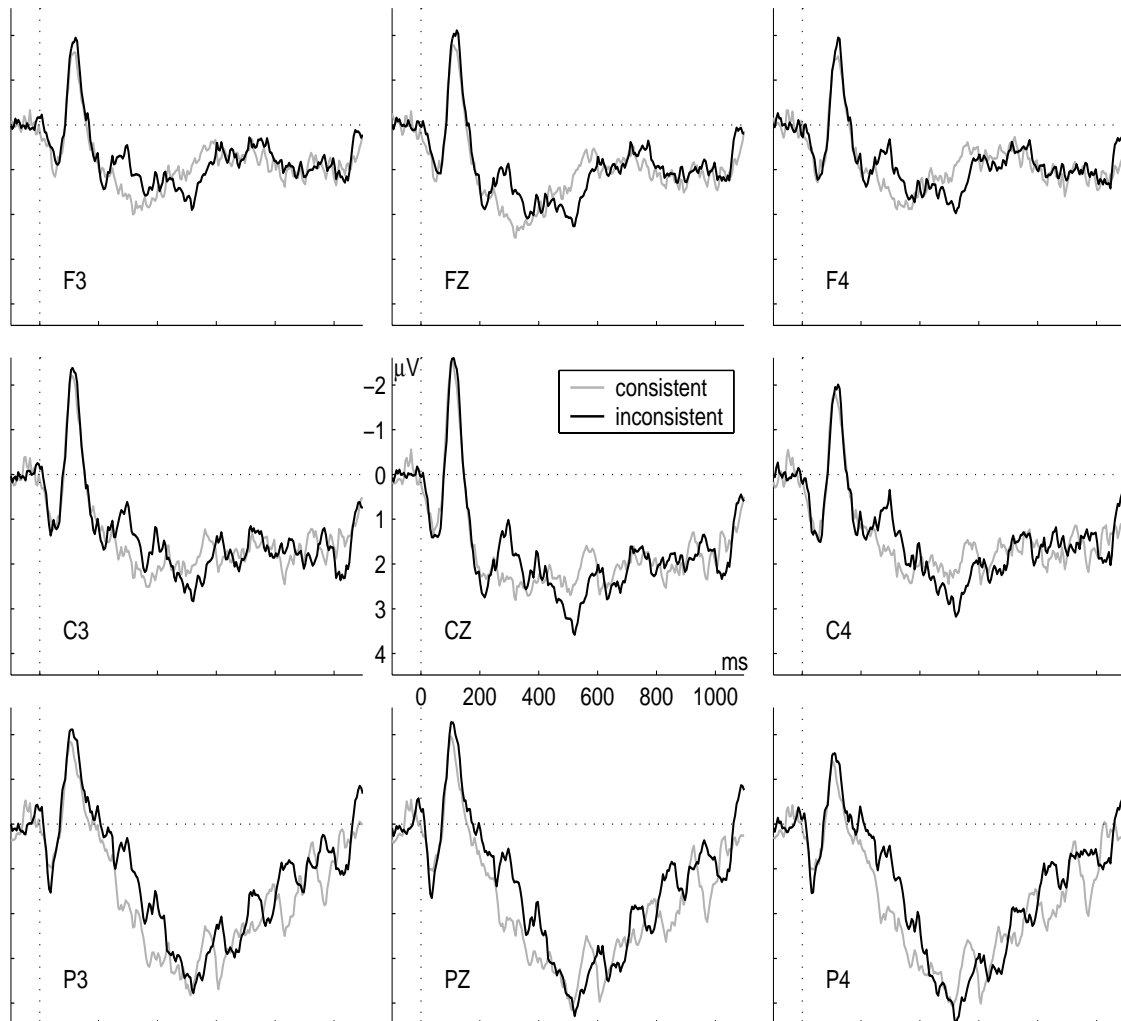


Figure 1.4: Grand average waveforms for the consistent and inconsistent exclamations in the priming experiment. Besides the P50 and the N100, two further ERP components can be seen: an N300 very similar to the one elicited by the sad exclamation in the oddball, and a P500 - a positive wave with latency $\approx 500\text{ms}$ and duration $\approx 150\text{ms}$.

1.5 Discussion

The verbal reports confirmed that participants identified all emotions correctly. The results of the acoustical analysis support the hypothesis that nonverbal vocalizations convey emotion primarily by voice quality.

An early negative ERP component, the N300, was elicited both by the deviant emotional exclamations in the passive oddball paradigm and by the inconsistent affect vocalizations in the priming experiment. The most plausible interpretation of the N300 is an early, well time-locked N400-like wave, which reflects cognitive processes related to context violations and broken expectations and thus implies emotion recognition from the prosodic features of the nonverbal stimuli. As discussed above (see Experiments), the finding of an N400-like response to the meaningful infrequent stimuli in a passive oddball paradigm is hardly surprising – apparently, the frequent joyful exclamations served as a context background for the incongruous, rare, doleful exclamation. The occurrence of the N300 in the ERP to both positive and negative inconsistent exclamations contradicts the hypothesis of valence specificity of this component. The short latency and duration of the N300 indicate rapid affect recognition based primarily on the voice quality of the first 100-150ms of the exclamations. This result is consistent with an earlier finding of emotion recognition from extremely short voice samples with duration $< 100\text{ms}$ (Pollack, Rubenstein, and Horowitz 1960).

The most similar design to the present priming experiment among the earlier N400 studies is the above-cited study of environmental sounds by Van Petten and Rieffelder (1995) (see Experiments). The N400 reported there had the same onset latency $\approx 200\text{ms}$, but a much longer duration ($\approx 700\text{ms}$) than the duration of the N300 ($\approx 100\text{ms}$) found by us. This may be due to a large latency variance of the negative wave elicited by incongruous sounds. As it has already been pointed out above, emotional exclamations are apparently recognized very rapidly by their onset voice quality, while other environmental sounds may be impossible to identify from

their onset acoustics. This was surely true for at least one of the sounds described by Van Petten and Rheinfelder (1995): "horse hooves striking pavement". (At least two consecutive strokes need to be processed in order to classify this stimulus as congruous or incongruous and perhaps even more are necessary for correct identification – nowadays this particular sound is not so commonplace in our environment). Such hard-to-recognize sounds would elicit a brain response with much later onset and peak latency than easy-to-recognize sounds like affective vocalizations. It is reasonable to suppose that the resulting ERP component in the average waveform of all different environmental sounds (which in the study by Van Petten and Rheinfelder (1995) were 99) would have onset latency determined by the easiest-to-recognize stimuli, whereas its duration would be determined by the hardest-to-recognize stimuli. The ERP peak latency would be determined by the most frequent stimulus recognition latency. Furthermore, the variance in the recognition point across stimuli is not the only possible source of latency variance – it is plausible that there was considerable variance in identification time across participants. With 99 different sounds presented, it is unlikely that all of them were equally familiar to all participants. Apparently, the opposite holds for emotional exclamations – the short duration of the N300 suggests that all affective stimuli presented in the present experiments were identified equally rapidly by all participants.

A more careful visual inspection of the ERP to incongruous environmental sounds in the study of Van Petten and Rheinfelder (1995) may give rise to another interesting interpretation. Actually, *two* separate negative waves can be seen in the waveform: a narrow peak at approximately 300ms, closely resembling the N300 found in the present study, and an immediately following later wave with much longer duration, like the N400 to incongruous words. When words were presented second in pair, preceded by related or unrelated sounds, the resulting N400 showed *no* such morphological differentiation. However, the authors reported and discussed in both cases only one, undivided N400 with onset latency ≈ 200 ms and long duration. Assuming that the ERP to incongruous environmental sounds really comprises two

distinct components, the following possible interpretation may be suggested: The early component is actually an N300 like the N300 to incongruous exclamations found in the present study. It reflects the rapid recognition of the easy-to-recognize sounds from their onset acoustics. The second wave reflects a more elaborate recognition process analysing the complicated temporal structure of the sound signals. In the case of simple emotional exclamations there is nothing more left to learn from the signals after the initial recognition of the expressed affect, and hence there is no second process and no second ERP component. Such interpretation is based on the notion that emotional exclamations are qualitatively different from most other environmental sounds – while exclamations usually convey a *static picture* of the speaker’s instant affective state, most other sounds represent some *time evolution* of the physical world, and every next sound fragment may potentially add new information about some changes in the environment. On the contrary, further processing of an emotional exclamation does not yield any information about the evolution of the affective state, but only unravels more subtle nuances. It should be noted also that the sound signals used in the study of environmental sounds had much longer duration = 2500ms than the emotional stimuli in the present study.

The interstimulus variation in ERP component latencies was systematically investigated for the case of spoken words by Woodward, Owens, and Thompson (1990): the peak latency of the negative wave ranged from 366ms to 670ms with a mean latency of 474ms. This value corresponds closely to the peak latency (≈ 460 ms) of the N400 to semantically incongruous last words in spoken sentences (McCallum et al. 1984). Interparticipant N400 latency variability is also high even in the case of the N400 to visually presented words. As Moreno, Federmeier, and Kutas (2002) reported in a recent study, the N400 peak latency ranged from approximately 300ms to 400ms even in participants of the highest English proficiency, thus resulting in an N400 latency ≈ 350 ms in the grand average ERP. Comparing these and other semantic priming results to the outcomes of the present experiments, we can summarize that all nonverbal affective vocalizations that were presented were recognized

by all of the participants approximately 150ms earlier than an average word takes to be recognized by an average person. If we assume that the recognition of emotional exclamations is also representative for the recognition of affective prosody of verbal material in general (see Introduction), and more specifically for the identification of onset emotional voice quality of spoken words, then we can conclude that *emotion is grasped considerably faster than meaning*. This conjecture can be proved (or disproved) by an experiment in which emotionally spoken words are presented instead of nonverbal vocalizations (see below). This is a very interesting point, because if the conjecture is true, it will mean that affect recognition could potentially bias or override further semantic processing, especially in case of contradiction between prosody and semantics (inadequate affect). For instance, communicating in a cheerful voice the news of the death of a beloved one, may result in distorting or ignoring the meaning of the message on the part of the listener. This notion may appear to be similar to the well known theory of fast emotional response based on rapid but superficial stimulus recognition by the amygdala, preceding the slower but precise recognition by the cortex (LeDoux 1993). Our concept, however, is essentially different. As already mentioned (see Introduction), a great body of evidence exists, which indicates that affective prosody is processed by the cortex and not by the amygdala. Furthermore, the present study is about *emotion recognition* from a stimulus and *not* about *emotional response* to that stimulus. Although an emotional response implies some minimal stimulus recognition, the opposite is not true: recognition of affective prosody does not necessarily imply emotional experience on the part of the listener: Imagine the extreme case of a psychopath who hears someone crying with terror: "Help!". If the above conjecture is correct, he will first recognize that the crying person is terrified and only then (just fractions of the moment later) will he understand that the person needs help. However, both discoveries will most likely leave him impassive.

In the study of environmental sounds, the reported N400 to incongruous sounds was significantly larger at the left than at the right scalp sites; the N400 to ver-

bal stimuli, on the contrary, showed a reverse, right-sided asymmetry (Van Petten and Rieffers 1995, see the works cited therein and the authors' discussion about "paradoxical lateralization"). In the present experiments, no significant lateralization effect was found in the scalp distribution of the N300 to emotional exclamations. This may indicate that different neural substrates are involved in the cognitive processing of affective vocalizations versus other environmental sounds. However the spatial resolution and source localization ability of an EEG from 9 scalp positions are too low to adequately address lateralization issues and these were not in the focus of interest of the present study.

An open question remains as to what extent the affect recognition process was facilitated by learning during the experiments. One possible scenario is that in the first, say, 10-20 trials the participants identified emotions by all prosodic features manifested by the whole vocalizations, and then quickly learned to discriminate the already familiar stimuli just from their onset. This problem stems from the great number of repetitions of the same few different exclamations. A further question is whether emotion recognition (or discrimination) from the first 100-150ms of the vocalizations was based solely on voice quality, or it also utilized intonation and accentuation patterns that could be perceivable even from such a short time interval and could be discernable but not definable from the spectrogram. Apart from a more elaborated acoustical analysis, there are at least three other ways to tackle these problems: first, by using a large variety of exclamations by many different people (different voices), thus avoiding repetitions; second, by using digitally manipulated sound signals derived from one and the same neutral vocalization by artificial selective variation of voice quality (Pihan et al. 2000); and third, using abridged vocalizations as stimulus material, e.g. the first 100ms (Pollack et al. 1960). Another important question is to what extent nonverbal vocalizations and emotional speech share the same prosodic features. Profound acoustic analyses are required to clarify this issue. Whatever the answer may be, however, the proposed abridgement technique can be utilized for the study of emotion recognition from ver-

bal utterances by presenting short word fragments, e.g. single vowels, as stimulus material.

The other ERP components that were found in the experiments demand further investigation to clarify their relation to emotion recognition. At this point, it should be noted that in the oddball paradigm, ERPs to different stimuli are compared to each other, while in the priming experiment, one and the same stimulus elicits different ERPs in two different experimental conditions. This might indicate that the ERP components that were found in the oddball but did not appear in the priming could possibly reflect physical differences between the stimuli, rather than their emotional meaning.

Chapter 2

Continuous Wavelet Measures of ERP Components

The classical measures, area and peak, used in Chapter 1 are the most popular estimators of ERP components. Although they are in standard use in ERP research, they are very rough estimators, which heavily rely on visual inspection of the waveforms, and are thus prone to experimenter bias. In this Chapter, a new ERP component detection and estimation method is introduced, which is based on the Continuous Wavelet Transform (CWT) and does not rely on visual inspection of the waveforms but allows more automated detection of peaks. At the same time it provides more precise estimation of ERP components and yields larger statistical difference effects than classical methods do.

2.1 Introduction

THE ASSESSMENT of the acquired EEG data is an important step in an ERP experiment. Especially in clinical applications, in which ERPs are used for cognitive diagnostics, it is vitally important to have a reliable assessment procedure, which would allow the clinician to make judgements at a high confidence level about the patient's status of neurocognitive functioning.

The purpose of the study presented in this chapter was to develop an estimation method based on the continuous wavelet transform (CWT) (Samar, Bopardikar, Rao, and Swartz 1999; Ende, Louis, Maass, and Mayer-Kress 1998), which would

reduce the variance of the obtained ERP measures, would make them suitable for single-trial assessments and would improve ERP component detection from average waveforms. Such procedure should be applicable in individual assessment, and in particular, in the diagnostics of severe neurological patients whose cognitive abilities cannot be expressed in their behavior (Kotchoubey et al. 2002). The method described in this chapter should be seen as a first step towards this goal. A mathematically and computationally more elaborate CWT method is presented in Chapter 3. The review of previous studies as well as the mathematical details of the CWT are also postponed for Chapter 3.

2.2 Continuous Wavelet Transform

CWT is a mathematical transformation which maps time curves onto smooth, two-dimensional surfaces called scalograms¹. Computationally the CWT is obtained from the cross-covariance of the ERP curve with a given template function called wavelet, which is systematically varied in width (scale) and position in time. The local extrema of the CWT give the template's scale and time positions of best match (highest covariance) with the ERP curve. Thus a peak in the curve is represented by a peak in its scalogram, where the two-dimensional position of the latter indicates the time position (component latency) and the width (component duration) of the former. In this sense, CWT is similar to some classical *template-matching* algorithms used in single-trial ERP analysis (Smulders, Kenemans, and Kok 1994). However, a crucial advantage of CWT is the special form of the wavelet template, which allows for optimal *scale separation* and hence better distinction of overlapping ERP components.

¹"CWT" denotes both the transformation and its result. As regards the second case, "CWT" and "scalogram" are further used interchangeably, as synonyms.

2.3 The total-average-CWT method

The total-average-CWT method introduced here was tested on the ERP data obtained from the experiments presented in Chapter 1. In the following, the procedure is described step by step.

CWTs were computed for each participant's average waveform as well as for each single trial and each electrode, using the "Mexican Hat" wavelet (Ende et al. 1998). Three different ERP measures were obtained from both participant averages and single trials: *area*, *fixed-wavelet* and *matched-wavelet*.

An area was computed as the average voltage in a time window determined by visual inspection of the grand average waveforms (see Chapter 1). The area can be seen as a kind of primitive template matching: computationally it is equivalent to the cross-covariance of each participant average and each single trial with a fixed rectangular template² (Figure 2.1) which is fitted visually to an ERP component (also detected visually) in the grand-average waveforms, without any additional fitting of the template's (window's) position and width to individual participant averages or single trials.

A wavelet measure was defined as a CWT value (the cross-covariance of the waveform with a wavelet) at certain time and scale according to the following procedure. First, a *total grand average* was calculated for each experiment. The total grand average included *all* trials of *all* participants of *both* experimental conditions subject to further statistical analysis. Then the CWT scalogram of the total grand average at Cz was computed. ERP components were identified as local extrema in this scalogram. This approach avoids the accumulation of chance involved in the detection of ERP components from difference waveforms (see Chapter 3 for details).

²Strictly speaking, this is only true under the assumption that the mean of an EEG curve is zero, which may be wrong for single (baseline-corrected) sweeps; however, this does not lead to loss of generality of our reasoning. In the case of a wavelet template, this assumption is unnecessary, because the mean of a wavelet is by definition zero.

The procedure is further illustrated on the example of the N300 wave found in the oddball experiment described in Chapter 1. The total average and its scalogram are displayed in Figure 2.1. The position of the N300 minimum was approximately: time

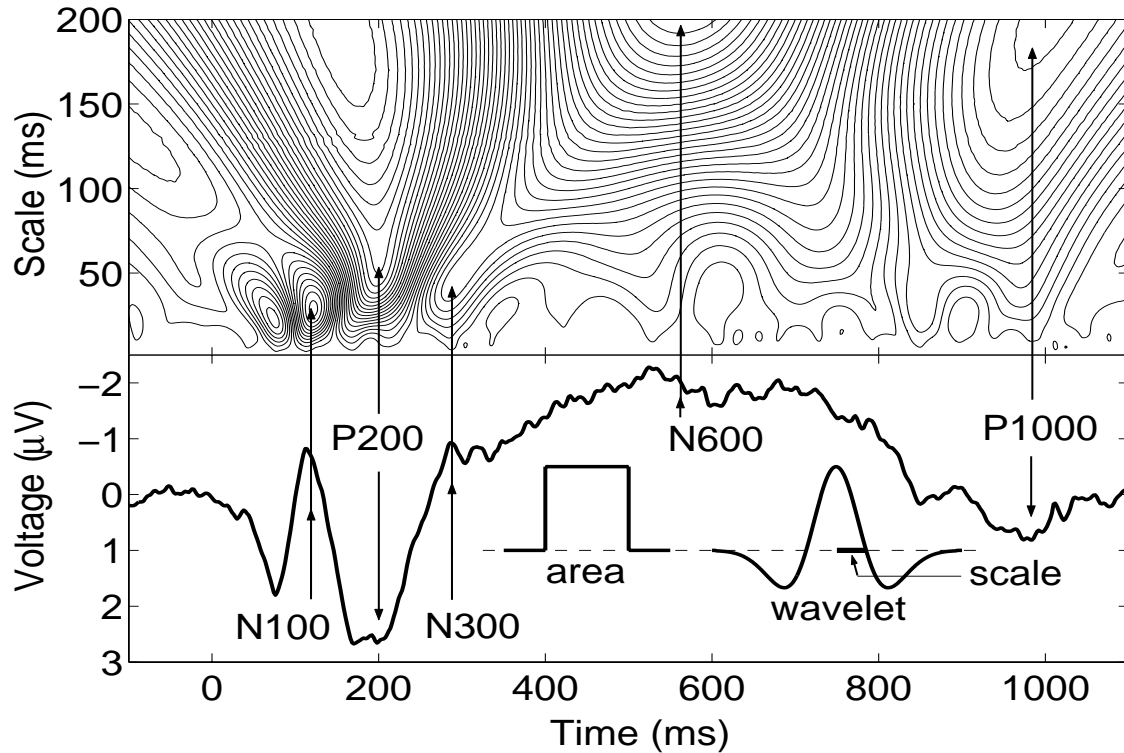


Figure 2.1: The rectangular area template, the wavelet template, the total grand average of both experimental conditions in the oddball experiment (lower plot), and its CWT scalogram (upper plot). Each point of the scalogram represents the cross-covariance of the total grand average with a wavelet of the corresponding time and scale. Extrema in the scalogram correspond to peaks in the waveform. Extremum scale gives a measure of peak width (ERP-component duration). The scale of the Mexican Hat wavelet (lower plot) is defined as the half-width between its zeros (thick horizontal line). Note that the area template has no minima under the zero line and hence cannot separate larger scales.

= 290ms, scale = 45ms. Accordingly, the N300 *fixed-wavelet* measure was defined as the CWT value at this same position for all participant averages, all single trials and all channels. The N300 *matched-wavelet* measure was defined as the minimum CWT value in the rectangular window $270\text{ms} < \text{time} < 310\text{ms}$, $35\text{ms} < \text{scale} < 55\text{ms}$ in each participant-average and each single-trial scalogram. Thereby, the matched

wavelet was partly adjusted to compensate for inter- and intraparticipant latency variations. Thus the fixed wavelet was fitted only once to the total grand average and then the cross-covariance with each participant average and each single trial was taken as a measure of the N300 effect, while the matched wavelet was additionally fitted to each participant-average and each single-trial waveform.

The other classical ERP measure, the *peak*-measure, was also obtained from the waveforms (see Chapter 1). The computation of peaks was similar to that of the areas with the only difference that the maximal or minimal voltage in the window of interest was taken instead of the average value.

Similarly to Chapter 1, all ERP measures were tested for statistical significance by three-way analyses of variance (ANOVA) with factors: condition(2) \times FCP(3) \times LMR(3). FCP and LMR were taken as within-participant or within-trial factors, whereas the experimental condition was taken as a within-participant factor for the participant-average ERP measures and as a between-trials factor for the single-trial ERP measures. Left/right and frontal/parietal asymmetries were studied by linear contrasts; predominance at the vertex Cz was studied by quadratic contrasts.

The wavelet computations were performed with MATLAB 6.0, and the ANOVAs – with SPSS 11.0.

2.4 Results

For the sake of brevity and clarity, only significant ANOVA probability values (*p*-values) and no effect sizes are reported. The results obtained with the fixed-wavelet measure are similarly not reported and the matched wavelet, which produced generally slightly better results, especially in the single-trial assessments, is further simply referred to as "wavelet". The results obtained with the peak measures were similar to those obtained with the area measures and are also not reported here. (See Chapter 1

for significant p -values obtained with the peak measure from the participant-average waveforms). Although, not all 4 ERP measures (area and wavelet, both single-trial and participant-average) are needed to characterize the components, their comparison is useful for testing the new total-average-CWT assessment method. Comparing (p -values) is equivalent to comparing effect sizes (cumulative probability distribution functions are monotonous).

2.4.1 Oddball

The ERP components found in the waveforms can be seen in Chapter 1, Figure 1.3. The direct comparison between "Oooh!" (woe) and "Oooh!" (joy) at Fz is shown in Chapter 1, Figure 1.2.

The N300 amplitude was significantly larger for "Oooh!" (woe) than for "Oooh!" (joy) by the main effect of condition with both participant-average measures: area ($p = 0.03$) and wavelet ($p = 0.001$), as well as with the single-trial wavelet ($p < 0.001$), but not with the single-trial area. The component was predominant at the vertex (Cz) by the significant quadratic contrasts of FCP, LMR and FCP \times LMR with both wavelet measures: participant-average ($p = 0.005$) and single-trial ($p < 0.001$), but not with the area measures.

The N600 amplitude was significantly larger for "Oooh!" (woe) than for "Oooh!" (joy) at frontal sites by the condition \times FCP interaction with both participant-average measures: area ($p = 0.03$) and wavelet ($p < 0.001$), as well as with the single-trial wavelet ($p < 0.001$), but not with the single-trial area. The component was predominant at mid-line sites by the significant quadratic LMR contrast with all of the measures: participant-average – area ($p = 0.02$) and wavelet ($p = 0.001$), and single-trial – area ($p = 0.008$) and wavelet ($p < 0.001$).

The P1000 amplitude was significantly larger for "Oooh!" (woe) than for "Oooh!" (joy) at frontal sites by the main effect of condition with both wavelet measures:

participant-average ($p = 0.03$) and single-trial ($p = 0.01$), but not with the area measures, as well as by the condition \times FCP interaction with both participant-average measures: area ($p = 0.01$) and wavelet ($p < 0.001$), as well as with the single-trial wavelet ($p < 0.001$), but not with the single-trial area.

No significant lateralization (linear-contrast effects of LMR or LMR \times condition) was found for any of the components.

2.4.2 Priming

The ERP components found in the waveforms can be seen in Chapter 1, Figure 1.4.

The N300 amplitude was significantly larger for the inconsistent vocalizations by the main effect of condition with both wavelet measures: participant-average ($p = 0.02$) and single-trial ($p = 0.04$); the effect was also nearly significant with the participant-average area ($p = 0.05$), but not with the single-trial area. The component was predominant at the vertex Cz by the significant quadratic contrasts of FCP, LMR, and FCP \times LMR, with both wavelet measures: participant-average ($p = 0.03$) and single-trial ($p < 0.001$), but not with the area measures.

The P500 amplitude was significantly larger for the inconsistent vocalizations by the main effect of condition with both wavelet measures: participant-average ($p = 0.04$) and single-trial ($p < 0.006$), but not with the area measures. The component was predominant at parietal sites by the significant linear FCP contrast with both wavelet measures: participant-average ($p < 0.001$) and single-trial ($p < 0.001$), but not with the area measures. Notwithstanding visual impression, the linear-contrast effect of FCP \times condition was not significant.

No significant lateralization (linear-contrast effects of LMR or LMR \times condition) was found for any of the components.

The significance of the N300 effect was also tested separately for each of the three

kinds of word-exclamation pairs: positive-[negative] ("joy"-[grief], "pleasure"-[rage], "surprise"-[disappointment]), negative-[positive] ("grief"-[joy], "fright"-[pleasure], "disappointment"-[surprise]), and negative-[negative] ("disgust"-[terror], "rage"--[fright], "terror"-[disgust]). Both wavelet measures were significant for the negative-[positive] pairs ($p < 0.05$) and nearly significant for the positive-[negative] and negative-[negative] pairs ($0.05 < p < 0.07$); the area measures were not significant.

2.5 Discussion

A new CWT-based method was developed for the better detection and assessment of ERP components from single trials as well as from participant averages. Essentially, it is a kind of template matching, where the template is a wavelet, and it is fitted to the signal not only in time (component latency) but also in frequency (scale, peak width). In all assessments, the wavelet measure yielded larger ANOVA effects than the traditional area and peak measures. This effect was particularly pronounced in the single-trial assessments – from all cases in which the wavelet measure was significant there was only one in which the area was significant. This result is not surprising, considering that the traditional area measure does not separate scales, and hence peaks of different widths, especially slow waves and DC components, are taken into account in a single component assessment, thus increasing variance and reducing statistical effect size and statistical power. This is best understood if the area measure is seen as template matching with a fixed rectangular template, which takes only positive values, while the wavelet template with its two deep minima extracts only peaks of nearly the same scale (Figure 2.1).

Furthermore, the total-average-CWT method has three important advantages upon its ancestors, the classical template-matching, single-trial algorithms (Smulders et al. 1994): First, the template is fitted to the signal not only in time, but also in scale. Second, the template is a wavelet, which guarantees better scale separation than non-wavelet templates, and hence better selective sensitivity of the wavelet mea-

tures to particular ERP components, as well as better separation of overlapping components (Samar et al. 1999). Third, the scale and time position of the template are determined from the total average of all trials (or all participant averages) entering the subsequent statistical analyses. The detection of ERP peaks from the total-average scalogram is blind to the separation of the sample into categories according to the experimental conditions (e.g. "congruous" versus "incongruous"). That makes it statistically more appropriate for testing difference hypotheses than traditional methods, which use the average of one category of trials or the difference waveform of the two categories for component detection. The post hoc visual selection of time windows of interest from the difference waveform may bias subsequent statistical significance tests by accumulation of chance. The total-average scalogram allows ERP component detection without visual inspection of the ERP curves, thus reducing experimenter bias. One systematic study using signal detection theory showed that computer-based methods may be more efficient in ERP detection than human observers (Valdes-Sosa et al. 1987). The issues of statistical power, statistical correctness and human-experts' fallibility are of particularly high relevance for individual assessment and diagnostics in clinical settings where only single-trial assessment is possible, the number of trials is limited, and individual patients' ERP components differ substantially from those observed in group averages (Kotchoubey et al. 2002). That is why, this growing field of ERP research can be seen as the primary and most natural area for application of the CWT method. Furthermore, in any ERP paradigm in which *new* or *small* components are examined (like in the present study), classical measures may prove unsatisfactory, and wavelet measures may yield much better and reliable results.

An obvious shortcoming of the total-average-CWT method comes with the relatively rare case when both experimental conditions elicit an ERP component of the same latency, width, and amplitude, but with different polarity, as such waves would cancel each other in the total average. An improved version of the CWT method, the τ -CWT, which provides a solution to this problem is presented in Chapter 3.

Chapter 3

ERP Assessment with the τ -CWT

Method

The usage of the ERP technique in clinical applications for diagnostic purposes, requires special methods of EEG data assessment, based on single trial analysis. The total-average-CWT method introduced in Chapter 2 is an example for one such method. In this chapter, a new, largely improved version of the method, the τ -CWT, is introduced. The τ -CWT is based on the CWT and Student's τ -statistics. The method was systematically tested in two prototypical ERP paradigms, oddball and semantic priming, which belong to basic tools for ERP based cognitive diagnostics. The method was compared to other assessment procedures based on Principal Component Analysis (PCA) and the Discrete Wavelet Transform (DWT). Similarity to clinical settings was achieved by the individual assessment of each participant's ERP data. Both whole waveforms and single ERP components were assessed by multivariate procedures including PCA data set reduction, Hotelling's T^2 -test and a randomization test. The assessment of the whole ERP waveforms is particularly relevant to the paradigms of the context violation class already introduced in Chapter 1. The results demonstrated the superiority of the τ -CWT to the other assessment methods.

3.1 Introduction

DETECTION AND QUANTIFICATION of ERP components has evolved to a separate interdisciplinary branch of ERP research. A variety of assessment methods at different levels of sophistication and practicality have been proposed. All of these methods have basically the same objective: to extract from the highly stochastic and chaotic EEG signal those patterns, which reflect the neurocognitive processes that

are investigated in the ERP experiments. In this chapter, it will be demonstrated how the total-average CWT method introduced in Chapter 2 can be improved to provide better extraction of such patterns for the purposes of both fundamental and applied-clinical research.

The standard assessment of an ERP component includes three basic steps (see Chapters 1 and 2). First, the average waveforms for the two experimental conditions are obtained from each individual participants data. Second, the grand-average waveforms, obtained by averaging across all participants, are visually inspected and the ERP component of interest is detected by simply finding the time window, in which the two ERP curves corresponding to the two experimental conditions diverge substantially. Third, the amplitude of the ERP component is measured in each participants average waveforms, by the mean voltage (area) or the peak value in the visually determined time window. This measure undergoes an analysis of variance (ANOVA) with, generally, two factors: topography (i.e., scalp distribution) and experimental condition. The single measurements are those obtained from the participants' averages and not from the single trials.

This standard assessment procedure is simple and plausible, but this is at the cost of important shortcomings. First, the within-participant variance is not utilized, it gets lost by the averaging procedure. This approach decreases the total variance in the data (only the between-participant variance remains), but this is at the cost of a dramatic reduction of the sample size by a factor equal to the number of single EEG trials performed with each participant in each experimental condition. Second, the visual inspection of the grand-average waveforms is a quite unreliable method of ERP component detection – it is imprecise and prone to experimenter bias. Third, the area and peak measures give very rough estimates of ERP component amplitudes. For instance, they are not applicable in single-trial analyses, because they cannot separate the scales (= wavelengths = inverse frequencies) of the highly chaotic oscillations present in the single-trial EEG signals (see Chapter 2). Thus, a good method for detection and quantification of ERP components

should provide: first, a clearly defined mathematical procedure for detection of ERP components, and second, a sensitive and discriminative mathematical measure for component quantification which should be applicable with single trials. "Sensitive and discriminative" denotes the ability to measure only a single ERP component defined by its position in time and its duration (position on the scale dimension). This includes good *scale separation*. The area and peak measures do not satisfy these two criteria. The time window in which they are applied is determined visually by the experimenter; the area measures all components at a scale larger than the window width and the peak does not discriminate between different scales at all (unless the signals are filtered priorly).

Several techniques have been proposed for ERP component extraction from single trials (Smulders, Kenemans, and Kok 1994). The classical ones use *template matching*, i.e., computing the cross-correlation (Woody 1967; Fabiani, Gratton, Karis, and Donchin 1987) or cross-covariance (Pfefferbaum and Ford 1988) between the EEG sweep and a predefined waveform, or peak picking after low-pass filtering of the signal (Ruchkin and Glaser 1978), which is virtually the same, since low-pass filtering is equivalent to a convolution (cross-covariance) of the signal with a certain template in the time domain. Recently, more sophisticated single-trial estimation methods based on template decomposition (Lange, Pratt, and Inbar 1997), subspace regularization (Karjalainen, Kaipio, Koistinen, and Vauhkonen 1999), maximum-likelihood solutions (Jaskowski and Verleger 1999), competitive neural networks (Lange, Siegelmann, Pratt, and Inbar 2000), etc. have been proposed.

With the rise of the wavelet era (Samar et al. 1999), single-trial algorithms based on the Discrete Wavelet Transform (DWT) came to existence (Demiralp, Yordanova, Kolev, Ademoglu, Devrim, and Samar 1999; Demiralp, Ademoglu, Schurmann, Basar-Eroglu, and Basar 1999). Such algorithms give a highly economical and structured representation of the information contained in EEG sweeps, but suffer the general disadvantage that this representation is predefined by the dyadic scheme (see Figure 3.2) and hence cannot be adjusted to a given ERP structure. More

complicated algorithms, using, e.g., wavelet packets (Raz, Dickerson, and Turetsky 1999) or wavelet networks (Heinrich, Dickhaus, Rothenberger, Heinrich, and Moll 1999; Heinrich, Moll, Dickhaus, Kolev, Yordanova, and Rothenberger 2001) have been proposed to remedy this situation. Another possible solution is to give up the representation of the whole signal and to concentrate on feature extraction. One such approach utilizes the CWT introduced in Chapter 3, which yields a highly redundant representation of the EEG signal as a function of two independent variables – time and scale (or frequency). Local extrema of this function indicate prominent features of the signal – transient peaks or oscillations.

Samar et al. 1999 commented that the CWT "is not very efficient". They pointed out that firstly, the information it displays is highly correlated and therefore unnecessarily redundant for analytic purposes, and secondly, it is very time consuming to compute directly.

In Chapter 2 a CWT-based method for detection and measurement of ERP components was introduced and its superiority to the classical area and peak methods was demonstrated. In this approach, ERP components are identified with local extrema in the CWT scalogram. Thereby, the redundancy of the CWT representation is turned into an advantage – it allows much better localization in both time and frequency than the fixed dyadic structure of the DWT does. The second disadvantage of CWT, pointed out by Samar et al. 1999 – the heavy computational load – can be overcome by a fast computational algorithm. One such algorithm developed in the present study is described in this chapter.

A major disadvantage of the total-average-CWT method proposed in Chapter 2 arises from the fact that the ERP components are detected from the *total grand average* waveform obtained by averaging over all trials from both experimental conditions. This procedure was chosen for reasons of statistical correctness – any detection procedure that utilizes the difference waveforms or is otherwise based on the separation of the EEG trials into groups according to the experimental conditions

involves accumulation of chance. For instance, in a 1000ms-long EEG epoch the probability for the occurrence of a 200ms-long wave by chance is roughly 5 times higher than in a 200ms-long epoch. If the statistical significance of such a wave is tested by ANOVA or some other test, the α -level should be corrected to take into account the chance accumulated by choosing out of 5 possible 200ms-wide time windows the one defined by the maximal difference between the two ERPs. The problem is that there is no adequate analytical correction procedure that takes into account the correlations between the data from the 5 windows and the fact that the windows are not actually fixed in time shift and length. The usage of wavelet measures, which do not rely on time windows, makes the problem even more complicated. Generally, any preselection of variables from the essentially multivariate ERP data sets faces the same problem, if information about the separation of the data into two samples according to the two experimental conditions is used to define a selection criterion. This holds for windows in the time domain, frequency bands, principal components, wavelet coefficients, etc.

Using the total average for ERP component detection in Chapter 2 was a way to avoid the accumulation of chance. In this approach, however, components from both experimental conditions are mixed together, and the method relies *only* on scale separation to discriminate between ERP waves of equal latency. This situation is clearly suboptimal, since in the extreme (but actually improbable) case of two ERP components of the same scale, latency and amplitude, but with different polarity, the method would fail to detect the waves because they would cancel each other in the total average waveform. Apart from this extreme and rare example, the total-average-CTW method generally does not provide the optimal, most "sensitive and discriminative" measure for detection and quantification of ERP components.

In the present chapter, an improved CWT-based feature extraction method, the t -CWT, is introduced. It was originally developed (by the author of the present dissertation) for classification of single-trial ERPs in brain-computer interfaces. The method was validated at the International BCI Classification Competition 2003

(Blankertz, Müller, Curio, Vaughan, Schalk, Wolpaw, Schlögl, Neuper, Pfurtscheller, Hinterberger, Schröder, and Birbaumer 2004), where it was a winner on two of the ERP data sets obtained in two different BCI paradigms – P300 speller (Farwell and Donchin 1988) and slow-cortical-potentials feedback (Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey, Kübler, Perelmouter, Taub, and Flor 1999). These results are presented in Chapter 4. The τ -CWT method is similar to the total-average-CTW method. The ERP components are identified again by the local extrema in a scale-time plot – the scalogram. The crucial difference is that this scalogram is not obtained from the total average, but from the difference waveform and from the variance in every scale-time sample point of the scalogram. The within-experimental-condition variance is computed from all the single-trial scalograms. Then Student's τ -value is computed for each scale-time point and the obtained surface, called τ -CWT or τ -value scalogram, is used for ERP component detection. The thus obtained local extrema are the points of maximal difference between the CWTs of the two experimental conditions. Thus, the τ -CWT method is certainly prone to accumulation of chance, which poses an obstacle to its application for testing of statistical hypotheses about the difference between two samples of ERPs. (However, there is no problem for discriminant analysis and classification.)

The aim of the study presented in this chapter was to adopt the τ -CWT method for use in ERP component significance testing and estimate its power. A randomization test is proposed as a simple solution to the accumulation-of-chance problem. Further, the strength of the method is demonstrated in *individual-participant* analyses, in which, on the one hand, the usage of single trials is inevitable, and on the second, previous group data can be used for ERP-component detection purposes without concern about accumulation of chance.

The present study also addressed another, widely neglected aspect of ERP assessment – the statistical comparison of *the whole waveforms*. This is particularly relevant for the wide class of *context violation paradigms* already mentioned in Chapter 1. In such paradigms, the presented stimuli have some *meaning* along with their

physical properties, and the the cognitive processing of this meaning is reflected by the measured ERPs. The experiments are accordingly designed to enhance the ERPs related to these higher cognitive processes and to neglect the processing of stimulus novelty, or occurrence frequency, or primitive physical features like sound frequency, picture color, stimulus duration, etc..

As already discussed in Chapter 1 an important property of meaningful stimuli is that they can build a context, which induces some expectations about the next-coming stimuli. The violation of these expectations is reflected by the N400 component in the ERPs to incongruous stimuli. However, the N400 amplitude and latency (whether quantified by classical area and peak measures or by more advanced wavelet methods) are an oversimplified representation of the ERP response to incongruous stimuli. The investigator may be interested in *any* significant difference between the ERPs to congruous and incongruous stimuli, because it would necessarily imply that the meaning of the stimuli as well as the context were adequately processed by the participant's cortex. This is not quite true for paradigms from the oddball class, in which various cognitive processes may be reflected in the difference ERP (Birbaumer and Schmidt 2003) – dishabituation of the orientation reaction reflected by increased N100 amplitude to deviant stimuli, physical mismatch reflected by the Mismatch Negativity (MMN), detection of novel, or rare, or target stimuli reflected by the P300, etc.. But in a context-violation paradigm, if the investigator is interested in the first place to find out, whether the brain can at all decipher, or not, the meaning encoded in certain stimuli, she or he might wish to first compare the whole ERPs elicited by the congruous and incongruous stimuli before concentrating on specific ERP components.

3.2 The τ -CWT Method

3.2.1 CWT

The CWT $W(s, t)$ of a signal $f(t)$, where t is the time variable (should not be confounded with Student's t !), is defined (Samar et al. 1999; Ende et al. 1998) by:

$$W(s, t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(\tau) \psi\left(\frac{\tau - t}{s}\right) d\tau, \quad (3.1)$$

where t denotes the time shift, s denotes the scale (also measured in time units) and ψ is the wavelet function, which has a zero mean:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (3.2)$$

Hence, the CWT is a kind of *template matching* (Smulders, Kenemans, and Kok 1994), i.e. a computation of the cross-covariance between the signal and a predefined waveform, which is shifted forwards and backwards in time and dilated and constricted in scale (see Chapter 2). The local extrema of $W(s, t)$ are the points of best match between the signal and the template in the time-scale (time-frequency) domain. They can be identified visually from the time-scale plot of $W(s, t)$, the scalogram Figure 3.5, 3.8, or automatically, by an appropriate algorithm. The advantage of the CWT over the classical template matching methods (Smulders, Kenemans, and Kok 1994) arises from the special properties of the wavelet template, which allow optimal *scale separation* of ERP components (see Chapter 2).

The choice of a particular wavelet function $\psi(t)$ in (3.1) depends on what kind of features should be extracted from the signal $f(t)$. Wavelets that are well localized in the time domain like the "Mexican Hat" are used for detection of ERP components and wavelets that are well localized in the frequency domain like the Morlet wavelet are used for detection of salient oscillations (Ende et al. 1998; Senkowski

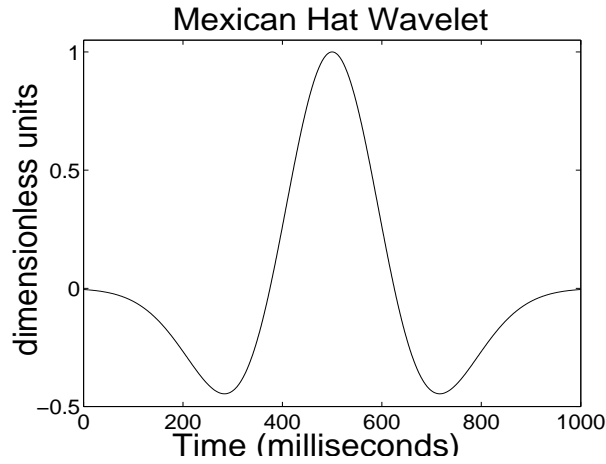


Figure 3.1: The Mexican Hat wavelet; $s = 500\text{ms}$, $t = 500\text{ms}$.

and Herrmann 2002). The application described here uses the Mexican Hat (Fig. 3.1) given by:

$$\psi(t) = (1 - 16t^2)e^{-8t^2}. \quad (3.3)$$

Note that (3.3) differs from the standard definition of the Mexican Hat: $\psi(t) = (1 - t^2)\exp(-t^2/2)$ used in Chapter 2 in the 4 times larger unity scale. In the standard definition the scale is the half width between the zeros of the Mexican Hat (Figure 2.1); in the definition given by (3.3) the scale is slightly larger than the distance between the minima (Fig. 3.1). This is done for convenience – thus defined, the scale is approximately equal to the wavelength (inverse frequency) of the Mexican Hat, which in ERP applications corresponds to the approximate duration of the ERP waves.

3.2.2 τ -CWT

The τ -CWT method for ERP detection and quantification uses the CWT and Student's two-sample τ -statistic. It is performed in 5 simple steps:

1. The CWT $W^{kn}(s, t)$ of the signal $f^{kn}(t)$ is computed for each channel k and

each single trial n according to (3.1).

2. The mean CWTs $\overline{W_g^k(s, t)}$ and the variances $\sigma_g^k(s, t)$ are obtained for each of the two samples (groups) of trials corresponding to the two experimental conditions:

$$\overline{W_g^k(s, t)} = \frac{1}{N_g} \sum_{n=1}^{N_g} W^{kn}(s, t), \quad (3.4)$$

$$\sigma_g^k(s, t) = \frac{1}{N_g - 1} \sum_{n=1}^{N_g} \left(W^{kn}(s, t) - \overline{W_g^k(s, t)} \right)^2, \quad (3.5)$$

where $g = 1, 2$ and N_g denotes the number of trials in sample g .

3. The two-sample t -statistic $\mathfrak{t}^k(s, t)$ is computed:

$$\mathfrak{t}^k(s, t) = \sqrt{\frac{N_1 N_2}{N_1 + N_2} \frac{d^k(s, t)}{\sigma_{pl}^k(s, t)}}, \quad (3.6)$$

where $d^k(s, t) = \overline{W_2^k(s, t)} - \overline{W_1^k(s, t)}$ is the difference average CWT and $\sigma_{pl}^k(s, t)$ is the pooled variance of the two samples.

4. The points (s^{ki}, t^{ki}) , at which the functions $\mathfrak{t}^k(s, t)$ have local extrema, are found. These are the points of (locally) maximal difference (variance taken into account) between the two samples in the time-scale domain and they have the following interpretation: t^{ki} is the latency and s^{ki} is the duration of ERP component i at scalp position k .
5. The CWT W^{kin} is computed for each point (s^{ki}, t^{ki}) and each single trial n according to (3.1):

$$W^{kin} = W^{kn}(s^{ki}, t^{ki}). \quad (3.7)$$

W^{kin} are the values of the single-trial CWTs at the points of maximal difference between the two samples and they have the following interpretation: W^{kin} is a measure of the amplitude of the ERP component i at scalp position k in trial n .

The importance of the variance for proper component detection was demonstrated by a comparison of the τ -CWT with the dCWT – a modification of the method, in which only the difference scalogram $d^k(s, t)$ is used for detection of local extrema and the variance $\sigma_{pl}^k(s, t)$ is not taken into account (3.6).

When all values of i are taken into account, W^{kin} give a representation of the whole single-trial EEG signals, which is not systematic, but is based on feature extraction, and the τ -CWT can be used for significance testing of the difference between the whole waveforms regardless of specific ERP components. The τ -CWT method was tested together with two known methods for systematic representation of whole waveforms – PCA and DWT.

In the case when a particular ERP component should be detected and quantified, only certain values of i are taken into account (e.g. only minima or only maxima, see below). By taking only certain values of i , it was tested how the τ -CWT method handles single ERP components and the obtained measures were compared to the classical area and peak measures.

In all applications, Hotelling's two-sample T^2 -test was used for comparison of multivariate means (Friston, Stephan, Heather, Frith, Ioannides, Liu, Rugg, Vieth, Keber, Hunter, and Frackowiak 1996; Rencher 1998, pp. 85–87). Whole waveforms and single components were tested by the same multivariate procedure.

3.2.3 Computational Algorithm

Since the CWT (3.1) is a linear transformation, it can be computed by multiplication of the signal with the corresponding transformation matrix. Both the signal and the matrix are represented in the time-frequency domain by DWT (Samar et al. 1999). This representation is much more economical than the original time domain representation, because in every real application, events at scales smaller than some cut-off scale s_c are ignored (low-pass filtered) and hence, the corresponding wavelet

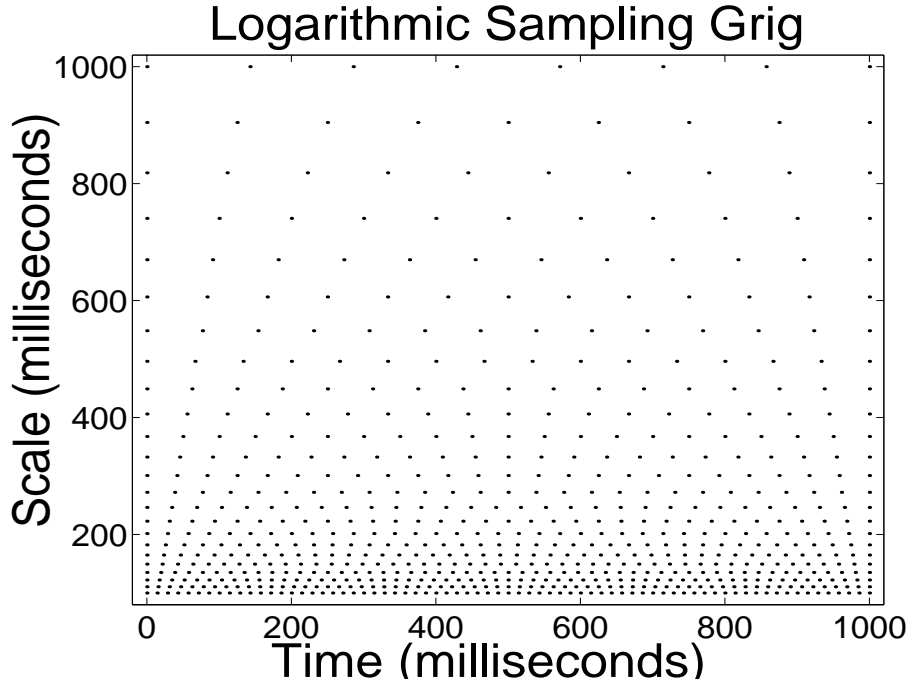


Figure 3.2: A logarithmic grid with 7×7 PPSPO (for illustration purposes only – 24×24 PPSPO were used in the applications). At $s = 200\text{ms}$ there are exactly 7 points every 200ms, at $s = 400\text{ms}$ there are exactly 7 points every 400ms, etc.. There are exactly 7 rows of points between $s = 100\text{ms}$ and $s = 200\text{ms}$, 7 rows between $s = 200\text{ms}$ and $s = 400\text{ms}$, etc.. The DWT with its dyadic representation scheme can be seen as digitized CWT sampled on the sparsest possible logarithmic grid with 1 PPSPO providing an exact and non-redundant representation of the signal.

coefficients can be deleted from the DWTs of the signals. The direct current (DC) component can also be filtered out by deleting the first wavelet coefficient.

The CWT is sampled on a logarithmic grid in the time-scale plane (Figure 3.2), which allows a constant, scale-invariant sample point density of $d_t \times d_s$ points per scale per octave (PPSPO). The CWT of all signals is computed in one single step – the matrix of the DWTs of the signals is multiplied by the CWT matrix. This algorithm is very fast – in ERP applications it typically takes from 1ms to 10ms per signal on an AMD 704MHz processor with 256MB of working memory (the computational speed depends on the cut-off scale s_{c0} and the number of PPSPO). Once the local extrema of $\tau^k(s, t)$ (3.6) are extracted, the CWT matrix is reduced

by deleting all rows except those corresponding to the points (s^{ki}, t^{ki}) (3.7). The features W^{kin} (3.7) are obtained by multiplication with this reduced CWT matrix.

The DWT and any other linear preprocessing transformations (resampling by interpolation, filtering, etc.; see below) are also computed by matrix multiplication. The advantage is that all consecutive operations are represented by one single transformation matrix, which is the product of the corresponding transformation matrices of the individual operations. Only the CWT is computed in a second step, one channel at a time, to avoid huge transformation matrices and to save working memory.

In the applications described below, all computations were performed with MATLAB 6.0. The DWT matrix was computed with the wavelet toolbox Uvi_Wave 3.0 using Symlets with 16 filter coefficients (González Sánchez, González Prelcic, and García Galan 1996).

3.3 Experiments

Two prototypical ERP paradigms were used to test and demonstrate the τ -CWT method. The first was the well known oddball paradigm: two different acoustic stimuli are presented repeatedly in a random order with different occurrence frequencies; the rare stimulus elicits the P300 component in the ERP (Johnson 1988). In a standard oddball task the participants are instructed to press a button on each occurrence of the infrequent (deviant) stimulus or simply to count how many times it occurs. Polich 1987; Polich 1989 showed that in a visual oddball, a P300 with a little smaller amplitude is elicited even without an active discrimination task. (Lang, Kotchoubey, Lutz, and Birbaumer 1997) replicated this finding also for the auditory oddball. In such "passive" oddball experiment, the participants are not given any instructions and the deviant stimulus apparently attracts their attention merely by its rare occurrence. (Note the difference between this paradigm, which utilizes only two simple stimuli, and the passive oddball paradigm with meaningful stimuli in-

troduced in Chapter 1.) In the present study, the t-CWT method was tested in a passive oddball experiment, because this design is more appropriate for diagnostic purposes than the active task and the obtained results are thus more relevant to clinical applications with paralyzed patients (Kotchoubey et al. 2002).

The second paradigm that was used was a member of the context violation class. It was the auditory semantic priming paradigm: spoken words are presented in pairs, half of which are semantically consistent, e.g. "man-woman", "strong-week", etc. and the other half are semantically inconsistent, e.g. "monk-ear", "aunt-branch", etc. (Hagoort, Brown, and Swaab 1996). The N400 component in the ERP is elicited by the incongruent second words. However this is not the only ERP response to semantic incongruity – an additional positive wave, the P600, often follows the N400. This positivity can be seen in ERP waveforms reported by many authors, but usually, it remains uncommented, while only the N400 is in the focus of the interest (McCallum, Farmer, and Pocock 1984). Münte, Heinze, Matzke, Wieringa, and Johannes 1998 studied the P600 effect systematically and confirmed its significance, thus challenging its traditional attribution to syntactic incongruences. They found no evidence for specificity of the syntactic positive shift and concluded that the P600 is evoked by linguistic violations of different kinds, including semantic and syntactic violations. In the present study, the P600 was taken into account by subjecting the whole ERPs elicited by the congruous and the incongruous stimuli to multivariate statistical tests. The semantic priming paradigm is one of the key diagnostic tools used with paralyzed patients in clinical applications (Kotchoubey et al. 2002).

3.4 Methods

3.4.1 Participants

36 healthy persons (20 female, mean age \approx 27.5 years) participated in both experiments. According to self report, all participants were right-handed and their mother tongue was German. Most of them were students and all were paid for their participation.

3.4.2 Stimuli

All stimuli were digitized at a sampling rate 22.05kHz/16bit. The rare and the frequent stimuli in the oddball experiment were 50ms-long, 75dB-loud sine tones, of frequency 0.8kHz and 1.3kHz respectively. All words for the priming experiment were recorded by a female, native German speaker.

3.4.3 Procedure

Stimulus presentation

In both experiments, the stimuli were presented in short sessions with 3-4min pauses between the sessions. The rare stimulus occurred 45 times and the frequent one occurred 255 times in each of the 4 sessions of the oddball experiment. In the priming experiment, 100 word pairs were presented in each of the 4 sessions. Half of the pairs were semantically consistent, and the other half were semantically inconsistent.

The two tones in the oddball experiment were presented in a randomized sequence at a constant rate of 1 tone every 0.85s. The word pairs in the priming experiment were presented in a randomized sequence at a constant rate of 1 pair every 1.7s. The stimulus onset asynchrony (SOA) between the first and the second word in a

pair was 0.5s.

In both experiments, the only instruction given to the participants was to listen attentively.

EEG recording

Digitized EEG (time resolution: 2ms/step (500Hz), voltage resolution: $0.1678\mu V$ /step) was continuously recorded from 9 scalp positions according to the 10-20 system: Fz, Cz, Pz, F3, F4, C3, C4, P3, P4. All electrodes were referenced to the linked mastoids. Electrical eye activity was also recorded by bipolar acquisition from the following sites: FP2 and a site below the right eye – for vertical eye movements and eye blinks; two sites laterally adjacent to the eyes – for horizontal eye movements. The EEG was filtered on-line (band-pass: 0.1Hz-70Hz, notch: 50Hz).

3.4.4 Data Analysis

Preprocessing

EEG epochs were created off-line and referenced to the 100ms prestimulus baseline. Eye-blink and eye-movement artifacts, both vertical and horizontal, were corrected off-line by a computerized procedure (Miller, Gratton, and Yee 1988; Gratton, Coles, and Donchin 1983). In the oddball experiment the 500ms-long epochs started at 100ms poststimulus. In the priming experiment the 1100ms-long epochs started also at 100ms poststimulus;

Four types of statistical outliers were excluded from the data according to the following procedure. The maximal and the minimal voltage as well as the maximal and the minimal value of the first derivative of the signal were found for each trial. Means and standard deviations were obtained for each of these four variables. Tri-

als, in which one of the values exceeded 3 standard deviations measured from the mean, were considered outliers and rejected. As a result from this procedure about 5% of all trials in each participant's data set were excluded from further analysis.

In the priming experiment, the data of the 4 sessions of each participant were pooled together. Thus there were about 380 EEG sweeps in each participant's data set after outlier rejection.

The statistical analysis of the oddball data showed saturation effects already with the data sets of the first session – all participants but one showed a significant P3 response with all wavelet methods. In order to compare the test power of the different methods, only 60 trials from the first session were retained (the first 51 standard and the first 9 deviant trials) and all other trials from the first session as well as the other sessions were dropped from further analyses.

Initial Filtering

The EEG signals were low-pass filtered in the time domain by convolution with the function:

$$\exp\left(-\left(\frac{4t}{s_c}\right)^2\right) \cos\left(\frac{2\pi t}{s_c}\right), \quad (3.8)$$

where t denotes the time and s_c – the cut-off scale. The thus defined cut-off scale is related to the cut-off frequency f_c by: $f_c = 1/s_c$. In both experiments $s_c = 100ms$ ($f_c = 10Hz$).

Log Time Sampling

After the initial filtering the EEG signals were resampled on a logarithmic time scale. The new sampling rate was a linear function of time – the interval between two adjacent time points was increased linearly. This was done to allow for *time dependent filtering* (see below) with a cut-off scale linearly increasing with time

(filtering frequency – linearly decreasing). The *scale dilation per second* λ defined how many times the cut-off scale was increased with each second:

$$s_c = s_{c0}(1 + \lambda t). \quad (3.9)$$

In both experiments $\lambda = 2$ units per second. The log time sampling was performed by spline interpolation.

Time Dependent Filtering

The EEG signals were low-pass filtered twice by the convolution procedure (3.8). In effect, waves of latency t and duration shorter than $s_{c0}(1 + \lambda t)$ were attenuated and peaks of latency t and duration shorter than $s_{c0}(1 + \lambda t)/2$ vanished altogether. This procedure takes into account the fact that late ERP components usually have longer duration than earlier waves. Thus a more economical presentation of the signal was obtained. In both experiments $s_{c0} = 200\text{ms}$.

The first $s_{c0}/4\text{ms}$ of the signals were multiplied by $(1 - \cos 4\pi t/s_{c0})/2$ to assure a smooth onset. Analogically, a smooth decay was accomplished by cosine attenuation of the last $s_{c0}(1 + \lambda T)/4\text{ms}$ of the $T\text{ms}$ long signal.

DWT

The EEG signals were represented in the time-frequency domain by DWT. This step was not intended as additional filtering, but as a convenient and economical representation, in which the higher frequencies (and the corresponding scales) that had already been filtered out in the previous steps were separated from the retained ones and deleted from the data set. Moreover, as described above, the CWT transformation matrix was computed in the time-frequency domain defined by the DWT and hence this representation was necessary for further analysis.

PCA

As an alternative to the DWT, Principal Components were extracted from the EEG signals in each channel (before DWT) using the total covariance matrix (not the pooled one) and retaining only the eigenvectors with eigenvalues greater than the average eigenvalue. Thus another economical representation of the signals was obtained, in which the already suppressed frequencies (scales) were excluded from the further analyses.

τ -CWT

The EEG signals were CW-transformed for scales ranging from $s_{c0}(1 + \lambda t)/2$ to $4T(1 + \lambda t)$, where T is the epoch length. (Note that, because of the log time sampling, the actual scale range was again time dependent.) The single-trial scalograms were sampled on a logarithmic grid with 24×24 PPSPO (Figure 3.3). ERP components were detected and by the τ -CWT method described above (3.5–3.7). The procedure was also performed on the pooled data sets obtained by excluding the data of one participant at a time from the pooled data of the whole group. The thus detected ERP components from each pooled data set were used for the statistical assessment of the individual data of the respective excluded participant. To show the importance of taking the variance into account, the procedure was repeated on the two-sample difference scalograms (dCWT). Because of the linearity of CWT, this was equivalent to peak detection from the scalograms obtained directly from the average difference waveforms, without computation of single-trial CWTs.

To test how the τ -CWT method handles single ERP components, the above procedure was repeated with taking into account only the local maxima in the oddball τ -scalograms (P300 wave) and the local minima in the priming experiment (N400).

Additionally, the traditional area and peak measures of the P300 in the oddball and the N400 in the priming were obtained by taking the average and maximum

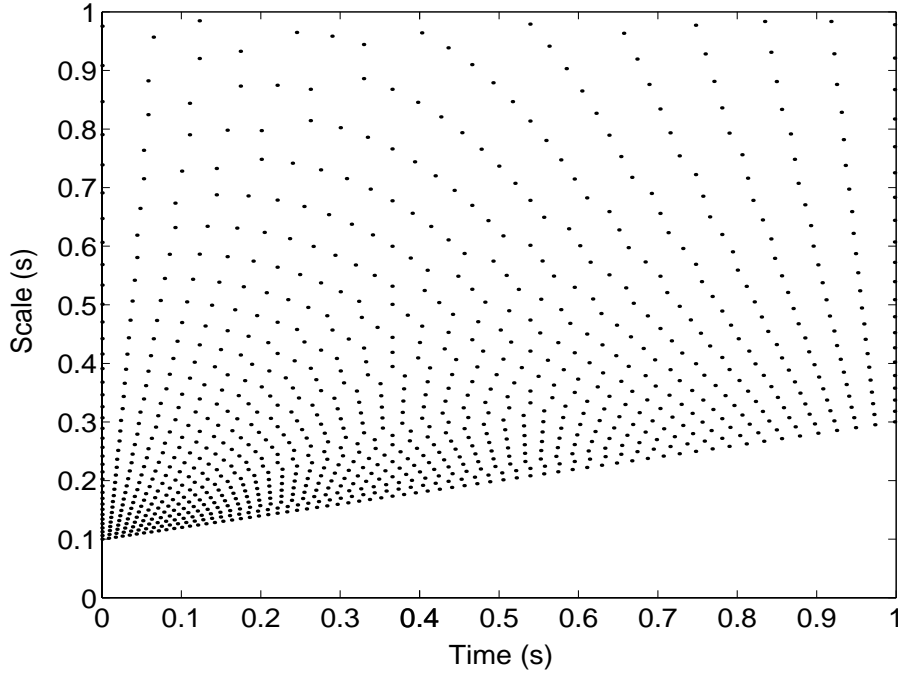


Figure 3.3: A CWT logarithmic grid with $s_{c0} = 200\text{ms}$, $\lambda = 2$ units per second and 12×12 PPSPO (for illustration purposes only – 24×24 PPSPO were used in the applications). Note that there are 12 sampling points between 0 and 0.1s at scale 0.1s, 12 sampling points between 0 and 0.2s at scale 0.2s and so on. There are $6=12/2$ sampling points between 0.5 and 0.7s at scale 0.1s according to the log time sampling formula. There are 12 rows of points between scale 0.1s and 0.2s, 12 rows of points between 0.2 and 0.4s and so on. The sampling point rows grow steeper with increasing scale in accordance with the log time sampling formula.

voltage values respectively in the time windows from 250ms to 350ms (P300) and from 200ms to 750ms (N400). These windows were determined by visual inspection of the grand-average waveforms.

Thus 9 sets of components were obtained for each participant in each of the two experiments: individual PCA, individual DWT, individual-t-CWT extrema, group-t-CWT extrema and group-dCWT extrema; individual-t-CWT maxima or minima, group-t-CWT maxima or minima, group area and group peak values. Note that the group measures were obtained from the group data, but applied on the individual data.

Statistical Analysis

Each of the individual-participant component sets was further reduced by PCA over all channels (from the total covariance matrix of the corresponding participant component set). The greatest eigenvalues that cleared 95% of the variance were retained. To compare the two experimental conditions, the two-sample T^2 -Hotelling test was performed on each of the 9 principal component sets for each participant. Additionally, the individual- τ -CWT method was incorporated in a *randomization test* (Edington 1987) with 2000 randomizations to correct the result of the Hotelling test for the accumulation of chance involved in the individual- τ -CWT procedure (preselection based on Student's t -test from the very same two samples which are compared).

The randomization test was performed as described in the following. The single trials of the given data set were randomly assigned to one of two samples of the same size as the original samples of the two experimental conditions. τ -CWT peak detection, PCA reduction and Hotelling test were performed as described above and the resulting p -value was stored. This procedure was repeated many times and finally the p -value obtained from the original samples was appended to the set of p -values obtained from the random samples. The percentage of p -values less than, or equal to, the original p -value was taken as the estimated probability that the difference between the two original samples had occurred by chance, i.e. that H_0 was true.

3.5 Results

3.5.1 Oddball

The grand average waveforms obtained from the oddball experiment are displayed in Figure 3.4. The corresponding dCWT and τ -CWT scalograms are displayed in

Figure 3.5 and Figure 3.6 respectively. The results from the statistical assessment of the individual-participant datasets are summarized in Table 3.1 and Table 3.2. Table 3.1 shows the results for the whole waveforms and Table 3.2 shows the results for the P300 component.

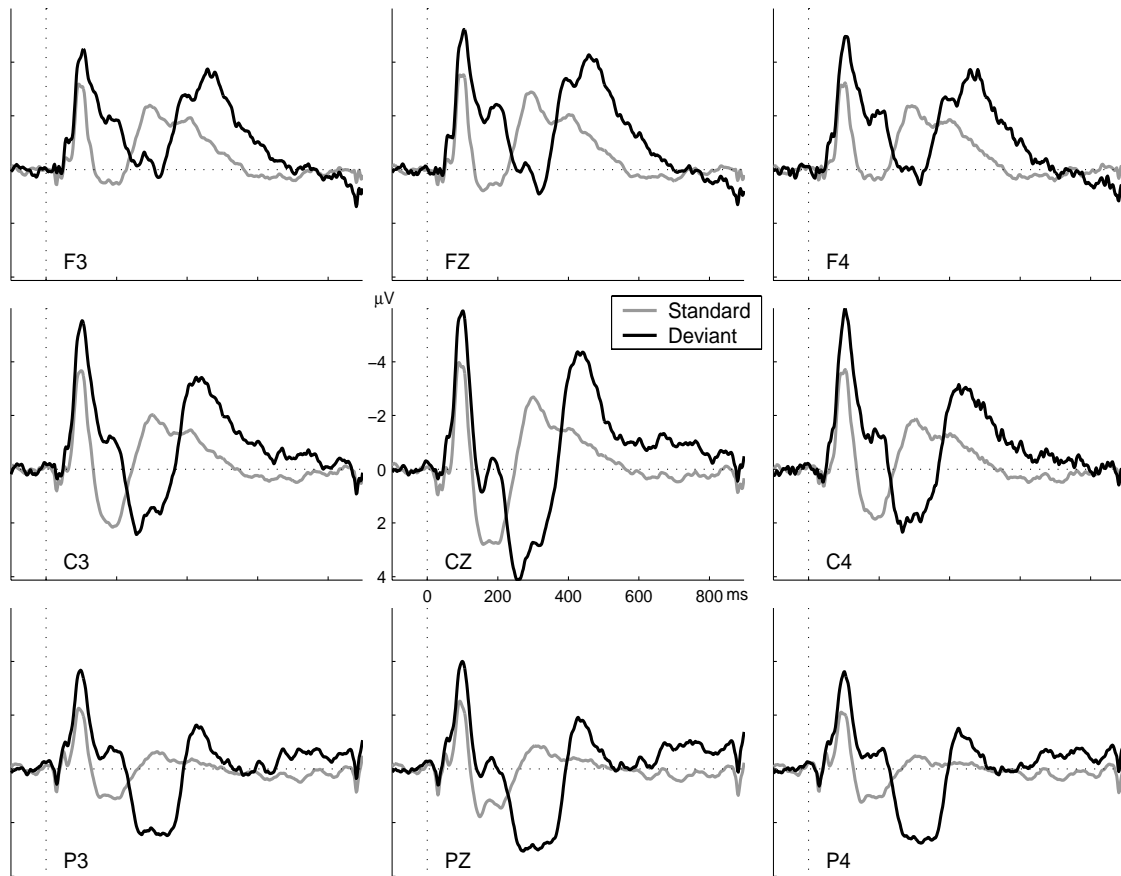


Figure 3.4: Grand average waveforms obtained in the oddball experiment. The following components in the ERP elicited by the deviant stimulus are readily discernable: N100 with increased amplitude, MMN with latency of about 200ms, P300 and late negativity with latency of about 450ms.

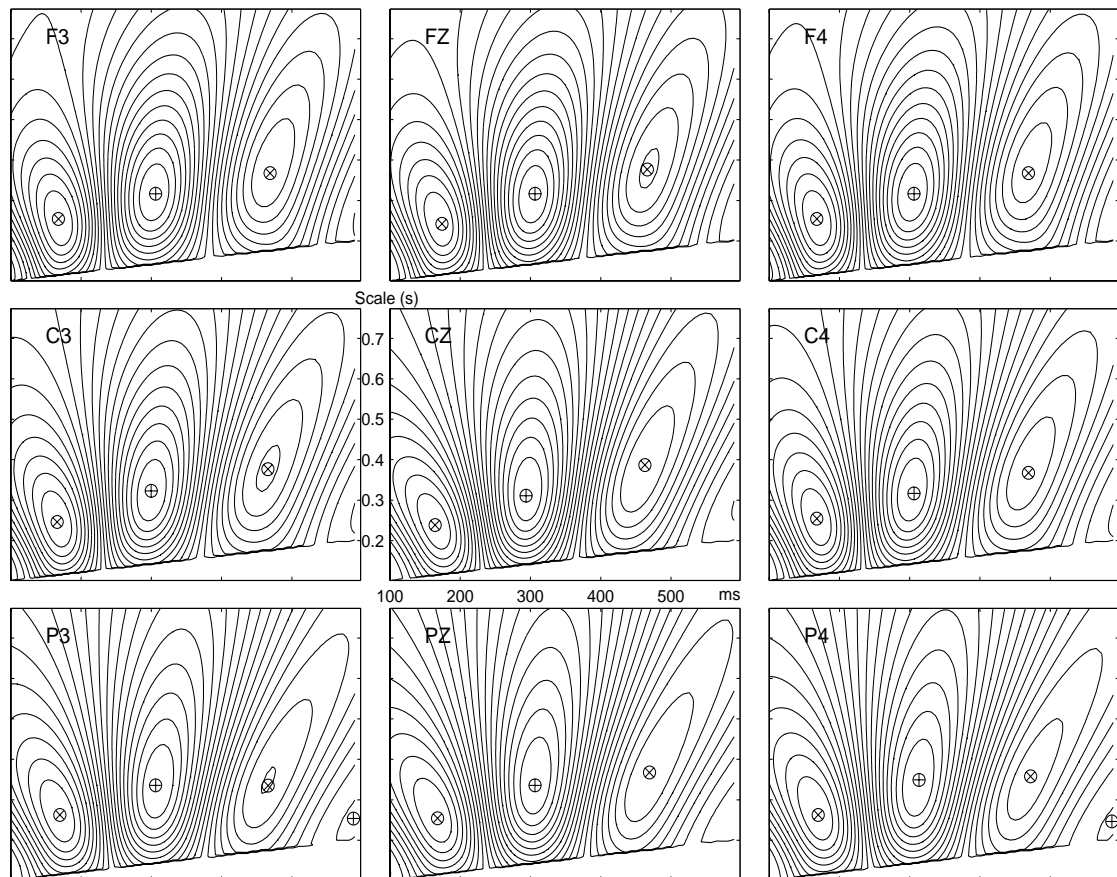


Figure 3.5: dCWT scalogram obtained in the oddball experiment. Local maxima are indicated by \oplus , local minima – by \otimes . The P300 maximum is detected at scale ≈ 320 ms.

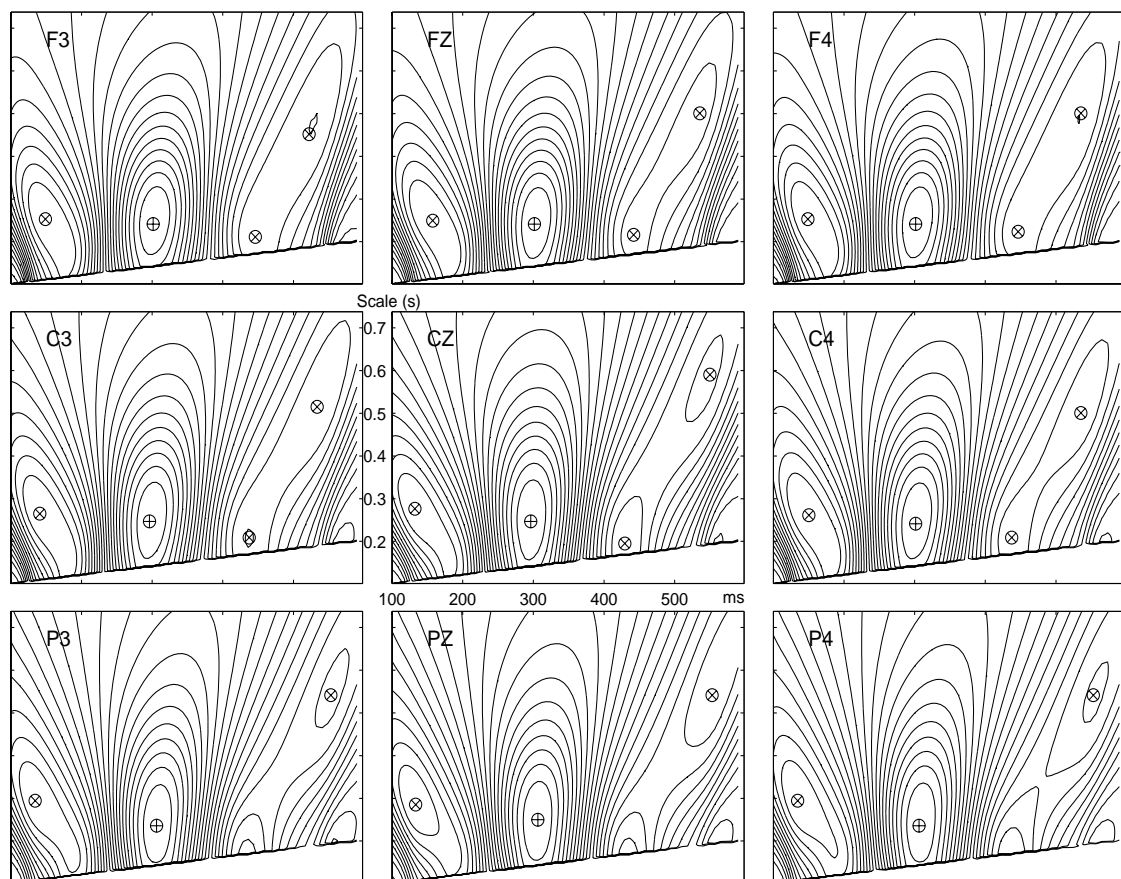


Figure 3.6: τ -CWT scalogram obtained in the oddball experiment. Local maxima are indicated by \oplus , local minima – by \otimes . The P300 maximum is detected at scale ≈ 250 ms. (note the difference to the dCWT scalogram).

Participant	<i>p</i> -values					
	group	group	individual	individual	individual	individual
	dCWT	t-CWT	PCA	DWT	t-CWT	t-CWT
	Hotelling	Hotelling	Hotelling	Hotelling	randomized	Hotelling ^{ac}
ALM	0.0069*	0.0001*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
BAG	0.0474*	0.0196*	0.0015*	0.0005*	0.0080*	0.0011* ^{ac}
BAS	0.0003*	0.0001*	0.0009*	0.0011*	0.0010*	0.0001* ^{ac}
BRA	0.0251*	0.0083*	0.0356*	0.0598	0.0105*	0.0020* ^{ac}
BRU	0.0213*	0.0445*	0.0145*	0.0161*	0.0055*	0.0004* ^{ac}
CIE	0.0010*	0.0009*	0.0001*	0.0001*	0.0060*	0.0005* ^{ac}
ENC	0.0004*	0.0001*	0.0006*	0.0007*	0.0010*	0.0001* ^{ac}
FEF	0.0542	0.0340*	0.2562	0.1576	0.0880	0.0171* ^{ac}
FIC	0.0591	0.0386*	0.0785	0.1155	0.1285	0.0287* ^{ac}
GIM	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
GOI	0.0002*	0.0004*	0.0000*	0.0000*	0.0015*	0.0000* ^{ac}
GUS	0.0888	0.0134*	0.1539	0.0336*	0.0050*	0.0008* ^{ac}
HAH	0.0147*	0.0207*	0.0073*	0.0112*	0.0115*	0.0014* ^{ac}
HEA	0.0005*	0.0000*	0.0028*	0.0041*	0.0030*	0.0003* ^{ac}
HII	0.0003*	0.0000*	0.0002*	0.0003*	0.0010*	0.0000* ^{ac}
JUE	0.1923	0.0278*	0.1744	0.1167	0.0135*	0.0018* ^{ac}
KAA	0.0001*	0.0002*	0.0005*	0.0005*	0.0020*	0.0002* ^{ac}
KAC	0.1750	0.1825	0.0141*	0.0175*	0.0100*	0.0007* ^{ac}
KUD	0.0004*	0.0004*	0.0003*	0.0004*	0.0005*	0.0002* ^{ac}
MAH	0.0013*	0.0022*	0.0004*	0.0003*	0.0005*	0.0000* ^{ac}
MAN	0.0000*	0.0006*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
MUV	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
NED	0.0154*	0.0060*	0.0088*	0.0152*	0.0040*	0.0003* ^{ac}
OTS	0.2936	0.2080	0.2421	0.2936	0.6960	0.3351 ^{ac}
RER	0.0014*	0.0002*	0.0003*	0.0007*	0.0055*	0.0009* ^{ac}
ROC	0.0001*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
ROM	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
SCA	0.1309	0.0013*	0.3814	0.2518	0.0075*	0.0008* ^{ac}
SCH	0.0002*	0.0001*	0.0000*	0.0001*	0.0005*	0.0000* ^{ac}
SCK	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
SCT	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
SCW	0.0055*	0.0021*	0.0001*	0.0001*	0.0010*	0.0000* ^{ac}
UMD	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
USD	0.0182*	0.0127*	0.0400*	0.0579	0.0115*	0.0007* ^{ac}
WIB	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*	0.0000* ^{ac}
ZIA	0.4431	0.1092	0.5548	0.5974	0.6495	0.3035 ^{ac}
Mean <i>p</i> -value	0.0444	0.0204	0.0547	0.0487	0.0466	0.0194 ^{ac}
Median <i>p</i> -value	0.0011	0.0005	0.0004	0.0005	0.0018	0.0002 ^{ac}
Number of *	28	33	29	28	32	34 ^{ac}
Percent *	78%	92%	81%	78%	89%	94% ^{ac}

* significant at level $\alpha = 0.05$.

^{ac} biased by accumulation of chance.

Table 3.1: Results of the statistical assessment of the whole waveforms obtained from the single-participant data sets from the oddball experiment.

Participant	<i>p</i> -values				
	group	group	group	individual	individual
	area P300 Hotelling	peak P300 Hotelling	t -CWT P300 Hotelling	t -CWT P300 randomized	t -CWT P300 Hotelling ^{ac}
ALM	0.1027	0.1536	0.0009*	0.0006*	0.0000* ^{ac}
BAG	0.3378	0.0649	0.0007*	0.0006*	0.0001* ^{ac}
BAS	0.0858	0.2094	0.0000*	0.0006*	0.0000* ^{ac}
BRA	0.1143	0.0584	0.0102*	0.0062*	0.0008* ^{ac}
BRU	0.9931	0.9102	0.0225*	0.0364*	0.0033* ^{ac}
CIE	0.3749	0.3845	0.0001*	0.0006*	0.0000* ^{ac}
ENC	0.0325*	0.0206*	0.0000*	0.0012*	0.0000* ^{ac}
FEF	0.0998	0.4819	0.0392*	0.1148	0.0204* ^{ac}
FIC	0.8464	0.9328	0.0116*	0.0519	0.0059* ^{ac}
GIM	0.0304*	0.0000*	0.0000*	0.0006*	0.0000* ^{ac}
GOI	0.0120*	0.0033*	0.0001*	0.0006*	0.0000* ^{ac}
GUS	0.6336	0.7215	0.0040*	0.0062*	0.0010* ^{ac}
HAH	0.6274	0.0835	0.0011*	0.0062*	0.0004* ^{ac}
HEA	0.8116	0.5949	0.0003*	0.0037*	0.0002* ^{ac}
HII	0.0076*	0.0007*	0.0000*	0.0006*	0.0000* ^{ac}
JUE	0.5351	0.6546	0.3189	0.0204*	0.0024* ^{ac}
KAA	0.0440*	0.0503	0.0000*	0.0006*	0.0000* ^{ac}
KAC	0.0297*	0.2188	0.0646	0.0086*	0.0004* ^{ac}
KUD	0.0438*	0.0490*	0.0001*	0.0006*	0.0000* ^{ac}
MAH	0.0016*	0.0602	0.0002*	0.0006*	0.0000* ^{ac}
MAN	0.0042*	0.2377	0.0000*	0.0006*	0.0000* ^{ac}
MUV	0.0001*	0.0001*	0.0000*	0.0006*	0.0000* ^{ac}
NED	0.2269	0.2940	0.0009*	0.0031*	0.0001* ^{ac}
OTS	0.6077	0.6089	0.1219	0.6241	0.2374 ^{ac}
RER	0.0321*	0.0169*	0.0000*	0.0012*	0.0001* ^{ac}
ROC	0.0113*	0.1233	0.0000*	0.0006*	0.0000* ^{ac}
ROM	0.0016*	0.0137*	0.0000*	0.0006*	0.0000* ^{ac}
SCA	0.1158	0.6977	0.0097*	0.0025*	0.0000* ^{ac}
SCH	0.0058*	0.0942	0.0000*	0.0006*	0.0000* ^{ac}
SCK	0.0003*	0.0001*	0.0000*	0.0006*	0.0000* ^{ac}
SCT	0.5238	0.0569	0.0000*	0.0006*	0.0000* ^{ac}
SCW	0.0027*	0.0029*	0.0649	0.0025*	0.0001* ^{ac}
UMD	0.0000*	0.0005*	0.0000*	0.0006*	0.0000* ^{ac}
USD	0.7263	0.9817	0.0213*	0.0154*	0.0012* ^{ac}
WIB	0.0011*	0.0079*	0.0000*	0.0006*	0.0000* ^{ac}
ZIA	0.4801	0.0331*	0.0768	0.5000	0.1608 ^{ac}
Mean <i>p</i> -value	0.2362	0.2451	0.0214	0.0393	0.0121 ^{ac}
Median <i>p</i> -value	0.0649	0.0742	0.0002	0.0006	0.0000 ^{ac}
Number of *	18	13	31	32	34 ^{ac}
Percent *	50%	36%	86%	89%	94% ^{ac}

* significant at level $\alpha = 0.05$.

^{ac} biased by accumulation of chance.

Table 3.2: Results of the statistical assessment of the P300 measures obtained from the single-participant data sets from the oddball experiment. The t -CWT measure shows clear superiority to the classical area and peak measures.

3.5.2 Priming

The grand average waveforms obtained from the priming experiment are displayed in Figure 3.7. The corresponding dCWT and τ -CWT scalograms are displayed in Figure 3.8 and Figure 3.9 respectively. The results from the statistical assessment of the individual-participant datasets are summarized in Tables 3.3 and 3.4. Table 3.3 shows the results for the whole waveforms and Table 3.4 shows the results for the N400 component.

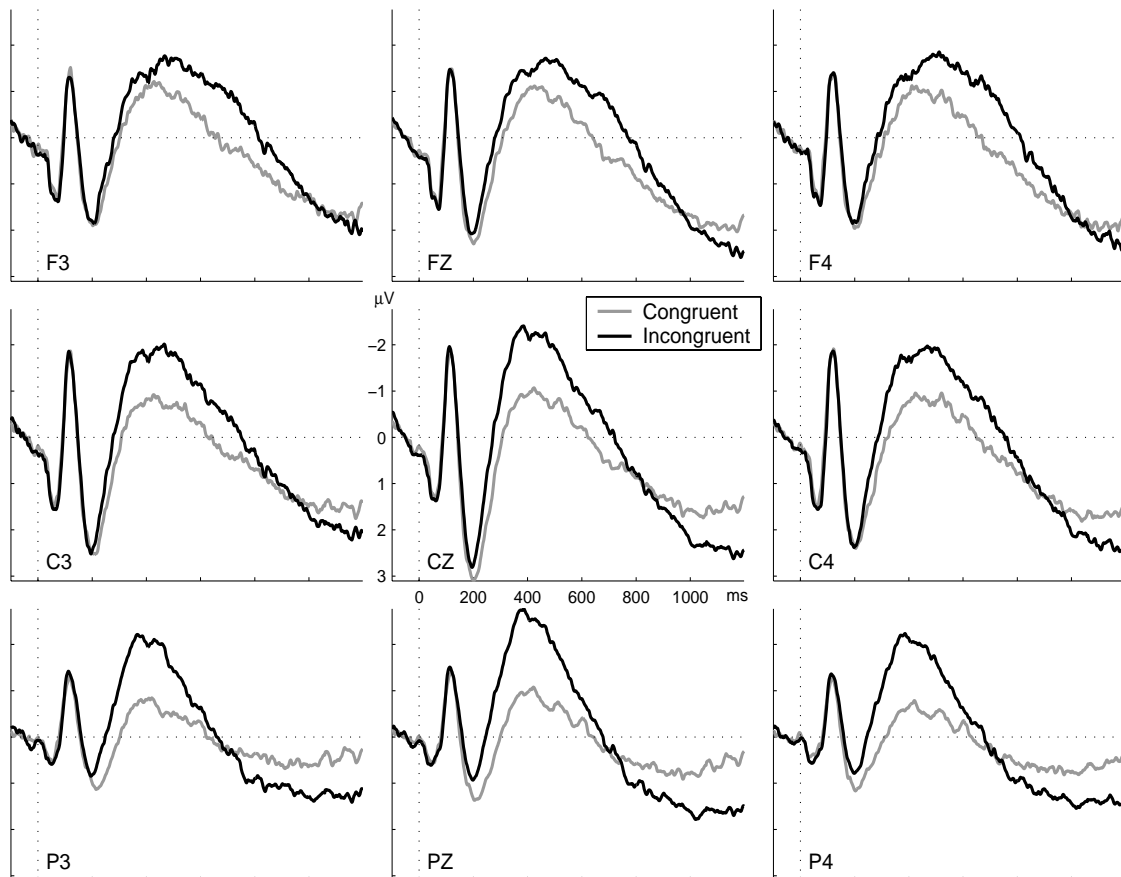


Figure 3.7: Grand average waveforms obtained in the priming experiment. Besides the N400, there is also a late positive wave elicited by the incongruous words. It can also be interpreted as an indicator that the incongruous stimuli were successfully recognized as words and their meaning was found inconsistent with the preceding words (first in pair).

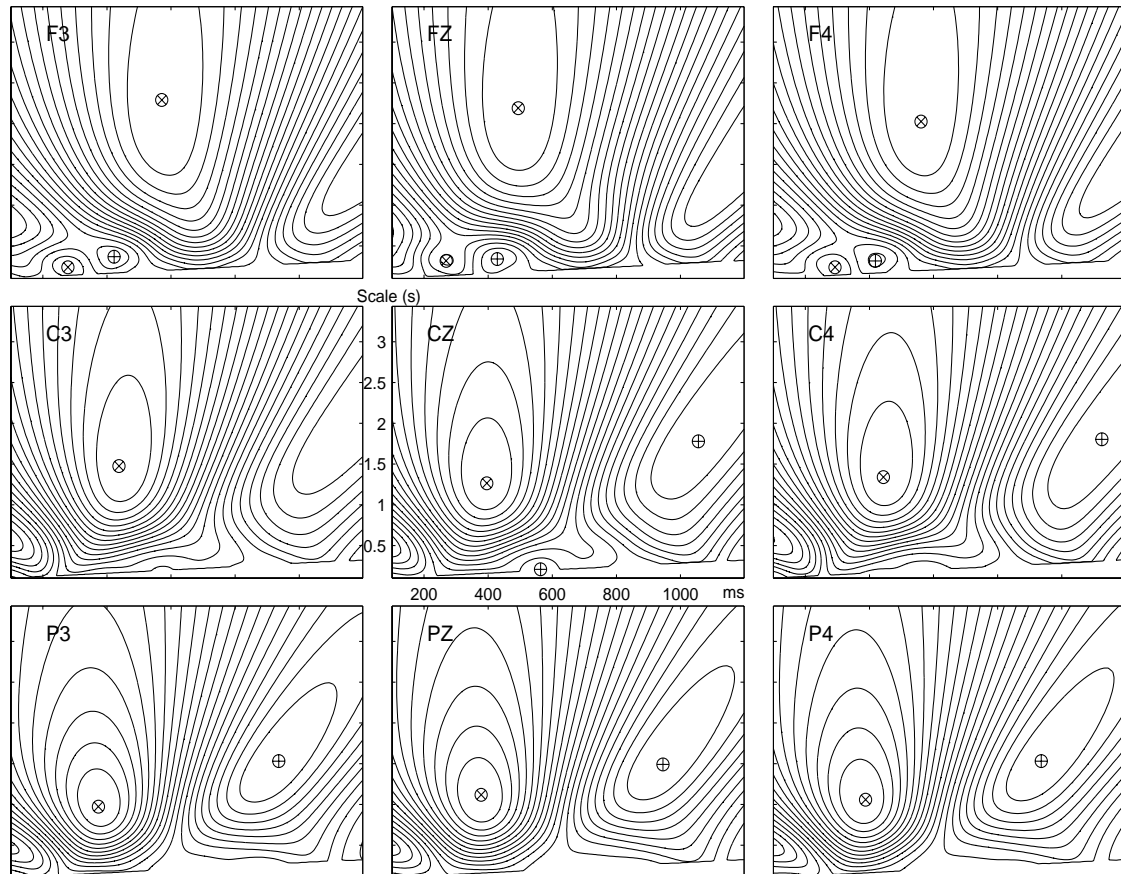


Figure 3.8: dCWT scalogram obtained in the priming experiment. Local maxima are indicated by \oplus , local minima – by \otimes . The late N400 minimum is detected at scale ≈ 2200 ms and time ≈ 450 ms at frontal sites. At central and parietal sites, the scale is smaller and the latency is approximately 400ms

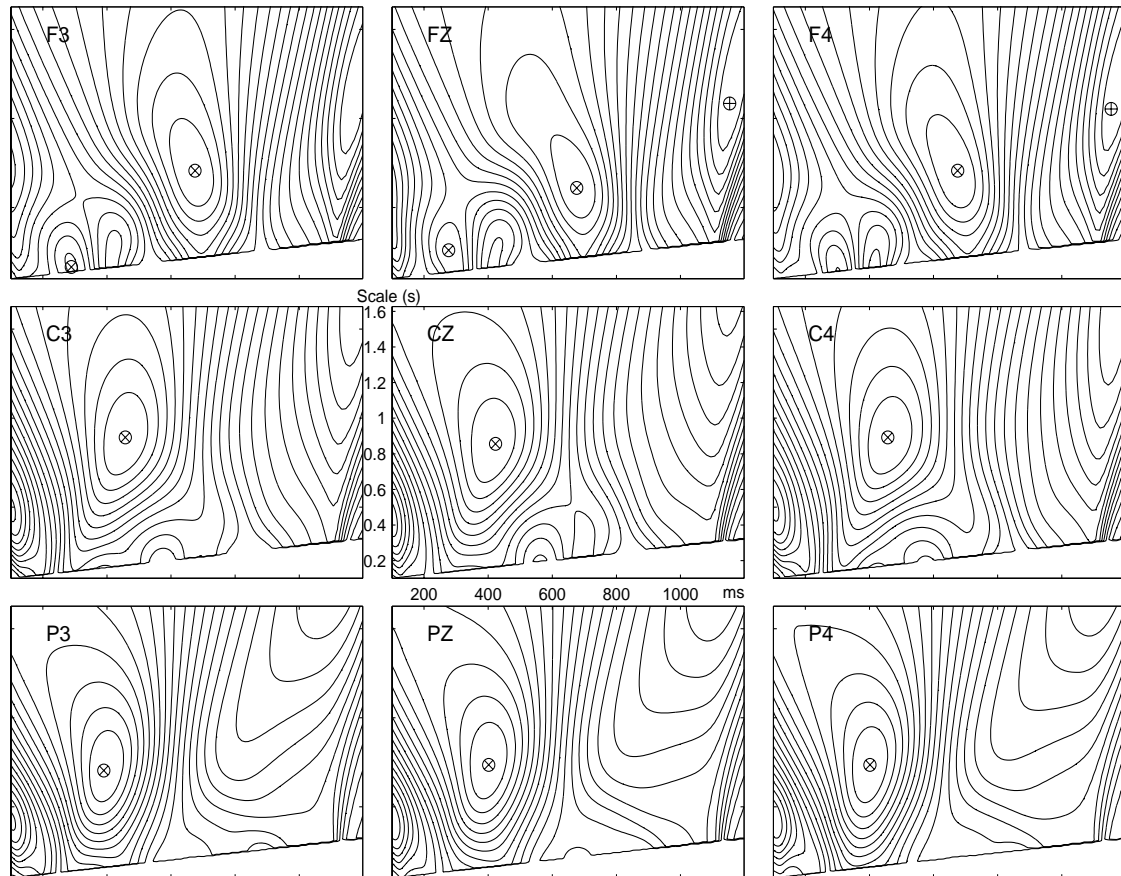


Figure 3.9: τ -CWT scalogram obtained in the priming experiment. Local maxima are indicated by \oplus , local minima – by \otimes . Note how different this scalogram is from the dCWT scalogram. The late N400 minimum is detected at scale ≈ 600 ms and time ≈ 650 ms at frontal sites. Also at central and parietal sites, the N400 scale is some 500ms smaller than the one obtained from the dCWT scalogram. The late positive wave is represented by a frontal maximum, while at the dCWT scalogram it is represented by central and parietal maxima with 200ms shorter latency.

Participant	<i>p</i> -values					
	group	group	individual	individual	individual	individual
	dCWT	t-CWT	PCA	DWT	t-CWT	t-CWT
	Hotelling	Hotelling	Hotelling	Hotelling	randomized	Hotelling ^{ac}
ALM	0.0070*	0.0047*	0.0316*	0.0438*	0.0543	0.0049* ^{ac}
BAG	0.0000*	0.0000*	0.0014*	0.0022*	0.0008*	0.0000* ^{ac}
BAS	0.0010*	0.0010*	0.0001*	0.0002*	0.0008*	0.0000* ^{ac}
BRA	0.0007*	0.0005*	0.0007*	0.0018*	0.0026*	0.0003* ^{ac}
BRU	0.0001*	0.0001*	0.0001*	0.0002*	0.0008*	0.0000* ^{ac}
CIE	0.5783	0.0463*	0.6021	0.6882	0.4823	0.1337 ^{ac}
ENC	0.1734	0.0354*	0.0409*	0.0345*	0.0139*	0.0007* ^{ac}
FEF	0.3728	0.4216	0.1331	0.1655	0.0398*	0.0044* ^{ac}
FIC	0.0005*	0.0044*	0.0019*	0.0006*	0.0012*	0.0000* ^{ac}
GIM	0.2121	0.1436	0.0012*	0.0016*	0.0024*	0.0001* ^{ac}
GOI	0.1251	0.0129*	0.0274*	0.0341*	0.0018*	0.0001* ^{ac}
GUS	0.0501	0.0278*	0.0088*	0.0037*	0.0006*	0.0001* ^{ac}
HAH	0.8870	0.6353	0.8664	0.9252	0.8928	0.4572 ^{ac}
HEA	0.0835	0.0363*	0.3024	0.2755	0.4446	0.1187 ^{ac}
HII	0.5552	0.0343*	0.0710	0.1544	0.0247*	0.0035* ^{ac}
JUE	0.2426	0.1874	0.3390	0.3271	0.0193*	0.0016* ^{ac}
KAA	0.0011*	0.0023*	0.0048*	0.0072*	0.0024*	0.0001* ^{ac}
KAC	0.0446*	0.0034*	0.0050*	0.0100*	0.0036*	0.0001* ^{ac}
KUD	0.0083*	0.0016*	0.0245*	0.0202*	0.0066*	0.0001* ^{ac}
MAH	0.0136*	0.0345*	0.1172	0.1225	0.1813	0.0271* ^{ac}
MAN	0.0089*	0.0000*	0.0000*	0.0001*	0.0006*	0.0000* ^{ac}
MUV	0.0000*	0.0000*	0.0000*	0.0000*	0.0012*	0.0000* ^{ac}
NED	0.0109*	0.0060*	0.0327*	0.0474*	0.0145*	0.0013* ^{ac}
OTS	0.0247*	0.0783	0.0867	0.0556	0.0795	0.0079* ^{ac}
RER	0.0033*	0.0187*	0.0281*	0.0489*	0.1946	0.0332* ^{ac}
ROC	0.2589	0.0076*	0.0005*	0.0008*	0.0006*	0.0000* ^{ac}
ROM	0.0173*	0.0077*	0.0293*	0.0333*	0.0241*	0.0021* ^{ac}
SCA	0.0043*	0.0130*	0.0152*	0.0144*	0.0145*	0.0010* ^{ac}
SCH	0.0088*	0.0226*	0.0391*	0.0834	0.1054	0.0158* ^{ac}
SCK	0.0723	0.0086*	0.0324*	0.0361*	0.0066*	0.0004* ^{ac}
SCT	0.0029*	0.0089*	0.0004*	0.0001*	0.0018*	0.0001* ^{ac}
SCW	0.1131	0.0349*	0.0414*	0.0181*	0.0193*	0.0020* ^{ac}
UMD	0.0000*	0.0000*	0.0001*	0.0001*	0.0006*	0.0000* ^{ac}
USD	0.0187*	0.0110*	0.0782	0.0438*	0.0018*	0.0000* ^{ac}
WIB	0.0000*	0.0000*	0.0000*	0.0001*	0.0006*	0.0000* ^{ac}
ZIA	0.0048*	0.0019*	0.1188	0.1560	0.0127*	0.0011* ^{ac}
Mean <i>p</i> -value	0.1085	0.0515	0.0856	0.0932	0.0737	0.0227 ^{ac}
Median <i>p</i> -value	0.0122	0.0088	0.0277	0.0267	0.0066	0.0003 ^{ac}
Number of *	23	31	26	26	28	33 ^{ac}
Percent *	64%	86%	72%	72%	78%	92% ^{ac}

* significant at level $\alpha = 0.05$.

^{ac} biased by accumulation of chance.

Table 3.3: Results of the statistical assessment of the whole waveforms obtained from the single-participant data sets from the priming experiment.

Participant	p -values				
	group	group	group	individual	individual
	area N400 Hotelling	peak N400 Hotelling	t -CWT N400 Hotelling	t -CWT N400 randomized	t -CWT N400 Hotelling ^{ac}
ALM	0.0394*	0.1385	0.0061*	0.0752	0.0050* ^{ac}
BAG	0.1443	0.3586	0.0000*	0.0011*	0.0000* ^{ac}
BAS	0.0879	0.3416	0.0019*	0.0006*	0.0000* ^{ac}
BRA	0.4871	0.9084	0.0002*	0.0025*	0.0001* ^{ac}
BRU	0.2538	0.5599	0.0001*	0.0012*	0.0000* ^{ac}
CIE	0.7900	0.8957	0.1043	0.0739	0.0058* ^{ac}
ENC	0.5881	0.7919	0.0324*	0.0205*	0.0006* ^{ac}
FEF	0.2615	0.3372	0.4319	0.0224*	0.0019* ^{ac}
FIC	0.2443	0.0856	0.0024*	0.0012*	0.0000* ^{ac}
GIM	0.6257	0.7642	0.1323	0.0019*	0.0000* ^{ac}
GOI	0.3250	0.2363	0.0067*	0.0348*	0.0025* ^{ac}
GUS	0.6315	0.9091	0.0159*	0.0006*	0.0000* ^{ac}
HAH	0.6408	0.9444	0.6009	0.8534	0.3172 ^{ac}
HEA	0.0380*	0.1893	0.0604	0.4323	0.0817 ^{ac}
HII	0.8730	0.8813	0.0333*	0.0714	0.0054* ^{ac}
JUE	0.2925	0.7476	0.1284	0.1789	0.0185* ^{ac}
KAA	0.1650	0.8975	0.0020*	0.0019*	0.0000* ^{ac}
KAC	0.6564	0.6233	0.0094*	0.0068*	0.0003* ^{ac}
KUD	0.9414	0.8971	0.0010*	0.0143*	0.0005* ^{ac}
MAH	0.1539	0.4540	0.0206*	0.1783	0.0163* ^{ac}
MAN	0.0150*	0.0670	0.0000*	0.0006*	0.0000* ^{ac}
MUV	0.0767	0.1860	0.0000*	0.0050*	0.0002* ^{ac}
NED	0.2475	0.0585	0.0031*	0.0112*	0.0008* ^{ac}
OTS	0.0088*	0.0562	0.0553	0.0795	0.0055* ^{ac}
RER	0.0403*	0.0176*	0.0028*	0.1329	0.0132* ^{ac}
ROC	0.2133	0.1559	0.0808	0.0087*	0.0003* ^{ac}
ROM	0.4579	0.5996	0.0040*	0.1043	0.0080* ^{ac}
SCA	0.6593	0.1614	0.0085*	0.0217*	0.0011* ^{ac}
SCH	0.5718	0.9044	0.0185*	0.0335*	0.0027* ^{ac}
SCK	0.8701	0.9855	0.0238*	0.2944	0.0348* ^{ac}
SCT	0.0041*	0.0495*	0.0085*	0.0180*	0.0007* ^{ac}
SCW	0.8032	0.7498	0.1296	0.0807	0.0063* ^{ac}
UMD	0.0093*	0.1077	0.0000*	0.0006*	0.0000* ^{ac}
USD	0.0160*	0.0063*	0.0065*	0.0012*	0.0000* ^{ac}
WIB	0.0011*	0.0001*	0.0000*	0.0006*	0.0000* ^{ac}
ZIA	0.0281*	0.0317*	0.0011*	0.0211*	0.0008* ^{ac}
Mean p -value	0.3406	0.4472	0.0537	0.0774	0.0147 ^{ac}
Median p -value	0.2507	0.3501	0.0076	0.0193	0.0007 ^{ac}
Number of *	10	05	27	24	34 ^{ac}
Percent *	28%	14%	75%	67%	94% ^{ac}

* significant at level $\alpha = 0.05$.

^{ac} biased by accumulation of chance.

Table 3.4: Results of the statistical assessment of the N400 measures obtained from the single-participant data sets from the priming experiment. The t -CWT measure shows clear superiority to the classical area and peak measures.

3.6 Discussion

By the study presented in this chapter, it was demonstrated how the \mathfrak{t} -CWT feature extraction method originally developed for BCI applications can be applied in the statistical assessment and significance testing of ERP components. The \mathfrak{t} -CWT method produced generally larger effects (reflected by smaller p -values) than the other methods it was compared with. This result was not unexpected, since the \mathfrak{t} -CWT was specially designed to extract features that best discriminate between the ERPs elicited in two different experimental conditions. To the best of our knowledge, this is the first ERP component detection and quantification method that is explicitly based on such criteria. Most other methods, which were briefly reviewed above, detect ERP components from one of the two average waveforms or from the difference waveform and usually do not take the variance into account. The PCA can be adopted to take into account both the average differences and the variance, by applying it to a sample of individual participants' average difference waveforms, but then it cannot use the within-participant variance contained in the single-trials and hence cannot be applied to individual data. The \mathfrak{t} -CWT uses both the mean difference and the total variance and thus utilizes the whole information contained in the single-trials and does not lose statistical test power by sample size reduction. In this study, the power of the \mathfrak{t} -CWT was tested with individual ERP data, but it can be applied in group assessments as well – the individual participants' data are simply pooled into one large data set, which is then processed according to the same procedures that was demonstrated on the individual data sets. Moreover, single ERP components detected and quantified with the \mathfrak{t} -CWT may be analyzed further by standard ANOVA procedures instead of being subjected to the Hotelling test.

The statistical test power of a given assessment method can be roughly estimated by the percentage of participants who showed a significant ERP effect as assessed with this method. The \mathfrak{t} -CWT method proved consistently more powerful than the

other methods tested here. Whether the whole waveforms or single ERP components were compared, the largest number of significant individual-participant effects was obtained with the τ -CWT method. It is important to emphasize the superiority of the τ -CWT to the dCWT, which demonstrates the importance of taking into account the variance for proper detection of components that best represent the differences between two ERPs.

In accord with the results from our previous study (see Chapter 2), the classical area and peak measures yielded much bigger p -values and much fewer significant effects than the wavelet measures (Tables 3.2 and 3.4) and thus proved quite inappropriate for single-trial analyses. The PCA and the DWT provided very similar results in terms of p -values and number of significant effects (Tables 3.3 and 3.1). In both experiments, the percent of individual participants' responses found significant with the τ -CWT method was higher than that obtained with the PCA and the DWT (Tables 3.3 and 3.1). However the superiority of the τ -CWT was not as pronounced as its advantage over the area and peak measures.

Both the individual- and the group-features based assessments of the whole waveforms in the priming experiment yielded a higher percentage of significant responses than the corresponding assessments of the N400 alone (Tables 3.3 and 3.4). Comparing the whole ERPs instead of single components certainly makes sense for the paradigms of the context-violation class, where any significant difference between the ERPs elicited by congruent and incongruent stimuli has a clear interpretation as reflecting successfully accomplished cognitive processes of stimulus recognition and context comprehension.

The present study addressed the issue of accumulation of chance, which is particularly important in the assessment of individual participants' data without any reference to previous group data. If the ERP components are detected from the difference waveforms of the same data samples, which are then subjected to significance testing, some kind of a correction should be performed to compensate for

the accumulation of chance involved in the procedure. This problem was solved by adopting a randomization test. The rightmost columns of Tables 3.1, 3.2, 3.3, and 3.4 show the underestimated p -values and the corresponding overestimated numbers of significant effects. The randomization test provides unbiased estimations of these quantities, but it is time consuming. An alternative is to use half of the trials for ERP detection and the other half for significance testing. This method has the advantage that it is much simpler and faster than the randomization test. It suffers, however, a loss of test power by the sample size reduction by half. Nevertheless, in group assessments, where a large number of trials from many participants is available, this much simpler and faster solution may be preferred to the randomization test.

In the present study, the focus was placed on the assessment of individual participants' ERP data, because this is exactly the case, in which single-trial analysis is indispensable, the data samples are small, and individual differences in ERP structure may be accounted for by individual component detection. The assessment of ERPs of single individuals is particularly important for clinical applications and ERP-based diagnostics of cognitive functioning. Since the ERPs of neurological patients may be quite different from those of the healthy population (Kotchoubey et al. 2002), the version of the τ -CWT method that detects ERP components from the individual EEG without using any other data is particularly suitable for diagnostic purposes.

A slight modification of the τ -CWT using the Morlet wavelet instead of the Mexican Hat (Ende et al. 1998) allows extraction of spectral features like e.g. phase-locked ("evoked") gamma activity (Senkowski and Herrmann 2002). For the case of non-phase-locked ("induced") activity, a complex wavelet should be used and the absolute value of the complex covariance (3.1) should be taken at every time-scale point of the single-trial CWT. More generally, the τ -CWT may be applied to other psychophysiological signals for extracting a pattern that best discriminates between two samples of measurement data.

Apart from its test power, the τ -CWT method is attractive with its conceptual simplicity, intuitive plausibility, scalogram visualization and clear interpretation of the extrema of the τ -value scalogram in terms of ERP components.

Thus far, only the advantages of the τ -CWT have been discussed. There is, however, at least one important limitation of the method, which should be mentioned. According to Heisenberg's uncertainty relation, a wave's good localization in the time domain is inevitably at the cost of poorer localization in the frequency domain and vice versa. Using the Mexican Hat wavelet for ERP component detection necessarily implies a compromise between scale separation and time localization. The point is that, while the central maximum of the Mexican Hat fits a given component, the two lateral minima fit adjacent components of the same scale and opposite polarity. Thus the obtained measure of the central component unavoidably contains some information about the amplitudes of the adjacent components. According to Heisenberg's uncertainty relation, there is no possible remedy of this situation. The time domain measures, area and peak, separate well the component of interest from the adjacent components, but produce quite poor results in single trials analyses, in which scale separation is crucial for proper component detection. This problem, however, ceases to exist when the comparison of the whole ERP waveforms, regardless of specific components, is addressed by the τ -CWT method.

Chapter 4

The τ -CWT as a Pattern Recognition Method in BCIs

In the study presented in this chapter, the τ -CWT method introduced in Chapter 2 was applied in two BCI paradigms – P300 speller and Slow Cortical Potential (SCP) feedback. This approach to BCI feature extraction and classification was validated in the International BCI Classification Contest 2003, where the τ -CWT method was a winner (provided best classification) on two ERP data sets acquired in these two paradigms. In the P300 speller paradigm the method provided results, which were as good as those obtained from simple and redundant features with a very powerful classifier based on machine learning – the Support Vector Machines (SVM). The τ -CWT method has the advantage that it is very simple, intuitively plausible, readily visualizable and the extracted features have clear interpretation as ERP components.

4.1 Introduction

BRAIN-COMPUTER INTERFACES (BCI) belong to the most innovative and exciting applications of ERP research. They allow completely paralyzed people to communicate with the world solely by means of their brain waves (Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey, Kübler, Perelmouter, Taub, and Flor 1999; Kübler, Kotchoubey, Kaiser, Wolpaw, and Birbaumer 2001). One bit of information is transferred from the brain to the computer by the discrimination between two different ERPs. This process includes feature extraction from the raw signals of

the electroencephalogram (EEG). The features that distinguish the ERPs may be extracted in the time, frequency or time-frequency domain.

In a standard BCI paradigm there are two data sets – the training data set and the test data set. Each data set contains two groups of ERP trials, reflecting two different brain responses to certain stimuli. Each trial comprises the EEG signals acquired from one or more scalp positions (channels). In the training data set the two groups of trials are known – each trial carries the label of the group, to which it belongs. The trials in the test data set are not labeled and have to be separated into groups according to the pattern that discriminates between the groups of the training set. The success of the classification generally depends on two factors. The first factor is how well the distinguishing pattern is represented by the features extracted from the test data. The second factor is the power of the classification method. It should be stressed that the two factors have independent impacts on classification results. Powerful classifiers like, e.g., Support Vector Machines (SVM), may yield very good results even using simple and redundant features (Meinicke, Kaper, Hoppe, Manfred, and Ritter 2003; Kaper, Meinicke, Grossekaethoefer, Lingner, and Ritter 2004). The present study shows that, contrariwise, efficient feature extraction may produce excellent results even with the most simple classification method, the classical Linear Discriminant Analysis (LDA) (Rencher 1998, pp. 230–231).

The most popular ERP features are the *areas* and the *peaks* of the ERP components (waves) defined by the mean and extremum voltage respectively computed in certain windows in the time domain determined by visual inspection of the ERP waveforms averaged across many trials (see Chapters 1 and 2). Although the quantification of ERP components by areas and peaks is the standard procedure in fundamental ERP research, these features are quite inappropriate for BCI purposes, because they are too rough and yield poor results in single-trial analysis (see Chapter 3). An alternative approach is to take simply *all* voltage values of the band-pass filtered signals as time domain features and feed them into a powerful classifier based on *machine learning* like SVM. This technique has proved extremely efficient in BCI

applications (Meinicke, Kaper, Hoppe, Manfred, and Ritter 2003; Kaper, Meinicke, Grossekaehler, Lingner, and Ritter 2004).

The τ -CWT method for detection and quantification ERP components, which was introduced in Chapter 3, has been originally developed as a feature extraction method for application in BCI paradigms. The efficiency of the method was demonstrated in the International BCI Classification Contest 2003 (Blankertz, Müller, Curio, Vaughan, Schalk, Wolpaw, Schlögl, Neuper, Pfurtscheller, Hinterberger, Schröder, and Birbaumer 2004), where it was a winner on two of the ERP data sets obtained in two different BCI paradigms – P300 speller (Farwell and Donchin 1988) and self-regulation of slow cortical potentials (SCPs) (Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey, Kübler, Perelmouter, Taub, and Flor 1999). The results are presented in this chapter.

4.2 BCI Study 1: P300 Speller

The P300 speller data sets were provided by the BCI group at the Wadsworth Center, Albany, NY, USA. This dataset represents a complete record of P300 evoked potentials recorded with BCI2000¹ using a paradigm described by Donchin, Spencer, and Wijensinghe 2000 and originally by Farwell and Donchin 1988. In this paradigm, the user is presented with a character display and instructed to focus attention on a particular character (target) of her/his choice. The characters start to flash randomly (see below for details) and each time the chosen character is illuminated, the user's brain reacts with an ERP response which is different from the one elicited by the illumination of nontarget characters. The experiment is so designed that each character in the display flashes exactly 2 times every 12 trials (see below). After every 180 trials (15 12-trial blocks) the computer performs a classification of the

¹BCI2000 is a flexible Brain-Computer Interface research and development platform. It supports a variety of brain signals, signal processing methods, and user applications, and is available free of charge for research purposes at <http://www.bci2000.org>.

single-trial ERPs into two groups – target and nontarget responses – according to a pattern extracted from previous (training) data acquired from the same user. The character, which has elicited the most pronounced target responses is designated a target character and written on the computer screen.

The training data is obtained in a slightly different design: at the beginning of each 180-trial block, the user is instructed which character she/he should focus at. Thus, the real target stimuli and consequently, the real target responses are known in the training data set, which allows the extraction of the pattern distinguishing target from nontarget ERPs. The most prominent component of this pattern is the P300 wave with peak latency $t \approx 300\text{ms}$ poststimulus and duration $s \approx 200\text{ms}$ (Figure 4.3).

4.2.1 Experimental Procedure

This description of the experimental design of the P300 speller paradigm follows closely the one provided by the organizers of the International BCI Classification Contest 2003 from the Wadsworth Center, Albany, NY, USA (Blankertz et al. 2004).

The user was presented with a 6 by 6 matrix of characters (see Figure 4.1). The users task was to focus attention on characters in a word that was prescribed by the investigator (i.e., one character at a time). All rows and columns of this matrix were successively and randomly illuminated at a rate of 5.7Hz. Two out of 12 illuminations of rows or columns contained the desired character (i.e., one particular row and one particular column). The responses evoked by these infrequent stimuli (i.e., the 2 out of 12 stimuli that did contain the desired character) are different from those evoked by the stimuli that did not contain the desired character and they are similar to the P300 responses previously reported (Farwell and Donchin 1988; Donchin, Spencer, and Wijensinghe 2000).

The EEG signals (digitized at 240Hz) were collected from 64 scalp positions (Fig-



Figure 4.1: The character matrix display in the P300 speller paradigm. In this example, the users task is to spell the word "SEND".

ure 4.2), from one participant in three sessions. The first two sessions were used as training data for feature extraction and the third session was used as a test data set for classification. Each session consisted of a number of runs. In each run, the participant focused attention on a series of characters forming one word. For each character, user display was as follows: the matrix was displayed for a 2.5 s period, and during this time each character had the same intensity (i.e., the matrix was blank). Subsequently, each row and column in the matrix was randomly illuminated for 100ms (i.e., resulting in 12 different stimuli – 6 rows and 6 columns). (After illumination of a row/column, the matrix was blank for 75ms.) Row/column illuminations were block randomized in blocks of 12. Blocks of 12 illuminations were repeated 15 times for each character (i.e., any specific row/column was illuminated 15 times and thus there were 180 total illuminations for each character). Each sequence of 15 blocks of illuminations was followed by a 2.5 s period, and during this time the matrix was blank. This period informed the user that this character was

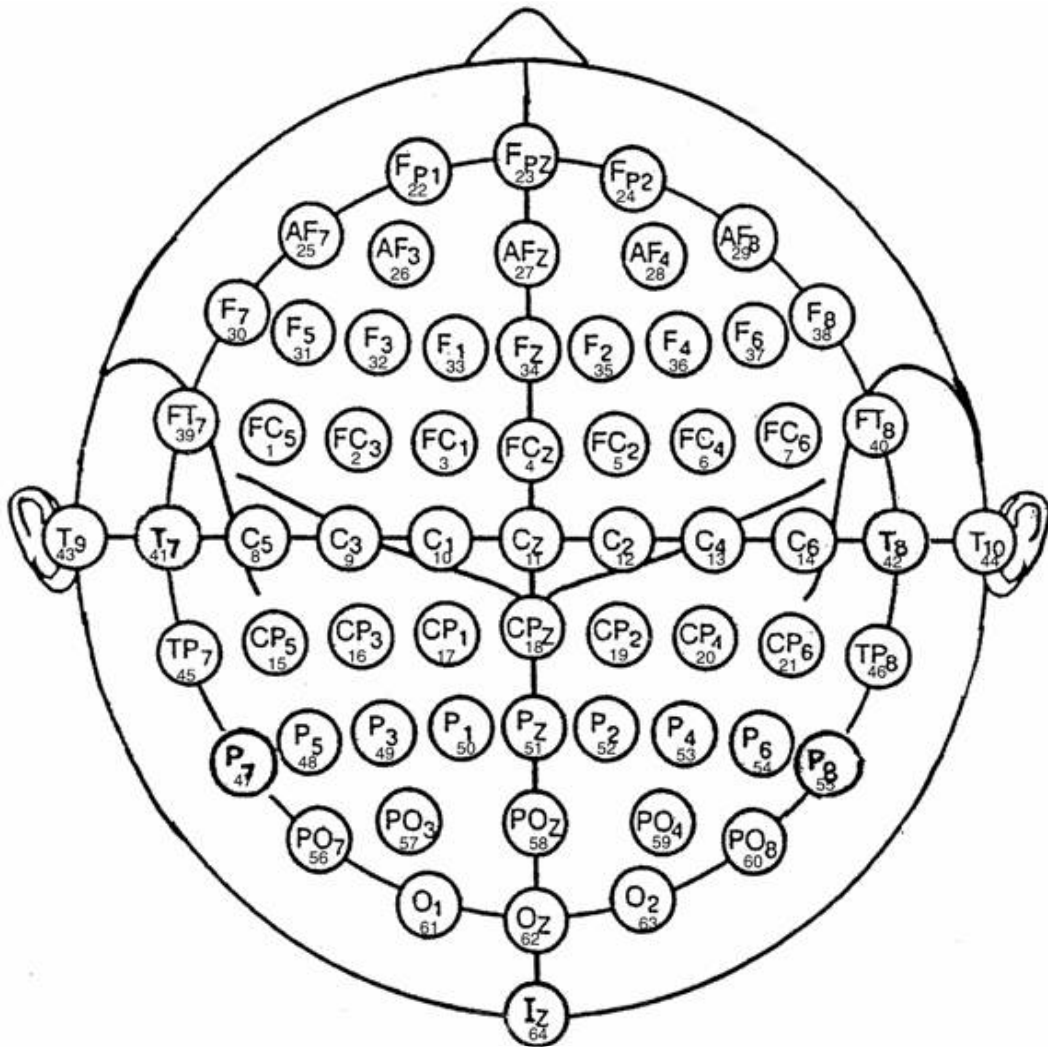


Figure 4.2: Electrode scalp positions, from which the EEG was recorded in the P300 speller paradigm.

completed and to focus on the next character in the word that was displayed on the top of the screen (the current character was shown in parentheses).

The training data set (sessions 1 and 2) contained 7560 trials for 42 characters and the test data set (session 3) contained 5580 trials for 31 characters. The objective in the International BCI Classification Contest 2003 contest was to recognize the correct characters in the training data set (session 3).

4.2.2 Data Analysis and Results

1000ms-long overlapping epochs starting at stimulus onset were extracted from the raw EEG signals. The signals were resampled on a logarithmic time scale by spline interpolation and low-pass filtered as described in Chapter 3 with scale dilation per second $d_s = 2$ and cut-off scale $s_0 = 75\text{ms}$. The filtered signals were represented in the time-frequency domain by DWT. Only the first 32 wavelet coefficients corresponding to scales $s > s_0(1 + d_s t)$ were retained; the first coefficient was also deleted (direct current (DC) correction). No baseline correction was performed.

Feature Extraction

Features were extracted from the training data set according to the τ -CWT procedure described in Chapter 3 with 12×12 PPSPO and 36×36 PPSPO. Additionally, the local extrema of the difference average waveform CWTs, $d^k(s, t)$ (3.6), were found for comparison purposes. The results for one scalp position are displayed at Figure 4.3. Besides the P300 maximum, several other extrema were found in the scalograms. Note the differences between the τ -value scalogram and the ordinary difference waveform scalogram. The P300 maximum was found at much smaller scale by the τ -CWT. Furthermore, the two late extrema at scale $s \approx 200\text{ms}$ in the difference CWT were found insignificant by the τ -CWT method; the contrary holds for the early extremum at $s \approx 200\text{ms}$, $t \approx 200\text{ms}$.

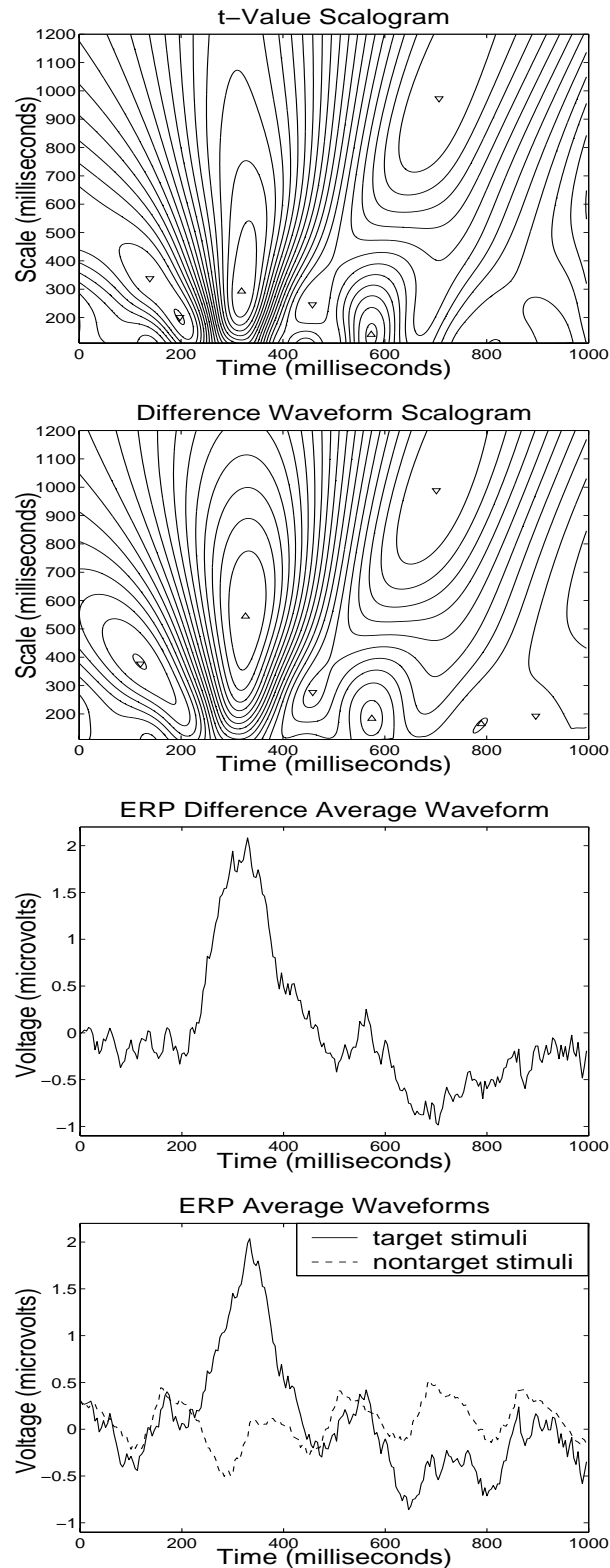


Figure 4.3: Average waveforms obtained from scalp position Cz (top of the head) and the corresponding scalograms with lines of constant altitude, local maxima marked by \triangle and local minima marked by ∇ . Note that, unlike in Chapters 1, 2 and 3, positivity is plotted upwards.

Classification

The classification of the test data was done by classical LDA (Rencher 1998, pp. 230–231) performed in the feature space (3.7). The scalar product ρ^n of each single-trial vector W^{kin} of the test data set with the coefficient vector c_{ki} of the discriminate function (Rencher 1998, pp. 201–202) was computed:

$$\rho^n = \sum_{k=1}^K \sum_{i=1}^{I_k} W^{kin} c_{ki}, \quad (4.1)$$

where K is the number of channels and I_k is the number of local extrema in the τ -CWT of channel k . The value ρ^n is a measure of how well trial n matches the target response pattern. The trials with the largest ρ in every 180-trial block determine the target character. This was done according to the following procedure. The "P300-score" of each trial was defined as the scalar product of that trial's feature vector with the coefficient vector of the discriminant function according to (4.1). The P300-score of each row and column was defined as the sum of the P3-scores of all trials in which this particular row or column was illuminated. The row and the column with the highest P3-scores were classified as target row and target column in this block of trials and the corresponding character according to the matrix display was classified as target character in this block. The results of this classification obtained in a feature space extracted from a 12×12 PPSPO τ -value scalogram are displayed in Table 4.2. The resulting words obtained from each run are displayed in Table 4.1.

Actually, the target characters were correctly identified from much smaller blocks of trials. The method was tested on the first 12, 24, 36, etc. trials of each 180-trial block (i.e. the first 1, 2, 3, etc. 12-trial blocks). All target characters were identified correctly with as few as 48 trials per character (4 12-trial blocks) using a 36×36 PPSPO τ -value scalogram (Table 4.3). The procedure was repeated with the features extracted from the ordinary difference CWT scalogram (36×36 PPSPO) – 84 trials per character (7 12-trial blocks) were needed for error-free classification

Run	Word
aas012r01:	FOOD
aas012r02:	MOOT
aas012r03:	HAM
aas012r04:	PIE
aas012r05:	CAKE
aas012r06:	TUNA
aas012r07:	ZYGOT
aas012r08:	4567

Table 4.1: The words obtained from the test data set in the P300 speller paradigm.

(Table 4.4). (This means that a paralyzed patient using this speller would need almost twice as much time to communicate a word, if the difference CWT is used, compared to the time needed with the τ -CWT method.)

The high scalogram resolution of 36×36 PPSPO also turned out to be important – the results presented at the International BCI Classification Contest 2003 showed that with a 12×12 PPSPO τ -value scalogram, 72 trials (6 12-trial blocks) were needed for perfect classification (Table 4.5) even after additional stepwise selection of variables (Rencher 1998, p. 217).

Column number:						Row number:					
1	2	3	4	5	6	1	2	3	4	5	6

Normalized column scores:	Selected column:	Normalized row scores:	Selected row:	Selected character:
aas012r01:				
17 8 0 8 22 45	6	52 28 10 0 7 3	1	F
5 15 48 15 0 17	3	7 30 47 14 0 2	3	O
9 23 41 17 11 0	3	17 19 37 20 8 0	3	O
7 0 21 55 15 1	4	45 21 11 10 0 12	1	D
aas012r02:				
66 26 0 1 7 1	1	1 14 56 18 11 0	3	M
2 12 69 10 7 0	3	9 23 49 14 5 0	3	O
8 17 50 14 0 11	3	9 25 54 10 1 0	3	O
21 52 15 8 4 0	2	0 15 13 48 20 4	4	T
aas012r03:				
8 69 10 6 0 7	2	23 51 17 0 9 0	2	H
51 22 8 16 0 3	1	61 16 9 11 2 0	1	A
64 12 8 10 7 0	1	22 18 35 20 5 0	3	M
aas012r04:				
1 9 5 80 5 0	4	19 14 50 14 3 0	3	P
12 18 56 11 0 4	3	29 43 16 11 0 1	2	I
14 0 6 10 51 19	5	50 21 0 12 9 8	1	E
aas012r05:				
7 20 58 12 4 0	3	46 21 12 11 0 11	1	C
65 10 7 0 10 8	1	65 22 4 8 0 1	1	A
2 3 0 17 62 16	5	22 54 11 0 13 1	2	K
6 0 20 15 48 11	5	56 21 10 7 6 0	1	E
aas012r06:				
23 52 9 5 0 11	2	3 5 3 70 19 0	4	T
7 31 51 8 0 3	3	16 0 19 53 4 7	4	U
8 69 10 6 0 8	2	5 18 56 20 0 2	3	N
71 14 7 5 3 0	1	71 12 7 0 7 3	1	A
aas012r07:				
22 66 3 2 7 0	2	5 17 8 20 50 0	5	Z
49 16 16 0 7 11	1	3 12 0 18 51 16	5	Y
75 19 2 0 2 3	1	22 42 19 11 5 0	2	G
14 21 39 16 9 0	3	21 16 40 21 0 2	3	O
29 49 7 9 0 6	2	0 11 21 48 19 1	4	T
aas012r08:				
2 11 0 20 6 60	6	10 0 6 7 60 18	5	4
47 21 11 9 0 13	1	1 16 0 6 5 72	6	5
23 50 19 5 2 0	2	6 10 0 13 22 49	6	6
19 17 57 6 2 0	3	8 0 7 6 13 65	6	7

Table 4.2: Normalized "P300 scores" for each row and column of the character matrix display obtained from the 180-trial blocks of the test data set. One such block (15 illuminations of each row and each column) was needed for selecting one character.

```

-----
Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

Selected character:

aas012r01:
  F  F  E  F  F  F  F  F  F  F  F  F  F  F  F
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  D  D  D  D  D  D  D  D  D  D  D  D  D  D  D

aas012r02:
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  N  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  N  T  T  T  T  T  T  T  T  T  T  T  T  T  T

aas012r03:
  H  I  I  H  H  H  H  H  H  H  H  H  H  H  H
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M

aas012r04:
  P  P  P  P  P  P  P  P  P  P  P  P  P  P  P
  I  I  I  I  I  I  I  I  I  I  I  I  I  I  I
  D  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r05:
  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  I  I  I  K  K  K  K  K  K  K  K  K  K  K  K
  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r06:
  T  T  T  T  T  T  T  T  T  T  T  T  T  T  T
  1  U  U  U  U  U  U  U  U  U  U  U  U  U  U  U
  H  N  N  N  N  N  N  N  N  N  N  N  N  N  N
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A

aas012r07:
  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z
  Z  S  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y
  B  A  G  G  G  G  A  G  G  G  G  G  G  G  G
  N  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  T  N  N  T  T  T  T  T  T  T  T  T  T  T  T

aas012r08:
  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4
  5  5  5  5  5  5  5  5  5  5  5  5  5  5
  B  6  6  6  6  6  6  6  6  6  6  6  6  6  6
  7  7  7  7  7  7  7  7  7  7  7  7  7  7
-----

Number of errors out of 31 characters:
 10  5  4  0  0  0  1  0  0  0  0  0  0  0  0
-----

Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

```

Table 4.3: Results obtained with 36×36 PPSPO in the τ -value scalogram for 1, 2, 3, etc. illuminations of each row and column of the character matrix display. Perfect classification (character identification) was obtained with as few as 4 illuminations.


```

-----
Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

Selected character:

aas012r01:
  F  F  K  L  L  F  F  F  F  F  F  F  F  F  F
  U  U  O  O  O  O  O  O  O  O  O  O  O  O  O
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  F  D  B  D  D  D  D  D  D  D  D  D  D  D  D

aas012r02:
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  N  P  O  O  O  O  O  O  O  O  O  O  O  O  O
  N  T  T  T  T  T  T  T  T  T  T  T  T  T  T

aas012r03:
  H  H  H  H  H  H  H  H  H  H  H  H  H  H  H
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M

aas012r04:
  P  P  P  P  P  P  P  P  P  P  P  P  P  P  P
  I  I  I  I  I  I  I  I  I  I  I  I  I  I  I
  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r05:
  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  H  I  K  K  K  K  K  K  K  K  K  K  K  K  K
  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r06:
  T  T  T  T  T  T  T  T  T  T  T  T  T  T  T
  1  U  U  U  U  U  U  U  U  U  U  U  U  U  U  U
  H  N  N  N  N  N  N  N  N  N  N  N  N  N  N
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A

aas012r07:
  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z
  Z  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y
  A  A  A  A  G  A  G  G  G  G  G  G  G  G  G
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  T  S  T  T  T  T  T  T  T  T  T  T  T  T  T

aas012r08:
  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4
  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
  B  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
-----

Number of errors out of 31 characters:
 10  5  3  2  1  1  0  0  0  0  0  0  0  0  0
-----

Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

```

Table 4.4: Results obtained with 36×36 PPSPO in the average difference scalogram for 1, 2, 3, etc. illuminations of each row and column of the character matrix display. Perfect classification (character identification) was obtained with 7 illuminations.

```

-----
Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

Selected character:

aas012r01:
  F  F  K  L  L  F  F  F  F  F  F  F  F  F  F
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  D  D  D  D  D  D  D  D  D  D  D  D  D  D  D

aas012r02:
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M
  U  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  O  O  O  O  O  O  O  O  O  O  O  O  O  O  O
  Z  T  T  T  T  T  T  T  T  T  T  T  T  T  T

aas012r03:
  B  B  H  H  H  H  H  H  H  H  H  H  H  H  H
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  M  M  M  M  M  M  M  M  M  M  M  M  M  M  M

aas012r04:
  P  P  P  P  P  P  P  P  P  P  P  P  P  P  P
  I  I  I  I  I  I  I  I  I  I  I  I  I  I  I
  D  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r05:
  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A
  G  I  I  K  K  K  K  K  K  K  K  K  K  K  K
  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E

aas012r06:
  T  T  T  T  T  T  T  T  T  T  T  T  T  T  T
  1  U  U  U  U  U  U  U  U  U  U  U  U  U  U  U
  H  N  N  N  N  N  N  N  N  N  N  N  N  N  N
  A  A  A  A  A  A  A  A  A  A  A  A  A  A  A

aas012r07:
  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z  Z
  Z  T  Z  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y  Y
  B  F  G  G  G  G  G  G  G  G  G  G  G  G  G
  N  O  O  U  O  O  O  O  O  O  O  O  O  O  O
  T  N  N  T  T  T  T  T  T  T  T  T  T  T  T

aas012r08:
  4  4  4  4  4  4  4  4  4  4  4  4  4  4  4
  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
  B  6  6  6  6  6  6  6  6  6  6  6  6  6  6
  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
-----

Number of errors out of 31 characters:
 11  5  4  2  1  0  0  0  0  0  0  0  0  0  0
-----

Number of illuminations:
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
-----

```

Table 4.5: Results obtained with 12×12 PPSPO in the t -value scalogram for 1, 2, 3, etc. illuminations of each row and column of the character matrix display. Perfect classification (character identification) was obtained with 6 illuminations.

4.3 BCI Study 2: Self-Regulation of SCPs

The τ -CWT method was also tested on two data sets provided by the BCI group at the Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Germany (Blankertz et al. 2004). The data were obtained from 5 scalp positions with the Thought Translation Device (TTD) (Kübler, Kotchoubey, Hinterberger, Ghanayim, Perelmouter, Schauer, Fritsch, Taub, and Birbaumer 1999; Hinterberger, Mellinger, and Birbaumer 2003) in a biofeedback experiment. In this paradigm the user learns to steer the course of her/his EEG in negative or positive direction, by willingly producing SCPs and getting immediate visual and/or acoustic feedback. This technique has been already successfully applied with paralyzed patients, who thus learned to make binary choices by means of their brain waves and thereby operate a speller or an internet browser (Hinterberger, Kaiser, Kübler, Neumann, and Birbaumer 2001).

4.3.1 Experimental Procedure

This description of the experimental design of the SCP biofeedback paradigm follows closely the one provided by the organizers of the International BCI Classification Contest 2003 from the Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Germany (Blankertz et al. 2004).

First Experiment

The EEG data (digitized at 256Hz) were collected from one healthy user from the following 5 scalp positions: Cz, 2 cm frontal of C3, 2 cm parietal of C3, 2 cm frontal of C4 and 2 cm parietal of C4. The user was asked to move a cursor up and down on a computer screen, while his cortical potentials were recorded. During the recording, the user received visual feedback of his SCPs (Cz-Mastoids). Cortical positivity lead

to a downward movement of the cursor on the screen. Cortical negativity lead to an upward movement of the cursor. Each trial lasted 6s. During every trial, the task was visually presented by a highlighted goal at either the top or bottom of the screen to indicate negativity or positivity from second 0.5 until the end of the trial. The visual feedback was presented from second 2 to second 5.5. Only this 3.5 second interval of every trial is provided for training and testing. The sampling rate of 256 Hz and the recording length of 3.5s resulted in 896 samples per channel for every trial.

Second Experiment

The EEG data (digitized at 256Hz) were collected from an artificially respired patient in the last stage of Amyotrophic Lateral Sclerosis (ALS) from the same scalp positions as in the first experiment. The patient was asked to move a cursor up and down on a computer screen, while his cortical potentials were recorded. During the recording, the subject received auditory and visual feedback of his slow cortical potentials (Cz-Mastoids). Cortical positivity lead to a downward movement of the cursor on the screen. Cortical negativity lead to an upward movement of the cursor. Each trial lasted 8s. During every trial, the task was visually presented by a highlighted goal at the top (for negativity) or bottom (for positivity) of the screen from second 0.5 until second 7.5 of every trial. In addition, the task ("up" or "down") was vocalised at second 0.5. The visual feedback was presented from second 2 to second 6.5. Only this 4.5 second interval of every trial is provided for training and testing. The sampling rate of 256 Hz and the recording length of 4.5s resulted in 1152 samples per channel for every trial.

4.3.2 Data Analysis and Results

Apart from the operant learning involved, there are two technical differences between the SCP Feedback and the P300 Speller paradigm. First, the SCPs are ERP components of much longer duration (scale) than the P300 and the other components found in the P300 Speller paradigm. The log time sampling described in Chapter 3 allows the τ -CWT method to handle early, short waves together with late, slow waves in 3-4s-long epochs. Second, the classification of SCP responses is done simply on a trial-by-trial basis according to a standard classification rule (Rencher 1998, p. 231), without using blocks of trials. This is related to the fact that SCPs are much easier to recognize from single trials than the P300.

The data were processed similarly to the P300 Speller data sets but with prestimulus baseline correction instead of DC correction and with scalogram resolution of 12×12 PPSPO.

At the International BCI Classification Contest 2003, the τ -CWT method provided the seventh best classification on the first Tübingen data set with 17.4% errors (the contest winner had 11.3%) and the best result – 45.6% errors – on the second data set (Blankertz et al. 2004). (All methods presented in the contest provided rather poor results on this, second data set, which suggested, that it was really very hard, if not impossible, to extract some useful discriminating pattern from it.)

4.4 Discussion

In the study presented in this chapter, the τ -CWT method introduced in Chapter 3 was applied as feature extraction procedure in two BCI paradigms. In the P300 speller paradigm it provided clearly better classification of the single-trial ERPs than the analogical procedure using the ordinary difference CWT. This result confirms the results obtained in Chapter 3 and demonstrates again the importance of taking

into account the variance in the data for proper ERP component detection.

The τ -CWT is a feature extraction method and it can be combined with any classification procedure. Its power was demonstrated by using the most simple LDA classification: the results were as good as those obtained from simple features but with a powerful machine-learning based classifier, the SVM (Kaper, Meinicke, Grossekaehler, Lingner, and Ritter 2004). The τ -CWT is attractive with its conceptual simplicity, intuitive plausibility, scalogram visualization and clear interpretation of the features in terms of ERP components. Moreover, the efficiency of the τ -CWT can possibly be increased further by using it in combination with such advanced classification methods as SVM.

Summary and Conclusions

Two major results were presented in this dissertation: first, the newly found electrophysiological brain response to inconsistent affective prosody marked by the N300 component in the ERP to incongruous emotional exclamations, and second, the newly developed t -CWT feature extraction method for fully automated detection and quantification of ERP components. Both results are important for both fundamental and applied-clinical psychophysiological research.

The two ERP paradigms introduced in Chapter 1 provide a means for tracking the cognitive process of voice emotion recognition in real time. It appears that this process is considerably faster than the analogical process of semantical-lexical recognition. The study presented in Chapter 1 is the first to address the recognition of emotional voice quality from a psychophysiological perspective. The found N300 effect is fairly small – about $1\mu V$ – and although it is big enough for the goals of fundamental psychological research, it might be insufficient for applied diagnostic purposes in clinical settings. Nevertheless, the passive oddball paradigm introduced in Chapter 1 has already been used with paralyzed patients and in single cases it has worked very well in proving that the corresponding patient showed unimpaired recognition of voice emotion (Kotchoubey, Lang, Bostanov, and Birbaumer 2003). Such patients not only showed a significant N300 effect, but their whole ERPs also resembled closely those of the healthy population as displayed in Chapter 1, Figures 1.3 and 1.4. However, as already mentioned in Chapter 1, the interpretation of the other components found in the ERP to the deviant emotional exclamation in

the oddball paradigm is somewhat problematic, since these later components may be attributed to simple acoustic mismatch cognitive processes. The N300 has a much more straightforward interpretation as psychophysiological measure of emotion recognition, because it appeared also in the ERP to the incongruous exclamations in the priming paradigm where the disparity between spoken words and exclamations excludes an explanation in terms of simple acoustical mismatch. Hence the priming paradigm should be more suitable for clinical diagnostics than the oddball. The problem is that the N300 effect found in the priming paradigm was even smaller than the one found in the oddball. As discussed in Chapter 1, the experiment can be largely improved by using more different emotional stimuli to avoid habituation and produce correspondingly larger statistical effects.

Another way to improve the diagnostic reliability of the N300 response in the prosody paradigms is to use better methods for ERP detection and quantification. The continuous wavelet measures of ERP components first introduced in Chapter 2 and then further developed and improved in Chapter 3 provide much larger statistical effects than the classical area and peak measures. The significance of the N300 effect was markedly improved even by the simpler total-average-CWT method presented in Chapter 2. The τ -CWT introduced in Chapter 3, which is essentially a pattern recognition method specially designed to extract the features (components) that best discriminate between two samples of ERP signals, was tested in two of the most popular ERP paradigms – oddball and semantic priming – which are also key instruments in the ERP-based cognitive diagnostics (Kotchoubey et al. 2002). It provided better results than all other methods it was compared with. The crucial role of taking into account the within- and between-participant variance was demonstrated. The conclusion is that the τ -CWT is fully capable of assessing single-patients ERP data on a single-trial basis. The method will be further tested on real coma patients in a large German-French study aimed at the investigation of the neurocognitive functioning in coma state.

The τ -CWT method was additionally validated in the International BCI Classifica-

tion Contest 2003, where it was a winner (provided best classification) on two ERP data sets acquired in two different BCI paradigms – P300 speller and SCP feedback. These results were presented in Chapter 4. By the application of the τ -CWT method, the P300 speller is immensely accelerated – from 5 to 10s are needed for writing one character instead 30s as in the original paradigm.

The τ -CWT method was implemented in a MATLAB 6.0 computer program, which was extensively tested in various settings: whole waveforms and single ERP components (Chapter 3), medium latency ERP components and SCPs (Chapter 4), few EEG channels and dense electrode array (Chapters 3 and 4), group and individual ERP data (Chapter 3). As a next step, it is planned to test the program on individual ERP data obtained from coma patients in the large German-French ERP coma assessment study.

Generally, the CWT may be seen as a conceptual framework not only for the τ -CWT, but also for a variety of other ERP assesment methods based on template matching. In Chapter 2 it has already been shown how the area measure can be formally defined as matching to a rectangular template. Peak picking after low-pass filtering (Ruchkin and Glaser 1978) can also be seen as component detection from a single scale slice corresponding to the cut-off frequency in a certain scalogram. Classical template maching algorithms also use a predefined scale, but adjust the time shift of the (non-wavelet) template to the single trials. (Of course, this single-trial fitting can be done also with the τ -CWT method – in the same way that it was done in Chapter 2 by the "matched-wavelet" measure.) Furthermore, the DWT with its diadic representation scheme (Samar, Bopardikar, Rao, and Swartz 1999) can be seen as a digitized CWT sampled on the sparsest possible logarithmic grid with 1 PPSPO to provide an exact and non-redundant representation of the ERP waveforms. Thus, thinking of ERP components as local extrema in some scalogram plot may provide us with a clearer picture of their properties and better understanding of their nature.

Bibliography

- Adolphs, R. and D. Tranel (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia* 37, 1285–1292.
- Adolphs, R., D. Tranel, H. Damasio, and A. Damasio (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672.
- Anderson, A. K. and E. A. Phelps (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport* 9, 3607–3613.
- Banse, R. and K. R. Scherer (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636.
- Bentin, S., M. Kutas, and S. A. Hillyard (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology* 30, 161–169.
- Bentin, S., G. McCarthy, and C. C. Wood (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalogr. Clin. Neurophysiol.* 60, 343–355.
- Bentin, S., Y. Mouchetant-Rostaing, M. H. Giard, J. F. Echallier, and J. Pernier (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience* 11, 235–260.
- Birbaumer, N., N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey,

- A. Kübler, J. Perelmouter, E. Taub, and H. Flor (1999). A spelling device for the paralysed. *Nature* 398, 297–298.
- Birbaumer, N. and W. Rief (Eds.) (2000). *Biofeedback-Therapie: Grundlagen, Indikation und praktisches Vorgehen*. Stuttgart: Schattauer.
- Birbaumer, N. and R. F. Schmidt (2003). *Biologische Psychologie* (fifth ed.). Berlin: Springer.
- Blankertz, B., K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer (2004). The BCI Competition 2003: Progress and perspectives in detection and discrimination of EEG single trials. *IEEE Trans. Biomed. Eng.* 51, 1044–1051.
- Bostanov, V. (2004). BCI Competition 2003 - data sets Ib and Iib: Feature extraction from event-related brain potentials with the continuous wavelet transform and the t -value scalogram. *IEEE Trans. Biomed. Eng.* 51, 1057–1061.
- Bostanov, V. and B. Kotchoubey (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology* 41, 259–268.
- Breitenstein, C., I. Daum, and H. Ackermann (1998). Emotional processing following cortical and subcortical brain damage: contribution of the fronto-striatal circuitry. *Behavioral Neurology* 11, 29–42.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge Univ. Press.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Murray. (Reprinted: HarperCollins, London, 1998).
- Demiralp, T., A. Ademoglu, M. Schurmann, C. Basar-Eroglu, and E. Basar (1999). Detection of P300 waves in single trials by the wavelet transform (WT). *Brain. Lang.* 66, 108–128.

- Demiralp, T., J. Yordanova, V. Kolev, A. Ademoglu, M. Devrim, and V. J. Samar (1999). Time-frequency analysis of single-sweep event-related potentials by means of fast wavelet transform. *Brain. Lang.* 66, 129–145.
- Donchin, E., K. M. Spencer, and R. Wijensinghe (2000). The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Rehab. Eng.* 8, 174–179.
- Edington, E. S. (1987). *Randomization tests* (second ed.). New York: Dekker.
- Emerson, C. S., D. W. Harrison, and D. E. Everhart (1999). Investigation of receptive affective prosodic ability in school-aged boys with and without depression. *Neuropsychiatry. Neuropsychol. Behav. Neurol.* 12, 102–109.
- Ende, M., A. K. Louis, P. Maass, and G. Mayer-Kress (1998). EEG signal analysis by continuous wavelet transform techniques. In H. Kantz, J. Kurths, and G. Mayer-Kress (Eds.), *Nonlinear analysis of physiological data*, pp. 213–219. Berlin, Heidelberg: Springer.
- Erhan, H., J. C. Borod, C. E. Tenke, and G. E. Bruder (1998). Identification of emotion in a dichotic listening task: Event-related brain potential and behavioral findings. *Brain Cogn.* 37, 286–307.
- Erwin, R. J., D. Van Lancker, D. Guthrie, J. Schwafel, P. Tanguay, and J. S. Buchwald (1991). P3 responses to prosodic stimuli in adult autistic subjects. *Electroencephalogr. Clin. Neurophysiol.* 80, 561–571.
- Fabiani, M., G. Gratton, D. Karis, and E. Donchin (1987). Definition, identification and reliability of measurement of the P300 component of the event-related brain potential. In P. Ackles, J. R. Jennings, and M. G. H. Coles (Eds.), *Advances in Psychophysiology*, Volume 2, pp. 1–78. Greenwich, CT: JAI Press.
- Farwell, L. A. and E. Donchin (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.* 70, 510–523.
- Fischer, C., D. Morlet, P. Bouchet, J. Luaute, C. Jourdan, F. Salord, et al. (1999).

- Mismatch negativity and late auditory evoked potentials in comatose patients. *Clin. Neurophysiol.* 110, 1601–1610.
- Fischer, C., D. Morlet, and M.-H. Giard (2000). Mismatch negativity and N100 in comatose patients. *Audiology and Neuro-Otology* 5, 192–197.
- Friston, K. J., K. M. Stephan, J. D. Heather, C. D. Frith, A. A. Ioannides, L. C. Liu, M. D. Rugg, J. Vieth, H. Keber, K. Hunter, and R. S. Frackowiak (1996). A multivariate analysis of evoked responses in eeg and meg data. *Neuroimage* 3, 167–174.
- Goffman, E. (1978). Response cries. *Language* 54, 787–815.
- González Sánchez, S., N. González Prelicic, and S. J. García Galan (1996). *Uvi-Wave. Wavelet Toolbox for use with Matlab*. Departamento de Tecnologías das Comunicacóns. Universidad de Vigo, Spain. <http://www.gts.tsc.uvigo.es/~wavelets>.
- Gott, P. S., A. L. Rabinovicz, and C. M. DiGiorgio (1991). P300 auditory event-related potentials in nontraumatic coma: Association with glasgow coma score and awakening. *Arch. Neurol.* 48, 1267–1270.
- Gratton, G., M. G. Coles, and E. Donchin (1983). A new method for off-line removal of ocular artifacts. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–84.
- Guerit, J. M., D. Verougstraete, M. Tourtchaninoff, D. Debatisse, and C. Witdoeck (1999). ERPs obtained with the auditory oddball paradigm in coma and altered states of consciousness: Clinical relationships prognostic value and origin of components. *Clin. Neurophysiol.* 110, 1260–1269.
- Hagoort, P., C. M. Brown, and T. Y. Swaab (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain* 119, 627–649.
- Heinrich, H., H. Dickhaus, A. Rothenberger, V. Heinrich, and G. H. Moll

- (1999). Single-sweep analysis of event-related potentials by wavelet networks – methodological basis and clinical application. *IEEE Trans. Biomed. Eng.* 46, 867–879.
- Heinrich, H., G. H. Moll, H. Dickhaus, V. Kolev, J. Yordanova, and A. Rothenberger (2001). Time-on-task analysis using wavelet networks in an event-related potential study on attention-deficit hyperactivity disorder. *Clin. Neurophysiol.* 112, 1280–1287.
- Hinterberger, T., J. Kaiser, A. Kübler, N. Neumann, and N. Birbaumer (2001). The Thought Translation Device and its applications to the completely paralyzed. In H. H. Diebner, T. Druckrey, and P. Weibel (Eds.), *Sciences of the Interface, Proceedings of the International Symposium, Karlsruhe, Germany, May 18–21, 1999*, Tübingen, pp. 232–240. Genista.
- Hinterberger, T., J. Mellinger, and N. Birbaumer (2003). The Thought Translation Device: Structure of a multimodal brain-computer communication system. In *Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering*, pp. 603–606.
- Ito, T. A., J. T. Larsen, N. K. Smith, and J. T. Cacioppo (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *J. Pers. Soc. Psychol.* 75, 887–900.
- Jaskowski, P. and R. Verleger (1999). Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Trans. Biomed. Eng.* 46, 987–993.
- Jemel, B., N. George, E. Olivares, N. Fiori, and B. Renault (1999). Event-related potentials to structural familiar face incongruity processing. *Psychophysiology* 36, 437–452.
- Johnson, R. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. In P. K. Ackles, J. R. Jennings, and M. G. H. Coles (Eds.), *Advances in psychophysiology*, Volume 3, pp. 69–138. Greenwich: JAI Press.

- Kane, N. M., S. R. Butler, and T. Simpson (2000). Coma outcome prediction using event-related potentials: P3 and mismatch negativity. *Audiology and Neuro-Otology* 5, 186–191.
- Kane, N. M., S. H. Curry, S. R. Butler, and B. H. Cummings (1993). Electrophysiological indicator of awakening from coma. *Lancet* 341, 688.
- Kane, N. M., S. H. Curry, C. A. Rowlands, A. R. Minara, T. Lewis, T. Moss, B. H. Cummins, and S. R. Butler (1996). Event-related potentials – neurophysiological tools for predicting emergence and early outcome from traumatic coma. *Intensive Care Med.* 22, 39–46.
- Kane, N. M., T. H. Moss, S. H. Curry, and S. R. Butler (1998). Quantitative electroencephalographic evaluation of non-fatal and fatal traumatic coma. *Electroencephalogr. Clin. Neurophysiol.* 106, 244–250.
- Kaper, M., P. Meinicke, U. Grossekhoefer, T. Lingner, and H. Ritter (2004). BCI Competition 2003 – data set IIB: Support Vector Machines for the P300 speller paradigm. *IEEE Trans. Biomed. Eng.* 51, 1073–1076.
- Karjalainen, P. A., J. P. Kaipio, A. S. Koistinen, and M. Vauhkonen (1999). Subspace regularization method for the single-trial estimation of evoked potentials. *IEEE Trans. Biomed. Eng.* 46, 849–860.
- Kotchoubey, B., S. Lang, R. Baales, E. Herb, P. Maurer, G. Mezger, D. Schmalohr, V. Bostanov, and N. Birbaumer (2001). Brain potentials in human patients with extremely severe diffuse brain damage. *Neurosci. Lett.* 301, 37–40.
- Kotchoubey, B., S. Lang, V. Bostanov, and N. Birbaumer (2002). Is there a mind? Electrophysiology of unconscious patients. *News Physiol. Sci.* 17, 38–42.
- Kotchoubey, B., S. Lang, V. Bostanov, and N. Birbaumer (2003). Cortical processing in Guillain-Barre syndrome after years of total immobility. *Journal of Neurology*. In press.
- Kübler, A., B. Kotchoubey, T. Hinterberger, N. Ghanayim, J. Perelmouter, M. Schauer, C. Fritsch, E. Taub, and N. Birbaumer (1999). The Thought

- Translation Device: A neurophysiological approach to communication in total motor paralysis. *Exp. Brain Res.* 124, 223–232.
- Kübler, A., B. Kotchoubey, J. Kaiser, J. R. Wolpaw, and N. Birbaumer (2001). Brain-computer communication: Unlocking the locked in. *Psychol. Bull.* 127, 358–375.
- Kutas, M. and S. A. Hillyard (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Lane, R. D., B. Sechrest, R. Riedel, V. Weldon, A. Kaszniak, and G. Schwartz (1996). Impaired verbal and nonverbal emotion recognition in alexithymia. *Psychosom. Med.* 58, 203–210.
- Lang, S., B. Kotchoubey, A. Lutz, and N. Birbaumer (1997). Was tut man, wenn man nichts tut? Kognitive EKP-Komponenten ohne kognitive Aufgabe. *Zeitschrift für Experimentelle Psychologie XLIV*, 138–162.
- Lange, D. H., H. Pratt, and G. F. Inbar (1997). Modeling and estimation of single evoked brain potential components. *IEEE Trans. Biomed. Eng.* 44, 791–799.
- Lange, D. H., H. T. Siegelmann, H. Pratt, and G. F. Inbar (2000). Overcoming selective ensemble averaging: unsupervised identification of event-related brain potentials. *IEEE Trans. Biomed. Eng.* 47, 822–826.
- Laureys, S., M. E. Faymonville, C. Degueldre, G. D. Fiore, P. Damas, B. Lambert, N. Janssens, J. Aerts, G. Franck, A. Luxen, G. Moonen, M. Lamy, and P. Maquet (2000). Auditory processing in the vegetative state. *Brain* 123, 1589–1601.
- Laver, J. (1980). *The phonetic description of voice quality*. Cambridge: Cambridge Univ. Press.
- LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behav. Brain Res.* 58, 69–79.
- Leinonen, L., T. Hiltunen, I. Linnankoski, and M.-L. Laakso (1997). Expression of emotional-motivational connotations with a one-word utterance. *J. Acoust.*

- Soc. Am.* 102, 1853–1863.
- Leinonen, L., I. Linnankoski, M.-L. Laakso, and R. Aulanko (1991). Vocal communication between species: Man and macaque. *Language & Communication* 11, 241–262.
- Linnankoski, I., M.-L. Laakso, R. Aulanko, and L. Leinonen (1994). Recognition of emotions in macaque vocalizations by children and adults. *Language & Communication* 14, 183–192.
- Lykken, D. T. (1991). Why (some) Americans believe in the lie detector while others believe in the guilty knowledge test. *Integr. Physiol. Behav. Sci.* 26, 214–222.
- Manassis, K., R. Tannock, and J. Barbosa (2000). Dichotic listening and response inhibition in children with comorbid anxiety disorders and ADHD. *J. Am. Acad. Child. Adolesc. Psychiatry* 39, 1152–1159.
- McCallum, W. C., S. F. Farmer, and P. V. Pockock (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 59, 477–488.
- McPherson, W. B. and P. J. Holcomb (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology* 36, 53–65.
- Meinicke, P., M. Kaper, F. Hoppe, H. Manfred, and H. Ritter (2003). Improving transfer rates in brain computer interfacing: a case study. In S. Becker, S. Thrun, and K. Obermayer (Eds.), *Advances in Neural Information Processing Systems*, Volume 15, Cambridge, MA. MIT Press.
- Miller, G. A., G. Gratton, and C. M. Yee (1988). Generalized implementation of an eye movement correction procedure. *Psychophysiology* 25, 241–243.
- Monnot, M., S. Nixon, W. Lovallo, and E. Ross (2001). Altered emotional perception in alcoholics: deficits in affective prosody comprehension. *Alcohol. Clin. Exp. Res.* 25, 362–369.
- Moreno, E. M., K. D. Federmeier, and M. Kutas (2002). Switching languages,

- switching palabras (words): an electrophysiological study of code switching. *Brain. Lang.* 80, 188–207.
- Müntze, T. F., H. J. Heinze, M. Matzke, B. M. Wieringa, and S. Johannes (1998). Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia* 36, 217–226.
- Mutschler, V., C. G. Chaumeil, L. Marcoux, N. Wioland, J. D. Tempe, and D. Kurtz (1996). Etude du P300 auditif chez des sujets en coma post-anoxique. Donnees preliminaires. *Neurophysiologie Clinique* 26, 158–163.
- Öhman, A. and N. Birbaumer (1993). Psychophysiological and cognitive-clinical perspectives on emotion: Introduction and overview. In N. Birbaumer and A. Öhman (Eds.), *The structure of emotion: Psychophysiological, cognitive and clinical aspects*, pp. 3–17. Seattle: Hogrefe & Huber.
- Pfefferbaum, A. and J. M. Ford (1988). ERPs to stimuli requiring response production and inhibition: effects of age, probability and visual noise. *Electroencephalogr. Clin. Neurophysiol.* 71, 55–63.
- Picton, T. W. and S. A. Hillyard (1988). Endogenous event-related potentials. In T. W. Picton (Ed.), *Handbook of Electroencephalographic Clinical Neurophysiology*, Volume 3, pp. 361–426. Amsterdam: Elsevier.
- Pihan, H., E. Altenmüller, and H. Ackermann (1997). The cortical processing of perceived emotion: a DC-potential study on affective speech prosody. *Neuroreport* 8, 623–627.
- Pihan, H., E. Altenmüller, I. Hertrich, and H. Ackermann (2000). Cortical activation patterns of affective speech processing depend on concurrent demands on the subvocal rehearsal system. A DC-potential study. *Brain* 123, 2338–2349.
- Polich, J. (1987). Comparison of P300 from a passive tone sequence paradigm and an active discrimination task. *Psychophysiology* 24, 312–320.
- Polich, J. (1989). P300 from a passive auditory paradigm. *Electroencephalogr. Clin. Neurophysiol.* 74, 41–46.

- Pollack, I., H. Rubenstein, and A. Horowitz (1960). Communication of verbal modes of expression. *Lang. Speech* 3, 121–130.
- Premack, D. (1971). Language in chimpanzee? *Science* 172, 808–822.
- Raz, J., L. Dickerson, and B. Turetsky (1999). A wavelet packet model of evoked potentials. *Brain Lang.* 66, 61–88.
- Rencher, A. C. (1998). *Multivariate Statistical Inference and Applications*. New York: John Wiley & Sons.
- Ross, E. D. (1981). The aprosodias. functional-anatomic organization of the affective components of language in the right hemisphere. *Arch. Neurol.* 38, 561–569.
- Ross, E. D., D. M. Orbelo, M. Burgard, and S. Hansel (1998). Functional-anatomic correlates of aprosodic deficits in patients with right brain damage. *Neurology* 50(suppl 4), A363.
- Ross, E. D., D. M. Orbelo, J. Cartwright, S. Hansel, M. Burgard, J. A. Testa, and R. Buck (2001). Affective-prosodic deficits in schizophrenia: profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J. Neurol. Neurosurg. Psychiatry* 70, 597–604.
- Ross, E. D., R. D. Thompson, and J. Yenkosky (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang.* 56, 27–54.
- Ruchkin, D. S. and E. M. Glaser (1978). Simple digital filters for examining CNV and P300 on a single-trial basis. In D. A. Otto (Ed.), *Multidisciplinary Perspectives on Event-Related Brain Potential Research*, pp. 579–581. Washington, DC: U.S. Government Printing Office.
- Samar, V. J., A. Bopardikar, R. Rao, and K. Swartz (1999). Wavelet analysis of neuroelectric waveforms: a conceptual tutorial. *Brain Lang.* 66, 7–60.
- Scherer, K. R. (1985). Vocal affect signalling: A comparative approach. In J. Rosenblatt, C. Beer, M.-C. Busnel, and P. Slater (Eds.), *Advances in the*

- study of behavior*, Volume 15, pp. 189–244. New York: Academic Press.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychol. Bull.* 99, 143–165.
- Scherer, K. R. (1994). Affect bursts. In S. H. M. van Goozen, N. E. Van de Poll, et al. (Eds.), *Emotions: Essays on emotion theory*, pp. 161–193. Hillsdale, NJ, US: Lawrence Erlbaum Associates.
- Scherer, K. R. and A. Kappas (1988). Primate vocal expression of affective states. In D. Todt, P. Goedeke, E. Newman, and D. Symmes (Eds.), *Primate vocal communication*, pp. 171–194. Berlin, Heidelberg, New York: Springer.
- Schlaghecken, F. (1998). On processing BEASTS and BIRDS: An event-related potential study on the representation of taxonomic structure. *Brain Lang.* 64, 53–82.
- Senkowski, D. and C. S. Herrmann (2002). Effects of task difficulty on evoked gamma activity and ERPs in a visual discrimination task. *Clin. Neurophysiol.* 113, 1742–1753.
- Seyfarth, R. M., D. L. Cheney, and P. Marler (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803.
- Signorino, M., S. D'Acunto, S. Cercaci, P. Pietropaoli, and F. Angeleri (1997). The P300 in traumatic coma: Conditioning of the oddball paradigm. *Journal of Psychophysiology* 11, 59–70.
- Smulders, F. T., J. L. Kenemans, and A. Kok (1994). A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalogr. Clin. Neurophysiol.* 92, 107–114.
- Speedie, L. J., N. Brake, S. E. Folstein, D. Bowers, and K. M. Heilman (1990). Comprehension of prosody in Huntington's disease. *J. Neurol. Neurosurg. Psychiatry.* 53, 607–610.
- Steinhauer, K., K. Alter, and A. D. Friederici (1999). Brain potentials indicate

- immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* *2*, 191–196.
- Twist, D. J., N. K. Squires, N. I. Spielholz, and R. Silverglide (1991). Event-related potentials in disorders of prosodic and semantic linguistic processing 1. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology* *4*, 281–304.
- Valdes-Sosa, M. J. et al. (1987). Comparison of auditory-evoked potential detection methods using signal detection theory. *Audiology* *26*, 166–178.
- Van Petten, C. and H. Rieffelder (1995). Conceptual relationships between spoken words and environmental sounds: event-related brain potential measures. *Neuropsychologia* *33*, 485–508.
- Woodward, S. H., J. Owens, and L. W. Thompson (1990). Word-to-word variation in ERP component latencies: spoken words. *Brain. Lang.* *38*, 488–503.
- Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Med. Biol. Eng.* *5*, 539–553.