University of Osnabrück

Department of Humanities

Master's Thesis

Cognitive Science Master Program

Neural Synchronization in the Crossmodal Integration of Camouflage Pictures and Sounds

Stefan Scherbaum

Matriculation Number.: 912 865

Schnatgang 49, 49079 Osnabrück

Stefan.Scherbaum@uos.de

Supervisors:

Prof. Dr. Peter König

Prof. Dr. Andreas Engel

Submitted: January 2006

Abstract

Temporal binding, a process based on the synchronization of neural signals, has been proposed as a mechanism for binding. The brain could use this mechanism to integrate information within single senses and between different modalities. Several studies have shown increased synchronization of brain signals during a variety of tasks that are associated with binding processes. In this EEG study, I investigate correlates of temporal binding, induced gamma (>24 Hz) energy and phase locking, in a crossmodal detection task.

17 subjects watched 150 camouflage pictures of natural scenes with animals or without animals while listening to a matching or non-matching auditory cue indicating the type of the animal. They had to detect and identify the animal. I analyzed data for time windows after stimulus onset and before the detection button press. Also, I compared different methods for performing this analysis.

Behaviorally, the auditory cue worked successful reducing the reaction times for congruent trials compared to incongruent trials. For EEG data, I found stronger induced gamma energy, time locked to stimulus onset, when an animal had been detected compared to animals not found. Similar patterns occurred for measures of global phase synchronization within the same time frame and frequency range. I also found significant differences in synchronization within the beta frequency band (13-18 Hz) before the recognition button press,the effect being stronger for incongruent picture-sound combinations compared to congruent stimuli.

The finding of stronger induced energy for detected animals supports the assumption that the measured oscillatory gamma energy is related to the binding of an object representation. While earlier studies showed increased energy for learned and recognized stimuli, this study indicates that previous experience is not necessary to enhance the gamma response.

The simultaneous increase in synchrony and energy in the gamma band supports theories that it might reflect a utilization or readout process. The increase in synchrony found within the lower beta band is interpreted as a correlate of crossmodal integration. As it was stronger for the incongruent sound condition, this enhanced beta band synchrony could be a correlate of the integration of conflicting information.

During analysis, different measures for synchrony were compared and turned out to influence the results in different ways. This indicates that conflicting results of other studies can be caused by differing methods.

Acknowledgements

There are a number of people I want to thank for their support and motivating enthusiasm.

Many thanks to:

My first supervisor *Peter König*, without whom this work would not have been possible, who inspired me with his scientific enthusiasm and spent time for my questions whenever I needed advise.

Everyone at the Neurobiopsychology Lab in Osnabrück, especially *Saskia Nagel* for sharing the experience of her own study, the fruitful discussions, and the curious and helpful glances on my monitor, *Robert Märtin, Hans-Peter Frey* and *Daniel Weiller* for the discussions about methods and their interest in upcoming problems.

Andreas Engel and *Stefan Debener,* who gave me the opportunity and their support to perform this study in their EEG-Lab in Hamburg.

Till Schneider and *Kriemhild Saha* for their support and shared time during the recording of the EEG data in Hamburg.

All my subjects.

Special thanks to:

Antje Petzold for enduring me as a thesis writer, for supporting and inspiring me with fruitful discussions, and everything else.

Martin Bleichner for the inspiring discussions, his patience, and his support.

Curd Reinert for being a flat mate and giving me a home during the time in Hamburg.

My parents *Hanne* and *Alfons* for their constant support during my studies and this thesis.

My friends and my brother *Martin* for their support and patience.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Information integration in the brain

The input into the brain, coming from the outer world, is complex: pictures, sounds, smells – all the information of the different senses is integrated by the brain into unified percepts of the world. To make things even more complex, incoming information of each single sense on its own has to be integrated by the brain to create a contiguous percept. How does it work that we see coherent objects and not only a heap of incoherent edges, colors, and movements? How can we listen to melodies instead of hearing single tones? How does it happen that we perceive objects that are structured following the classical Gestalt principles?

I will address these questions in this thesis. There are several theories how to answer these questions and the EEG study presented here will try to find evidence for one of theses theories: binding by neural synchrony.

1.1.1 The Binding Problem

The question how the parts of incoming information are build to a coherent perception has been formulated first by von der Mahlsburg (von der Malsburg, 1981) as the Binding Problem. According to the classical view of the visual system (Hubel and Wiesel, 1968), there is a hierarchy of processing levels, each with neurons for different features (see Figure 1).

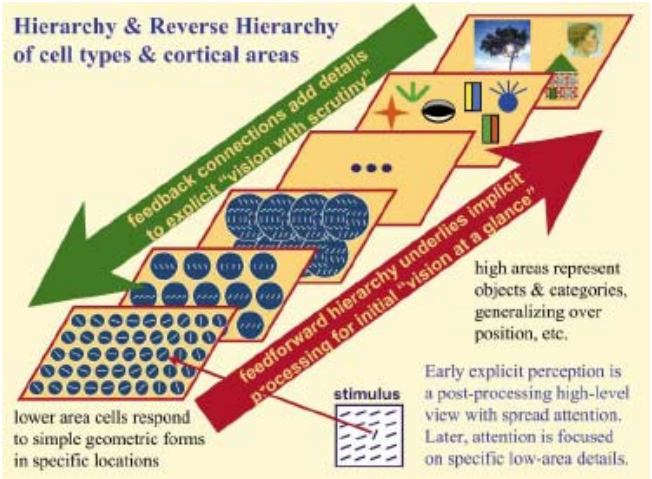


Figure 1 - The hierarchy of feature processing in the visual system:
Higher levels of processing process more complex features (from Hochstein and Ahissar, 2002).

The neurons at the first levels represent simple features like edges of specific angles. They also have the smallest receptive fields. Going along the different levels, the features represented by

each neuron become more and more complex and the width of the receptive fields gets larger and larger. To perceive an object as a coherent entity, it is necessary to combine all these represented features to a united representation. As long as there is only one object, the task is easy: The active neurons signal the features of this object. However, being confronted with two or more objects produces the so-called superposition catastrophe (von der Malsburg, 1999): It is not clear anymore, whether the active neurons signal features of the one or the other object. There is a need for a mechanism that signals the relationship of certain features and the relatedness to a common object that evoked the activity of these neurons. The single features of the one object must be bound together and separated from the features of the other objects.

Roskies (Roskies, 1999) calls this kind of binding perceptual binding. It is necessary for perception in all modalities. Following Roskies, another kind of binding is also of importance: cognitive binding. How are percepts (what we perceive, e.g. seeing an apple) and stored concepts (what we know, e.g. what an apple tastes like) bound together? How is information from different senses combined (e.g. identify an object by touching it if one had experienced it only visually)?

It is not clear, whether all these kinds of binding (perceptual and cognitive) rely on the same neural mechanisms. There are three possible mechanisms proposed for binding which are not mutually exclusive: attentional, combinatorial, and temporal binding.

Attentional binding (Treisman and Schmidt, 1982; Treisman, 1998; Reynolds and Desimone, 1999) is assumed to work by sharpening the receptive field of the neurons representing one object. The response of the system therefore focuses only on the attended object and features of other objects are ignored at that specific moment.

Combinatorial binding is based on the possibility to combine and code all incoming signals with the classical hierarchical solution without getting problems with the combinatorial explosion (Ghose and Maunsell, 1999). Even if models have shown that this is possible, there is the question, whether models, trying to solve single tasks, will also work on the large-scale problem (Roskies, 1999).

Temporal binding or binding by synchrony (Singer, 1999) is the approach I will follow in this work. The basic idea is that groups of neurons, so called assemblies, signal their relatedness by firing with the same rhythm (Varela, Lachaux, Rodriguez and Martinerie, 2001). This would allow distributed representations of objects to be constructed and destructed as needed from one

moment to the other. There exists a lot of evidence for this mechanism to be involved in binding. Freiwald and colleagues (Freiwald, Kreiter and Singer, 1995) found synchronized neurons in the striate cortex of the cat when a coherent bar was shown to the animal in contrast to incoherent bars. And Fries and colleagues (Fries, Roelfsema, Engel, König and Singer, 1997) found synchronization for the perceived stimulus in an binocular rivalry task. Further experiments have shown that synchronization followed the classical Gestalt grouping rules (for a review, see Singer, 1999). Bichot and colleagues (Bichot, Rossi and Desimone, 2005) have shown that synchrony occurs when monkeys do conjunction searches for objects with combined features (comparable to looking for someone with blonde hair and a blue suit within a crowd of people).

Though there is a lot of evidence (for more detailed reviews see Singer, Engel, Kreiter, Munk, Neuenschwander and Roelfsema, 1997; Singer, 1999), there are many questions left unanswered. Shadlen and Movshon (Shadlen and Movshon, 1999) claim that temporal coding lacks the capacity to code all possible relations. Additionally, they show that even if neural synchrony would be a mechanism to signal the relatedness of neural signals, this would still not answer how the system decides that specific features/ neurons are related and should become synchronized.

The evidence of synchrony occurring in processes where binding is assumed to happen indicates that synchronization of neural signals is, by some means, involved in the binding process. Therefore, it would be a step ahead to discover its nature. This could contribute to further research about the other steps of the binding process. To find the constraints of temporal binding and to show its occurrence in humans is therefore an important issue for understanding the involved brain processes.

1.1.2 Cross Modal Processing

After focusing on the study of brain systems for single modalities, the research on crossmodal integration and the interaction of these systems received more and more attention.

Several studies have shown that reaction times on a stimulus become shorter if other stimuli are presented simultaneously in another modality (Miller, 1982; Frens, Van Opstal and Van der Willigen, 1995). Many other findings included the influence of stimuli in one modality on the detection threshold of stimuli in another modality and the conditions in which crossmodal effects occur (for a review, see Calvert and Thesen, 2004). Interestingly, there have also been findings showing that different subjects might have different dominating senses (Giard and Peronnet, 1999). This dominance could be an important factor when studying crossmodal effects.

Regarding these behavioral effects, the question in dispute is how this interaction of multiple senses works. The theory of late integration dominated the debate for a long time. It states that signals are processed in brain areas for each modality first and are integrated later. Behavioral findings of effects at times lower than 40 ms contradict this theory (Giard and Peronnet, 1999). Additionally, there are findings showing direct connections between brain areas that have been assumed to be single modality areas (Wallace, Ramachandran and Stein, 2004; Zimmer, Lewald, Erb, Grodd and Karnath, 2004). Therefore, the strictly hierarchical view of multi sensory processing gives way to theories that allow an early interaction of inputs.

Looking at crossmodal effects from the perspective of the research in the binding problem, one possible mechanism for the interaction and integration of senses could be neural synchrony. In fact, von Stein and colleagues (von Stein, Rappelsberger, Sarnthein and Petsche, 1999) found neural synchronization in a task where subjects had to use the visual and the haptic sense to identify objects. Closing the gap between crossmodal research and the investigation on mechanisms for binding, studies on crossmodal binding gain insight for both fields of research.

1.2 Bringing Pieces Together: Neuronal Synchronization in Humans

Some of the specific properties of synchronization determine the study of this process in humans. In the following, I will shortly describe these properties. Following this, I will present methodological issues and results concerning synchronization in humans.

1.2.1 Properties of Synchronization

Though neural synchrony is the firing of neurons with the same rhythm, this does not imply that synchronized neurons fire simultaneously. But the interval between the spikes of two neurons (the so-called phase lag) has to be constant. This allows synchronization within one area (short range or local) and over longer distances (long range or large scale).

Considering the brain as a dynamic system of ongoing synchronization and desynchronization (Bressler and Kelso, 2001) opens the question for the temporal scale of these processes. In an important review, Varela and colleagues (Varela, Lachaux et al., 2001) propose synchronization intervals of 100-300 ms, based on several empirical findings. This time window might also represent a basis for the time windows of consciousness (e.g. Crick and Koch, 2003).

1.2.2 Studying Synchrony in Humans

How can research on synchronization in humans be done noninvasively? The short timescale and the general nature of the processes make Electro Encephalography (EEG) and Magneto Encephalography (MEG) the methods of choice. They measure ongoing processes of the brain directly by the produced electrical fields on the scalp (EEG) or the induced electro magnetic fields over the scalp (MEG). The time resolution of these methods goes down to several milliseconds allowing the study of ongoing dynamics at a great detail. On the other hand, these methods offer only low spatial resolution, due to volume conduction. However, with modern methods like Independent Component Analysis (ICA) and source localization algorithms, it is possible to increase the spatial resolution. As the primary interest for studies about neural synchrony is on the temporal dynamics of the signal, the remaining drawbacks are acceptable.

Can we study a process with large-scale methods like EEG and MEG, if the process is based on the synchronization of spikes of several neurons in one area or between several areas? Several studies suggest that it is possible to find correlates of synchrony noninvasively. Müller and colleagues (Müller, Bosch, Elbert, Kreiter, Sosa, Sosa and Rockstroh, 1996) used an experimental setup for humans that had been used successfully for monkeys before (Kreiter and Singer, 1996) and found strong oscillations that are assumed to be related to (local) synchrony. In the following, I will describe further studies that show different ways to study short and long range synchrony.

1.2.2.1 Local Integration

Local integration means binding of features between neurons in the same area, with distances of about 2 mm to about 1 cm (Varela, Lachaux et al., 2001) and latencies of about 4-6 ms. Von Stein and colleagues (von Stein and Sarnthein, 2000) showed evidence that synchronization over these short distances should occur primarily in the gamma band frequency range (25-70 Hz and higher). On the scalp, this kind of synchronization should produce strong oscillating signals, as a whole assembly of neurons is firing almost simultaneously below one electrode (compare Varela, Lachaux et al., 2001).

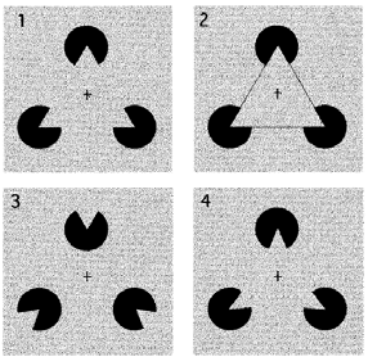


Figure 2 - Illusory Kanizsa figures.
These pictures induce oscillations, when people percieve the illusory white triangle (from Tallon-Baudry, Bertrand, Delpuech and Pernier, 1996)

In fact, several studies (Tallon-Baudry, Bertrand et al., 1996; Tallon-Baudry, Bertrand, Delpuech and Permier, 1997; Keil, Müller, Ray, Gruber and Elbert, 1999; Herrmann and Mecklinger, 2000) have found oscillations for subjects watching illusory figures, where binding is supposed to occur (see Figure 2).

Other studies (Rodriguez, George, Lachaux, Martinerie, Renault and Varela, 1999; Lachaux, George, Tallon-Baudry, Martinerie, Hugueville, Minotti, Kahane and Renault, 2005) have found strong oscillations for the recognition of so called mooney faces (see Figure 3). Rose and Buchel (Rose and Buchel, 2005) showed an involvement of oscillatory activity in the perception of apparent motion. Pulvermüller (Pulvermüller, 2001) reports oscillations at different locations for different words (binding of concepts), and Joliot and colleagues (Joliot, Ribary and Llinas, 1994) reported oscillations related to auditory binding.

Figure 3 – Mooney faces.
These stimuli were used by (Rodriguez, George et al., 1999) and (Lachaux, George et al., 2005) and induced oscillations when recognized (from Lachaux, George et al., 2005)

With all the evidence, the precise role of these gamma band oscillations is still unclear. On the one hand, they are affected by attention (Müller, Gruber and Keil, 2000) and previous experience with the material (Tallon-Baudry, Bertrand et al., 1997). This indicates that these oscillations might not reflect the pure binding process. Herrmann and colleagues (Herrmann, Munk and Engel, 2004) propose a model in which gamma oscillations are the reflection of a utilization reponse following a memory match: incoming stimuli signals match preactivated patterns and these patterns are read out by following processes. This could explain the results of studies that found gamma oscillations for degraded pictures of animals (Tallon-Baudry, Bertrand et al., 1997; Goffaux, Mouraux, Desmet and Rossion, 2004). These oscillations where stronger for those pictures the subjects had been trained to recognize.

Thus, the interpretation of gamma oscillations remains an open field of research. The relatedness to some kind of binding seems very probable, but does not seem to be the whole story.

1.2.2.2 Large Scale Integration

Large scale integration is defined as synchrony over polysynaptic pathways and distances greater than 1 cm (Varela, Lachaux et al., 2001). The time lag between two synchronized signals is greater than 8 ms. Von Stein and Sarnthein (von Stein and Sarnthein, 2000) showed evidence that this kind of integration happens in lower frequency bands (theta, alpha and lower beta, 4-18 Hz). Accordingly, Mima and colleagues (Mima, Oluwatimilehin, Hiraoka and Hallett, 2001) found synchronization of signals in the alpha band for the perception of objects. Singh and colleagues (Singh, Barnes, Hillebrand, Forde and Williams, 2002) have shown that strong synchronization and desynchronization in the alpha and beta band are related to the activation of certain brain areas detectable with functional magnetic resonance imaging (fMRI). And von Stein and colleagues (von Stein, Rappelsberger et al., 1999) found an increase of coherence in a multimodal identification task in the beta band.

Contradictory to these results, many other studies have found long range synchrony in the gamma band. Rodriguez and colleagues (Rodriguez, George et al., 1999) showed an increase of global synchrony for the perception and categorization of mooney faces. Knyazeva and colleagues (Knyazeva, Kiper, Vildavski, Despland, Maeder-Ingvar and Innocenti, 1999) found interhemispheric synchrony for the perception of gratings that covered both visual hemifields. And Lutz and colleagues (Lutz, Lachaux, Martinerie and Varela, 1999) correlated strong patterns of frontal synchronization with the conscious perception of 3D objects within noise.

The reasons for these different findings have not been resolved, yet. Therefore it remains unclear whether large scale integration occurs in higher or lower frequency bands and studies concerning long range synchrony should explore both ranges.

1.3 Consequences of a previous study and aim of this study

A previous pilot study of our lab, performed by Saskia Nagel, already investigated the subject of multimodal integration and neural synchronization.

In this study, pictures of natural scenes with animals were shown to the subjects and sounds of these animals were played to the left or the right ear. Analyzing the resulting data, we were not able to find consistent effects that matched the synchronization hypothesis (to our great disappointment). As mentioned in Nagel's report about the experiment, the problems during the analysis of the data were partly caused by the complex setup of the experiment. Evading some of these problems should enable the current study to be based on clearer assumptions and to produce better results.

The setup of the last experiment included several lateralization conditions and no time locking event. Additionally there was no reliable baseline time to compare synchronization with. This produced a number of unspecific hypotheses to look for and gave an explorative character to the analysis of data. My first aim was to reduce this amount of explorative research and to construct a study that was more hypotheses based.

I decided to abandon the lateralization condition of the previous setup, as it was not possible to detect any lateralization effects in the previous study. This might have been caused by two reasons. First, for basic processing, there are no strong lateralization effects in the auditory system (Zatorre, Belin and Penhune, 2002), compared for example to the visual system. Second, visual lateralization is hard to control during a stimulus presentation time of more than 6 seconds due to possible saccades.

Second, I wanted a setup that required the subjects to use the sounds and to integrate the information of both senses. In the previous experiment, the pictures were easily detectible so that subjects did not really have to attend to the sounds and therefore could have ignored them.

Finally yet importantly, the long time interval of each trial made analysis difficult for the data of the previous study. As there was no specific time for the subjects to focus their attention on the area where the animal had been placed, it was unclear when effects should have occurred. We

realized that there might just be a huge variability in the timing of these attentional actions (and therefore the corresponding effects we were looking for). So I wanted a specific event to time lock my analysis on.

I chose a setup similar to previous studies (Rodriguez, George et al., 1999; Goffaux, Mouraux et al., 2004), camouflage pictures of animals as stimuli (compare Tallon-Baudry, Bertrand et al., 1997; Goffaux, Mouraux et al., 2004), and to instruct subjects to find the hidden animal. As the animals are hard to find on this kind of picture, subjects could use the sound of the animal as a good hint. To control that they really attended to the sounds and that they had seen the animal if they stated so, I included catch trials with animals not matching the sound or no animal on the picture at all. To have an event to time lock my analysis on, I wanted the participants to press a button at the moment they recognized the animal. To have a reliable baseline for comparison, I included a pre stimulus interval between each trial.

The aim of my study was to investigate local and global synchronization effects directly after stimulus onset (staying in line with Tallon-Baudry and Bertrand, 1999 and Rodriguez, George et al., 1999) and before the recognition of the animal (directly before the button press). Other possible methods for the analysis of the resulting data will be proposed in section 2.4.2.6.

2 Methods

In this section, I will first describe the stimuli and the participants of the study. Following, I will present the setup of the experiment and technical details of data acquisition. Afterwards, I will describe the data analysis and the used statistics.

2.1 Stimuli

In the following, I will describe the basic properties of the stimuli used in this study. A precise description of the construction of the stimuli can be found in Appendix A.

2.1.1 Visual Stimuli

I decided to use pictures of animals and natural scenes, as I wanted to have stimuli that show natural objects avoiding the oversimplification of many artificial stimuli. To stay in line with previous experiments and to make the detection of the animals harder, I wanted to alter the images in a way that they look like the famous Dalmatian used by Tallon-Baudry and colleagues (Tallon-Baudry, Bertrand et al., 1997) or the degraded pictures used by Goffaux and colleagues

(Goffaux, Mouraux et al., 2004). These so-called camouflage pictures presumably demand a high amount of binding but still resemble normal images. I also needed sounds of animals to present them together with these images.

In general, I modified the contrast of pictures of animals to obtain the camouflage effect (see Figure 4). This way, I produced 120 pictures of 26 animals and 30 pictures of scenery without animals.

Figure 4 - One of the stimulus images used in this study.
Left: original image; right: modified camouflage image. In principle, I maximized the contrast so that the animal merges with the background.

2.1.2 Auditive Stimuli

I wanted the sounds to be distinct cues for the subjects. Therefore, I ordered the animals on the pictures into 21 categories and created an animal sound for each of these categories. The sounds had a duration of 5.25 seconds and were as continuous as possible. I normalized the sounds with a B-Weighting to obtain equal loudness (IEC 61672, www.iec.org).

2.1.3 Consequences of a Pretest of the Stimuli

To examine the properties of the stimuli, I executed a behavioral pretest (see Appendix B). Building on the conclusions of this pretest, I modified the original setup of the experiment to compensate some of the discovered effects:

- The pretest had shown that there is a big variance in the difficulty of the pictures. As I planned to align the data around subject's button presses that signaled detection, I assumed the high variance in reaction time of each picture to be acceptable. I also decided to split the data into subjective conditions (see section 2.4.2.4) to avoid unrecognizable pictures distorting the group statistics.

- In the pretest, there was a strong advantage for pictures that had been presented without any sound. I assumed there was a majority of picture sound combinations in which the sound was no reliable cue to the animal in the picture. This might have caused the subjects to ignore the sound as much as possible. Therefore I decided to increase the proportion of congruent combinations to more than 50% to make the sound a valid and reliable cue.

2.2 Subjects

Subjects were 17 candidates (4 female, mean age 24.1), chosen from the subjects pool of the neurophysiology laboratory. Most of them were medical students who had registered as subjects during an obligatory practical training seminar. All participants were healthy, had normal or corrected to normal vision and full auditory capabilities. 2 subjects were treated for mild depression at the time of the experiment, one with serotonin reuptake inhibitors and one with herbal medicine (hypericum). Both subjects showed no conspicuous behavioral results in the experiment and were included into the analysis.

All subjects gave informed consent with the conditions of the experiment and filled in a form about their current medical state (duration of sleep, consumption of drugs, current medical treatments) and about their handedness (Edinburgh handedness inventory (Oldfield, 1971)). All subjects were paid.

2.3 Setup and Data Acquisition

The following paragraphs describe the experimental environment, the detailed experimental procedure (for a complete list of steps, see Appendix H), the setup of the experiment, and the technical details of the EEG recording system.

2.3.1 Environment

I had the opportunity to conduct the experiment in the EEG laboratory of Dr. Stefan Debener. It is hosted at the Institute for Neurophysiology of Prof. Dr. Andreas Engel in the University Hospital Hamburg-Eppendorf .

The laboratory has a shielded EEG chamber providing practically noise free EEG signals. As the whole electrical system within the chamber is battery driven, the signal is also free of typical line noise. Other sources of electrical signals are also disabled as the stimulus monitor (CRT) is

mounted outside the chamber behind a shielded screen, the audio signals are transmitted via air tubes, and the EEG data transmission is done via glass fiber cables.

The subject is seated in a comfortable chair in the middle of the chamber (see Figure 5), at about 2 meters distance from the screen. During the experiment, the door of the chamber is closed, light is switched off and the temperature is kept constant by an air conditioning system to avoid sweat artifacts. The subject is monitored all the time by a small surveillance camera. The experimenter can be contacted all the time via an intercom system.

Figure 5 - The EEG chamber.
The shielded CRT is visible on the left side.

2.3.2 Procedure

Subjects filled in all forms, including the informed consent form. After washing their hair, they were prepared for the experiment (see Appendix H) and introduced to the experimental environment (EEG chamber, intercom system, etc.).

They had the opportunity to watch there own EEG waves on the stimulus monitor to see the influence of different movements on the signal. At this point I also emphasized that it is important for them to fixate the crosshairs on the screen during stimulus presentation. By showing them the possible artifacts, I wanted to motivate them to be relaxed but also to sit as still as possible. When subjects had no more questions, I started the experiment.

I subdivided the experiment into 4 stages:

1. A test of the reaction time
2. A tutorial about the main experiment
3. The sound learning phase

4. The main experiment with the presentation of the animal pictures and sounds

For each subject, a left hand or a right hand setup was chosen for the whole experiment in a pseudo randomized way so that I had 9 right hand oriented and 8 left hand oriented runs in the end.

In the following, I will describe the different experimental parts in detail.

2.3.2.1 Test of reaction time

This first task consisted of 10 trials. The subject saw a white square in the middle of the screen and had the instruction to press the main button (according to the handedness condition) as fast as possible.

I had two reasons for this task. First, I wanted to get baseline reaction time data. As I had seen in the pretest, there was a huge variability between subjects. Therefore, I wanted to have the possibility to use this for the analysis if necessary. Second, this was a good opportunity for the subjects to settle down and get used to the electrode cap and the whole environment while doing a relatively simple task that still resembled some demands of the main experiment.

As I was only interested in the timing data no EEG data was recorded during this task.

2.3.2.2 Tutorial

The tutorial was an opportunity for subjects to get familiar with the setup of the main experiment. They were shown each possible stimulus combination (animal-sound, congruent; animal-sound, incongruent; scenery, random sound) and received an explanation of the right reaction after each trial. The stimuli were only used once for the tutorial and did not occur during the experiment to avoid memory effects etc.

I wanted to give subjects the opportunity to get hands-on experience before producing real data. Hence, I judged a tutorial as a better way than the usual abstract description of the procedure.

2.3.2.3 The sound learning phase

After the tutorial and a short break, subjects were able to listen to all animal sounds as often as they wanted. Together with the sound, the name of the corresponding animal/ category was presented.

The pretest had shown that a few sounds (e.g. zebra, whale…) were not very familiar to many subjects. Thus, I wanted to enable the subjects to distinguish between the sounds to avoid bad behavioral data caused by false categorizations. I also wanted to compensate the advantage in familiarity of typical European animals compared to less familiar animals.

2.3.2.4 The main experiment

The main experiment consisted of 6 blocks of picture presentations. 40 pictures were presented in each block in pseudo randomized order, so that each block presented 27 animal pictures with a congruent sound, 13 with an incongruent sound, and 10 pictures of scenes without animals and a randomly chosen animal sound. The first 3 blocks contained only new pictures, the second 3 blocks were newly shuffled repetitions of the first 3 blocks. I decided to include this repetition to see weather there are memory effects (behavioral and in the EEG) or not. In the latter case I could use the additional trials for statistical evaluation of the EEG effects.

After each block, there was a short break of about 2 minutes. The door of the EEG chamber ws opened and the subject was instructed to use the time to relax actively.

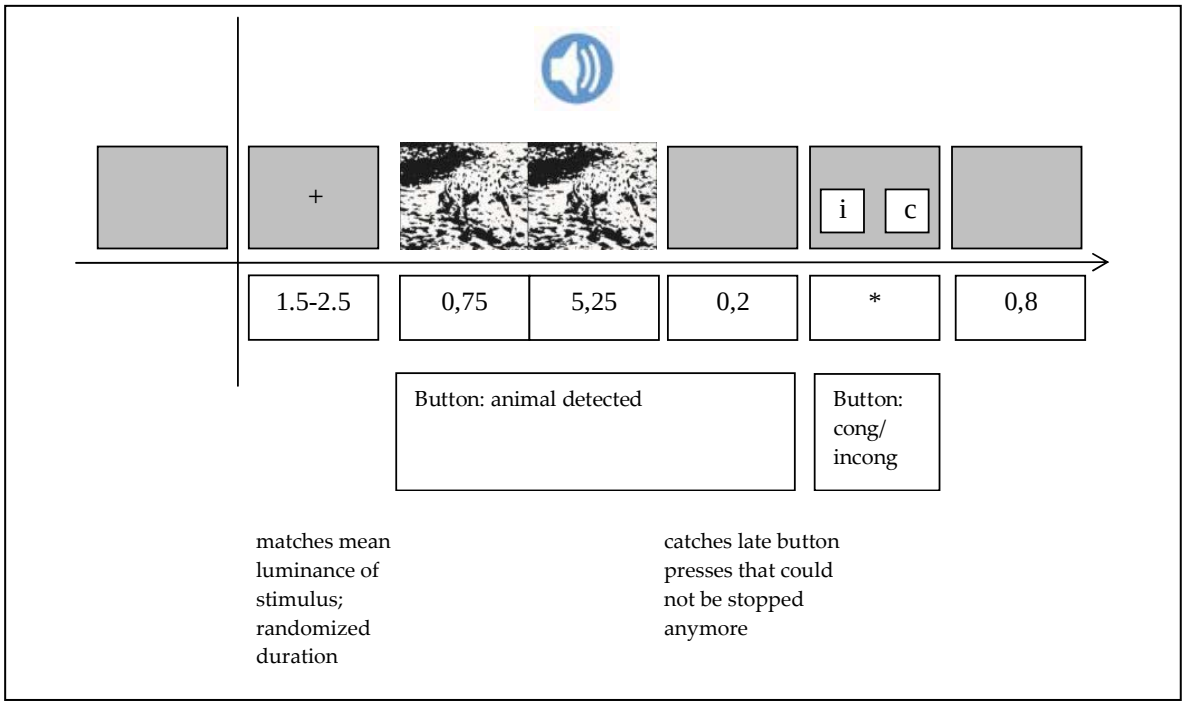


Figure 6 - The setup of the experiment.
After a variable prestimulus interval, the camouflage picture is shown for 6 seconds. 750 ms after stimulus onset, the sound of an animal (matching/not matching the picture) is played for 5.25 seconds. There is ashort grey screen afterwards to catch late responses. If subjects indicated to see an animal, they were asked if the sound matched the animal. After a 800 ms pause, the next trial began.

The detailed setup of each trial can be seen in Figure 6. For the prestimulus interval, I used a grey background, matching the mean luminance of the following stimulus, together with crosshairs. I randomly varied the prestimulus time between 1.5 and 2.5 to avoid expectation effects (compare Rodriguez, George et al., 1999). During the tutorial, subjects had been instructed to fixate the crosshairs shown during the prestimulus time and the presentation of the stimulus.

After this screen, the picture stimulus (animal or scene) was presented for 6 seconds, together with the crosshairs plotted on it. 750 ms after picture stimulus onset, the auditory stimulus (sound of animal) was presented for the remaining time of 5.25 seconds. I have chosen this latency of the sound to have a short unimodal interval. As I wanted to analyze the EEG data later for the effects found by previous studies (Tallon-Baudry, Bertrand et al., 1997; Rodriguez, George et al., 1999; Goffaux, Mouraux et al., 2004), I wanted to have conditions as similar to these experiments as possible. Subjects were instructed to press the main button (according to the lateralization condition) as soon as they recognized an animal on the picture. To ensure that a very late button press still counted, there was a 0.2 seconds grey screen after the stimulus presentation during which a button press was still possible.

If subjects pressed the button, a congruency question was presented after stimulus presentation. Subjects had to press the main button to state that the animal they had seen and the sound had been congruent. Otherwise, they had to press the other button.

After this question screen, there was a grey screen of 0.8 seconds before the next trial began.

2.3.3 Technical Setup

Stimulus presentation was controlled by a PC (Intel P3, 512 MB RAM) running Windows 98 and Presentation (Neurobehavioural Systems, Version 9.12). Presentation enables the experimenter to control the stimulus presentation with a precision of milliseconds. It offers a simple scripting language to build complex setups.

EEG was recorded from 128 electrodes fixated on the head via a flexible electrode cap (see Figure 7). 2 electrodes recorded EOG (Electrooculogram) and were placed below each eye. The tip of the nose of the subject served as reference.

Figure 7 - The EEG cap.
128 electrodes are arranged in an extended 10-20 layout (see Appendix G).

Electrode impedances were kept below 10 kΩ during the whole experiment. The electrodes were connected to battery driven BrainAmps Amplifiers (Brain Products). The signal was recorded by a PC (Intel P4, 1GB RAM) running Windows XP and Brain Vision Recorder (Brain Products, Version 1.02). The PC for stimulus presentation also sends event signals (e.g. stimulus onset or button press) to the PC for EEG recording via parallel port.

EEG data was recorded in continuous mode with a sampling rate of 1000Hz and 0.1 Hz low cutoff and 250 Hz high cutoff frequency.

2.4 Data Analysis

In the following paragraphs, I will describe the analysis process of the data, including technical details, preprocessing steps, artifact rejection, and methods for signal analysis and their evaluation (as some of them produced different results than others).

2.4.1 Technical Details

The EEG data analysis was performed on an Apple Macintosh (Motorola G5, 4GB RAM) running MAC OS X. As software I used Matlab 7, (The Mathworks, version R14, SP2). For preprocessing and artifact rejection, I used the open source software EEGLab (Delorme and Makeig, 2004), version 5.412, running within Matlab. All functionality of this software is accessible via a GUI and via Matlab scripting. Hence, it is possible to automate many processing steps.

2.4.2 Methods

The analysis of the EEG data included several steps: preprocessing, the time frequency transformation, and the analysis of the EEG data. As almost every step included several possibilities to choose from, there were a huge number of permutations of possible ways in the end. To stay within a reasonable timeframe, I had to ignore (or discard after comparison) several possibilities (for proposals on further analysis see section 4.7.1). An overview of the most important decisions is shown in Table 1, with chosen possibilities shaded in gray.

Importing Data into EEGLab and downsampling			
Downsampling with EEGLab		Downsampling with Brain Vision Analyzer	
Filtering, manual rejection of bad data, epoching			
ICA artifact rejection			
Extraction of conditions			
A priori, objective		A posteriori, subjective	
Event for alignment of data			
Stimulus Onset		Button press	
Signals to analyze			
channels/ electrode		ICA components	
Method for Time Frequency Transformation (Extraction of Phase and Energy)			
Wavelet Analysis: short wavelets	Wavelet Analysis: long wavelets	Fourier Transformation	Hilbert Transformation
Measurements			
Induced Energy	Phase Locking Value (Lachaux, Rodriguez, Martinerie and Varela, 1999)	Event related Phase Cross Coherence (Delorme and Makeig, 2004)	Mean Squared Coherence
Grand Average over Subjects			
z-transformed		Not normalized	

Table 1 - The main decisions on the way of data analysis.
For each step of analysis, the possible methods are shown. The chosen methods a shaded grey.

2.4.2.1 Importing Data: Converting and Downsampling

After comparison of different methods (see Appendix C), I decided to import the raw data directly into EEGLab. First, I converted the endian order of the data (there is a different byte order for PCs and Macs). Afterwards, I imported the data into EEGLab channelwise,

downsampled it from 1000 Hz to 250 Hz and concatenated the channels to a complete EEGLab dataset.

2.4.2.2 Filtering, Rejection of Bad Data, Epoching

After importing the data, I looked for long time periods of bad data (due to movements or breaks) and deleted them.

For the next steps, I made two attempts. On the first attempt, I separated the data into epochs after cleaning it. Then I calculated the independent components of the data (see 2.4.2.3.1). The resulting components were of very bad quality. After looking at the components of all subjects, I tried to find a reason why only few subjects showed good components (for a detailed description of this process see Appendix D). In the end, I realized that I had to high pass filter the data before epoching. It seemed that the reason for this was the low cutoff frequency during recording that had been set to 0.1 Hz, which was too low and left large slow waves in the signal. The waves disrupted the Independent Components Analysis I executed later, resulting in bad components. Therefore, I stepped back to the imported data, filtered the whole signal with a high pass filter (low cutoff frequency of 1 Hz), and had to do the epoching and ICA again in the second attempt.

After filtering, I separated the continuous data into epochs. As subjects often began to move heavily when the presentation of the animal picture ended and the congruency question was presented, I did not want to include this artifact loaded data into the epoched data. On the other hand, I needed the event data of the congruency question as the button press of the subject defined the epoch as "perceived congruent" or "perceived incongruent". I solved this conflict by writing the data of the congruency question into the data of the "animal detected" button press. Afterwards, I could abandon the data that was not necessary for EEG analysis.

I separated the data into epochs of -750 ms to 6400 ms relative to stimulus onset (leaving a little of overhead that could be clipped if edge effects occur during the time/frequency transformation later on; see section 2.4.2.5).

2.4.2.3 Rejection of Artifacts

For artifact rejection, I followed the steps recommended by Delorme and Makeig (Delorme and Makeig, 2004), using EEGLab. First I deleted epochs that were disturbed heavily by movements or other artifacts (see Figure 8). I looked especially for artifacts that occurred only one time

(compared to artifacts occurring frequently, e.g. eye blinks) as this category of artifacts is problematic for the following Independent Component Analysis.

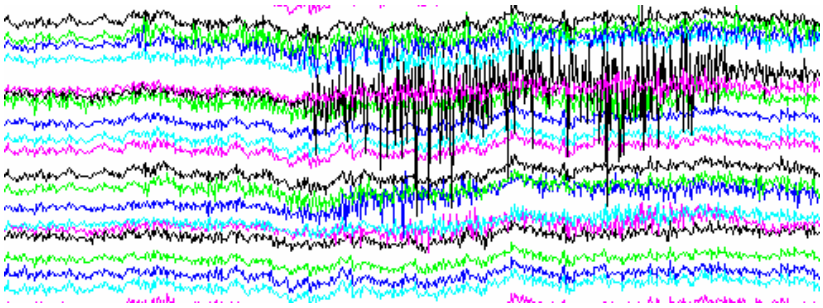


Figure 8 – An artifact within the EEG waves.
Strong artifactual signals, produced e.g. by movements of the body, have to be removed before further processing of the data.

After cleaning the epochs, the next step was to extract the independent components from the channel data. I will describe the Independent Component Analysis (ICA), its role in EEG artifact rejection, and my experience with it in the following paragraphs.

2.4.2.3.1 Independent Component Analysis and EEG

Independent Component Analysis (ICA) is a method to extract hypothetical source signals from the recorded mixture of signals without knowing anything about the environment. For EEG, this means that ICA can separate different brain sources and artifacts like eye blinks (see Figure 9) or muscle tension from the recorded electrode data. The artifacts can then be deleted and the electrode data can be reconstructed without the artifactual signals (Jung, Makeig, Humphries, Lee, McKeown, Iragui and Sejnowski, 2000).

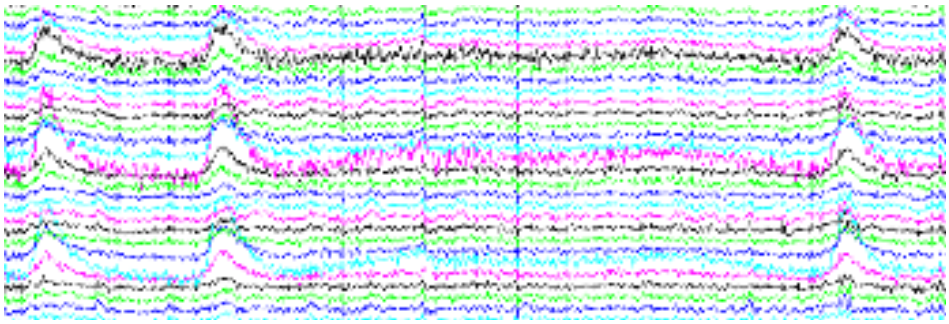


Figure 9 – Eye blinks artifacts.
These artifacts show specific properties: Strong peaks in the signal are pronounced for frontal channels.

As ICA does a blind source separation, it does not require to know anything about the special nature of the data and the environment. ICA only requires three assumptions that are mostly easy to justify (Makeig, Bell, Jung and Sejnowski, 1996):

1. The mixing process is linear and instantaneous (no delays),
2. the number of sources is equal to the number of mixed signals, and
3. the sources are really independent of each other.

For EEG data, assumption 1 is fulfilled, as it is one of the basic assumptions of EEG that the signals recorded by the electrodes are linear and instantaneous mixtures of electrical brain signals (and artifactual signals like eye blink muscle activity) (Makeig, Bell et al., 1996). Assumption 2 is not easy to fulfill, as the real number of sources within the brain is not known. Still, numerical simulations have shown that ICA is practically able to extract high quality EEG components that have reasonable topographies and time courses (Jung, Makeig et al., 2000; Delorme and Makeig, 2004), especially for artifactual components. Finally, Assumption 3 is fulfilled as artifactual signals and several brain processes are not stimulus locked to brain signals and therefore independent.

In principle, ICA is a family of algorithms that iteratively calculate a transformation or mixing matrix for the data, so that the resulting source signals become statistically independent in the end. In comparison to other (and older) methods, based on assumptions theoretically not justifiable for EEG data (e.g. principal components analysis), ICA maximally separates artifactual information from the rest of the data. Therefore, it allows removing the artifactual parts without loss of original signal information. Afterwards, the electrode data can be reconstructed completely, except for the removed artifactual information, using the mixing matrix (Iriarte, Urrestarazu, Valencia, Alegre, Malanda, Viteri and Artieda, 2003).

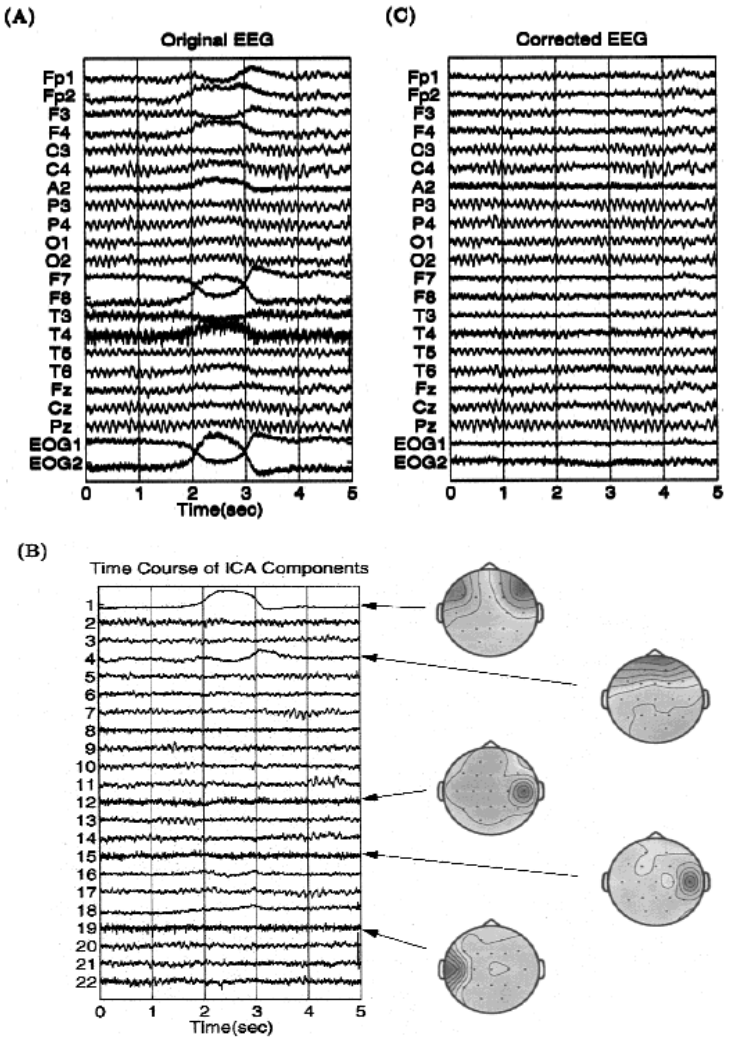


Figure 10 - How ICA works.
The algorithm takes the original data (A) and calculates components that are statistically independent (B) and can be mixed together to derive the original signal. Artifacts, separated from brain signals by this process, show typical properties, e.g. the topographies shown in (B). These artifactual components can be removed. Remixing the remaining components, the original data is restored without the artifacts (C). (from Jung, Makeig et al., 2000).

As EEGLab is shipped with a version of the so-called infomax ICA algorithm (Bell and Sejnowski, 1995), I used this neural network based algorithm to decompose and clean my data.

2.4.2.3.2 Clustering of Independent Components

The process of rejecting artifactual ICA components is highly based on personal experience (Jung, Makeig et al., 2000) and could only be partly automated as shown by Delorme and Makeig (Delorme, Makeig and Sejnowski, 2001). To identify artifacts, several properties have to be considered, e.g. the components spectra or the topography.

With 128 components to check, it is hard to keep a constant quality of judgment. To support this process, I decided to build a tool that categorizes components automatically (for a detailed description of the resulting EEGLab plugin, see Appendix E. I was inspired by Delorme and colleagues (Delorme, Makeig et al., 2001), who suggest a statistics based, semi automatic rejection procedure for EEG trial data. Looking for a similar solution for EEG ICA Components, I derived statistical measures of the important component properties. Then I used clustering to categorize this multidimensional statistical data. The result were groups of components that had several of these properties in common. Some of these groups contained mainly artifactual components. A picture of a two dimensional plot of component clusters can be seen in Figure 11.

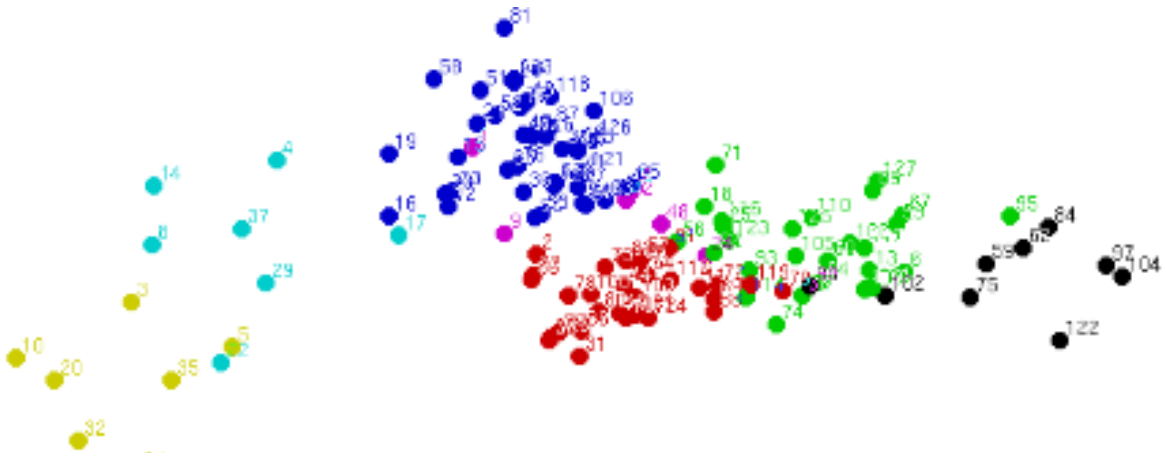


Figure 11 - A plot of the first two PCA dimensions of the clustered ICA components.
Artifactual components can be found mostly in the lower right clusters.

In general, in this two-dimensional view, there is a strong tendency for components placed on the right lower side to be artifacts. Figure 12 shows the topographies of the components of the rightmost cluster.

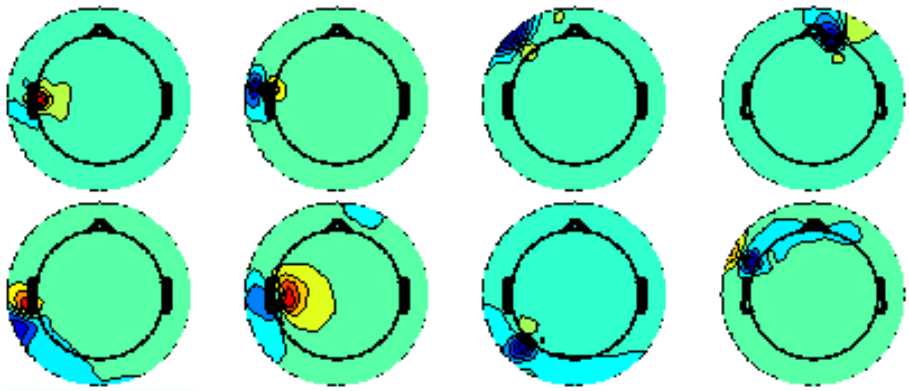


Figure 12 - A cluster of artifactual components found by the algorithm.
The topographies, centered on single electrodes, are typical for muscle tension artifacts.

Additionally, the process was able to separate the eye blink/ movement artifacts (components 1 and 9 in Figure 11) into one artifact cluster.

Summarizing, the clustering built quite reliable clusters, usable for the suggestion of rejectable components. The classification did not reach 100% hits, but it helped to get an easy overview about each component's properties, a gist for the whole data. Thereby, it made the rejection of components much easier.

2.4.2.4 Extraction of Conditions

Before executing the time frequency transformation and the calculations, based on this transformation, I had to split the data into conditions. I defined the following 4 sets of trials as different conditions for analysis:

- C1: Animal on picture, sound congruent or incongruent
    - C11: Animal with congruent sound
    - C12: Animal with incongruent sound
- C2: No animal with random sound

There were 2 possibilities to define these conditions: by the objective definition (the real stimulus combinations) of the trials or by the subjective categorization (the congruency judgement of the subjects). I preferred the latter one, as it posed fewer assumptions on the properties of the stimuli. As the pretest of the stimuli was not too extensive I had not much prior knowledge about reactions to each stimulus. For example, during the experiment, subjects sometimes told me that they were sure there was a butterfly on a specific image even if there was no stimulus showing a butterfly in the set. Therefore I wanted to define the conditions based on subjective results and not on prior assumptions about the stimuli. To say it colloquially: If subjects are sure that there is a dog, then I accept that a dog can be seen in the image.

Theoretically, this could be supported by the following assumption: If a dog is perceived, single features within the image are bound together to build up the percept of a dog. As I looked for correlates of binding, I had to categorize the trial as one where I could find correlates of binding.

Practically, the behavioral result of both kinds of conditions turned out to be comparable (see section 3.1). Because of a high percentage of animals not found, I finally decided to discard the objective categorization and to continue only with the subjective categorization.

2.4.2.5 Extraction of Energies and Phase Locking Values

After finishing all preprocessing steps, I had to transform the data into the time frequency domain for further analysis. Afterwards, I wanted to calculate measures for local (energy) and global synchronization (phase locking). There are several methods available for all these steps. To decide which ones to use, I had to evaluate these methods, first. The results of this evaluation are described in the following paragraphs.

2.4.2.5.1 Comparison of Methods for Time Frequency Transformation

As my analysis was based on the phase and the energy of the signal, I had to transform the data from the time to the time frequency domain. In the literature I had reviewed, this had been done in three ways: Wavelet Analysis, Hilbert transform, and Fourier transform. Each of these ways had its advantages and disadvantages so that I had to choose one that suited my demands best. Additionally, I wanted to produce data that was comparable to previous studies of interest.

First I planned to choose the wavelet based methods used by previous studies (e.g. Tallon-Baudry, Bertrand et al., 1997; Lachaux, Rodriguez et al., 1999). Both convolve families of complex Morlet's Wavelets with the signal to extract a complex signal.

The wavelet is defined as

$$w(t, f_0) = (\sigma_t \sqrt{\pi})^{-1/2} \exp(-t^2 / 2\sigma_t^2) \exp(2i\pi f_0 t)$$

with

$$\sigma_f = 1/2\pi\sigma_t .$$

σ_f and σ_t denote the length of the wavelet in the frequency and time domain, *t* denotes time and f_0 denotes the frequency of interest (called center frequency). The wavelet has a Gaussian shape around its center frequency. Both, Lachaux et al. and Tallon-Baudry et al. emphasize that the ratio of f_0 / σ_f should be chosen greater than 5. Tallon-Baudry and colleagues chose a ratio of 7, while Lachaux et al. effectively choose a ratio of $7 \times 2 \times \pi$ leading to much larger wavelets. Therefore, the short wavelets used by Tallon-Baudry et al. have a spectral bandwidth of 5.8 Hz and a duration of 111.4 ms at 20 Hz while the long wavelets of Lachaux et al. are more specific for each frequency and therefore have a longer duration of 699 ms (see Figure 13 for comparison).

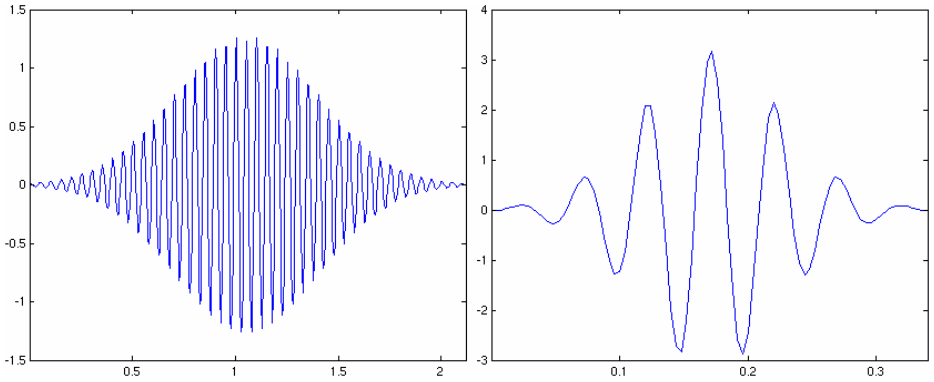


Figure 13 - The real parts of 2 wavelets at 20 Hz.
Long wavelets (left) are used by (Lachaux, Rodriguez et al., 1999) and short wavelets (right) are used by (Tallon-Baudry, Bertrand et al., 1997). Due to different parameters, the length of the two wavelets differs dramatically (X-axis: time in seconds, Y-axis: power).

While Tallon-Baudry et al. convolve the signal directly with the wavelets, Lachaux and colleagues filter the data before convolution with a bandpass filter of +/- 2 Hz, arguing that otherwise the method is prone to artifacts. This is astonishing as the selectivity of their wavelets in the frequency domain is already very specific.

The long wavelets are not suitable for low frequencies, as the length of one trial becomes too short for the long wavelets duration. Even the short wavelets are not usable below 20 Hz. Therefore, Palva and colleagues (Palva, Palva and Kaila, 2005) propose to use the Hilbert transform for the extraction of lower frequencies.

Extracting the time frequency spectrum of a signal via the Hilbert transform is a two step process (Bruns, 2004): First, the signal is filtered with a bandpass filter. Second, the signal is transformed as described in (Bruns, 2004). The resulting complex signal is called the analytic signal with a complex part assigned to each real part of the signal. The instantaneous phase of the signal is the angle of each complex number.

Finally, Delorme and Makeig (Delorme and Makeig, 2004) use the so-called short term Fourier analysis, a fast Fourier transform that is calculated for each step in time. The signal around each sample point is multiplied by a window (its length and type has to be chosen well to avoid side effects as spectral leakage (Bruns, 2004)) and then Fourier transformed. This results in a quasi-continuous time/ frequency spectrum.

Comparing these methods, Bruns (Bruns, 2004) had shown that they produce comparable results (for the right range of frequencies and parameters). As I wanted to stay in line with the methods of preceding studies, I preferred the two wavelet approaches. To save processing time, I had to

choose one of both methods. On the first glance, they seemed quite similar. Hence, to have decisive information, I calculated the time frequency transformation of original EEG data with both methods and additionally with the Hilbert transformation to have a third comparison.

The phase coherences of the resulting data of the different methods are shown in Figure 14. The results calculated with the long wavelets show low correlations with the results of the short wavelets and the Hilbert transform. The correlations between the short wavelets and the Hilbert transform are almost 1, indicating that the derived results are similar. Only for low frequencies, the correlation between the short wavelets and the Hilbert transformation are relatively low, too. This is caused by the wavelet problems below 20 Hz (see the description of Wavelet Analysis above).

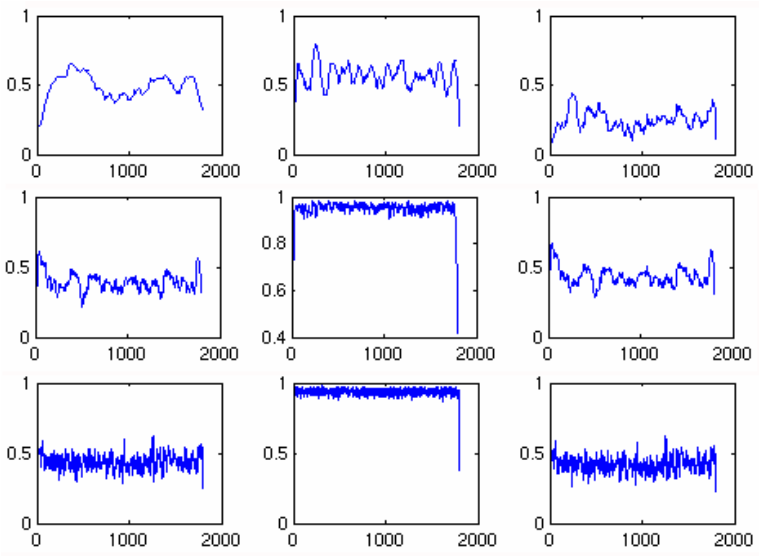


Figure 14 - Phase coherence of signals resulting from different transformation methods.
Left: short wavelets and long wavelets; Middle: short Wavelets and Hilbert transform; Right: Hilbert transform and long Wavelets. Upper row: 4 Hz; Middle row: 15 Hz; Lower row: 55 Hz.
Only the Wavelet Analysis with short Wavelets and the Hilbert transform produce compareable results for middle and high frequencies. (X-axis: time, Y-Axis: correlational value)

To see, whether the differences in phase cancel out by further processing steps, I calculated the Phase Locking Value (see 2.4.2.5.2) between two channels, each pair calculated with the three methods. The results resemble the previous ones (see Figure 15).

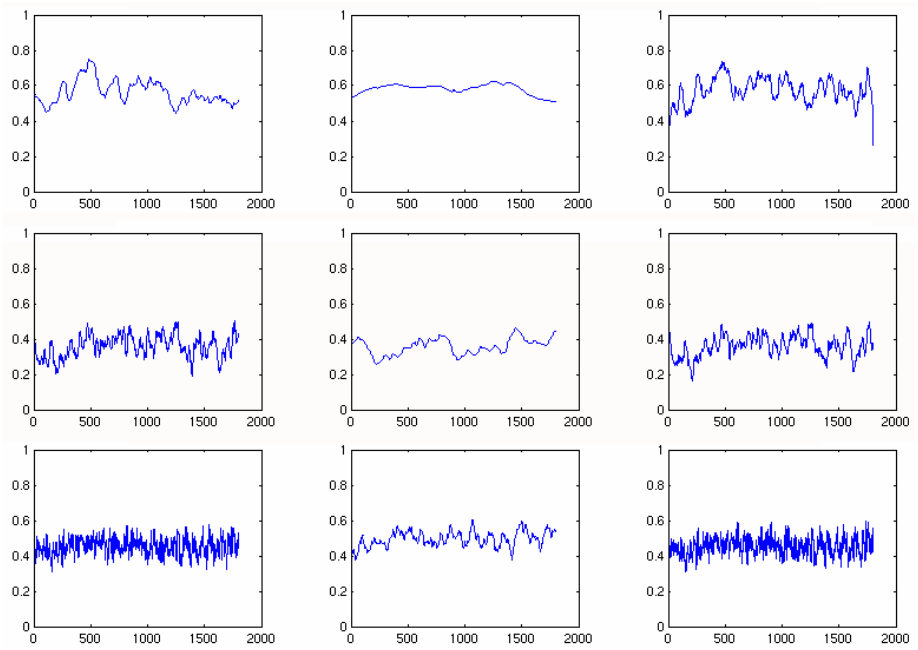


Figure 15 - Phase Locking Values for two signals transformed with different methods.
Left short wavelets; Middle: long wavelets; Right: Hilbert transform.
Upper row: 4 Hz; Middle row: 15 Hz; Lower row: 55 Hz.
For middle and high frequencies, the short wavelets and the Hilbert transform produce comparable results (X-axis: time, Y-Axis: PLVs).

The most noticeable difference is the extremely low time resolution of the long wavelets, resulting from the long duration of the used wavelets. While the PLVs of the short wavelets are similar to those of the Hilbert transform, the long wavelets produce very different results, resembling the extremely smoothed result of the short wavelets. A possible advantage of the low resolution of longer wavelets is mentioned by Trujillo and colleagues (Trujillo, Peterson, Kaszniak and Allen, 2005): As these wavelets integrate over longer time ranges, they are more powerful to detect phase synchronization. The drawback is a loss of time resolution so that two "real" synchronization events could melt into one measured event.

Finally, I visually compared the results of the Hilbert transform with the results produced by EEGLab via the Short Term Fourier Transformation. Again, the results were quite similar, leaving only the method of Lachaux and colleagues producing different data.

Therefore I decided to use the wavelet method of (Tallon-Baudry, Bertrand et al., 1997) to extract the energies of the signal (for 20 to 70 Hz, see section 2.4.2.5.3) and the Hilbert Transformation for the extraction of the phase of the signal (for 4 to 50 Hz). Additionally, to stay in line with previous research, I used the long wavelets of Lachaux et al. (Lachaux, Rodriguez et al., 1999) for higher frequencies (35 to 50 Hz) for comparison. I judged this to be the best compromise to use

the advantages in time frequency resolution of the wavelet method for the gamma energy and to use the low frequency abilities of the Hilbert transform for the analysis of long range synchronization.

2.4.2.5.2 Comparison of Phase Locking Measures

After obtaining the time frequency data, I wanted to analyze the signal with measures for synchronization. I found several methods in the literature to calculate synchronization between two signals: the Phase Locking Value (PLV) (Lachaux, Rodriguez et al., 1999), the Event Related Phase Cross Coherence (ERPCOH) (Delorme and Makeig, 2004) and Magnitude Squared Coherence (MSC). Additionally, the measurement of induced energy (see 2.4.2.5.3) is used in several studies (see 1.2.2.1) as measurement of local synchronization within one channel.

In principle, the PLV measures the stability of the phase lag between two signals for a given frequency over trials. After extracting the instantaneous phases $\phi_1(f,t,n)$ and $\phi_2(f,t,n)$ (where *f* is frequency, *t* is time and *n* trials) of the signals (see 2.4.2.5.1) for the frequency of interest, the phase difference is defined as

$$\theta(f,t,n) = \phi_1(f,t,n) - \phi_2(f,t,n).$$

To obtain the PLV, the phase differences of all trials are added up and averaged over trials:

$$PLV(f,t) = \frac{1}{N}\left|\sum_{n=1}^{N} \exp(i\theta(f,t,n))\right|$$

N denotes the total number of trials. The result is a value between 0 and 1, defining the stability of the synchronization between the two signals.

Graphically, one could imagine each phase difference of each trial as a vector of length 1 and the angle of the phase difference. Adding up all vectors of all trials and dividing the length of the summed vector by the number of trials results in the PLV (see Figure 16).

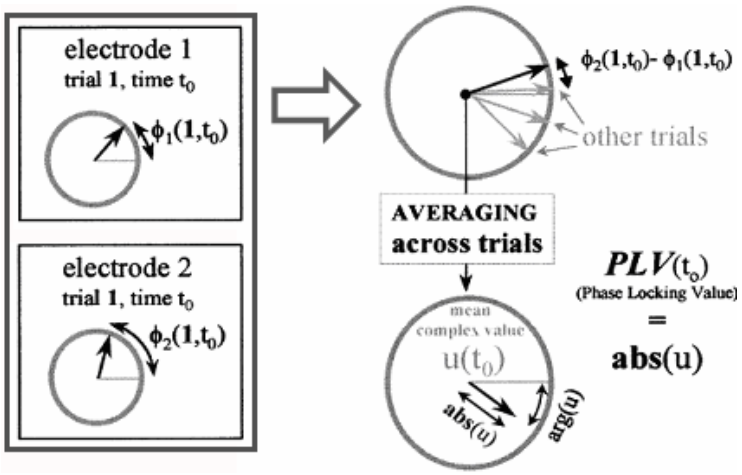


Figure 16 - A graphical description of calculating the PLV.
The phase differences of all trials (left) are averaged as vectors on the complex number plane. The PLV is the length of the resulting average vector (modified from (Lachaux, Rodriguez et al., 1999)).

Another measurement for synchronization is the Event Related Phase Cross Coherence. This measurement is implemented in EEGLab and is defined by:

$$ERPCOH^{a,b}(f,t) = \frac{1}{N}\sum_{n=1}^{N}\frac{F_n^a(f,t)F_n^b(f,t)^*}{\left|F_n^a(f,t)F_n^b(f,t)\right|}.$$

$F_n(f,t)$ denotes the Fourier spectrum of trial *n*, at frequency *f* and time *t*, while

$F_n^a(f,t)^*$ denotes the complex conjugate of this spectrum.

Comparing both formulas for the PLV and the ERPCOH, the measurements are practically equivalent. In fact, visually comparing ERPCOH and the PLV on my data for a randomly selected pair of channels revealed no difference.

Finally, the method of Mean Squared Coherence (MSC) is defined by

$$MSC^{a,b}(f,t) = \frac{\left|\sum_{n=1}^{N}F_n^a(f,t)F_n^b(f,t)^*\right|^2}{\sum_{n=1}^{N}\left|F_n^a(f,t)\right|^2\cdot\sum_{n=1}^{N}\left|F_n^b(f,t)\right|^2}$$

(compare Bruns, 2004).

According to Lachaux and colleagues (Lachaux, Rodriguez et al., 1999), this measurement has the drawback that the amplitude of the signal influences the coupling value. Therefore, it does not represent pure phase locking.

I decided to use the PLV for my analysis, as it stays in line with previous research (Lutz, Lachaux et al., 1999; Rodriguez, George et al., 1999), is practically similar to ERPCOH, and it avoids the drawbacks of the MSC.

2.4.2.5.3 Calculation of Induced and Evoked Energies

As measure of the synchronization of local neural assemblies, several studies (Tallon-Baudry, Bertrand et al., 1997; Keil, Müller et al., 1999; Rodriguez, George et al., 1999; e.g. Kaiser, Hertrich, Ackermann, Mathiak and Lutzenberger, 2005) use the energy of the signal at a specific electrode (for the mathematical definition, see below). They argue that high energy can only be measured if many neurons below the electrode fire together so that each neurons signal is not cancelled out.

There are two types of energy that can be calculated: evoked and induced energy. Evoked energy is phase locked to the stimulus onset (or another event) and is basically the energy of the event related potential (ERP). The ERP is calculated by first substracting the mean baseline from each trial, and second averaging the resulting data over trials. The drawback of this approach is that it eliminates peaks, not phase locked to stimulus onset. The problem is shown in Figure 17.

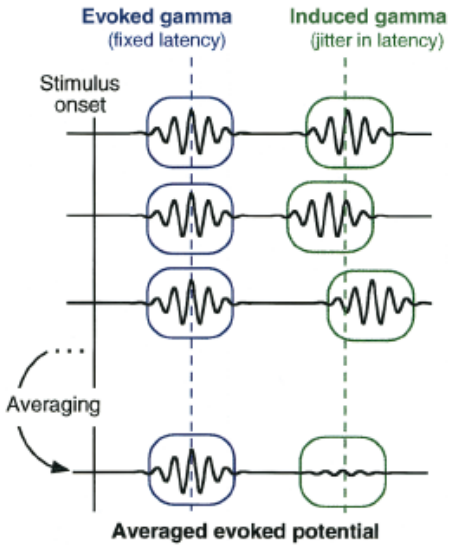


Figure 17 - The difference between evoked and induced energy.
Activity not phase-locked to stimulus onset can cancel out (from Tallon-Baudry and Bertrand, 1999).

To avoid this disadvantage, one has to calculate the induced energy: first measuring energy of each trial and then averaging the results over trials. This way, activity not phase locked to stimulus onset stays in the averaged result if it is not noise that occurred only within a few of the trials.

I calculated both, evoked energy, and induced energy.

Energy in general is defined as:

$$E_n(f,t) = \left\lfloor F_n(f,t) \right\rfloor^2$$

$F_n(f,t)$ denotes the spectrum (in this case the complex result of the Wavelet Analysis) of one trial (or the ERP). Energy is calculated for each time bin *t* and frequency bin *f* of each trial *n*.

Calculating energies using the Wavelet Analysis (see section 2.4.2.5.1), I found edge effects at the begin and the end of a signal in time (an effect also mentioned elsewhere (Herrmann and Mecklinger, 2000)). To avoid these effects, I calculated energies with an overhead of 250 ms at the edges. This overhead was removed afterwards.

For induced energy, there are two ways to normalize the data of each trial: One way is to substract the baseline of each trial (analogue to calculating the ERP) (Tallon-Baudry, Bertrand et al., 1997). Another ways is to also substract the mean of the baseline each energy value and devide the result by the standard deviation of the baseline (Rodriguez, George et al., 1999).

$$E_n^{normalized}(f,t) = \left[E_n(f,t) - mean(Base(f))\right] / std(Base(f))$$

A problem of these approaches are noisy baselines influencing the final result. It is possible that a baseline with an accidental peak in energy lowers the normalized result. As the relatively high mean of this baseline is subtracted from the stimulus related data the result is much lower than data of other trials, even if it contains the same basic processes. If there is a relatively low number of trials, this could influence the final result. Analyzing the data, I realized that about 15% of the baselines are far above average (>2 std.). Therefore, I calculated two sets of results, one using all trials and another one using only the lower 85% of the trials.

As a final step, Rodriguez et al. (Rodriguez, George et al., 1999) averaged the induced energy over electrodes to get some kind of global energy measurement. As energy in the meaning described above is a measurement of local synchronization, I judged this last step as questionable. Therefore, I also calculated mean energies for pooled electrode sites (topographical regions of interest (Herrmann and Mecklinger, 2000)), additionally to the global measurement of Rodriguez and colleagues. These regions were defined over occipital (visual) and temporal (auditory) channels.

2.4.2.5.4 Calculation of Synchronization

I calculated the synchronization between each pair of electrodes with the PLV (see 2.4.2.5.2). To normalize the PLVs of each pair and to weight the PLV by the distance between electrode pairs, I substracted the mean of the baseline and devided the result by the standard deviation of the baseline (compare Rodriguez, George et al., 1999).

As for energies, this baseline normalization bears some risks if there is a low number of trials in a condition. The baseline PLVs reflect the mean synchronization over trials during the baseline. As the baseline duration is randomized, there should be no specific effect in this interval, and any effects of single trials should zero out by averaging. The baseline should only reflect the static level of synchronization between two electrodes due to proximity. If there are not many trials to calculate the PLVs from, this assumption might be violated, because the noisy effects of single trials could not zero out completely. Therefore, some baselines might show extremely high or low PLVs that do not only reflect the static synchronization. To solve this problem, I used a common baseline for all conditions: I calculated the PLVs of all trials of one subject and used this as baseline for the PLVs of each condition. This approach is based on two assumptions. First, there cannot be a condition-dependent effect in the baseline data. Second, the calculation of the data over all trials provides enough trials to calculate a representative baseline.

The result of these calculations represents a huge amount of data (channel x channels x frequencies x time) 4.6.3. Therefore, I calculated several averaged measurements on this data:

- The mean PLV between
    - Occipital ('visual') and temporal ('auditory') pairs: Assuming that the integration of information is done in areas for early stimulus processing, synchrony could occur between these electrodes.
    - Occipital and parietal pairs: Assuming that synchrony is occurring between areas involved in higher levels of processing, synchrony for detecting the animal could be found here.
    - Temporal and parietal pairs: Synchrony for identifying the sound and integrating this information might be found here (and has been found previously (von Stein, Rappelsberger et al., 1999)).

- A global PLV by averaging the PLV of all electrode pairs (compare Lutz, Lachaux et al., 1999; Rodriguez, George et al., 1999).

The last measurement generally indicates how synchronized many sources are at a specific time, but it is unclear what this means in detail. To understand this, the topography of synchrony at this moment has to be considered, too.

2.4.2.5.5 Reduction of Data Volume

Without a reduction of data volume, a dataset for one subject, one condition, and one interesting event would have a size of about 5.4 GB (for a detailed calculation and a description of how to reduce data volume, see Appendix F).

I reduced the volume of data by lowering the precision (double to single), optimizing use of storage (cutting symmetric matrices into half), and using an exponential stepping for the selection of interesting frequencies. This resulted in the analysis of the following frequencies:

- Energy: 20,21,…,25,26; 28,30,…,54,56; 59,62,…,78,80
- PLV: 4,5,…,24,25; 26,28,…,48,50

Finally, the amount of data for one data set was about 1.1 GB.

2.4.2.6 Statistics and Hypotheses

The statistics I calculated apply to the conditions defined in section 2.4.2.4.

2.4.2.6.1 Behavioral Data

As I expected the auditive cues to improve the performance of subjects, I checked the difference between the reaction times of the congruent (C11) and the incongruent (C12) trials with a Wilcoxon Signed Rank test (as reaction times are not normally distributed, I could not use parametric tests like an ANOVA). I calculated the same statistics for the percentage of found animals in each condition.

Additionally, I checked the correlation between reaction time differences (C11-C12) of objective and subjective conditions to see if there are differences in a behavioral measure.

I also checked for trial to trial effects. To see if the type of the previous trial influences the reaction time in the current trial, I calculated mean reaction times for each combination C11-C12, C11-C11 etc.). For each non-matching pair (e.g. C11-C12), I compared the reaction times with the

corresponding matching pair (e.g. C12-C12). Taking the possible high inter subject variance in reaction times into account, I tested for significance of differences with a Wilcoxon Signed Rank Test.

To see how good subjects judged the stimuli and how good objectively and subjectively defined conditions match, I calculated descriptive data.

2.4.2.6.2 EEG Data

Basically, the statistical analysis consisted of two parts: a hypothesis based part and an explorative part. As test for significance for all comparisons, I used the Wilcoxon Signed Rank test, as it is robust against inter subject variability and can be applied to data not normally distributed. The statistics of the explorative part can only be of descriptive nature, as it is heavily based on multiple comparisons that cannot simply be corrected by alpha correction (compare von Stein, Rappelsberger et al., 1999). I will first present the analysis based on prior findings and theories. Then I will present the explorative, descriptive part.

In general, there were two approaches to check the data for effects based on hypotheses:

- checking the relevant time windows and frequency ranges defined by prior studies
- looking for specific effects in all conditions (like the strongest peak) described by previous studies

These are the hypotheses based questions I investigated statistically:

- Are there differences in induced energies directly after stimulus onset between C1 and C2? If gamma energy was a correlate of binding, one could observe a difference between pictures with an animal and without an animal. According to several authors (Tallon-Baudry and Bertrand, 1999; Herrmann, Munk et al., 2004; Kaiser and Lutzenberger, 2005), the induced gamma response should happen in a time window of 200-350 ms and within a frequency range of 40-70 Hz. Therefore, I looked at three time windows of 50 ms in that range.
- Are there differences in induced energies directly after the auditive stimulus onset between C1 and C2 or C11 and C2? According to Tallon-Baudry and Bertrand (Tallon-Baudry and Bertrand, 1999), there are differences between active and passive listening as early as 80 ms. Therefore, I was looking for differences with time windows of 100 ms from 80-600 ms after onset of the auditive stimulus in a frequency range of 30-50 Hz. As

the effect might be stronger compared to matching stimuli, I compared this for C1/C2 and for C11/C2.

- Are there differences in synchrony after stimulus onset between C1 and C2? A relationship between induced gamma energy and synchronization has been proposed by Hermann and colleagues (Herrmann, Munk et al., 2004). Therefore, I looked at the point of highest difference of gamma energy between C1 and C2 for effects in synchronization. Another way to look for an effect is using time frequency windows. Rodriguez and colleagues (Rodriguez, George et al., 1999) found increased synchrony from about 150-300 ms in the gamma band (35-45 Hz) as proposed by von Stein and Sarnthein (von Stein and Sarnthein, 2000) for local feature binding. Therefore, I also looked at windows of 50 ms from 150-350 ms to check for differences between C1 and C2 in this frequency range.
- Are there differences in synchrony after onset of the auditive stimulus between C1 and C2 or C11 and C2? The relationship between the visual gamma response and synchrony could also hold for auditive stimuli. Therefore, I looked at time frequency windows of 100 ms, 100-600 ms after stimulus onset in the gamma band (35-45 Hz). Additionally, von Stein and Sarnthein (von Stein and Sarnthein, 2000) propose synchrony in the lower beta band (13-18 Hz) for supra modal processing. Therefore, I also looked at these frequencies within the time window. As the effect might only hold for congruent stimuli, I looked for effects in C1/C2 and C11/C2.
- Is there a difference in synchronization between C11 and C12 before the recognition button press? Looking at the lower beta frequency range (von Stein and Sarnthein, 2000), there could be a difference between matching and not matching stimuli. I looked at the average of an interval of 1 second before button press.

The second part of the analysis was of explorative nature. One way to look for interesting effects was to look at specific points (e.g. the largest difference) in the data. Another way was to look for patterns by browsing the whole data. Both ways deliver results that are not statistically reliable. They are not based on prior hypotheses and not include multiple comparisons. Therefore, the answers to the following questions are of descriptive nature.

- Are there differences in induced energies after the auditive stimulus onset between C11 and C12? Differences could be caused by a different degree of attention to the auditive cue or insight into the mismatch of stimuli. I was interested in the difference within time

windows of 100 ms from 100-600 ms after onset of the auditive stimulus in a frequency range of 20-50 Hz (see above).

- Are there differences in induced auditive energies around the detection button press between C11 and C12? As for the last question, I was interested in the difference within time windows of 100 ms in a frequency range of 20-50 Hz (see above).
- Are there differences in synchrony before the button press between C11 and C12 for different frequency bands (other than the lower beta band) over time? I used the bands defined by von Stein and Sarnthein (von Stein and Sarnthein, 2000) (also, see above) to look for differences in frequency bands. To look at the dynamic aspect within the bands, I used time windows of 100 ms (for each frequency bands).

Additionally, to get an overview on the global properties of possible effects, I calculated contrasts and statistical significance maps over subjects for each point in the time-frequency windows of analysis.

Especially for the choice of points for comparison in the explorative part, normalization of data might be a question of importance. As I used paired tests for the statistics, subject variability should not have influenced the results this way. But the choice of points of interest was based on the grand average of the data. Without normalization, subjects with a stronger baseline level could have a stronger influence on the result. Therefore, I also compared the results of z-transformed energy data to check for differences.

I also analyzed ERPs to get an overview of the data and to get an impression of the time-windows for analysis. As ERPs are not the focus of this thesis, I will only present an overview of this data. To remove residual noise from the ERPs, I used a low pass filter with a 40 Hz high cut off.

3 Results

In this section, I will present the results of the data analysis. After describing the behavioral effects in the experiment, I will present the effects for EEG energy, synchrony, and ERPs.

3.1 Behavioral

First, I checked if the task with its conditions worked successfully. In fact, I found a significant difference for reaction times between the congruent and incongruent condition (see Figure 18). Subjects found animals in the congruent condition much faster (Wilcoxon Signed Rank test, $p<0.01$). For the detection of animals, there was only a trend in the data for congruent stimuli being recognized more often (Wilcoxon Signed Rank test, $p=0.062$).

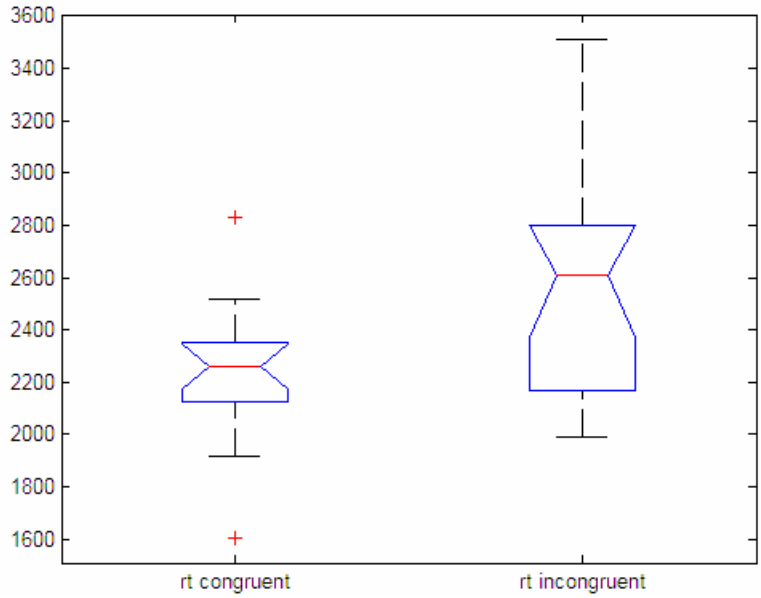


Figure 18 - Reaction times for the congruent and the incongruent condition.
Congruent stimuli are detected significantly faster than incongruent ones. Crosses represent outliers, box include the first and third quartile and the median (red line). Whiskers extend to 1.5 interquartile range. Notches represent an estimate of the incertainty of the median (X-axis: condition, Y-axis: time in ms).

Checking for differences between objective and subjective conditions, there was a strong correlation ($r=0.93$) between reaction time differences (C11-C12) for both ways to define the conditions.

An overview of categorizations by subjects can be seen in Figure 19. The majority of congruent stimuli were correctly judged as congruent (mean=66%) or categorized as no animal (mean=23%), and the majority of incongruent stimuli were correctly judged as incongruent (mean=68%) or categorized as no animal (mean=27%). Only few pictures without an animal were categorized as congruent or incongruent (mean=5.3%). Overall, there was a high percentage of stimuli judged

correctly (mean=72.7%). This shows that for found animals, objective and subjective categorizations are quite compareable.

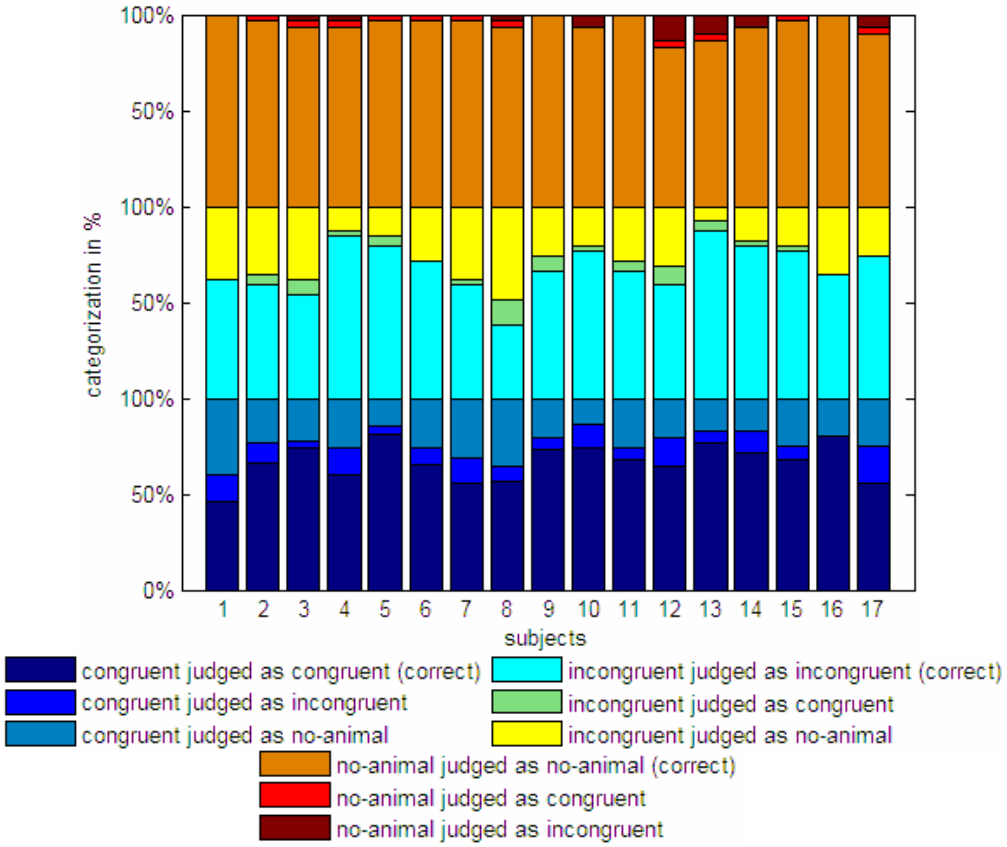


Figure 19 - Judgments by subjects in percent.
Subjects judged most images right (lowest bar of each group). High variance between subjects is visible in this graph, especially for the incongruent group (middle group).

Additionally, I analyzed trial-to-trial effects. There were no significant effects for previous trials influencing the next trial.

Summarizing, there were significant effects for the congruency condition influencing the speed and the reliability of the judgments of subjects. For congruency conditions, reaction time effects were similar for objective and subjective classification of stimuli. For found/not found conditions, subjective conditions could be more adequate, because of the high percentage of animals not found.

3.2 EEG Data

I will present the results of the EEG data analysis in two steps. For each feature (energy, and synchrony), I will show the statistical results of the tests, described in section 2.4.2.6.2) and the

according time frequency graphs. Following this, I will present the topographical properties of the significant results.

3.2.1 Induced Energies

I will first present the hypothesis based results. After that, I will describe the explorative results. I checked, if the results might be influenced by different methods of calculation (e.g. normalization, see section 2.4.2.6.2, or high baselines, see section 2.4.2.5.3). I will also present these results at the end of this section.

The first question concerned differences in gamma energy between C1 and C2 after stimulus onset. For the first 3 time windows, energy in C1 was stronger than in C2 (200-350 ms, $p<0.01$) (see Figure 20). The point of strongest difference was at 288 ms after stimulus onset at 56 Hz. Within C1, there was no difference between C11 and C12.

The next event was the onset of the auditive stimulus, 750 ms after the onset of the visual stimulus. I looked at gamma energy at temporal electrodes. Energy in C2 was significantly stronger than in C1 for 3 time windows, becoming stronger over time (1030-1130 ms, $p<0.05$; 1230-1330 ms, $p<0.01$; 1330-1430 ms, $p<0.001$) (see Figure 21). The comparison between C11 and C2 showed a similar pattern, but only for 2 windows (1230-1330 ms, $p<0.05$; 1330-1430ms, $p<0.01$).

There were two explorative results for induced energy: For the time after the onset of the auditive stimulus, 1 window showed energy in C11 stronger than in C12 (1030-1130 ms, $p<0.05$). For the time before the recognition button press, there was also 1 window showing values for C11 significantly higher than for C12 (-500 to -400 ms, $p<0.05$) (see Figure 22). There was no significant time window for occipital energies within this interval.

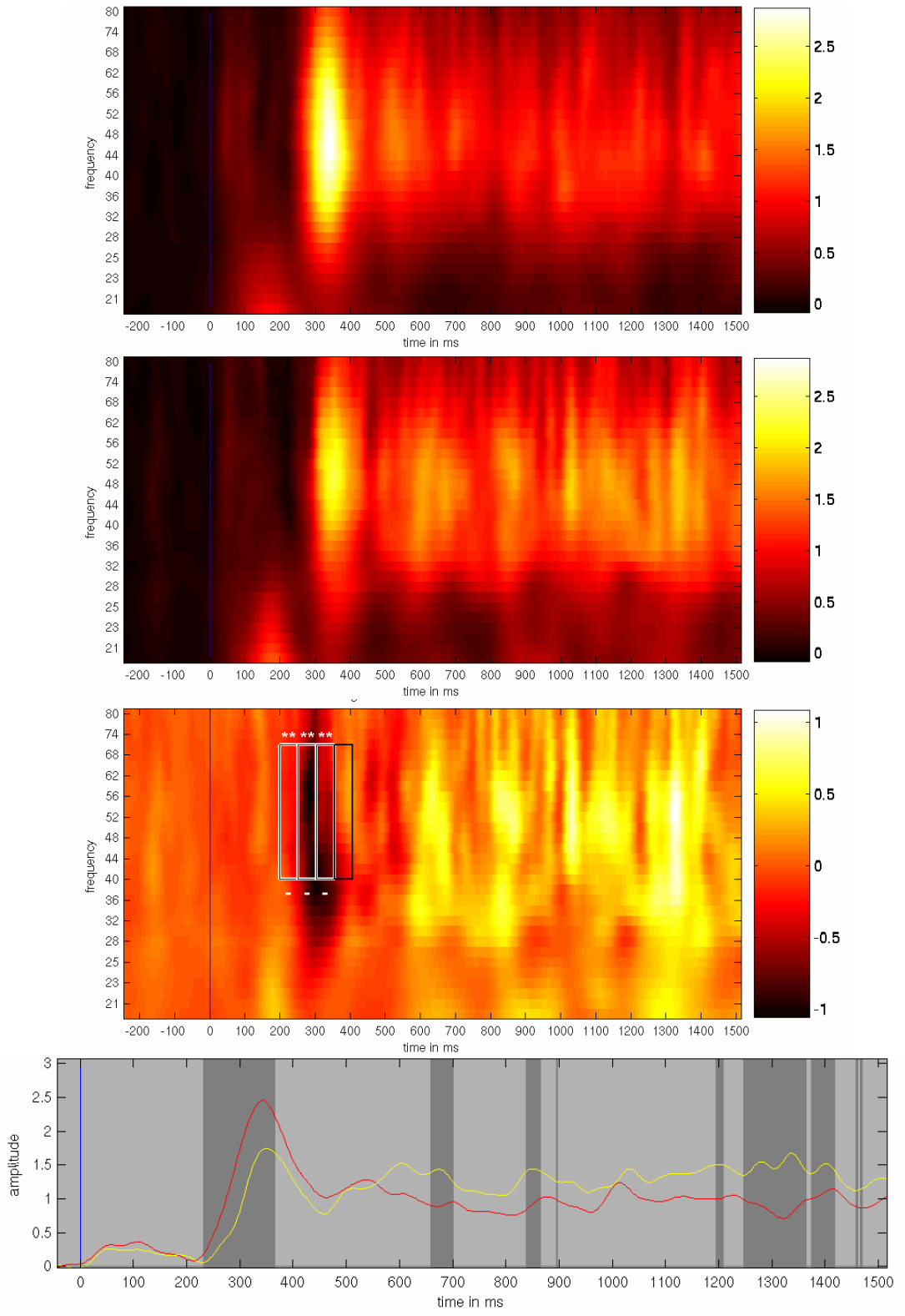


Figure 20 – Occipital energy for the time after stimulus onset.
For C1, C2, and the contrast C2-C1 (top to bottom), values are in standard deviations from the 500 ms baseline: 3 time windows show a highly significant contrast for C1 > C2. In the line plot (bottom) for the classical 40 Hz frequency, the difference in energy for C1 (red) and C2 (yellow) is clearly visible. Significant intervals are shaded in dark grey.

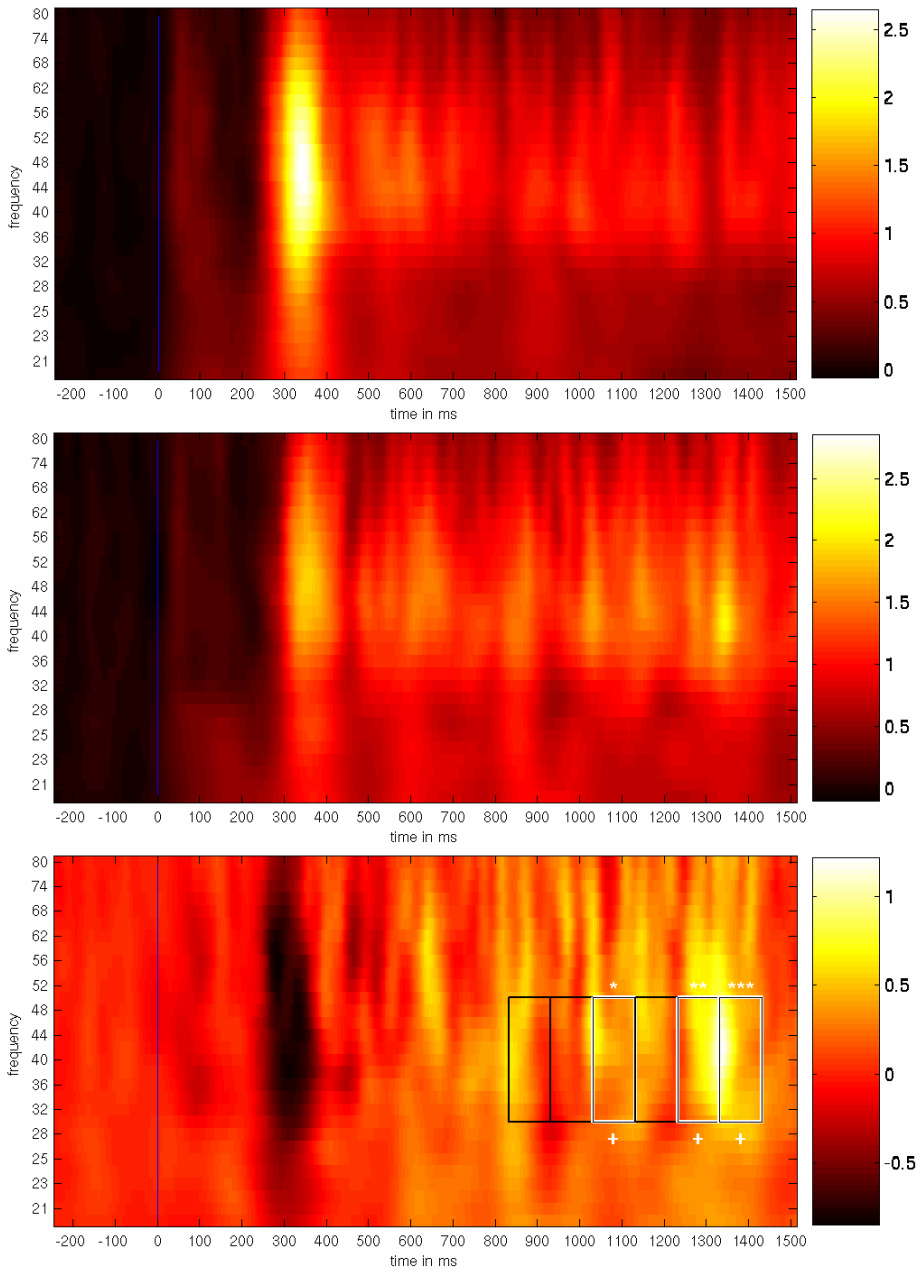


Figure 21 – Temporal energy for the interval after auditive stimulus onset.
For C1, C2 and the contrast C1 - C2 (top to bottom) , values are in standard deviations from the 500 ms baseline: 3 time windows show a significant difference with C2 > C1.

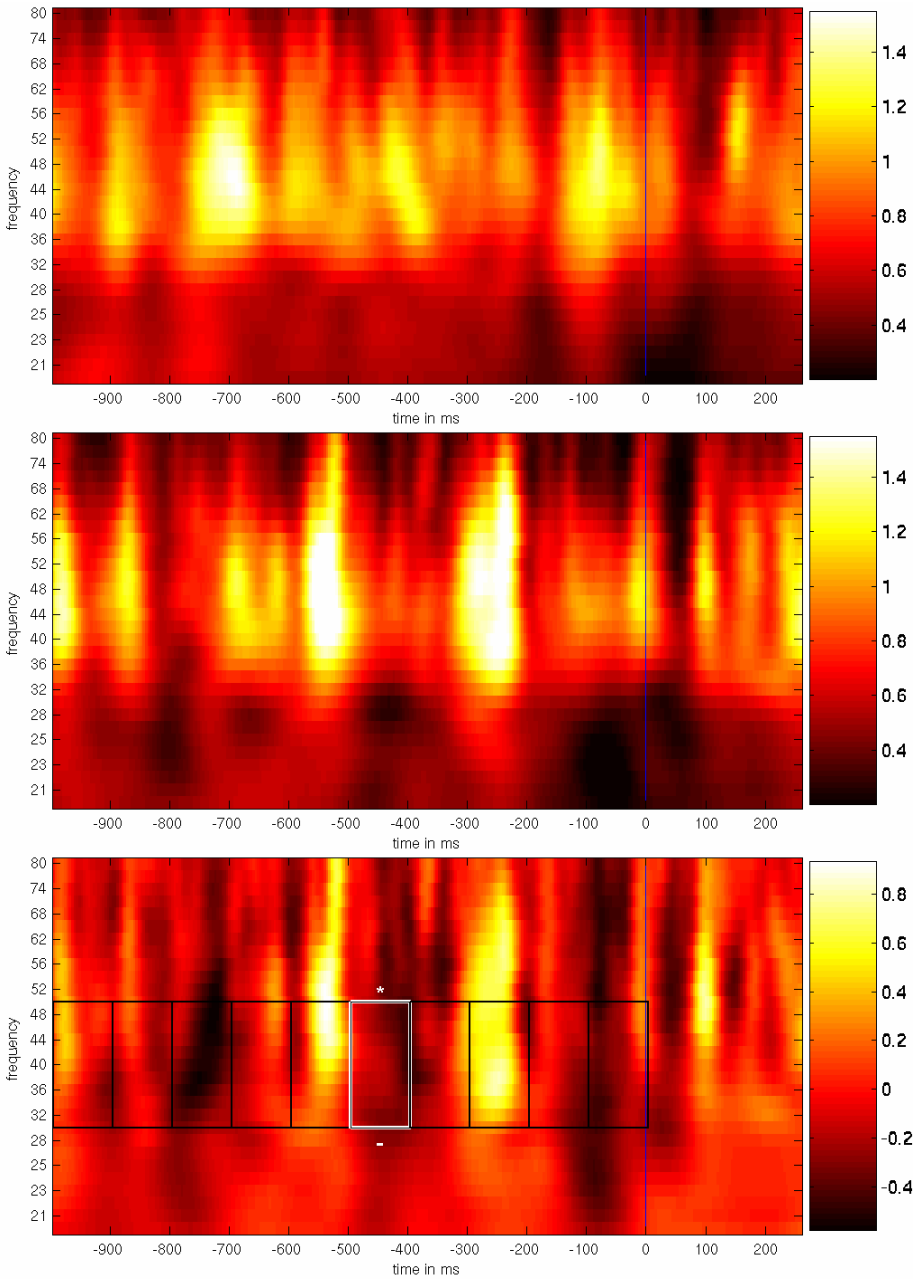


Figure 22 – Temporal energy for the time before the button press.
For C11, C12, and contrast C11-C12 (top to bottom), values are in standard deviations from the 500 ms baseline: Only one time window shows a hardly significant difference for C1 > C2.

As indicated in section 2.4.2.6.2, there could be a difference in the averaged results, if the subject's data had not been normalized before grand averaging over subjects. The hypotheses based results remained constant. For explorative results, there where no differences, either. Summarizing, the analysis of the normalized data showed no differences to the analysis of data not normalized.

A further possibility of inducing artifacts in the data are trial baselines with extraordinary high means. Analyzing the data revealed that on average, 15% of the trials showed high baselines. Therefore, I created data sets for each condition containing 85% of the original trials with the lowest baselines. I tested these data sets for the effects found above. For both, hypotheses based tests and explorative tests, the effects stayed constant.

Finally, to rule out the possibility of unbalanced conditions influencing the results, I created sets with the same number of trials for each condition and calculated the data for the mean of all electrodes. This revealed no visible difference to the unbalanced conditions.

Looking at the summed significance maps (see Figure 23) for C1-C2 after stimulus onset for visual electrodes, a stable global pattern of significance over different approaches is visible. This explains why the results show no differences even though there were minor changes for the energy values in each condition.

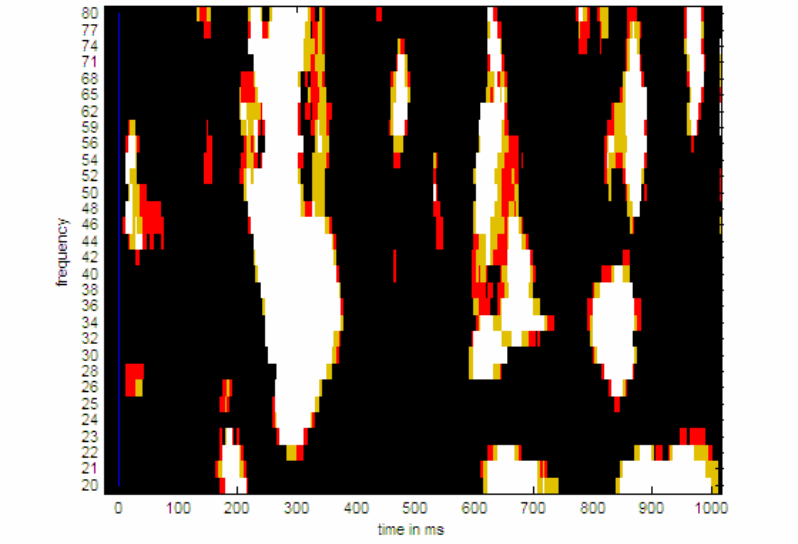


Figure 23 - Summed significance maps of energy for different methods.
The results are very similar (black: not significant; red: significiant with 1 method; yellow: significant with 2 methods; white: significant with all methods)

In the following, I will present the topographies (for 33 electrodes) of the found effects. Figure 24 shows the topography of the strongest peak in energy in C1 after stimulus onset (340 ms, 46 Hz) for C1 and C2. It is strongest under occipital and parietal electrodes.

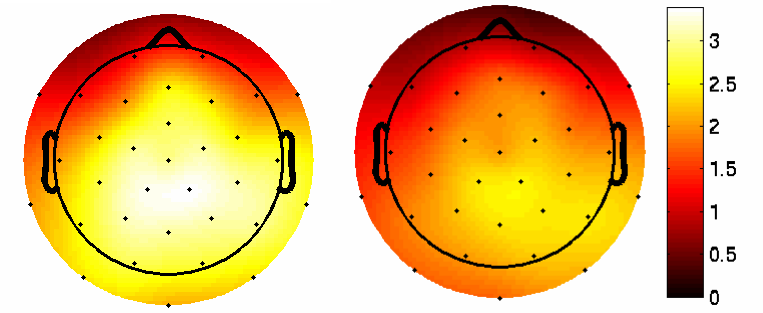


Figure 24 – Topographies of energy for the strongest peak after stimulus onset.
For C1 and C2 (left to right), , values are in standard deviations from the 500 ms baseline.

The topography of the most significant difference between C1 and C2 after onset of the auditive stimulus (1330-1430 ms) is shown in Figure 25.

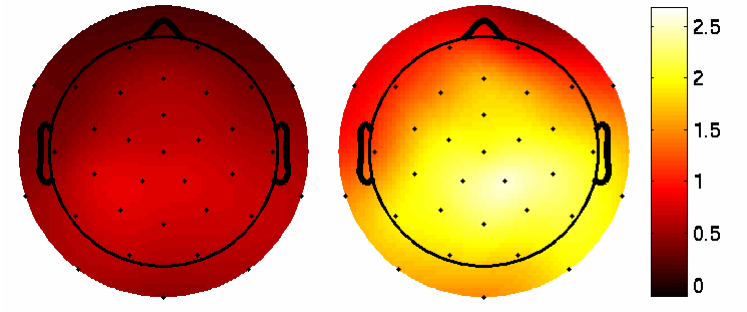


Figure 25 – Topographies of energy for the most significant window after onset of the sound.
For C1 and C2 (left to right), , values are in standard deviations from the 500 ms baseline.

Figure 26 shows the difference in topographies of C11 and C12 for the (barely) significant window before the button press (-500 - -400 ms).

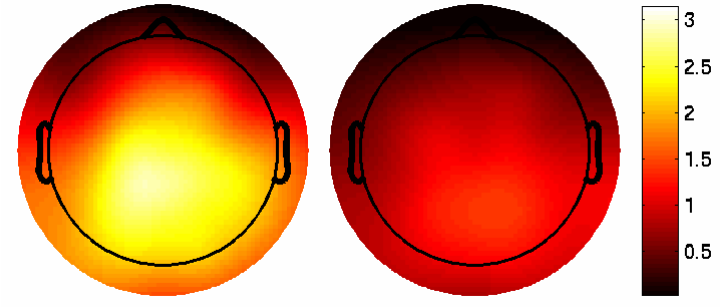


Figure 26 – Topographies of energy for the only significant window before the button press.
For C11 and C12 (left to right), , values are in standard deviations from the 500 ms baseline.

3.2.2 Synchrony

For synchrony, I first looked for effects directly after stimulus onset in mean synchrony values.

I found significantly higher synchrony for C1 near the point of greatest difference in gamma energy between C1 and C2 (288 ms, 50 Hz, $p<0.01$). As phases had been extracted only for 4-50 Hz, I could not check the original point at 56 Hz.

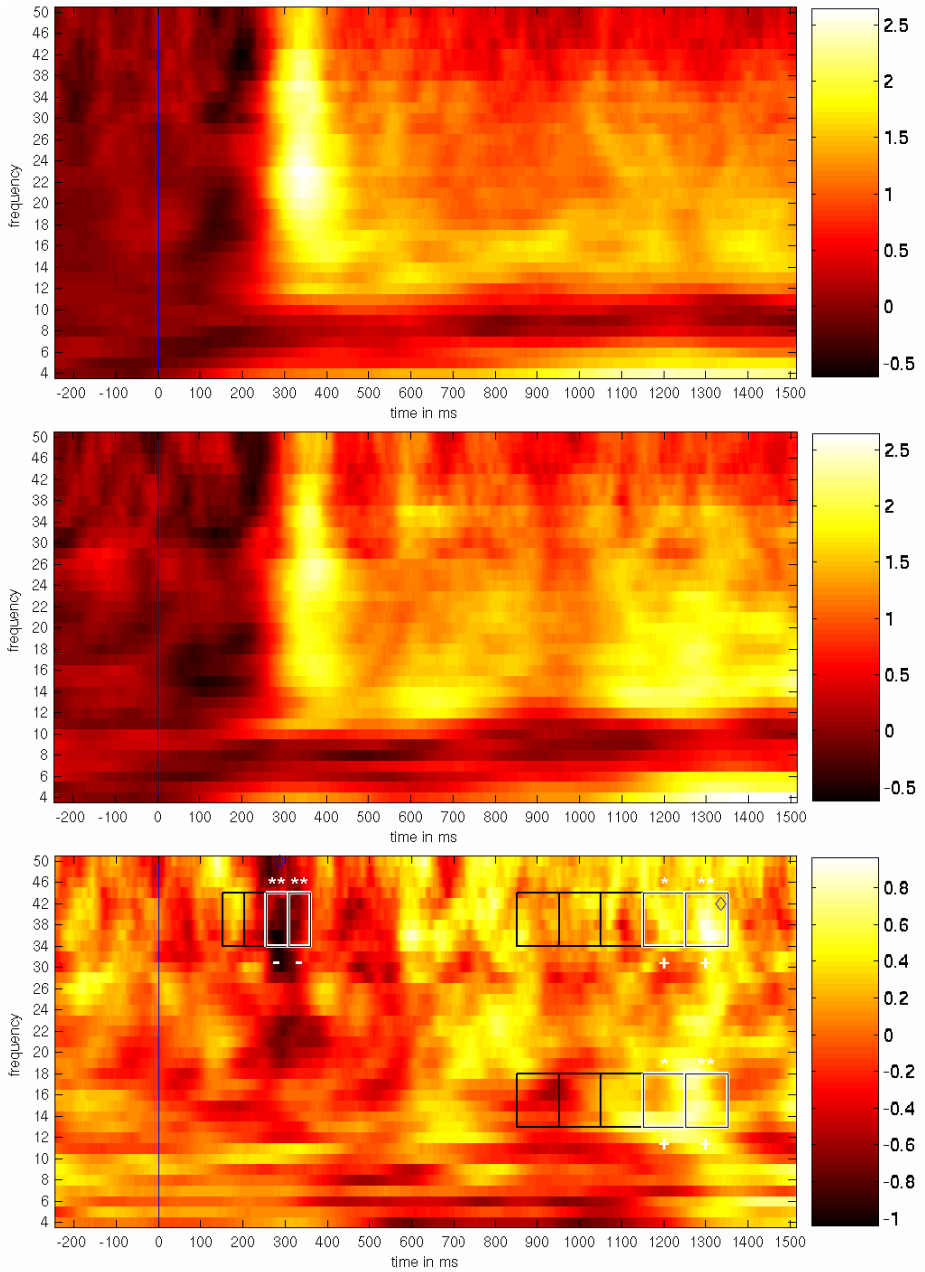


Figure 27 - Synchrony over all electrodes after stimulus onset.
For C1, C2, and contrast C1-C2 (from top to bottom), values in standard deviations from baseline: The first peak after stimulus onset is stronger for C1, while C2 shows higher values for the time after auditive stimulus onset.

2 time windows after stimulus onset also showed a significant difference for C1-C2 (250-350 ms, 35-45 Hz, $p<0.01$), with C1 being stronger (see Figure 27).

After the onset of the auditive stimulus, 2 windows showed a significant difference in the gamma range for C1-C2 (1150-1250 ms, 35-45 Hz, $p<0.05$; 1250-1350 ms, $p<0.01$), with C2 being stronger (see Figure 27). No window was significant for C11-C2. Within the lower beta band, the same 2 time windows showed the same significant effect for C1-C2 (1150-1250 ms, 13-18 Hz, $p<0.01$; 1250-1350 ms, $p<0.001$) (see Figure 27). Checking the gamma range for crossmodal (occipital and temporal) electrodes, there was only 1 window showing a significant difference for C1-C2 (1250-1350 ms, 35-45 Hz, $p<0.05$) and no window for C11-C2. For the lower beta frequency range and crossmodal electrodes, the same 2 windows as for mean synchronies also showed a significant difference for C1-C2 (1150-1250 ms, 13-18 Hz, $p<0.05$; 1250-1350 ms, $p<0.01$) and again no significant window for C11-C2.

Comparing frequency bands for the interval before the recognition button press, I also looked at mean and at crossmodal electrode pairs. For mean values, I found a significant difference within the beta1 band, with C12 stronger than C11 (13-18 Hz, $p<0.05$). For crossmodal pairs, there was a similar difference, even more significant (13-18 Hz, $p<0.01$) (see Figure 28).

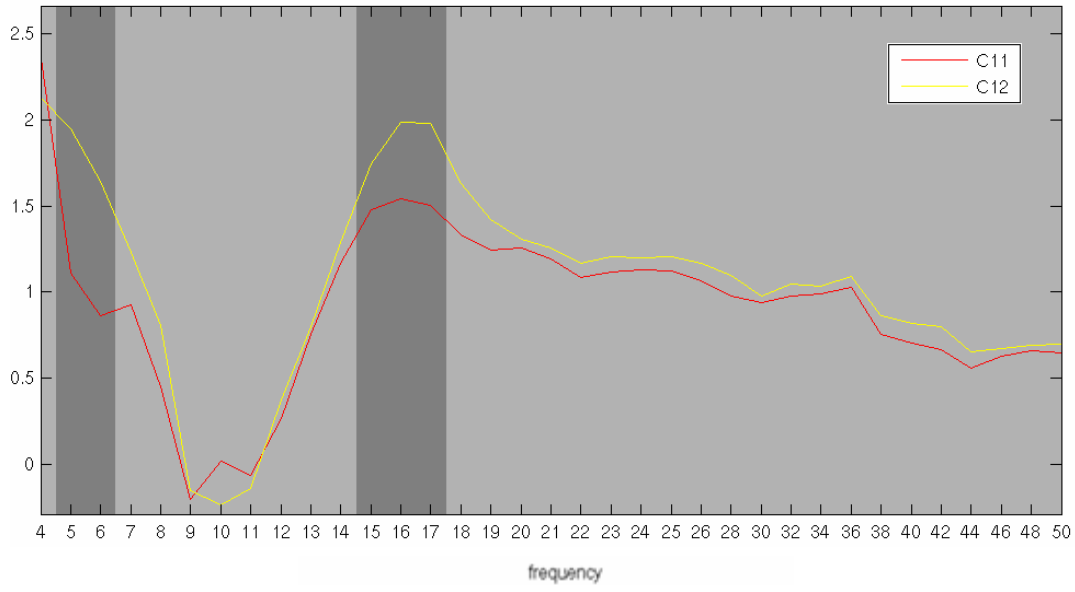


Figure 28 - Average crossmodal synchrony for frequency bands before button press.
C12 shows higher synchrony in the theta and the lower beta frequency band (significant differences are shaded in dark gray) than C11. (Values are in standard deviations from the 500 ms baseline, X-axis: frequencies; Y-axis: synchronization)

Looking at the explorative results, there where several significant differences for the time frequency windows before the button press (see Figure 29). Within the theta band, C2 was stronger than C1 in 2 windows (400-600 ms, 4-8 Hz, $p<0.05$). As expected, there where 2 significant windows in the lower beta band, with C12 stronger than C11 (300-400 ms and 700-800 ms, 13-18 Hz, $p<0.05$). In the higher beta band and the gamma band, there was one significant window, again with C12 stronger than C11 (1-100 ms, 19-23 Hz and 24-35 Hz, $p<0.05$).

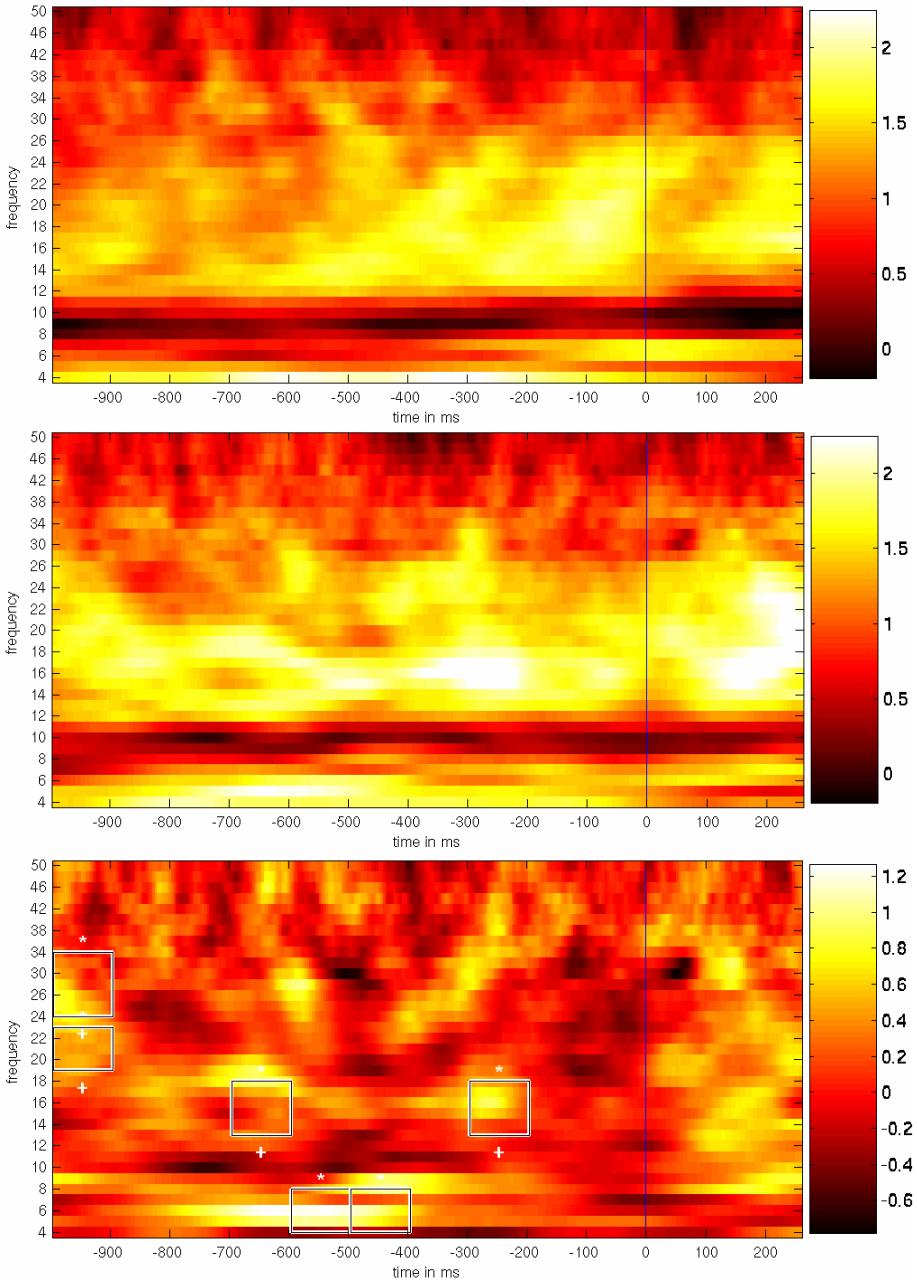


Figure 29 - Synchrony over all electrodes before the button press.
For C11, C12, and contrast C11-C12 (from top to bottom), values are in standard deviations from the 500 ms baseline: The first after stimulus onset peak is stronger for C1, while C2 shows high synchronization for the time after auditive stimulus onset.

Additionally to these results, I computed the synchrony values with the long wavelets used by Lachaux and colleagues (Lachaux, Rodriguez et al., 1999). Comparing synchrony plots (see Figure 30) reveals similar structures, but stronger effects than for the Hilbert transform.

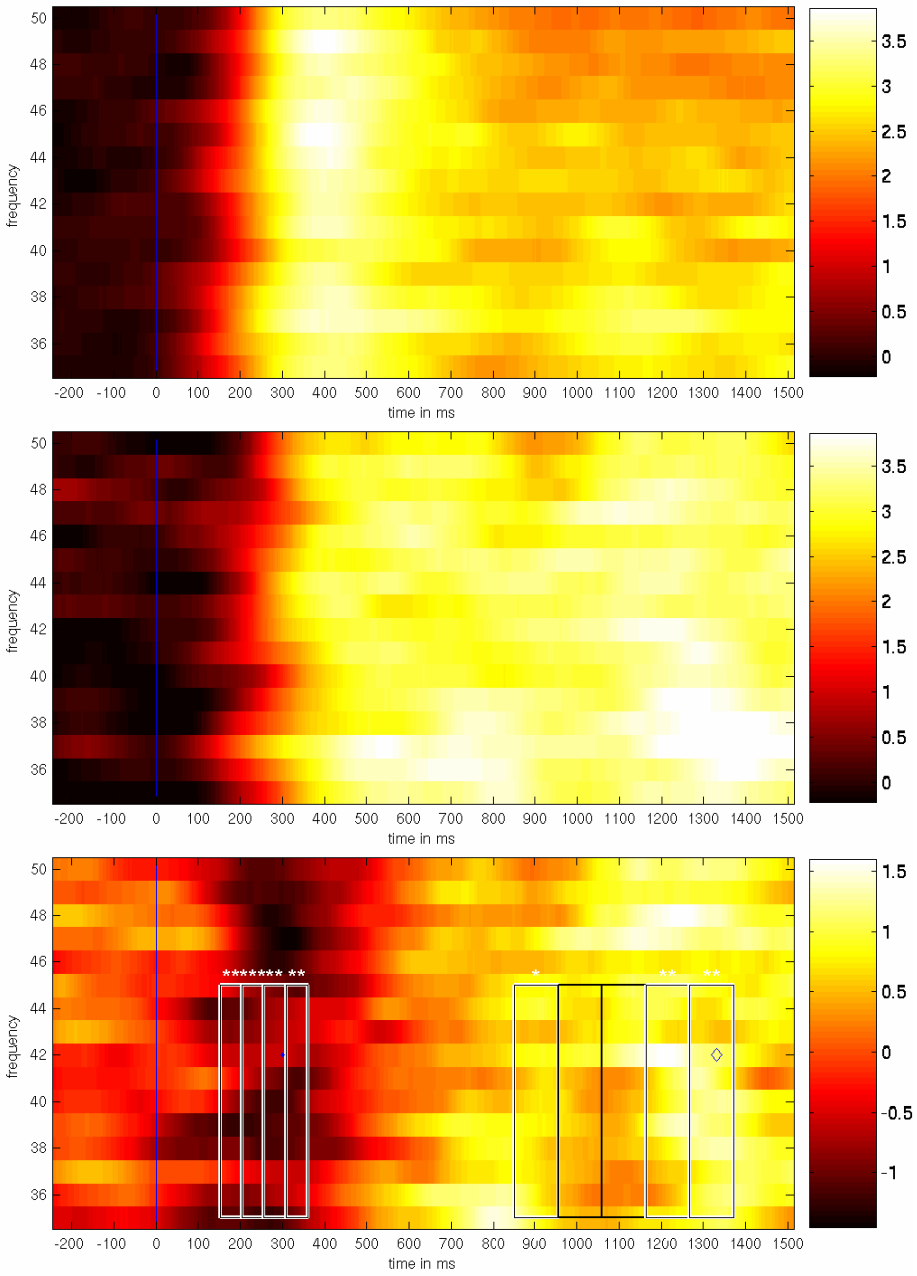


Figure 30 - Mean synchronies for long wavelets, after stimulus onset.
For C1, C2, and contrast C1-C2 (from top to bottom), values are in standard deviations from the 500 ms baseline: After stimulus onset, the difference between C1 and C2 is even stronger than for short wavelets. The results after onset of the auditve stimulus are compareable.

Therefore, I had a closer look at the results of these calculations. As the long wavelets were not applicable to frequencies lower than 30 Hz (see section 2.4.2.5.1), I could only look at gamma band synchrony.

All 4 time windows after stimulus onset showed significantly higher values for C1 compared to C2 (150-350ms, 35-45 Hz, $p<0.01$) (compare Figure 30). For the time after the auditory stimulus onset, C2 was significantly stronger in 3 time windows (850-950 ms, 35-45 Hz, $p<0.05$; 1150-1350 ms, $p<0.01$) (see Figure 30). In 2 windows, C2 showed higher values than C11, but only hardly significant (850-950 ms, 1250-1350 ms, 35-45 Hz, $p<0.05$). For crossmodal electrode pairs, the same windows were significant for C1/C2 (850-950 ms, 1150-1250 ms, 35-45 Hz, $p<0.05$; 1250-1350 ms, $p<0.01$). There was also a hardly significant window for C11/C2 (1250-1350 ms, 35-45 Hz, $p<0.05$). Surprisingly, there were no significant windows for any frequency band comparisons in the time window before the button press.

In the following, I will present the topographical properties of the significant events. As the classical head/line plots miss the temporal dynamics of synchronization, I decided to use an 8x8 time frequency matrix of 8 electrodes of interest: 25 (front l), 24 (temp l), 5 (pari l), 22 (occi l), 21 (occi r), 4 (pari r), 19 (temp r), 18 (front r) (compare Appendix G).

Figure 31 shows the topography of the PLV dynamics of C1 and C2 after stimulus onset. In general, occipital electrodes are the most synchronized electrodes with other areas. The first peak in synchrony is stronger for C1, while C2 shows a stronger ongoing pattern of synchrony. Strongest synchrony can be found for occipital-frontal and occipital-temporal pairs.

Figure 32 shows the topography of C11 and C12 after stimulus onset. Again, the coherence of occipital electrodes is strongest. After the onset of the auditive stimulus, C12 shows higher synchrony in low frequency bands than C1 for occipital-temporal and occipital-frontal pairs.

Finally, Figure 33 shows the time window before the button press for C11 and C12. Differences occur mainly in the beta band at early times and later in the theta band for occipital-frontal and occipital-temporal pairs

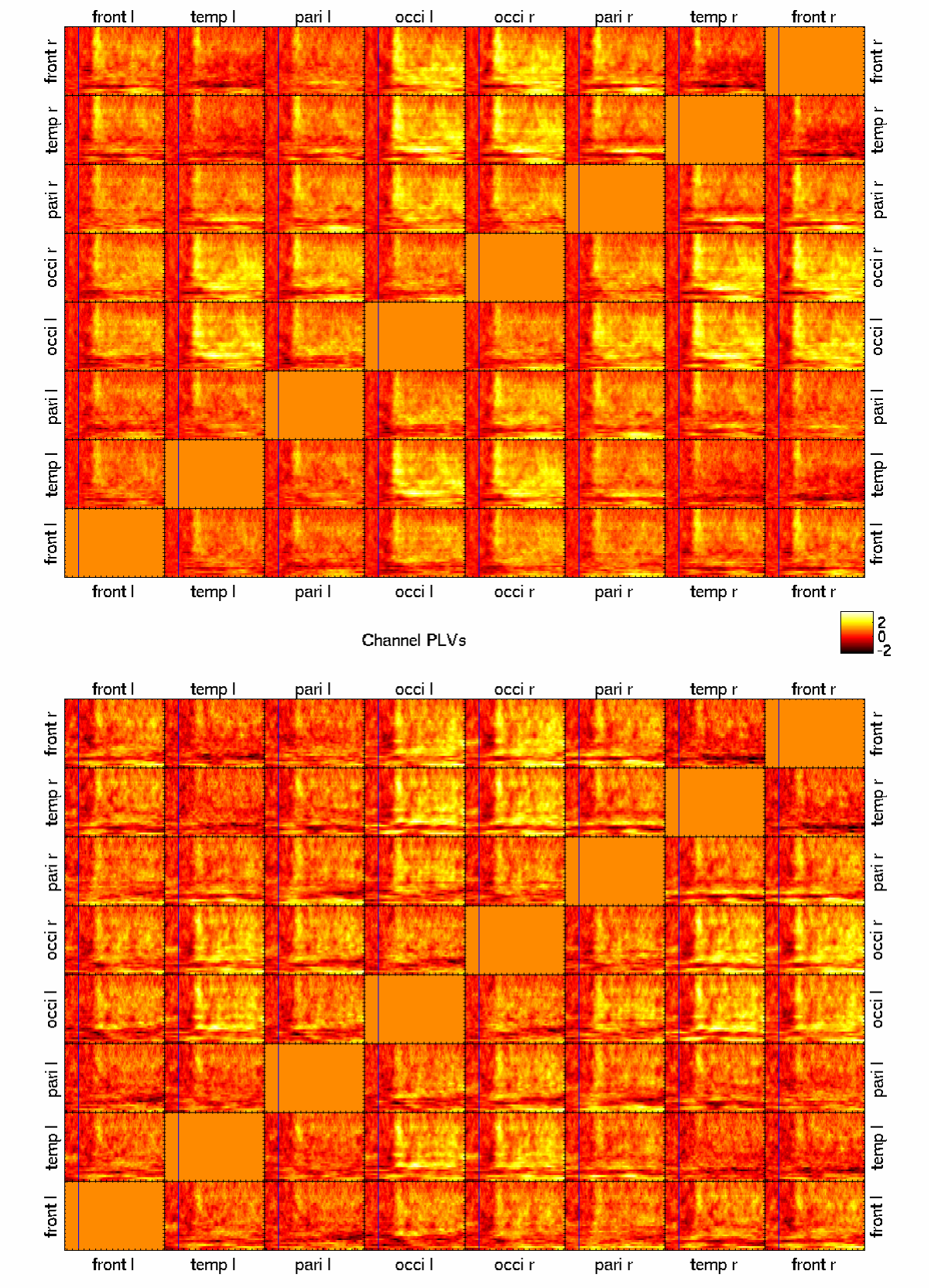


Figure 31 - Synchrony matrix for the time after stimulus onset.
For C1 (top) and C2 (bottom), values are in standard deviations from the 500 ms baseline: High synchrony can be found for occipital-frontal and occipital-temporal pairs.

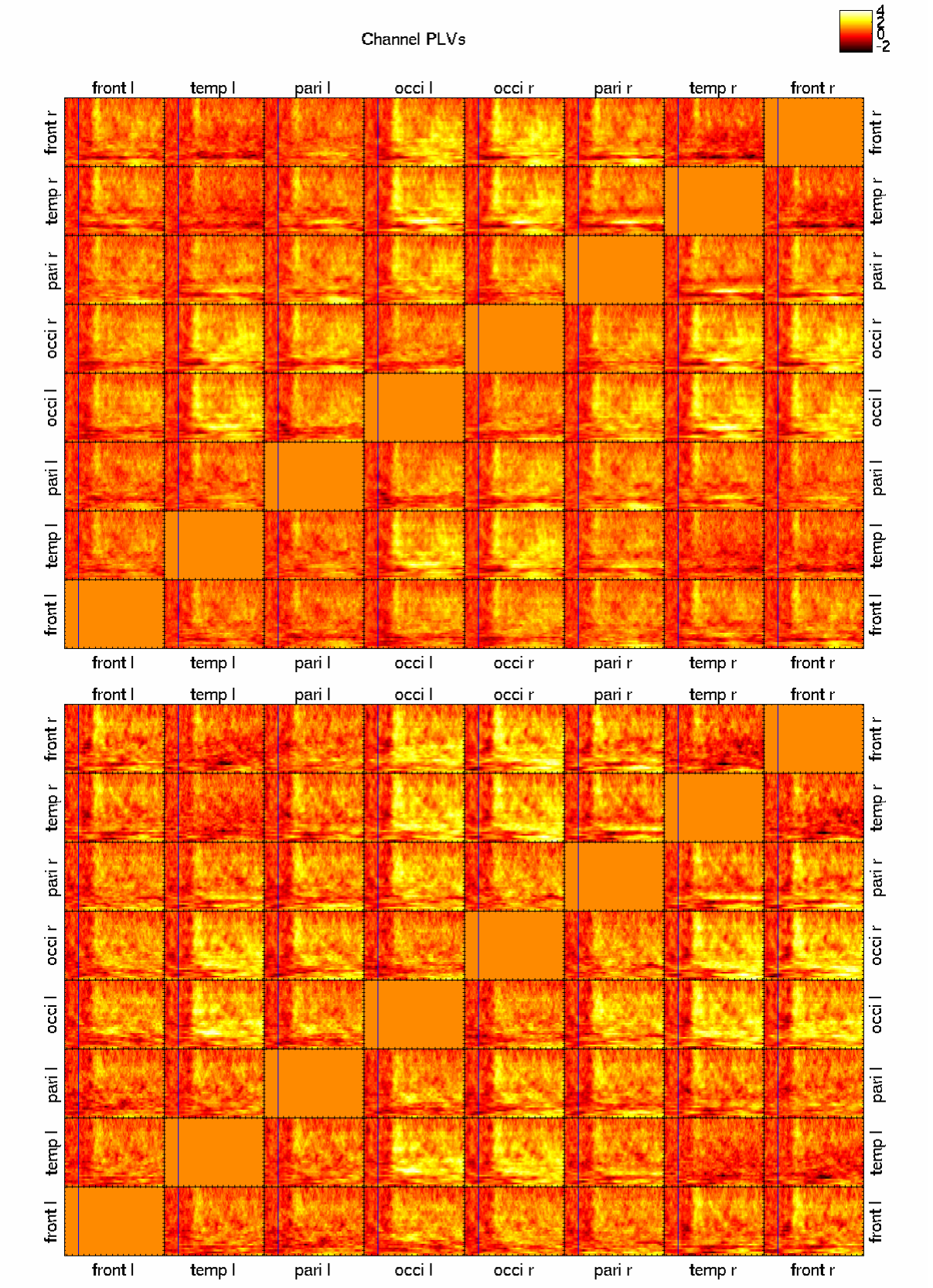


Figure 32 - Synchrony matrix for the time after stimulus onset.
For C11(top) and C12 (bottom), values are in standard deviations from the 500 ms baseline: After onset of the auditive stimulus, differences can be seen especially for lower frequency bands.

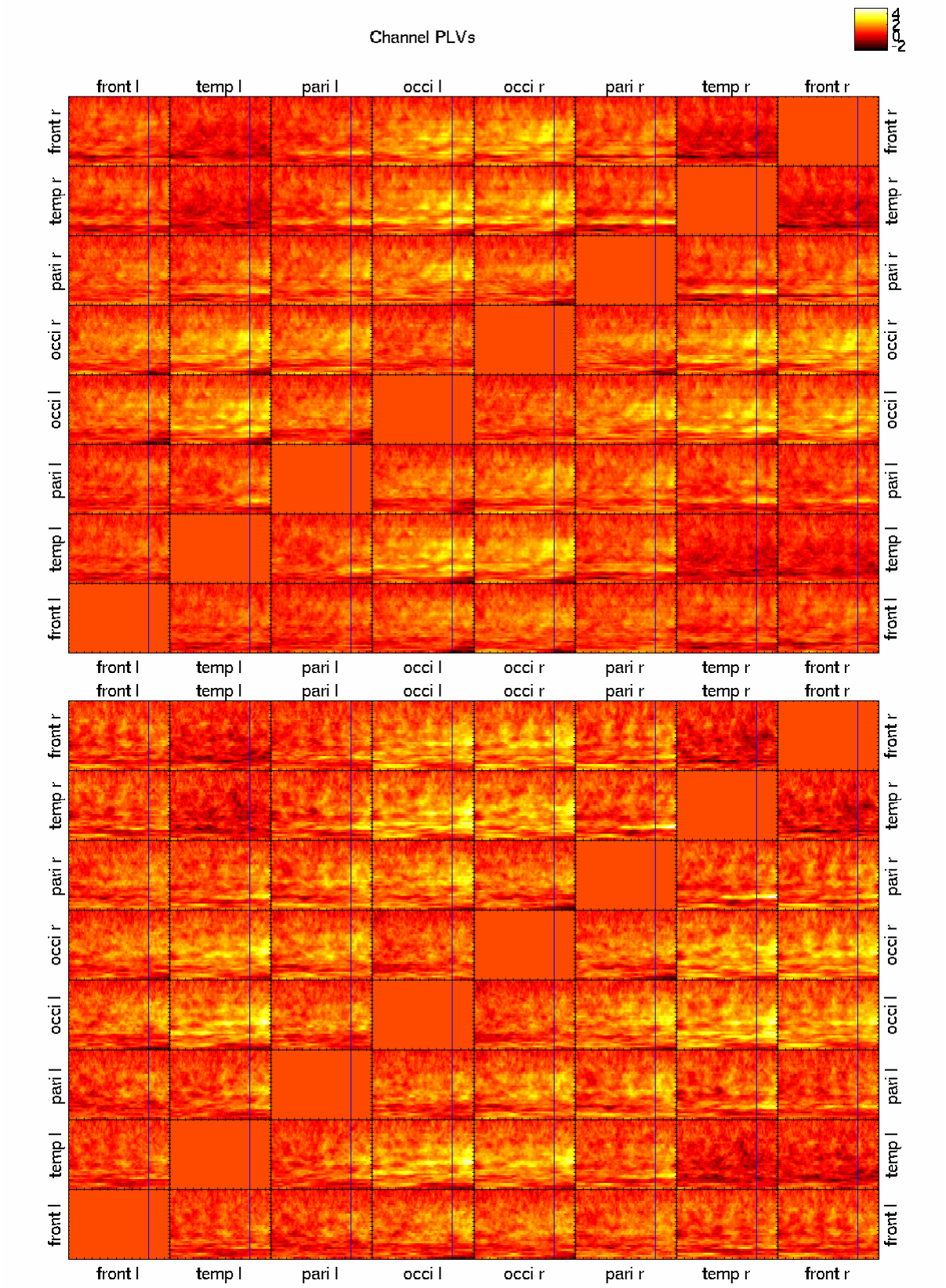


Figure 33 - Synchrony matrix for the time before recognition button press.
For C11 (top) and C12 (bottom), values are in standard deviations from the 500 ms baseline: Higher synchrony can be found for C12 especially for lower frequency bands and for occipital-temporal pairs.

3.2.3 ERPs and Evoked Energies

Looking at ERPs and evoked energies, I was first interested in the time after stimulus onset and occipital electrodes. The ERPs of C1 and C2 showed no differences for the N100 (negative peak at about 100 ms, (compare Herrmann and Knight, 2001)) and P200 (positive peak at about 200 ms) components, but a stronger N200 and P300 for C2 ($p<0.05$). As shown in Figure 34, there is a significant area for the difference of the N100 component that could not be found by looking at overall ERPs, explaining the missing result.

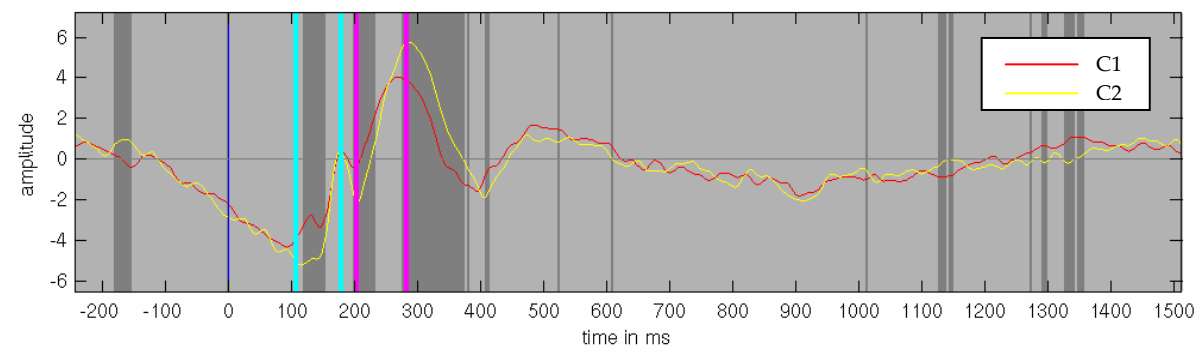


Figure 34 – Occipital ERPs after stimulus onset.
For C1 (red) and C2 (yellow), values in µV: C2 shows a stronger N200 and P300 component.

For the onset of the auditive stimulus, I analyzed the central electrode 1 as the topographies were most prominent there. There was no significant difference in ERP energies between any condition. For ERPs, the N200 is significantly stronger for C2 than for C1 ($p<0.001$) (see Figure 35) and the same was found for C11/C2 ($p<0.01$). The P300 is also stronger for C2 compared to C1 and C11 ($p<0.01$). As can be seen in Figure 35, a possibly significant N100 was again missed as it was not strong enough in the averaged ERP.

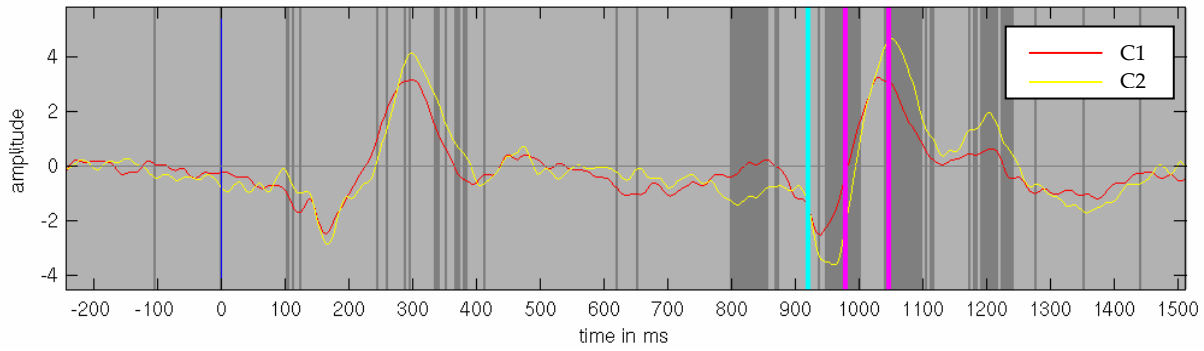


Figure 35 – Central ERPs after auditive stimulus onset.
For C1 (red) and C2 (yellow), values in µV: C2 shows a stronger peak for the N200 component.

A comparison of C11 and C12 shows a significantly stronger P300 for C12 ($p<0.01$) (see Figure 36).

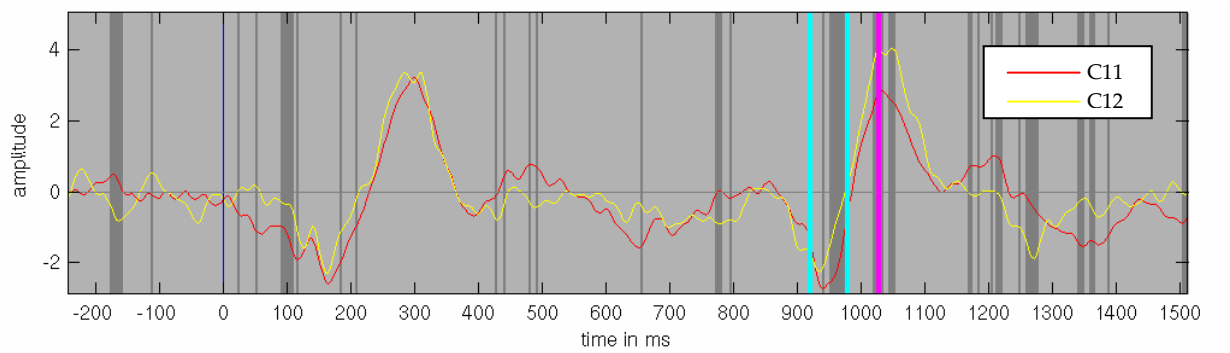


Figure 36 – Central ERPs after auditive stimulus onset.
For C11 (red) and C12 (yellow), values in µV: C12 shows a stronger P300 component.

Looking at the ERPs and evoked energies around the recognition button press, there is no significant difference between the energies of C11 and C12 for any time window. The same holds for the ERP.

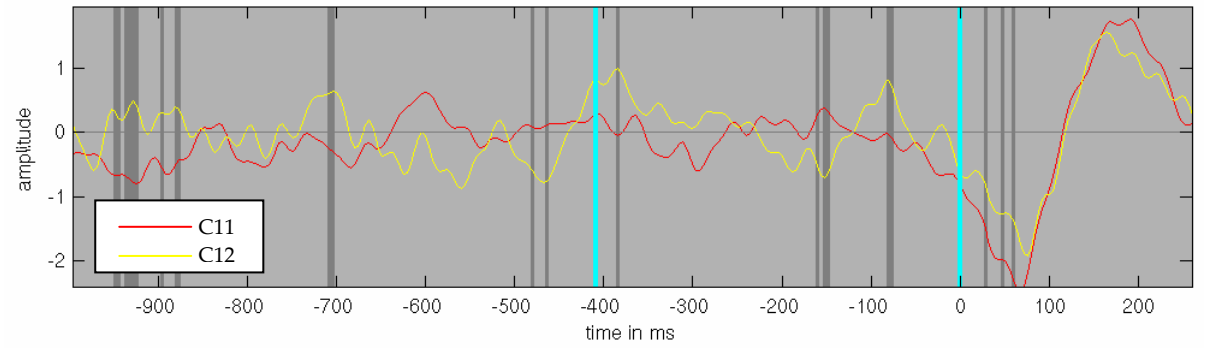


Figure 37 – Central ERPs for the time window before the recognition button press.
For C1 (red) and C2 (yellow), values in µV: There are no specific components or timewondows of significance visible, but at the strongest point of the overall ERP, C1 is stronger than C2.

Looking at the curves (see Figure 37), there seems to be no specific peak and significant differences cancel each other out in the overall average, so the selection of any point for a significance test seems to be arbitrary. Additionally, the scale of the peaks is very low.

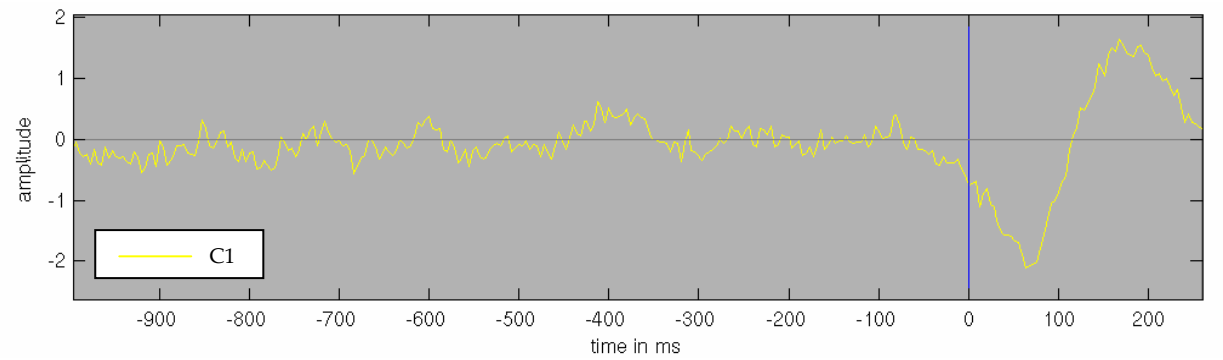


Figure 38 – Medio-frontal ERP of the peaks after the button press.
Data for C1, values in µV: Under frontomedial electrodes, a negative-positive wave occurs directly after the recognition button press.

Finally, a surprising result was a clear ERP response immediately after the button press under medio-frontal electrodes (see Figure 38 for C1).

I will now present the topographies of the found effects.

The ERP topography of the N200 and P300 components in C2 can be seen in Figure 39.

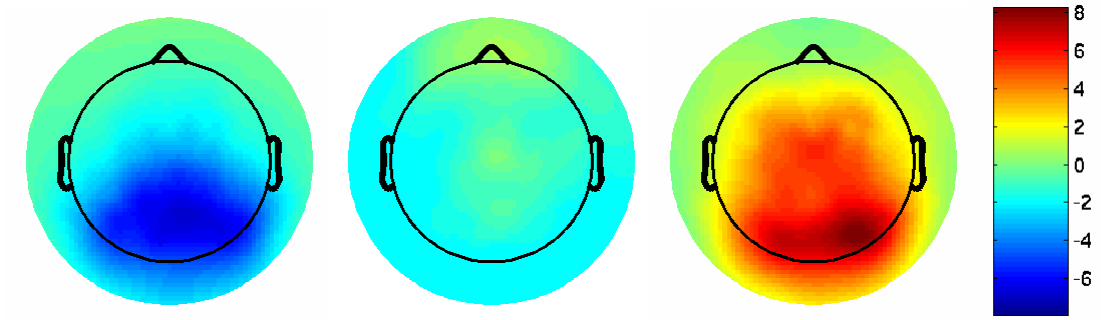


Figure 39 - Topographies of the ERPs after stimulus onset.
For C2, N100, N200, P300, values in µV: While the N100 is focused under occipital-parietal electrodes, the P300 extends to frontal electrodes.

The negative peak is strong under occipital-parietal electrodes, while the positive peak extends from occipital-parietal to more frontal electrodes.

The topography of the auditive N200 of C2 is shown in Figure 40. It is strongest under parietal electrodes.

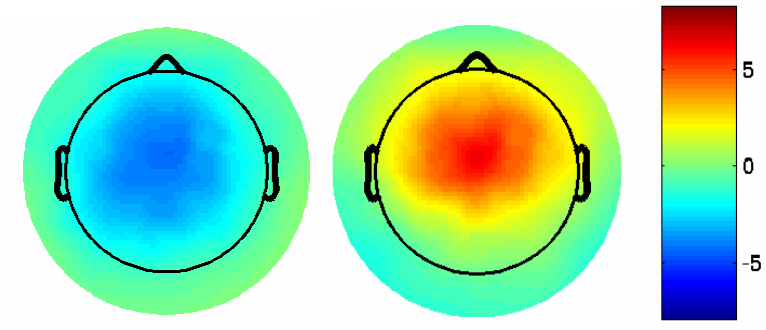


Figure 40 - Topographies of the ERP after auditive stimulus onset.
For C2, N200, P300, values in µV: The responses are almost cemtered under Cz (electrode 1), with the P300 extending to frontal electrodes.

The P300 of C12 can be seen in Figure 41. It is focused under parietal and frontal electrodes.

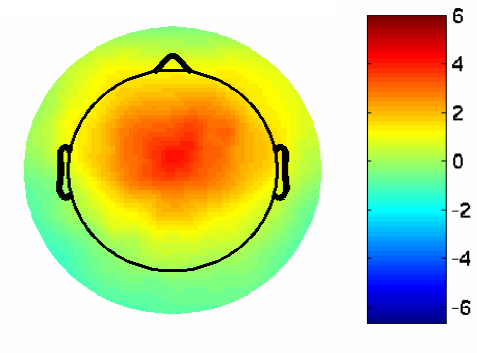


Figure 41 - Topography of the ERP after auditive stimulus onset.
For C12, P300, values in µV: The peak is centered below parieto-frontal electrodes.

The topography of the strong negative and positive ERP response after the button press is shown in Figure 42.

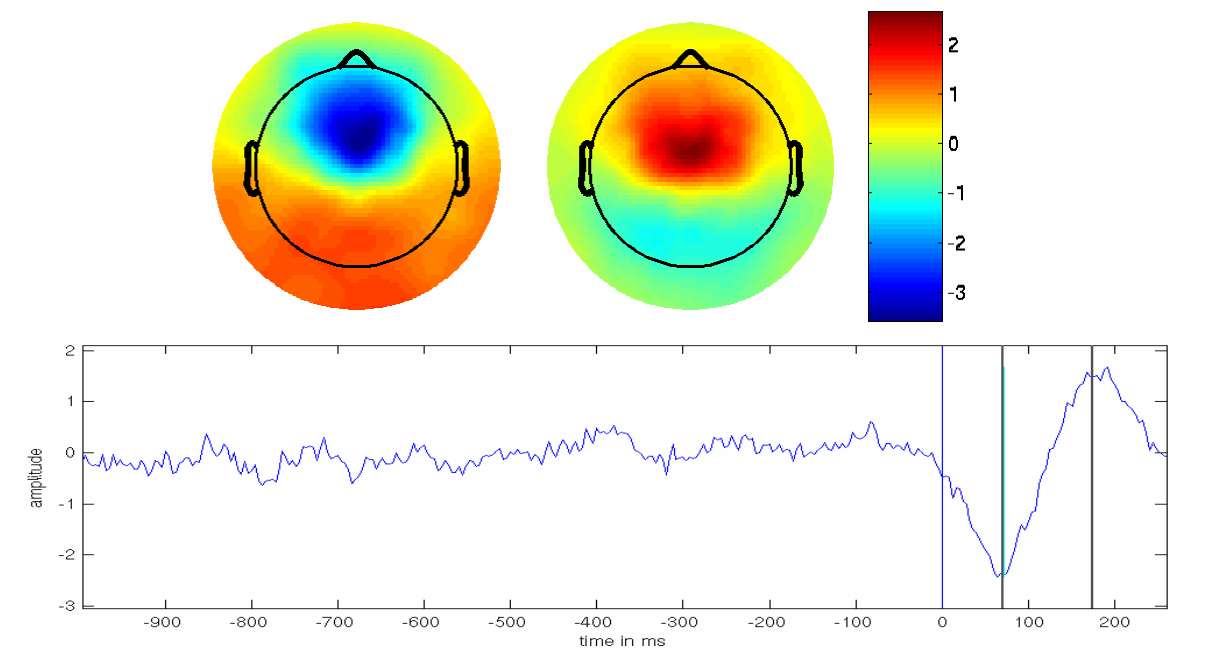


Figure 42 - Topography of the ERP after button press.
For C1, 80 ms and 190 ms (left to right), values in µV: The negative peal focuses below fronto-medial electrodes and combines with a strong positive occipital-parietal-temporal field, while the positive peak stands alone.

While the negative peak under fronto-medial electrodes is accompanied by positive activation under posterior electrodes, the positive peak totally focuses under fronto-medial electrodes.

4 Discussion

Summarizing the hypothesis based analysis, I have found significant differences in induced energy and in synchrony between the animal-detected and animal-not-detected condition. Further, I found differences in crossmodal synchrony for congruent and incongruent sounds. I will discuss these results in detail in the following paragraphs.

4.1 Behavioral Effects

The Behavioral data show a clear advantage for pictures presented with a congruent sound (see Figure 18). This could lead to two different conclusions: The difference could show that subjects used the offered auditory cue to speed up their performance, which did not work out for incongruent cues. Or participants were impeded by the conflict induced by the incongruent sound and could not answer as fast as for the congruent stimulus. Subjective reports after the experiment support the first interpretation that the congruent cue immediately helped to detect

the animal. Nevertheless, finding the right interpretation could only be done by investigating the behavior of subjects compared to a no-cue condition.

As the mean reaction time was long after the onset of the auditory stimulus and as there was a difference between the two conditions, an influence of the sound can not be neglected. Participants had enough time to integrate the additional information in either way.

The fact that only 5% of the no-animal pictures were judged as animal pictures shows that the control condition was reliable. The overall rate of correctly judged stimuli of 78% shows that the experimental detection task was demanding but feasible. The lower percentage of correctly judged congruent and incongruent pictures shows the difficulty of the categorization task. The difference between correct detection and categorization (congruent/ incongruent) might result from subjects, reacting on typical patterns of animals (and therefore detect the animal) without seeing exactly, which animal they perceive (and therefore catagorize badly). Still, the better reaction times for congruent stimuli show that the sound speeded up subjects reactions and that participants used it to categorize the stimuli.

4.2 The Gamma Band Response

4.2.1 The Response after Onset of the Visual Stimulus

The results demonstrate that an induced gamma response can be caused by an existing or missed Gestalt (compare Figure 20). Comparing this result with those found by previous studies (Tallon-Baudry, Bertrand et al., 1997; Goffaux, Mouraux et al., 2004) for similar stimuli, it is quite surprising. In those studies, subjects had learned the stimuli before and the enhanced response was visible for the learned stimuli. Therefore, it had been attributed to a top down binding of the learned representation in a process of experience based perception. The same has been found by a further study (Gruber, Müller and Keil, 2002) for incomplete figures. The data of my study suggest that previous learning is not necessary to elicit the enhanced response. In line with the results of Rodriguez and colleagues (Rodriguez, George et al., 1999), the detection of the Gestalt was sufficient. This might indicate a dissociation between detection (integrating the features to a Gestalt) and recognition (memory match). As an objection, one could assume that the detection of the animal already represents some kind of recognition, as participants had been instructed to detect “an animal”. As the range of animals was quite large, ranging from mammals to birds, this objection would assume very abstract representations for expectation.

Still, there remains the question what the induced gamma response means? Peaks in induced gamma energy have been found in many previous studies (Joliot, Ribary et al., 1994; Tallon-Baudry, Bertrand et al., 1996; Tallon-Baudry, Bertrand et al., 1997; Keil, Müller et al., 1999; Herrmann and Mecklinger, 2000; Pulvermüller, 2001; Gruber, Müller et al., 2002; Kaiser, Hertrich et al., 2005; Lachaux, George et al., 2005; Rose and Buchel, 2005) for as different tasks as figure perception, perception of apparent motion, auditive deviance detection, and word recognition. An interpretation trying to integrate many of these findings is offered by Herrmann and colleagues (Herrmann, Munk et al., 2004). They propose the late induced gamma response to be a correlation of a utilization mechanism. This allows the redirection of attention or the read out of found information by response processes. Therefore, they predict also increased synchrony for the time of the gamma peak. There are several reasons, why this interpretation might not hold for the data presented here: First, Herrmann and colleagues assume a memory match or a classification according to pre-activated patterns to precede the utilization response. This memory match should be reflected in a peak in evoked gamma energy. As stated above, a memory match is very improbable for the stimuli used in this study. Accordingly, I have not found enhanced evoked energy in the data. Second, the mean reaction time was much larger compared to the time scale of the induced gamma peak. Third, the onset of the informative auditory cue was much later than the induced response. This excludes the read out by a direct response process. Still, a redirection of attention might happen after the first impression of the picture. Subjects stated that they used this cue to recognize the animal, so that a response process should not have started before this cue, but the direction of attention to the auditory cue combined with some kind of expectation might have happened. The strong increase in synchronization indicates that neural assemblies indeed exchanged information, making a process similar to the one described by Herrmann et al. (Herrmann, Munk et al., 2004) appearing likely.

I have excluded several methodological explanations, concerning the normalization of data, the exclusion of trials with abnormal baselines, and the balancing of the number of trials in each condition. All these variations have produced the similar results. The practically not significant difference between the congruent/incongruent condition at that time is also evidence for the validity of the results. I can also exclude a biasing influence of properties of the stimulus material, e.g. animal pictures being brighter or more complex. Busch and colleagues (Busch, Debener,

Kranczioch, Engel and Herrmann, 2004) have shown that the induced gamma is not influenced by these differences and seems to reflect more cognitive properties of the material.

There is a final objection raising the question, whether the increased gamma energy just represents a state, in which subjects detect figures more easily. As I used the subjective categorization for the definition of conditions, I cannot fully exclude this possibility. The majority of stimulus categorizations was correct, supporting the conclusion that the found gamma response is caused by differences in the stimuli and not by internal states of the participiants. A further analysis of the data with objective conditions or the analysis of animal/detected vs. no animal/not detected conditions could further support my interpretation, given the difference between responses remains.

4.2.2 The Response after Onset of the Auditive Stimulus and before the Button Press

The analysis has shown that induced energy for the no-animal condition is significantly stronger after the onset of the auditory stimulus (compare Figure 21). This could be explained by increased attention, as subjects might try to get a cue which animal they have to find on the picture. As several studies (Tiitinen, Sinkkonen, Reinikainen, Alho, Lavikainen and Naatanen, 1993; Tallon-Baudry and Bertrand, 1999) have shown, the gamma band activity can be enhanced by attention to the specific stimulus. There are several objections against this interpretation. First, it is not possible to find a distinct peak as for the visual response to the stimulus. The most significant time windows start long after the stimulus onset and should therefore not reflect early stimulus related processes modulated by attention. Second, the first significant time window directly after the onset of the sound does not show the strong negative peak as found by another study (Tallon-Baudry and Bertrand, 1999) for the first window before the strong positive peak. Third, looking at temporal and at occipital electrodes, there is generally stronger tonic gamma energy for the no-animal condition. This might indicate that adding the normal fluctuations in energy to this general increase leads to significant differences. Following this argument, there is a new question: what does this tonic increase in energy mean? Does it reflect the ongoing search for the (hypothetic) animal? As the mean reactions time is about two seconds, the search should have lasted longer than the first few milliseconds for both conditions, ruling out this explanation. Possible explanations for the found effect therefore remain hypothetical and speculative.

I could not find any convincing differences between the congruent and the incongruent condition after the auditory stimulus onset. The significant differences hardly reached the level of

significance and as there was no distinct peak within these windows, one must be careful to accept any differences there. This missing result has a positive side effect: According to the match and utilization model (Herrmann, Munk et al., 2004), one could have expected a gamma peak for the match of picture (and a possibly produced expectation for the sound) and the played sound. As there was no specific peak, neither in evoked nor in induced energy, this indicates that participants had not fully detected the animal at that time but needed the additional information of the cue to find the Gestalt.

The fact that there was also no difference in the interval before the recognition button press (see Figure 22) could have three reasons: First, the difference might have been too subtle to find it with the number of trials, subjects and perhaps also with this setup. The data for this time does not provide any recognizable structure that just nearly missed significance. Hence, this explanation does not seem very probable. Second, the alignment on the button press might not match the different individual reaction times and cognitive processes in that interval. This might be indicated by the missing pattern in the data, where peaks seem to be totally unrelated to any process like the initiation of the button press. Third, there might simply be no difference, as the audio-visual integration mainly happens during a short interval after the auditory stimulus onset. As there was also no difference within that time interval, this explanation does not seem satisfactory, either.

As I have no data for the no-animal condition for that interval, due to the missing button press, I could not look at this data for further comparisons. Therefore, further explanations stay speculative.

4.3 Synchronization Effects

4.3.1 Synchronization after Onset of the Visual Stimulus

The animal-detected condition shows significantly stronger global synchronization after the onset of the visual stimulus (compare Figure 27). The time scale of this peak could be compared to the effects found by Rodriguez and colleagues (Rodriguez, George et al., 1999) for perceived mooney faces. It seems that increased gamma synchrony plays an important role for the perceptual process. Several authors suggest it as a mechanism for the binding of elementary visual features into a coherent percept (Bressler and Kelso, 2001; Varela, Lachaux et al., 2001). Findings indicating this role in bottom-up processing have been presented by König and colleagues (König,

Engel and Singer, 1995) for cats, and von Stein and colleagues (von Stein, Chiang and König, 2000) observed increased gamma synchrony also locally and especially for novel stimuli.

Looking at the topography of the found synchrony, the first peak mostly shows coherence between occipital, temporal and frontal electrodes. A similar pattern of long-range synchrony has also been reported by Rodriguez et al. (Rodriguez, George et al., 1999). This seems to contradict the local binding role of gamma synchronization and stands in line with other human studies that have found long range synchronization in the gamma band (Knyazeva, Kiper et al., 1999; Lutz, Lachaux et al., 1999; Haig, Gordon, Wright, Meares and Bahramali, 2000; Bhattacharya, Petsche and Pereda, 2001). These studies provide first evidence that differences in gamma synchrony could be caused by the task relevance of the stimulus and by the experience, subjects had with the stimulus material.

As the animal pictures had more task relevance than the pictures of pure scenery, this might be a reason for the found difference. Still, the functional role of gamma synchrony needs further investigation and explanations for the found effect have to stay speculative.

4.3.2 Synchronization after Onset of the Auditory Stimulus and before the Button Press

I found two significant time windows for the gamma band and for the beta band between the animal-detected and the no-animal condition after the onset of the auditory cue. The difference for the beta band was even stronger for crossmodal electrode pairs while the gamma band difference was much weaker for these pairs.

Bhattacharya and colleagues (Bhattacharya, Petsche et al., 2001) also found high gamma synchrony for the perception of music. Staying in line with other authors (see above), they attribute it to the binding of musical features. Synchronization in the beta band has been proposed as a candidate for supra-modal integration (von Stein and Sarnthein, 2000; Varela, Lachaux et al., 2001). Accordingly, von Stein and colleagues (von Stein, Rappelsberger et al., 1999) have found increased beta band synchrony for the multimodal processing of objects. Surprisingly, the difference found in my data is stronger for the no-animal condition. This excludes an interpretation as low level crossmodal binding as this should have been stronger for the congruent condition. On the opposite, it could reflect the crossmodal exchange of information between the two sensory areas, trying to integrate the information of one modality to support the perceptual process of the other one. Again, this interpretation is speculative and further evidence will be necessary to draw stronger conclusions on this question.

4.3.3 Synchronization before the Recognition Button Press

Looking at the interval before the button press, there is a difference within the beta band accompanied by a significant difference in the theta band, with the incongruent condition being stronger than the congruent condition. Both bands have been associated with medial and long range integration processes (von Stein and Sarnthein, 2000; Bressler and Kelso, 2001; Varela, Lachaux et al., 2001). Theta band synchrony has been shown to be involved especially in working memory processes (Sarnthein, Petsche, Rappelsberger, Shaw and von Stein, 1998; von Stein, Rappelsberger et al., 1999; Schack, Vath, Petsche, Geissler and Moller, 2002). Again, the beta band synchrony can not simply be explained by low level crossmodal intergration as this should have been stronger for the congruent condition. In combination with the theta band synchrony, the two increases for the incongruent condition might reflect an interplay of crossmodal attempts to integrate conflicting information (beta band) and an involvement of other areas within this task (theta band). This effort could be mediated by cross frequency relationships as proposed and found by (Varela, Lachaux et al., 2001; Palva, Palva et al., 2005; Schack, Klimesch and Sauseng, 2005).

4.3.4 Long or Short Wavelets? Answers to an Important Question

The comparison of the results produced by the Hilbert transform and the results produced by the long wavelets (Lachaux, Rodriguez et al., 1999), revealed striking differences. The long wavelets produced larger effects that were much more significant for the time after the onset of the stimuli. This matches the explanation made by Trujillo and colleagues (Trujillo, Peterson et al., 2005), showing the tradeoff between a gain in significance and the related loss of level of detail. Looking at the details revealed by the Hilbert transformed results, it remains questionable if the long wavelets integrate over too much time and "produce" better results. As Varela et al. (Varela, Lachaux et al., 2001) expected episodes of synchronization to happen on a time scale of about 300 ms, the approach might be justifiable. The drawback is to loose lower frequencies that must be analyzed with different methods (Palva, Palva et al., 2005). Summarizing, the found difference in significance shows that a comparison of studies using these different methods must be performed with great care and awareness of these differences.

4.4 ERP effects

As ERPs are not the focus of this work, I will only discuss the results shortly.

There are two main results:

First, for all comparisons, there was a stronger P300 component for the no-animal condition compared with the animal-detected condition and for the congruent/ incongruent comparison. Several studies have shown that this component is associated with cognitive processes (like novelty registration or stimulus discrimination) and is not influenced by external factors as stimulus luminance (Herrmann and Knight, 2001). Therefore, these differences might reflect different cognitive processes as a consequence of the different qualities (e.g. congruent/ incongruent) of the stimuli.

Second, there is a strong negative-positive waveform directly after the time of the recognition button press. This wave could be interpreted as an error related negativity (ERN) wave, reflecting the difficulty of the task and the uncertainty for the subjects, as they did not receive any feedback (Botvinick, Cohen and Carter, 2004; Ridderinkhof, Ullsperger, Crone and Nieuwenhuis, 2004). This interpretation fits quite well with the fact that, when asked to report freely after the experiment, subjects tended to judge their own performance much lower than it really was. It also matches the real difficulty of the task, reflected in the behavioral recognition data.

4.5 A Merged Picture of Gamma and Sync

A striking finding is the co-occurrence of increased gamma energy and increased gamma synchronization for the animal-detected condition compared to the no-animal condition. This has also been found by Rodriguez and colleagues (Rodriguez, George et al., 1999) in a similar setup. As they state, this indicates that the peak in synchrony is functionally involved in the perception process. Another study (Lutz, Lachaux et al., 1999) also showed synchrony in specific areas occurring together with increased gamma energy. As König and colleagues (König, Engel et al., 1995) showed for cats, gamma oscillations accompany long range synchronization while they do not have to occur for short range synchronization processes. This poses the question, if the found gamma synchrony reflects only short range synchrony or the synchronization in a more distributed neural network accomplishing higher cognitive processes as proposed by Varela and colleagues (Varela, Lachaux et al., 2001).

A model, predicting this coincidence, is propsed by Herrmann and colleagues (Herrmann, Munk et al., 2004). They judge the induced gamma response as a utilization response (compare section 4.2.1), accompanied by strong synchronization when processes exchange information to prepare an adequate response or to redirect attention. As explained above (see section 4.2.1), the present

data does not totally fit these predictions. I have not found the predicted evoked energy peak and reaction times were much higher than the time scale of the gamma response. Still, regarding the peak in gamma synchrony and the assumed role of synchrony in general, the predicted exchange of information for further processing seems a viable explanation for my findings. Further investigations in this field will be necessary to clarify this issue (see also section 4.7.2.2 for a proposal).

4.6 Critical Points of This Study

4.6.1 Methodological Problems

As for the previous study (see section 1.3), there are critical points of the current study that could be improved in further studies.

First, the choice of the time window was a tradeoff between separating the post stimulus from the pre button press interval and looking at as much time as possible. First, I was looking at a time window of 1 second after stimulus onset. Later I realized that extending this range to 1.5 seconds captures the dynamics caused by the onset of the auditive stimulus much better. Still, as no one knows exactly when synchronous events should occur, the choice of the time window stays difficult.

Though a lot of time had been invested in the construction of the stimuli, the overall number of trials was still very low. This had two consequences. First, I could not find significant effects within each subject. Therefore, I concentrated on differences over subjects. Second, it was practically impossible to balance the conditions, as I needed to have a majority of congruent trials (to make the sound a valid cue) and could not include trials to be thrown away later. However, I investigated the effects of unbalanced trials and this revealed no severe differences for this study.

Finally, I lacked data for the pre button press interval for the condition without animal. Using surrogate data for statistical analysis only offers a weak alternative. On the other hand, as subjects could recognize the animal until the last second, I could not instruct them to press a button when they where sure that there is no animal. Still, the comparison between the congruent and the incongruent condition revealed interesting results.

4.6.2 The Problem of Multiple Approaches

As could be seen in Table 1, there are many different ways to analyze data for neural synchrony. Some of these different approaches do not produce the same results (compare section 2.4.2.5.2) or differ in significance values they produce (compare 4.3.4). As there is no standardized approach for this field of research, yet, the results of different studies are hardly comparable. On the other hand, the combinatorial explosion of applicable methods makes it impossible to exploit several approaches for one study just to produce comparable results.

I tried to stay in line with relevant previous studies but I also had to change several steps on my way of analysis due to a different setup and different questions.

Efforts are needed to further investigate the differences of methods, to see if some of them can be exchanged easily (see e.g. Bruns, 2004) and others are too different and produce different results. On the long run, a unification of approaches seems necessary to a specific degree to ensure continuity and coherence of studies and results.

4.6.3 The Problem of Complexity

This work included two kinds of analysis, hypothesis based and explorative analysis.

Especially the latter one is time consuming and as the data gets more and more complex, the process gets less and less manageable. There is a high risk to overlook possibly interesting effects or to invest too much time in a dead end. In fact, it was not possbile to look at all thinkable combinations of data.

The problem of complexity reaches a high level for the research on synchronization. In principle, one has to look at each pair of electrodes implicating a quadratic factor for the number of comparisons. Adding the upcoming research on cross channel cross frequency synchronization (e.g. Palva, Palva et al., 2005), this combinatorial explosion makes explorative research even more a blind flight.

A solution for this problem might be to use statistical methods like clustering of data by statistical measurements. Specific clusters could give hints, which combination of channels and frequencies are worth further investigation for effects.

The amount of work to get a possibly usable tool certainly adds up to another possible thesis. Still, it could be a way to handle the complexity of the data in cases where there are no strong hypotheses and standardized methods.

4.7 Proposals for Further Analysis and Studies

As the timeframe of this work was rather limited, I had to focus on a set of questions and methods for analysis of the data. In the following, I will describe further questions and methods to analyze the data. I will also propose further possible experiments to go on with the search for neural synchronization.

4.7.1 Further Analysis

Even after applying several methods of analysis, there are many others methods left that could be applied on the data without producing redundant results.

4.7.1.1 Selection of Conditions

To split the data into conditions, I used the subjective categorization of each subject. There are several other possibilities.

- Select by objective categorization
- Select the intersection of both sets: only right positives as Condition 1 and right negatives as Condition 2. This might even increase the difference in the gamma response (compare the selection of Rodriguerz et al. (Rodriguez, George et al., 1999)). Additionally, it could reject the objection that the detection of animals was dependent of a subjective state reflected by different levels in energy (compare section 4.2.1).

4.7.1.2 Further Analysis of Synchronization Data

There are several ways to calculate averaged data of synchrony: One can select specific frequencies for each subject by the strength of the induced gamma response (compare Rodriguez, George et al., 1999) or by the temporal patterns or strength of synchrony (compare Trujillo, Peterson et al., 2005) and mean only these frequencies over subjects. As there is no standardized way to determine individual frequencies, there are several reasonable methods available.

4.7.1.3 Cross Channel Cross Frequency Analysis

As proposed by Plava and colleagues (Palva, Palva et al., 2005), there might be interesting effects of synchronization between signals of different frequencies. Different frequencies are hypothesized to represent different scales of cortical integration (von Stein and Sarnthein, 2000). Hence, it is possible that cross frequency synchrony occurs within single electrodes or between pairs of electrodes. In the first case, one should check if possible synchronization might be caused

by harmonics of the lower frequency (compare Palva, Palva et al., 2005). Looking for synchrony between frequencies at one electrode and between two electrodes at one of the synchronous frequencies might reveal interesting process of local and global information integration.

4.7.1.4 ICA Components and SCD

The signals recorded by the electrodes are very blurry due to volume conduction of the underlying tissue. Therefore, it would be interesting to analyze two things:

- Instead of analyzing the time frequency transform of channels, one could analyze the components derived from the ICA. These components represent distinct neural sources and synchronization effects between these independent components (independent on the long run, so that synchrony could still occur for short periods (Meinecke, Ziehe, Kurths and Müller, 2005)) could be more informative than between channels that are innately synchronized due to volume conduction. Additionally, source localization for interesting components could reveal where synchronized neural sources are in the brain.
- Another method to reduce the blurring of the recorded data is the source current density transformation (Nunez, 1989). I tried to remove synchronization caused by the blurring of the recorded data by substracting the PLV baseline of each channel combination. It could be interesting to compare this to results of other methods like SCD, performing the deblurring before the time frequency transformation.

4.7.1.5 Memory effects

After the first 3 blocks of trials in the experiment, I had appended 3 blocks that were shuffled repetitions of the first ones. As subjects indicated recognition of most of the pictures, I could not use the data of these blocks for my analysis. Still, one could analyze this data for memory effects. These could be found in a change in energy or as different patterns of synchronization (for memory effects and synchronization see e.g. (Sarnthein, Petsche et al., 1998; von Stein and Sarnthein, 2000; Schack, Vath et al., 2002; Arbib, 2003), for memory and energy see e.g. (Tallon-Baudry, Bertrand, Peronnet and Pernier, 1998; Herrmann, Munk et al., 2004)).

4.7.1.6 Rereference Data

An issue brought up by Trujillo et al. (Trujillo, Peterson et al., 2005) is the reference of the EEG data. Following their argumentation, synchrony could occur as an artifact caused by a change at

the reference electrode. Studying the effects of different types of reference transformation on the results of the analysis could be interesting in this context.

4.7.2 Further Experiments

In the following, I will propose two experiments, one continuing the line of research that has been started with the study by Saskia Nagel and the study presented in this thesis, and the other one based on a theory about the relationship of synchrony and induced gamma energy.

4.7.2.1 Binding in animated movies

Looking at the different levels of binding (perceptual and cognitive), as described in section 1.1.1, a rising question could be, which kind of (crossmodal) binding had been investigated in this study. If crossmodal binding is defined as the creation of a unified perception of an object, there could be doubt that this could have been found in this study. An example that came to my mind is that of a television reporter: If TV stations get no video link to there reporters, they only show a picture of the person and play her telephone report. In these situations, I always realize that I cannot link the voice and the picture together. Instead, if there is a normal reporter shown, speaking into her microphone, I never doubt that this is the person speaking.

Normally, people do not to bind still pictures with sounds, as this is not the contingency learned throughout their life. Therefore, to study crossmodal binding in a strict (perceptual) sense, one should study moving objects that make noises. Movement (e.g. of natural animals), of course, adds a lot of uncertainty to an experimental setup. A possible solution for complete control over the setup might be rendered videos. To stay in line with the experiment presented here, one could produce animated animals ('moving their mouth') together with matching or not matching sounds. As the videos are animated, one could totally control the time locking events to study (e.g. the onset of the sound). This kind of setup would also mean to stay with complex and somehow natural stimuli, compared to moving dots and bars.

4.7.2.2 Induced gamma energy and synchronization

An interesting effect I found in this study was the relationship between gamma synchrony and induced gamma energy. In their model about gamma energy, Herrmann and colleagues (Herrmann, Munk et al., 2004) predict a connection between these two phenomena. For them, early gamma energy reflects a match between incoming stimuli and preactivated templates: a

memory match. The late gamma response reflects utilization behavior for the response the subject has to execute.

Combining the setup of the experiment presented in this thesis with the predictions of this Memory and Utilization Model (MUM) results in the following setup: During a grey screen, the sound of an animal is played to activate the template for this animal. After a short interval, the camouflage picture of the animal is shown. Again, subjects have to press a button, if they recognize the animal.

The MUM would predict early gamma oscillations (the template match) for the visual stimulus and an increase of synchronization in combination with the second gamma response, reflecting the broadcasting of the response information. Therefore, the theory provides clear predictions for the proposed setup that could be tested.

Appendix A Construction of Camouflage Pictures and Sound

All stimuli for this experiment have been newly created for this study. I will describe the non standardized procedure in the following paragraphs.

A.1 Video

Using Google Picture Search (http://www.google.de/imghp), I looked for pictures of animals and equivalent pictures of scenery without animals.

I soon had to realize that there is no 100% standardized routine to build the camouflage pictures. Therefore it was not possible to build a script that creates the stimuli automatically. Using Adobe Photoshop 6.0, I roughly applied the following procedure to build a camouflage picture:

- Find a picture with Google matching these criteria:
    - The animal is not totally in the foreground and, preferably, it has a strong camouflage pattern on the coat. For example, it could stand in front of a structure that somehow resembles the pattern on the coat of the animal (see Figure 43).
    - The background should not be flat but have a strong pattern like bushes or trees. Additionally, strong contrast (e.g. by sunlight and shadows) helps hiding the animal.

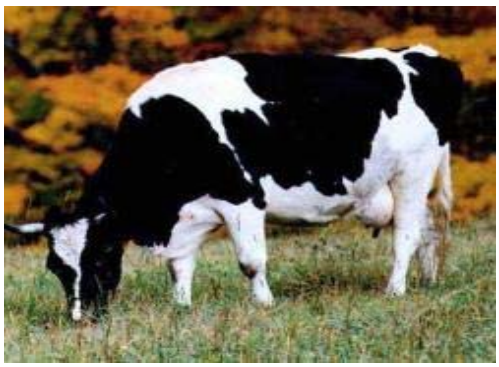


Figure 43 - A typical image to become a camouflage picture.
The animal is in front of a patchy background and the patters of coat and background are similar

- Convert the image to grey scale.
- Optional, but often recommended: Blur the image with a Gaussian filter so that the contours of the animal vanish lightly into the background.

- Increase contrast until background and animal merge together.
- In the end, convert the picture to binary scale to save memory and enable the experimental software to preload all images. The result is a black and white camouflage picture (see Figure 44)

Figure 44 - The completely degraded picture.
For naive subjects, it is hard to see the animal without any further cues to the cow

For several reasons, it is possible that this process does not work in a single iteration. For example, many pictures have a sharpness gradient, so that the background far away loses sharpness. Other pictures have a strong background/ foreground contrast. Therefore, I recommend executing the described steps in several iterations (see Figure 45).

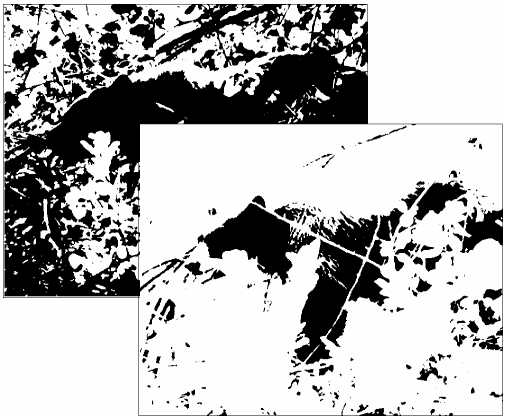


Figure 45 – Several iterations to produce a camouflage picture.
Using several layers with different contrast settings helps merging the animal into complex backgrounds. Still, it preserves the basic properties of the image.

This could be done by using several layers, one for each iteration for different parts of the image with different blurring and contrast settings. The resulting layers can be merged by erasing different parts of the layers. The result is a semitransparent stack of image parts that becomes one camouflage image.

With this method I built 120 camouflage pictures of animals and 30 pictures of scenery, matching the background of the animal pictures (trees, sea, bushed, woodland). I used images of 26 animals: bee, wasp, bear, cat, chicken, cow, dog, donkey, duck, eagle, elk frog, giraffe, horse, monkey, owl, parrot, sheep, goat, snake, tiger, leopard, lynx, dolphin, whale, and wolf.

A.2 Audio

The sounds of the animals needed to be specific and easily to recognize to provide a usable cue. After the pretest (see Appendix B), I decided to put similar animals into categories and to use one sound for each category (instead of each type of animal). This resulted in 21 sounds for these categories: insect, bear, cat, chicken, cow, dog, donkey, duck, eagle, elk, frog, giraffe, horse, monkey, owl, parrot, sheep/ goat, snake, cat of prey, whale, and wolf.

I found matching audio files by using the free preview files of the sound portal AudioSparx (www.audiosparx.com). For some animals (e.g. elks or whales), distinct sounds were hard to find. Especially the giraffe turned out to be an impossible case as it only generates sound in the infrasonic range. As the giraffes provided perfect camouflage pictures, I decided to use them together with the unusual sounds of zebras instead, assuming that hardly any german subject knows these sounds (which proved true during both, the pretest and the experiment).

I used Cool Edit Pro 1.0 (Syntrillium Software, now Adobe Audition 1.0) to edit the sounds. They were upsampled to 44100 kHz, 16 Bit, mono. Afterwards I edited them, so that there were as few brakes in between the single roars, screeches, or clucks as possible. I wanted to build an auditory stimulus as constant over time as possible. The clips were cut so that they had an immediate and significant onset and lasted for exactly 5.25 seconds.

As different animal sounds have different spectral distributions and therefore differ in subjective loudness, I applied a B-Weighting of loudness to normalize the volume of the clips according to the subjective experience of the listener (IEC 61672, www.iec.org).

Appendix B Pretest of Stimuli

To examine the effects of the visual stimuli for the EEG study, I conducted a behavioral experiment. In particular, I was interested to get information about the difficulty of the task, the probabilities of false positives and false negatives, and the crossmodal effects on reaction times.

B.1 Methods

B.1.1 Stimuli

The visual stimuli were the ones constructed for the EEG study as described in Appendix A. I used a set of 120 animal pictures and 15 landscape pictures without animals. The sounds were similar to those used in the EEG study, but at the time of the pretest, there was a different sound for each type of animal while later in the EEG study, I put animals into categories and had different sounds for these categories of animals. In addition to the animal sounds, I had a brown noise sound to test for the effects of a "no sound" condition. Opposite to the sounds in the EEG study, the sounds in the pretest had a duration of only 2 seconds, with small deviations for the animals sounds due to variations in the length of the sounds of different animals.

B.1.2 Subjects

Subjects were 12 students of the Neurobiopsychology Lab (3 female) with a mean age of 23.3 years. They had never seen or heard the stimuli before the experiment. The participation was voluntarily and unpaid.

B.1.3 Setup

The setup of the study was analogue to the planed EEG experiment with slight modifications to provide us with deciding data for the EEG experiment.

As presentation software I used the psychophysics toolbox (Brainard, 1997; Pelli, 1997) and MATLAB 6.5 (The Mathworks Inc.) running on a PC (Intel Celeron, 512MB RAM).

Subjects were seated in front of a CRT monitor, stereo speakers and a keyboard. The instruction was given to them by a tutorial that showed them all situations that could occur in the course of the experiment. Subjects were instructed to press a button, as soon as they perceived and were

able to identify the animal in the picture. They were informed that some pictures might not show an animal and that some of the sounds were not matching the animal on the picture. They were told to look on the fixation cross all the time while watching the stimuli images.

After the tutorial, subjects were presented all the sounds of the animals that could occur during the experiment. Subjects had the opportunity to listen to the sounds several times until they felt confident being able to recognize them.

The experimental trials had the structure as illustrated in Figure 46.

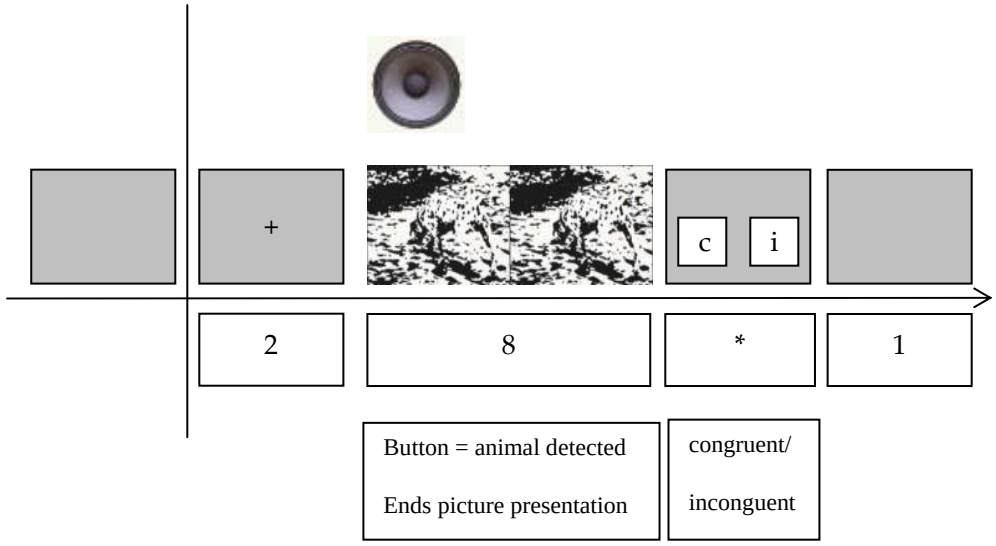


Figure 46 - Setup of the Pretest experiment
After a constant pre stimulus interval, the pictures were shown for 8 seconds. If subjects pressed a button within this interval, the trial ended immediately. After recognizing an animal, subjects had to answer a congruency question.

After a short grey screen, a fixation cross appeared, then the stimulus picture was presented for 8 seconds simultaneously with the sound (duration about 2 seconds). A button press, indicating that the subject identified an animal, finished the presentation immediately and the question about the matching of the animal picture and the sound was shown. Otherwise the trial lasted the full duration of 8 seconds without asking for congruency afterwards.

There were four possible combinations of stimuli that could occur in each trial:

- animal picture with matching animal sound (40 trials)
- animal picture with non matching random animal sound (40 trials)
- animal picture with noise (40 trials)

- landscape picture with random animal sound (15 trials).

Pictures were randomly distributed into these conditions with each picture being presented one time per subject. Trials of each condition were shown alternately in random order.

B.1.4 Analysis of Data

I analyzed the data with Matlab 6.5 (The Mathworks Inc.) running on a PC (Intel Celeron, 512 MB RAM) with Windows XP. I calculated statistics for each subject, for all subjects together, and for each picture.

Additionally to the objective measurements, I also collected and analyzed the reports of subjects about their recognition performance and possible reasons for difficulties.

B.2 Results and Discussion

B.2.1 Recognition Performance

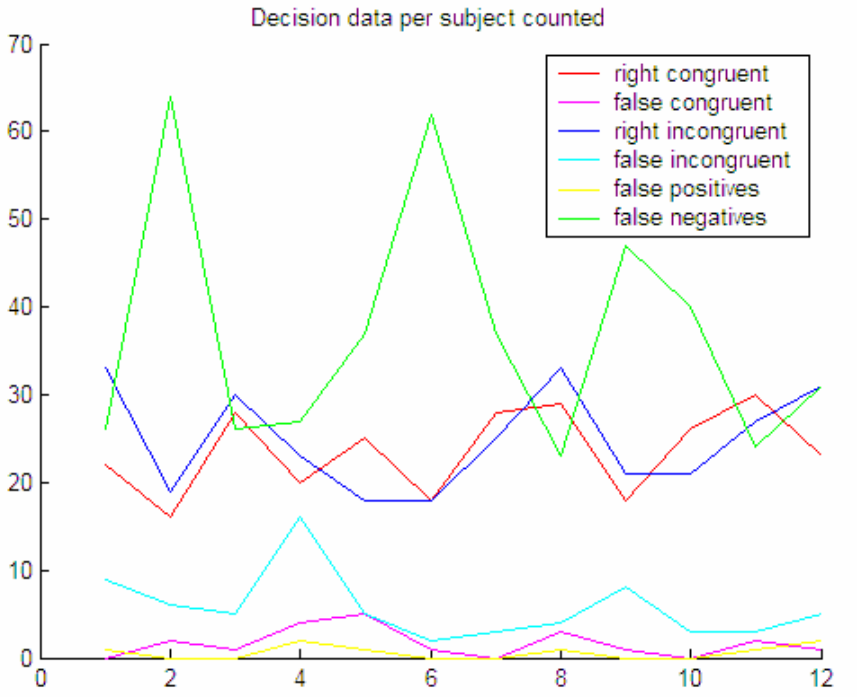


Figure 47 - Number of Images categorized right or wrong by each subject
(X-axis: subjects, Y-axis: number of categorizations)

As shown in Figure 47, the number of correctly judged congruent (red) and incongruent (blue) stimuli is stable and relatively high for all subjects (congruent: mean=4.8, std=23.58; incongruent: mean=24.92, std=5.74). The transfer of stimuli between the congruent and incongruent category by false judgment is small (pink and cyan) (congruent: mean=1.67, std=1.61; incongruent:

mean=5.75, std=3.84). The number of false positives (detected animals on pictures showing no animal) is very low (yellow) (mean=0.67, std=0.78) while the number of unrecognized animals is relatively high and shows a high variance over subjects (mean=37, std=14.16).

Investigating this relatively high level of missed animals, a look at Figure 48 shows that this is caused by a group of pictures which seem to be hard to recognize (n=18, p<0.4) while most of the pictures are recognized with a probability of over 50% (n=87). The figure also shows that pictures presented in the congruent condition are easier to detect than in the incongruent condition (percentage congruent: mean=0.72; percentage incongruent: mean=0.65). The conditional probability for detecting a congruent picture, given the incongruent has been detected, is higher than vice versa (P(detect-C|detect-I)=0.68; P(detect-I|detect-C)=0.62).

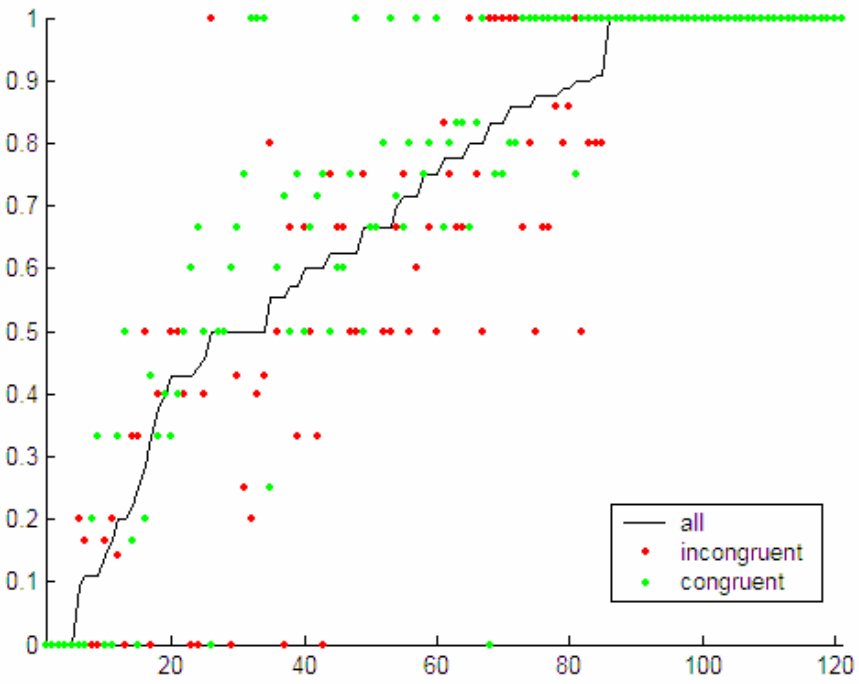


Figure 48 - The percentage of hits per picture per condition
(X-axis: picture, Y-axis: percentage of hits)

Looking at the features of the difficult pictures I realized most of them were subjectively hard to recognize and many of them had decisive features in areas far away from the fixation cross.

Overall, these results indicate that there might be a big difference in subjects performance in the recognition of the animals. There was a slight advantage in recognition of congruent pictures. The transfer of stimuli between congruent and incongruent categories was low while many animal pictures were not recognized and transferred to the no animal condition. The latter partly seemed to be dependent of specific features of the pictures.

B.2.2 Reactiontimes

Most of the pictures were recognized within an interval of six seconds (96.36%) while the mean reaction time for detection of an animal is more than two seconds (mean=2.25, std=144).

As Figure 49 indicates, reaction times show high variance over subjects. There is a clear advantage for pictures presented with no sound (mean=2.04, std=0.24), while only a small advantage can be recognized for congruent pictures (mean=2.33 sec, std=0.44 sec) compared to incongruent pictures (mean=2,42, std=0.39).

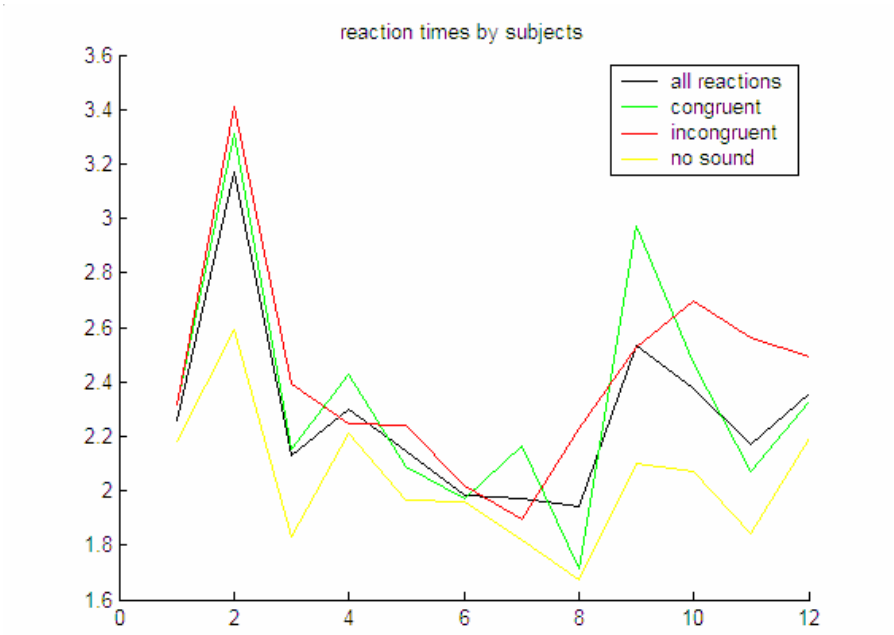


Figure 49 - Reaction times by subjects
(X-axis: subjects Y-axis: reaction time)

Figure 50 and Figure 51 show the reaction times per picture sorted by the difference between the presentation with no sound or with a congruent sound. While Figure 50 shows the absolute reaction times per picture, Figure 51 shows reaction times that have been normalized with each subjects mean reaction time to make times comparable between subjects. Comparing both figures one can see only minor differences indicating that normalization is not necessary to compare reaction times over subjects.

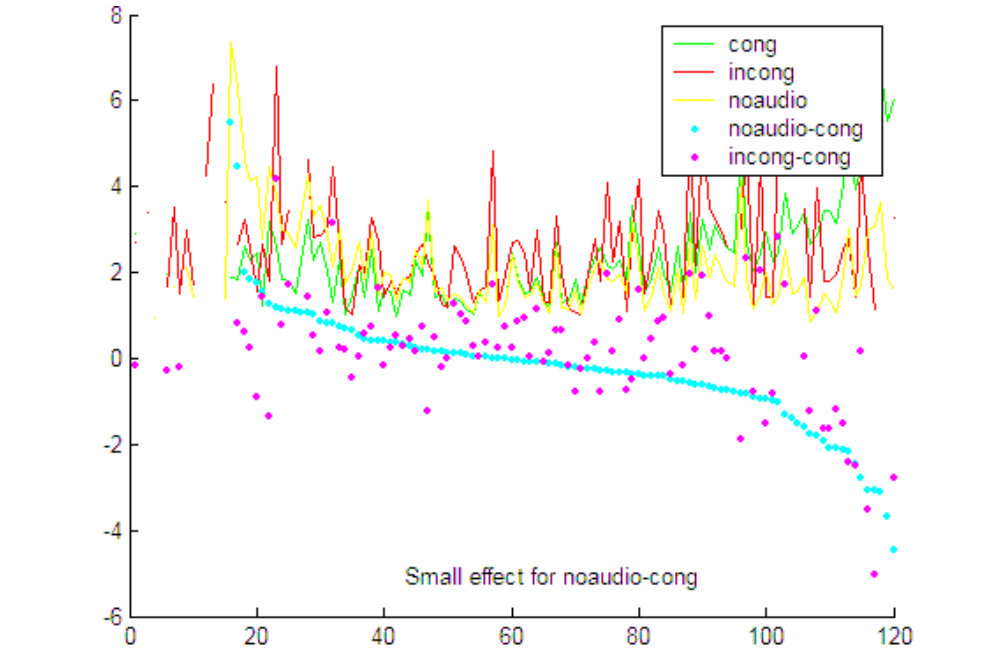


Figure 50 - Reaction times per picture.
Sorted by the difference between the no sound and congruent sound conditions (X-axis: picture, Y-axis: reaction time)

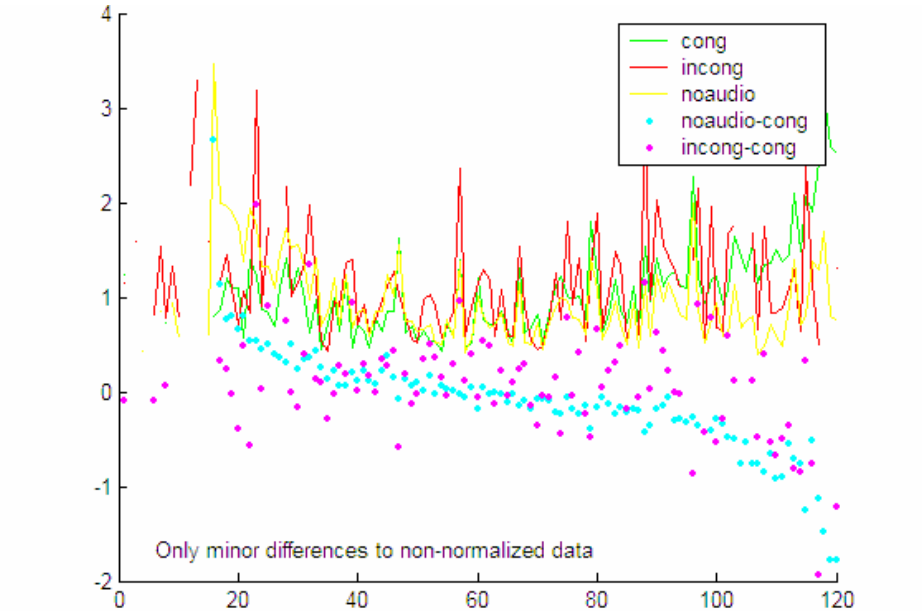


Figure 51 - Reaction times per picture.
Sorted by the difference between the unnormalized times of the no sound and congruent sound conditions (X-axis: picture, Y-axis: reaction time)

Therefore I will focus on the data visible in Figure 50.

The first twenty pictures show very bad detection rates, indicated by the gaps in the data. This fits the data presented in Figure 48 indicating that some pictures are too hard to be recognized.

There is a visible small advantage for several of the congruent pictures compared with the no audio pictures. This advantage holds for about one half of the pictures. The other half shows a disadvantage that is pronounced for a small group of pictures, indicating that for these pictures the sound was diverting. Especially these last pictures seem to cause the overall advantage of the no sound condition.

B.2.3 Learning Effects

Looking at learning effects, I compared the first 10 pictures of one condition with the last 10 pictures. There is no reliable learning effect visible for reaction times for neither condition, as shown in Figure 52 for the congruent condition (difference: mean=0.12 sec, std=0.78). Differences between subjects could be attributed to different levels of difficulty of the pictures in the first and last 10 trials.

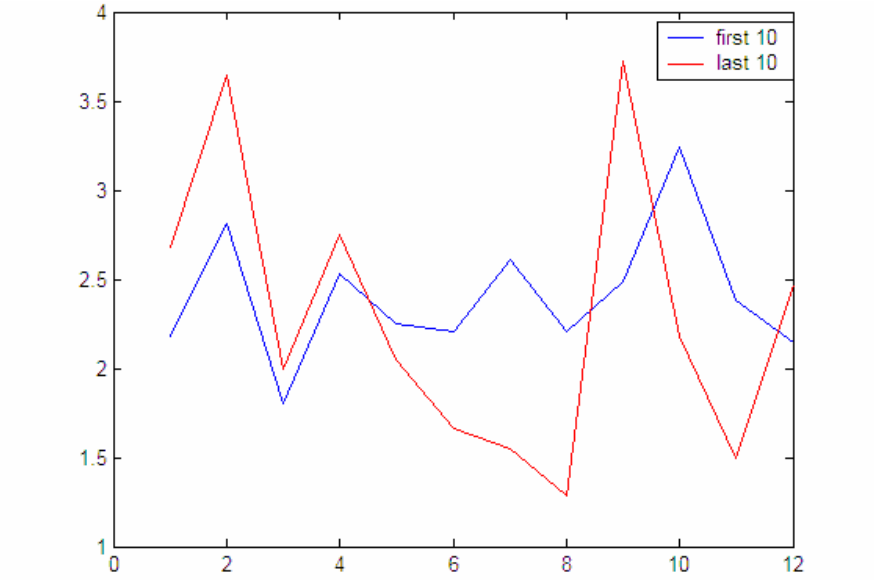


Figure 52 - Mean reaction times for first and last 10 pictures of the congruent condition

This indicates that for all conditions, reaction times seem to be constant over time.

B.2.4 Subjective reports

I asked subjects for their subjective experience during the experiment. They reported that

- Some sounds were not distinctive enough to clearly distinguish different animals
- they ignored the sounds as they were a bad cue, and
- some of the sounds were very unintuitive and therefore were no big help in recognizing the animals.

This indicates that the chosen ratio of congruent stimuli to incongruent and no animal stimuli might have been chosen wrong (40 pictures/95 pictures = 0.42). Overall, there are more situations where the sound is misleading and therefore subjects prefer to ignore the sound. This could also explain the low crossmodal effect in reaction times and the huge advantage of pictures presented without sound over the congruent and incongruent pictures.

Many subjects also reported that while the experiment was quite interesting and fun for them, they still felt quite tired at the end, caused by the monotony of the task.

B.3 Conclusion

From this pretest, I drew the following conclusions for the design of the main experiment:

- As about 25 animal pictures were almost never recognized, these pictures could be excluded from the experiment. Another solution could be to keep the pictures within the set of stimuli (to have as many trials as possible) and expect that only few subjects will find the animals and so the pictures would count as scenery in the subjective categorization.
- As the low ratio of helping sounds to misleading sounds drove subjects to ignore the sounds, the number of stimuli in the congruent condition had to be increased. It was necessary to have a ratio greater than 1 to enable the sound to provide a reliable cue for the subjects.
- There was only a small migration between the congruent and the incongruent stimuli groups. Therefore the recognition of the animals seemed to be very reliable (e.g. not many people saw a cow on a picture showing a cat). This allowed me to reduce the number of the incongruent catch trials. Migration to the no animal group seemed to be high (many animals had not been recognized). Therefore, the no animal group could be designed small, as it would receive enough stimuli caused by wrong judgments (this effect will be increased by including the 25 worst recognized pictures).
- Mean reaction times were about 2 seconds. Therefore, the onset of the animal sound could be placed about 750 ms after the onset of the picture to provide me with enough data to analyze for pure visual synchronization first and then for audio-visual synchronization.

- Most pictures were recognized within a reaction time of about 6 seconds. To make the experiment as short as possible (to avoid tiredness), I shortened the stimulus presentation to this interval.
- The sounds of the animals had to be improved. First, they had to be disambiguated by choosing more distinct sounds and by putting animals together into categories, if they sound too similar (e.g. goose and sheep). Second, the presentation of the sounds was prolonged to the full time of the picture presentation to provide more information to the subjects.
- As there seemed to be no learning effects in the task, I ignored this kind of effect in the analysis of the EEG data.

Appendix C Converting Data from Brain Vision Analyzer (PC) to EEGLab (Mac)

In the Neurophyiology Lab in Hamburg, I recorded EEG data with a sampling rate of 1000 Hz, resulting in a huge amount of data. Therefore the data was not directly readable with EEGLab and I had to down sample it before importing the data. Another technical problem was that I had recorded the data on a PC (Intel architecture) but wanted to analyze it on a Apple Macintosh (Motorola architecture). As the byte order of the big and little endian is different on these systems, I needed to convert the data before reading it with the Mac.

First, I tried to solve these problems by downsampling and converting the data with Brain Vision Analyzer (Brain Products, Version 1.05), still in the laboratory in Hamburg. This software offers an efficient template mechanism for Macros and therefore the task could be done easily.

Back in Osnabrück, I loaded the data into EEGLab and compared one channel of this data with a channel of the raw data converted on the Mac and down sampled by EEGLab. I saw clear differences between the two versions (I checked for filters and other parameters to be as similar as possible). I discussed this topic with the producers of both software packages (BVA and EEGLab) and both stated that their method is working flawless and the diverging data results from different ways of filtering and downsampling. In the end, I decided to discard the files converted with BVA and to import the raw data channel wise into EEGLab and down sample the data there. As I planned to use EEGLab for the further analysis, I wanted to stay with only one software package.

I did the endian conversion using the Unix command *dd*:

dd conv=swab if=[inputfile] of=[outputfile]

Afterwards I used a modified script from the colleagues in Hamburg to read in each channel, down sample it (FIR low pass with a cutoff at 112.5 Hz) to 250 Hz sampling rate and append it to the rest of the data set.

Appendix D Influence of EEG Data Quality on ICA Components

As ICA finds a component for each independent 'feature' in the data, it also creates components for artifacts that occur only in one epoch. Therefore it is important to delete epochs that are heavily artifact loaden before executing the ICA to have as many components as possible for the regularly occurring effects.

I detected a special version of this issue during my first steps: As the low cutoff frequency during recording of the EEG data had been too low, I had slow movements of data in each epoch. As these slow waves were cut into pieces by epoching the continuous data, ICA tried to find a new component for each of these movements of the signal. Therefore, the resulting components topography and spectra were extremely unusual. Except for eye blinks and strong muscle activity, there were almost no other components that were not related to these slow waves (see Figure 53). After filtering the data with a 1 Hz high pass filter, the effect disappeared.

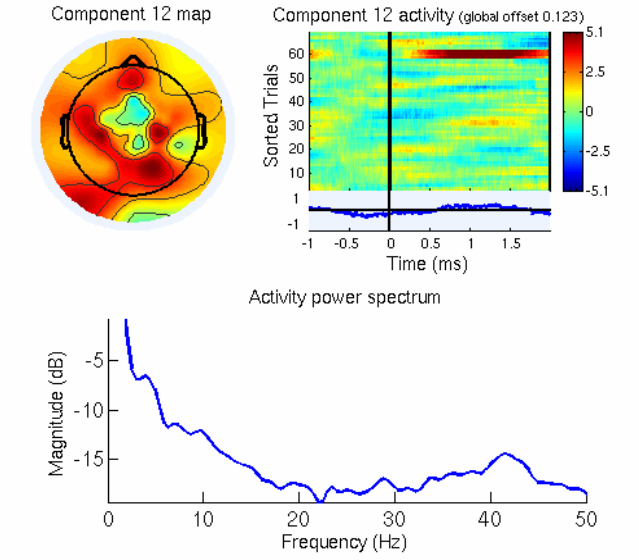


Figure 53 - A bad ICA component produced by missed highpass filtering.
The whole component is dominated by a strong slow wave (around trial 60).

Recapitulating, to retrieve high quality ICA components, it is important

- to remove strong single artifactual events from the data before ICA and
- to remove trends in epochs by high pass filtering the data.

Appendix E Clustering of ICA Components for artifact detection

Independent Component Analysis is a statistical method that does blind source separation. Therefore, there is practically no implied meaning within the results and no easy interpretation. Neither the order of the components nor the seemingly arbitrary weights in the transformation matrix tell anything about the nature of each component per se.

Therefore, classification of components is a task that cannot be automated easily (Jung, Makeig et al., 2000) and that demands personal experience. The nature of this task has several drawbacks:

- It demands hard and erroneous learning from newcomers, until they achieve a reasonable level of experience.
- Basing the classification solely on personal experience makes decisions somewhat arbitrary.
- Manual classification of components is time consuming and there might be a reasonable loss of attention during the task.
- It is hard to keep constantly track of all features of each component that have to be cared about for classification.

To address these issues, I decided to write a clustering tool that introduced some statistical measures to classify the components automatically. This way, it enables the user to get an overview or gist of each group of components. I was inspired to do this by the semi-automatic trial rejection mechanism included in EEGlab (Delorme, Makeig et al., 2001). This mechanism makes trial rejection less arbitrary by proposing trials to reject, based on a classification by several measures as kurtosis, probability of data etc.

I used the following properties to classify components as artifacts (compare Delorme and Makeig, 2004):

1. The frequency spectrum: High power in high frequencies is a distinctive feature of muscle artifacts. Eye blinks also have a distinctive spectrum that is different to normal EEG spectra.
2. The components topography (the weight matrix): Muscle artifacts often concentrate around few electrodes, while eye blinks concentrate around the frontal electrodes.

3. The distribution of the activation of the component over time: muscle artifacts often occur only for one block of trials. Therefore, the overall activation of such a component is high for one block of trials and then low for the rest of the trials.
4. The strength of the component's ERP: a distinctive ERP at stimulus onset or around events indicates brain signals.

I measured these properties in the following way:

1. To measure the spectrum, I used several ratios of frequency samples:
   40/20 Hz, 50/10 Hz, 50/20 Hz, 50/30 Hz, 20/10 Hz, 10/5 Hz.
   Additionally, I included the overall power of the spectrum.
2. To measure the component's topography, I counted the weights that were greater than 2 standard deviations of the whole weight matrix.
3. To search for blocked activity of a component over trials, I summed the absolute activity of all trials over time, convoluted the resulting signal with a 30x1 block, and used the maximum value of the resulting signal as an indicator of blocked activity.
4. To measure the ERPness, I calculated the kurtosis of the components ERP.

These measures provided 10 dimensions describing each component. I used the kmeans method to cluster the data with 100 iterations to derive stable clusters. Clustering the components into 7 groups was the best tradeoff between clearness of classification and reliability of classification.

The screenshot Figure 54 shows the result of such a clustering in a 2D view (showing the 2 main dimensions of a PCA on the clustered data). Rejected components are marked with a red square. For the data shown, there is strong tendency for artifactual components to be on the right side.

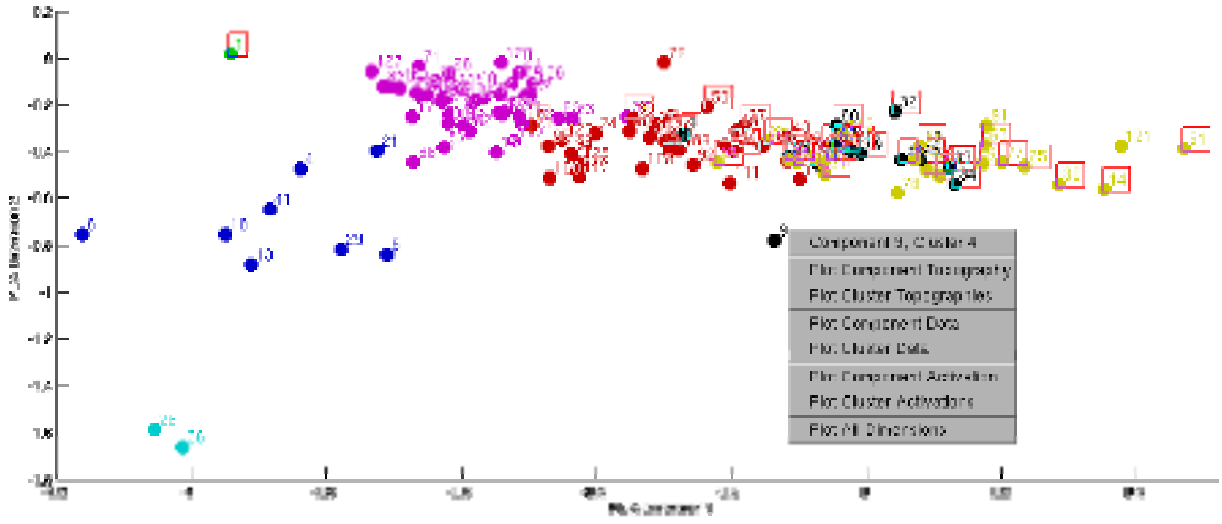


Figure 54 - The results of the clustering process, mapped on the first two PCA dimensions. Artifacts can be found mainly on the right. Components marked for rejection are marked with a read square. A context menu for each component offers further views.

The only one not following this rule is component 1, which represents a cluster on its own and contains eye blinks (see Figure 55).

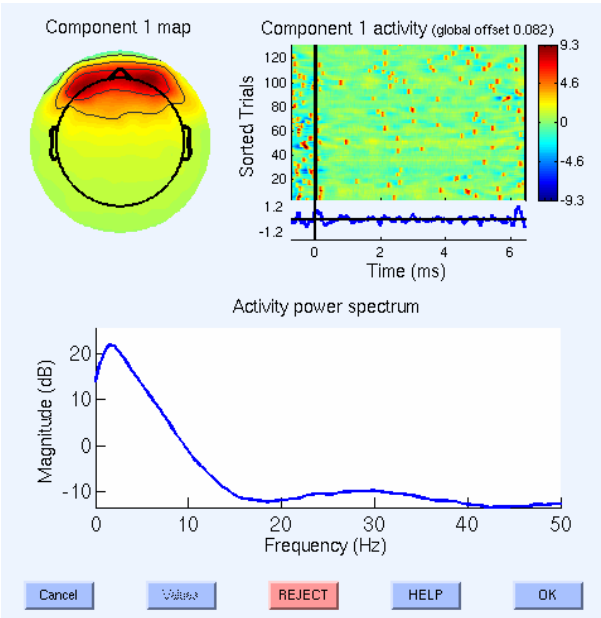


Figure 55 - Plot of the components properties, using built in functions of EEGLab. Component 1 shows a typical topography and spectrum for eye blink artifacts.

I integrated the tool into EEGLab as plugin. It offers the user to plot component topographies (see Figure 56), to scroll component activations (see Figure 57), or to view the component's statistical data (Figure 58).

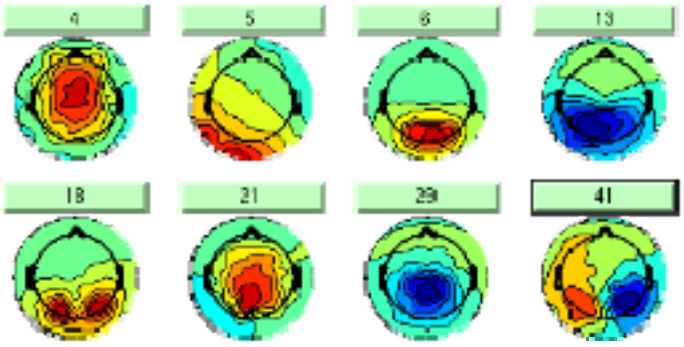


Figure 56 - Topographies of all components in one cluster.
Topographies can be plotted using the built in functions of EEGLab.

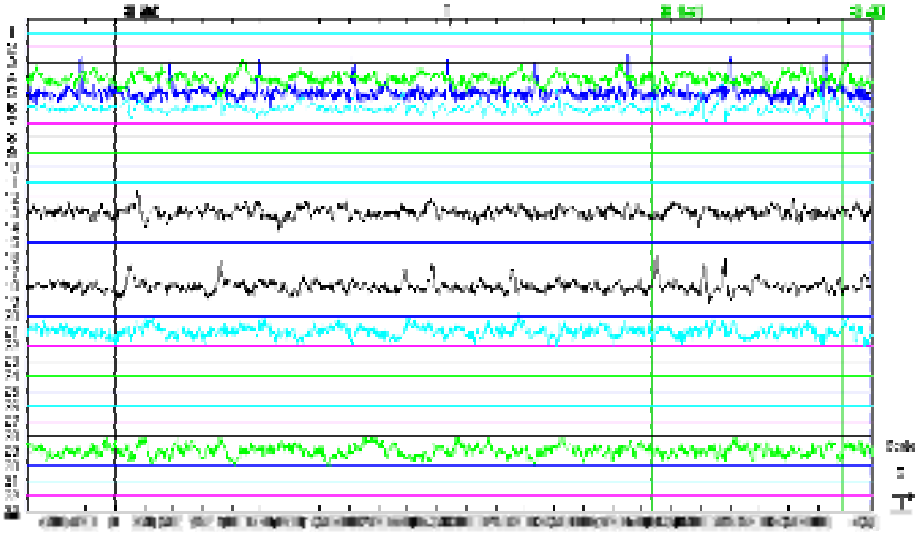


Figure 57 – Waveform of a cluster of components.
To show specifically the activations of the components in one cluster, the tool removes all components, not included in this cluster, and then uses the built in EEGLab function.

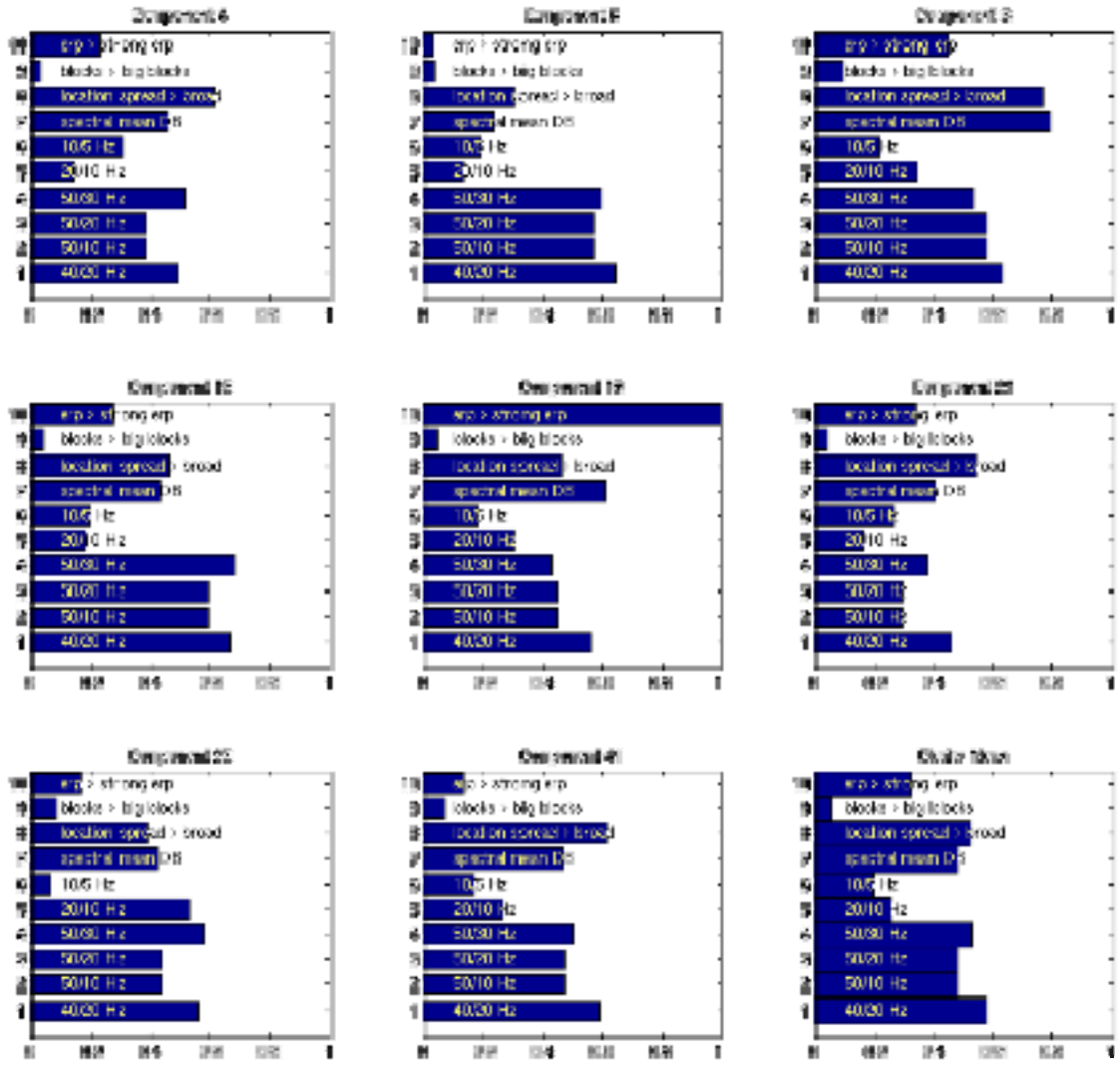


Figure 58 – Overview of cluster data.
To see, why the tool clustered components together, the user may have a look at the statistical data fed into the clustering process.

All views are available for only one component or for the components of a whole cluster. It is also possible to view 2D plots of all clustered dimensions (see Figure 59).

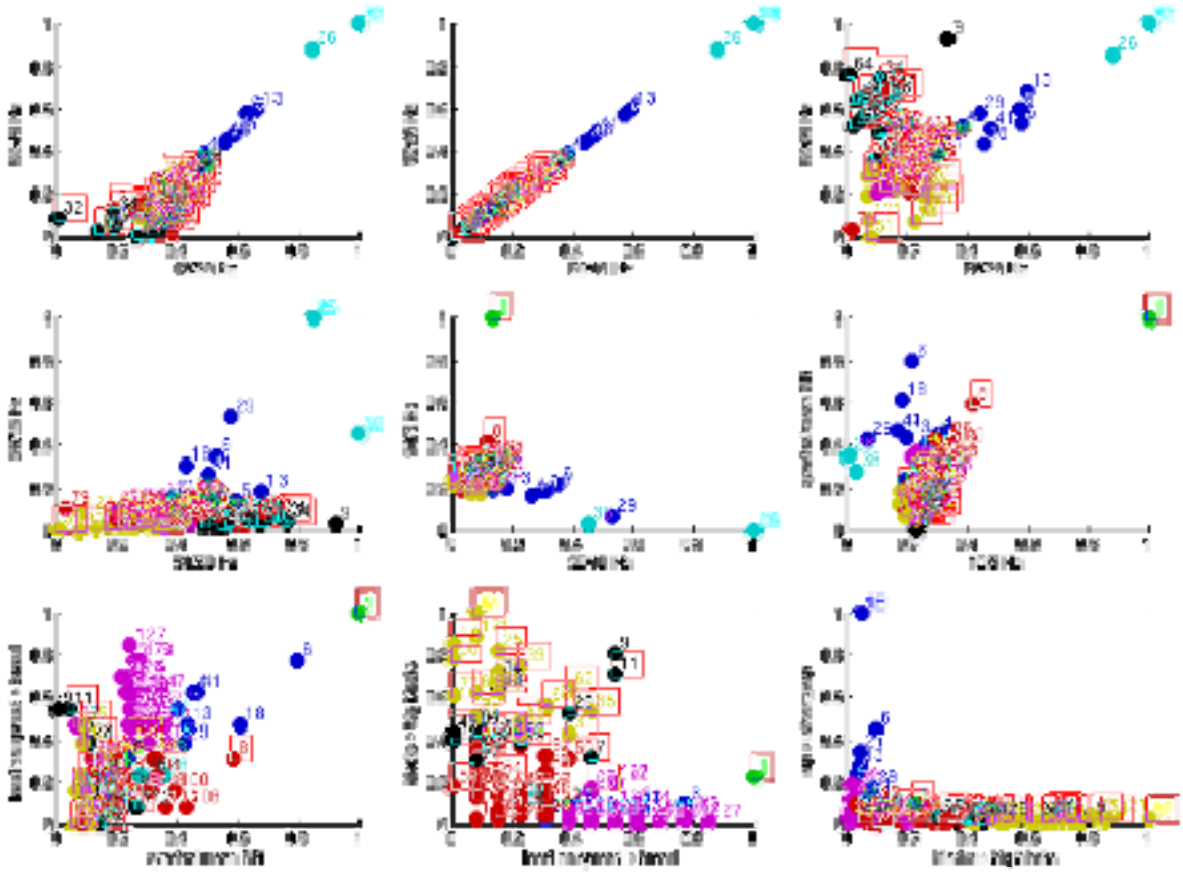


Figure 59 – View of all clustering dimensions.
As the PCA view of the clusters is a very reduced view on the data, the user can view all dimensions underlying the clustering process.

Summarizing, the tool offers a quite reliable method to suggest components for rejection. It does not automate the process of artifact rejection, but it helps to solve or reduce the problems of the process described above.

Appendix F Reduction of Data Volume

The time frequency transformation produces a large amount of data that is too large for practical use on current computers. In the following paragraphs, I will describe the steps I took to reduce the data sets to a bearable size in detail.

For the following calculation, I am assuming a dataset with 128 channels, 150 trials, frequencies from 4 to 80 Hz, and a timeframe of 1 second (250 samples at 250 Hz sampling rate). This results in a volume of data of 128 x 150 x 76 x 250 = 364.8 million values. For energy, all these values have to be stored. As the PLV is a measurement over trials, the factor of 150 is reduced to 1. On the other hand, as the PLV is calculated for each pair of electrodes, there is again a factor of 128, resulting in 311.296 million values.

Calculating the space in bytes, this results in (364.8 + 311.296) x 2 x 8 Byte = 5.408 GB.

This is only the volume for one subject, one condition and one event to look at. Therefore, I had to reduce the amount of data. There were two easy steps to cut down the amount of data:

- Reduction of the precision from double (64 bit) to single (32 bit), which is also done by EEGLab and has not revealed any visible differences in a comparison on randomly chosen real data.
- Cutting the PLV matrix (channels x channel) into half, as the PLV is symmetrical and therefore the matrix is also symmetrical and one side can be thrown away.

To save further space, I decided to increase the steps between frequencies. Figure 60 shows the similarity (PLVs of phases) of signals of neighboring frequencies. The higher the frequency, the more similar become the signals. Therefore I decided to use an exponential stepping between the frequencies, taking almost all low frequencies and less and less higher frequencies (see Figure 61).

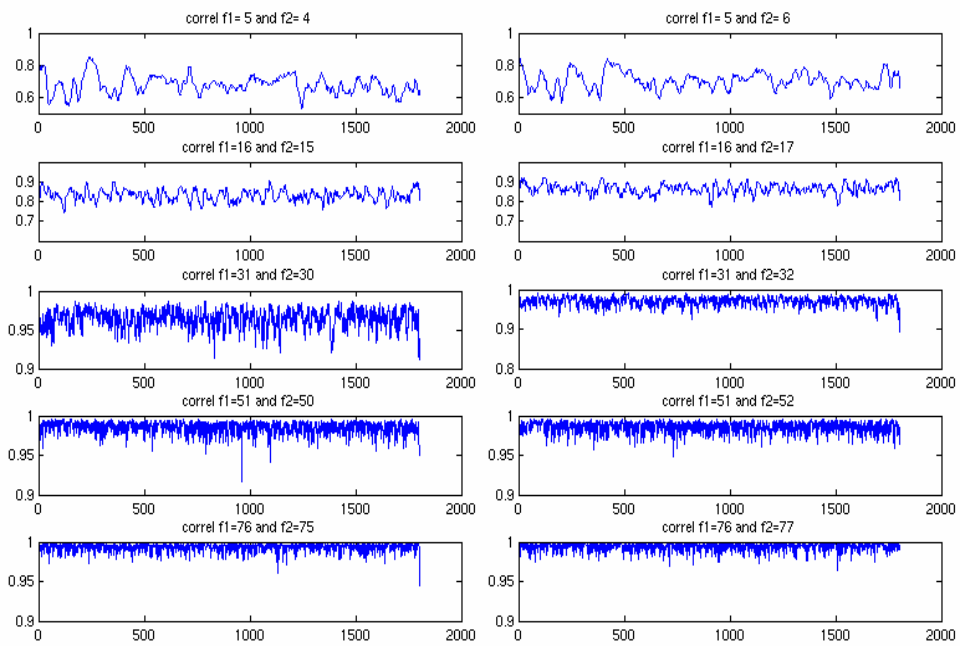


Figure 60 - Similarity of phases for different center frequencies and neighboring frequencies.
Left: lower neighbor; Right: higher neighbor: The higher the frequency, the better is the correlation between center and neighboring frequency (X-axis: time, Y-Axis: PLVs).

This resulted in a reduction factor of about 1.6 and in these two ranges of frequencies for energies and phase synchronization:

- Energy: 20,21,…,26; 28,30,…,56; 59,62,…,80
- Synchronization: 4,5,…,25; 26,28,…,50

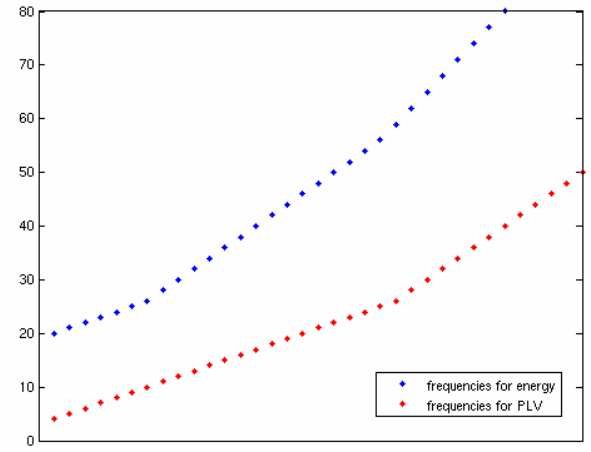
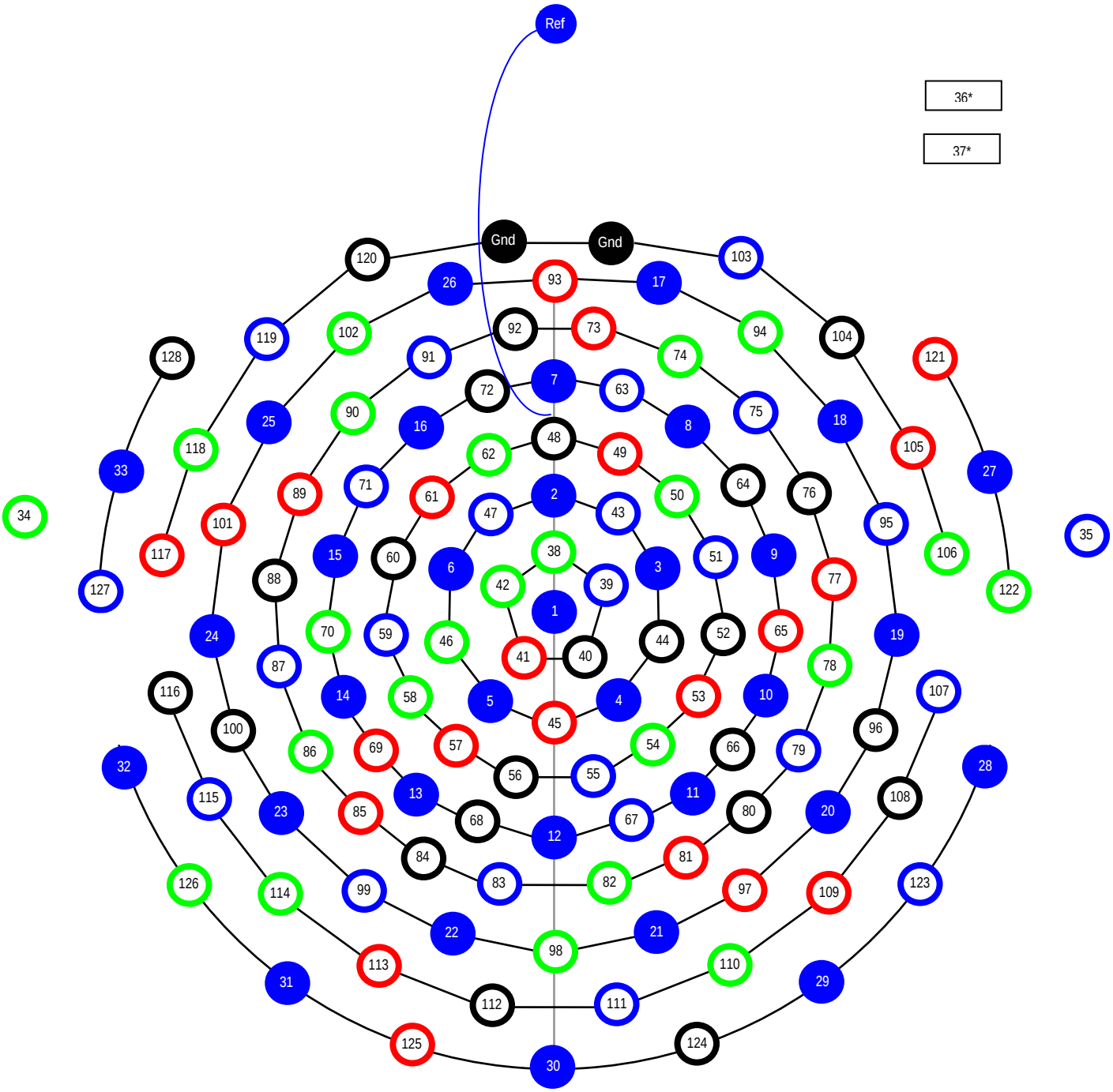


Figure 61 - The exponential stepping between analyzed frequencies.
Dropping frequency data helped to reduce the overall amount of data (X-axis: frequency bin; Y-Axis: assigned frequency).

Finally, the amount of data for one set was about 1.1 GB, which was still a huge amount of data for further processing but bearable in comparison to the first estimation.

Appendix G Layout of EEG channels



Appendix H Checklist for EEG experiment

Preparation:

- Check clean cap and electrodes, fill tubes.
- Start computers, check sound (loudness and loose contacts)
- Prepare forms

Welcome procedure:

- Introduction: show washing area, EEG chamber, recording area
- Show intercom and camera to establish save atmosphere

Forms:

- Let subject sign forms: consent, EHI, FAL, account data
- Vision test (right distance! 3 of 4 matches)
- Audition test (30 dB, all frequencies, left and right)

Washing hair: also wash forehead, dry hair

Preparation:

- Size of head, Distance: nasion, enion
- Place cap (no hair between ears and cap!)
- Check symmetry and enion/nasion (Cz on middle of line)
- Nose reference and EOG: Clear skin with alcohol, clear skin with paste. Adjust cables of electrodes into right direction (should not impair vision!)
- Other electrodes: devide hair below electrode, clean skin with alcohol, fill in paste (put tube right on skin and reject tube slowly while filling)

Entering cabin:

- Place subject on chair
- Switch on ampifier
- Plug in cables
- Fixate cables on chair
- Check impedances on screen, rub in paste (impedances < 10K)
- Refill paste

- Instruction
- Hand over ear plugs with phones

Instruction

- Sit still, relax, expecially relax your shoulders
- During breaks: relax and move to relax
- Show EEG: clench teeth, move body and shoulders, move eyes, blink
- Fixate crosshairs during stimulus presentation
- Speak into Intercom if there are any problems or questions
- Explain experiment (roughly, as tutorial is detailed)
- Announce breaks
- Check Sound on both ears

Starting experiment

- Switch screen: impedances → stimulation
- Create recording file
- Start Presentation run
- Be aware of breaks

After experiment:

- Remove ear plugs, EOG and nose reference
- Unplug EEG
- Remove cap

Switch off amplifier!!!

References

Arbib, M. A. (2003). The handbook of brain theory and neural networks. Cambridge, Mass., The MIT Press.

Bell, A. and T. J. Sejnowski (1995). "An information-maximization approach to blind separation and blind deconvolution." Neural Computation **7**: 1129-59.

Bhattacharya, J., H. Petsche, et al. (2001). "Long-range synchrony in the gamma band: role in music perception." J Neurosci **21**(16): 6329-37.

Bichot, N. P., A. F. Rossi, et al. (2005). "Parallel and serial neural mechanisms for visual search in macaque area V4." Science **308**(5721): 529-34.

Botvinick, M. M., J. D. Cohen, et al. (2004). "Conflict monitoring and anterior cingulate cortex: an update." Trends Cogn Sci **8**(12): 539-46.

Brainard, D. H. (1997). "The Psychophysics Toolbox." Spat Vis **10**(4): 433-6.

Bressler, S. L. and J. A. S. Kelso (2001). "Cortical coordination dynamics and cognition." Trends in Cognitive Sciences **5**(1): 26-36.

Bruns, A. (2004). "Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?" Journal of Neuroscience Methods **137**(2): 321-332.

Busch, N. A., S. Debener, et al. (2004). "Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response." Clin Neurophysiol **115**(8): 1810-20.

Calvert, G. A. and T. Thesen (2004). "Multisensory integration: methodological approaches and emerging principles in the human brain." J Physiol Paris **98**(1-3): 191-205.

Crick, F. and C. Koch (2003). "A framework for consciousness." Nat Neurosci **6**(2): 119-26.

Delorme, A. and S. Makeig (2004). "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis." J Neurosci Methods **134**(1): 9-21.

Delorme, A., S. Makeig, et al. (2001). "Automatic artifact rejection for EEG data using higher-order statistics and Independent Component Analysis." Proc of 3rd International Independent Component Analysis and Blind Source Decomposition Conference.

Freiwald, W. A., A. K. Kreiter, et al. (1995). "Stimulus dependent intercolumnar synchronization of single unit responses in cat area 17." Neuroreport **6**(17): 2348-52.

Frens, M. A., A. J. Van Opstal, et al. (1995). "Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements." Percept Psychophys **57**(6): 802-16.

Fries, P., P. R. Roelfsema, et al. (1997). "Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry." Proc Natl Acad Sci U S A **94**(23): 12699-704.

Ghose, G. M. and J. Maunsell (1999). "Specialized representations in visual cortex: a role for binding?" Neuron **24**(1): 79-85, 111-25.

Giard, M. H. and F. Peronnet (1999). "Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study." J Cogn Neurosci **11**(5): 473-90.

Goffaux, V., A. Mouraux, et al. (2004). "Human non-phase-locked gamma oscillations in experience-based perception of visual scenes." Neurosci Lett **354**(1): 14-7.

Gruber, T., M. M. Müller, et al. (2002). "Modulation of induced gamma band responses in a perceptual learning task in the human EEG." J Cogn Neurosci **14**(5): 732-44.

Haig, A. R., E. Gordon, et al. (2000). "Synchronous cortical gamma-band activity in task-relevant cognition." Neuroreport **11**(4): 669-75.

Herrmann, C. S. and R. T. Knight (2001). "Mechanisms of human attention: event-related potentials and oscillations." Neurosci Biobehav Rev **25**(6): 465-76.

Herrmann, C. S. and A. Mecklinger (2000). "Magnetoencephalographic responses to illusory figures: early evoked gamma is affected by processing of stimulus features." Int J Psychophysiol **38**(3): 265-81.

Herrmann, C. S., M. H. J. Munk, et al. (2004). "Cognitive functions of gamma-band activity: memory match and utilization." Trends in Cognitive Sciences **8**(8): 347-355.

Hochstein, S. and M. Ahissar (2002). "View from the top: hierarchies and reverse hierarchies in the visual system." Neuron **36**(5): 791-804.

Hubel, D. H. and T. N. Wiesel (1968). "Receptive fields and functional architecture of monkey striate cortex." J Physiol **195**(1): 215-43.

Iriarte, J., E. Urrestarazu, et al. (2003). "Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study." J Clin Neurophysiol **20**(4): 249-57.

Joliot, M., U. Ribary, et al. (1994). "Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding." Proc Natl Acad Sci U S A **91**(24): 11748-51.

Jung, T. P., S. Makeig, et al. (2000). "Removing electroencephalographic artifacts by blind source separation." Psychophysiology **37**(2): 163-78.

Kaiser, J., I. Hertrich, et al. (2005). "Hearing lips: gamma-band activity during audiovisual speech perception." Cereb Cortex **15**(5): 646-53.

Kaiser, J. and W. Lutzenberger (2005). "Human gamma-band activity: a window to cognitive processing." Neuroreport **16**(3): 207-11.

Keil, A., M. M. Müller, et al. (1999). "Human gamma band activity and perception of a gestalt." J Neurosci **19**(16): 7152-61.

Knyazeva, M. G., D. C. Kiper, et al. (1999). "Visual stimulus-dependent changes in interhemispheric EEG coherence in humans." J Neurophysiol **82**(6): 3095-107.

König, P., A. K. Engel, et al. (1995). "Relation between oscillatory activity and long-range synchronization in cat visual cortex." Proc Natl Acad Sci U S A **92**(1): 290-4.

Kreiter, A. K. and W. Singer (1996). "Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey." J Neurosci **16**(7): 2381-96.

Lachaux, J. P., N. George, et al. (2005). "The many faces of the gamma band response to complex visual stimuli." Neuroimage **25**(2): 491-501.

Lachaux, J. P., E. Rodriguez, et al. (1999). "Measuring phase synchrony in brain signals." Hum Brain Mapp **8**(4): 194-208.

Lutz, A., J. P. Lachaux, et al. (1999). "Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. PG - 1586-91." Proc Natl Acad Sci U S A **99**(3): 1586-1591.

Makeig, S., A. Bell, et al. (1996). Independent Component Analysis of Electroencephalographic Data. Advances in Neural Information Processing Systems 8. D. Touretzky, M. Mozer and M. Hasselmo. Cambridge MA, MIT Press**:** 145-151.

Meinecke, F. C., A. Ziehe, et al. (2005). "Measuring phase synchronization of superimposed signals." Phys Rev Lett **94**(8): 084102.

Miller, J. (1982). "Divided attention: evidence for coactivation with redundant signals." Cognit Psychol **14**(2): 247-79.

Mima, T., T. Oluwatimilehin, et al. (2001). "Transient interhemispheric neuronal synchrony correlates with object recognition." J Neurosci **21**(11): 3942-8.

Müller, M. M., J. Bosch, et al. (1996). "Visually induced gamma-band responses in human electroencephalographic activity--a link to animal studies." Exp Brain Res **112**(1): 96-102.

Müller, M. M., T. Gruber, et al. (2000). "Modulation of induced gamma band activity in the human EEG by attention and visual information processing." Int J Psychophysiol **38**(3): 283-99.

Nunez, P. L. (1989). "Estimation of large scale neocortical source activity with EEG surface Laplacians." Brain Topogr **2**(1-2): 141-54.

Oldfield, R. C. (1971). "The assessment and analysis of handedness: the Edinburgh inventory." Neuropsychologia **9**(1): 97-113.

Palva, J. M., S. Palva, et al. (2005). "Phase synchrony among neuronal oscillations in the human cortex." J Neurosci **25**(15): 3962-72.

Pelli, D. G. (1997). "The VideoToolbox software for visual psychophysics: transforming numbers into movies." Spat Vis **10**(4): 437-42.

Pulvermüller, F. (2001). "Brain reflections of words and their meaning." Trends in Cognitive Sciences **5**(12): 517-524.

Reynolds, J. H. and R. Desimone (1999). "The role of neural mechanisms of attention in solving the binding problem." Neuron **24**(1): 19-29, 111-25.

Ridderinkhof, K. R., M. Ullsperger, et al. (2004). "The role of the medial frontal cortex in cognitive control." Science **306**(5695): 443-7.

Rodriguez, E., N. George, et al. (1999). "Perception's shadow: long-distance synchronization of human brain activity." Nature **397**(6718): 430-3.

Rose, M. and C. Buchel (2005). "Neural coupling binds visual tokens to moving stimuli." J Neurosci **25**(44): 10101-4.

Roskies, A. L. (1999). "The binding problem." Neuron **24**(1): 7-9, 111-25.

Sarnthein, J., H. Petsche, et al. (1998). "Synchronization between prefrontal and posterior association cortex during human working memory." Proc Natl Acad Sci U S A **95**(12): 7092-6.

Schack, B., W. Klimesch, et al. (2005). "Phase synchronization between theta and upper alpha oscillations in a working memory task." Int J Psychophysiol **57**(2): 105-14.

Schack, B., N. Vath, et al. (2002). "Phase-coupling of theta-gamma EEG rhythms during short-term memory processing." Int J Psychophysiol **44**(2): 143-63.

Shadlen, M. N. and J. A. Movshon (1999). "Synchrony unbound: a critical evaluation of the temporal binding hypothesis." Neuron **24**(1): 67-77, 111-25.

Singer, W. (1999). "Neuronal synchrony: a versatile code for the definition of relations?" Neuron **24**(1): 49-65, 111-25.

Singer, W., A. K. Engel, et al. (1997). "Neuronal assemblies: necessity, signature and detectability." Trends in Cognitive Sciences **1**(7): 252-261.

Singh, K. D., G. R. Barnes, et al. (2002). "Task-related changes in cortical synchronization are spatially coincident with the hemodynamic response." Neuroimage **16**(1): 103-14.

Tallon-Baudry, C. and O. Bertrand (1999). "Oscillatory gamma activity in humans and its role in object representation." Trends Cogn Sci **3**(4): 151-162.

Tallon-Baudry, C., O. Bertrand, et al. (1997). "Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans." J Neurosci **17**(2): 722-34.

Tallon-Baudry, C., O. Bertrand, et al. (1996). "Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human." J Neurosci **16**(13): 4240-9.

Tallon-Baudry, C., O. Bertrand, et al. (1998). "Induced gamma-band activity during the delay of a visual short-term memory task in humans." J Neurosci **18**(11): 4244-54.

Tiitinen, H., J. Sinkkonen, et al. (1993). "Selective attention enhances the auditory 40-Hz transient response in humans." Nature **364**(6432): 59-60.

Treisman, A. (1998). "Feature binding, attention and object perception." Philos Trans R Soc Lond B Biol Sci **353**(1373): 1295-306.

Treisman, A. and H. Schmidt (1982). "Illusory conjunctions in the perception of objects." Cognit Psychol **14**(1): 107-41.

Trujillo, L. T., M. A. Peterson, et al. (2005). "EEG phase synchrony differences across visual perception conditions may depend on recording and analysis methods." Clin Neurophysiol **116**(1): 172-89.

Varela, F., J. P. Lachaux, et al. (2001). "The brainweb: phase synchronization and large-scale integration." Nat Rev Neurosci **2**(4): 229-39.

von der Malsburg, C. (1981). "The Correlation Theory of Brain Function." Max Planck Institute for Biophysical Chemistry **Internal Report 81**(2).

von der Malsburg, C. (1999). "The what and why of binding: the modeler's perspective." Neuron **24**(1): 95-104, 111-25.

von Stein, A., C. Chiang, et al. (2000). "Top-down processing mediated by interareal synchronization." Proc Natl Acad Sci U S A **97**(26): 14748-53.

von Stein, A., P. Rappelsberger, et al. (1999). "Synchronization between temporal and parietal cortex during multimodal object processing in man." Cereb Cortex **9**(2): 137-50.

von Stein, A. and J. Sarnthein (2000). "Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization." Int J Psychophysiol **38**(3): 301-13.

Wallace, M. T., R. Ramachandran, et al. (2004). "A revised view of sensory cortical parcellation." Proc Natl Acad Sci U S A **101**(7): 2167-72.

Zatorre, R. J., P. Belin, et al. (2002). "Structure and function of auditory cortex: music and speech." Trends Cogn Sci **6**(1): 37-46.

Zimmer, U., J. Lewald, et al. (2004). "Is there a role of visual cortex in spatial hearing?" Eur J Neurosci **20**(11): 3148-56.

Eidesstattliche Erklärung

Hiermit erkläre ich, Stefan Scherbaum, die vorliegende Arbeit "Neural Synchronization in the Crossmodal Integration of Camouflage Pictures and Sounds" selbstständig verfasst zu haben und keine anderen Quellen und Hilfsmittel als die angegebenen verwendet zu haben.

Osnabrück, den 31.12.2005

Stefan Scherbaum

Matrikelnummer; 912 865

Affirmation

I, Stefan Scherbaum, hereby confirm that I composed the thesis at hand, entitled "Neural Synchronization in the Crossmodal Integration of Camouflage Pictures and Sounds" independently and that I did not use any other resources or auxiliary means than the ones stated.

Osnabrück, December 31, 2005

Stefan Scherbaum

Matriculation Number: 912 865