The Time Course of Negative Priming

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von

Hendrik Degering, geb. Schrobsdorff

aus Göttingen

Göttingen, 2009

D7
Referent: Prof. Dr. Theo Geisel
Korreferent: Prof. Dr. Marcus Hasselhorn
Tag der mündlichen Prüfung:

Contents

1 Introduction

Theoretical life in psychology seemed just a forever-long sequence of dichotomies.
(Newell, 1994)

The present thesis reports on an interdisciplinary attempt at explaining the negative priming effect, a characteristic of selective attention, by a combination of behavioral fundamental research, neuroimaging studies, theoretical psychology and computational modeling. The negative priming effect is one of a very few measures for ignoring. It is revealed as a slowing down of responses to stimuli that were ignored recently as compared to those that are new. Since the discovery of the negative priming effect in 1966 a vivid debate on the cognitive mechanisms underlying the deceleration has evolved, without arriving at a conclusive consensus.

Over the years, a large number of negative priming experiments have been conducted, mostly focusing on a special aspect of the effect by the use of a unique paradigm. Regrettably, the results of each of these studies show a unique pattern as well. Only the bare slowing of responses to previously ignored stimuli is found in most of the experiments, but virtually any manipulation of a paradigm also affects negative priming.

In the introduction we will explain the negative priming phenomenon in section 1.1, giving a condensed overview of the exhaustive presentation of the field of negative priming in chapter 2. We then expound the importance of computational modeling for theoretical psychology in section 1.2. The structure of the thesis is presented in section 1.3 which also describes our research on negative priming as a whole. Finally, in section 1.4 we will conclude the introduction by listing the original contributions included in this thesis.

1.1 Negative Priming

Selective attention enables goal-directed behavior despite the permanent, immense input to the sensory system. The downside of this ability involves the problem of how information is ignored. Contradicting early speculations of an active attending and passive ignoring, a special situation revealed the active nature of ignoring. In the original experiment by Dalrymple-Alford and Budayr (1966), subjects had to process lists of Stroop tasks. While in the original Stroop task no systematic repetition of color and color words was implemented, these authors composed the stimulus cards in a special way, namely the ignored meaning of a color word always became the to be named color in which the next word was shown on some of the lists, on others there was no relation between two succeeding words. The experiment showed that people were slower in responding to the related lists compared to unrelated stimulus colors. Even if the semantic meaning of the words has been ignored, it must have entered the cognitive system as it showed the characteristic interference.

Since then, several standard negative priming paradigms have emerged, each featuring a certain dimension on which priming happens, e.g. the identity of stimulus objects or their location on the presentation screen. The set of stimuli also varies enormously, e.g. pictures, shapes, words, letters, sounds, colored dots. The common denominator of all paradigms is the classification of stimuli in targets that have to be attended to, and distractors that are to be ignored. Stimulus repetitions are considered in dependence of the role of the repeated object as target or distractor in two related trials. Variations of this basal setting include the manipulation of experimental parameters like the time between two related trials, the number of distractors from zero in some trials to multiple over the entire experiment, and the saliency of the distractor. For a detailed listing of the sometimes contradictory results, see section 2.3.

Because of the controversial nature of the negative priming effect, a variety of different theoretical accounts have been developed. But until now none of the theoretical accounts is able to explain all aspects of the negative priming effect, they all have their strong points as well as their shortcomings. All theories assume different mechanisms to be responsible for negative priming. In order to clarify the situation of diverging explanatory accounts, the time course of negative priming is crucial. The mechanisms postulated by the different theories act in different stages of trial processing.

1.2 Computational Modeling of Negative Priming

The theories to explain negative priming can be categorized roughly as memory-based and activation-based accounts. The first group assumes the memorization of a trial and eventually a retrieval of the information in the next trial. The latter group assumes negative priming to be caused by interference of trial processing with persistent activation from former trials. Both directions produced a variety of small branches, many limited to a single appearance in order to explain a certain, singular pattern of results. But due to the lack of comparability and concreteness, there is no solution of the debate on the level of argumentative theories in sight.

In the face of such a sensitive phenomenon it is understandable that no comprehensive explanation has been found so far. Because a satisfying theory should be less complex than the data it explains, it seems reasonable to focus on the interaction of the underlying causes instead on ad hoc defined data features. However, a main reduction of complexity is already achieved by the design of experiments. Nevertheless, a theoretical approach is based on the assumption that the complexity of experimental data can be further reduced by identifying repeating patterns in the data. Our first attempt, the implementation of a simplified but still biologically motivated model of target selection has proven too simple by our experiments. Although it provided us with a tangible account of several dependencies of negative priming. A crucial point in the specification of mechanisms producing negative priming seems to be the exact time course of processing a trial where a previously ignored stimulus has to be attended in comparison with the processing of an unprimed stimulus. Therefore, we faced the problem to reveal temporal information about negative priming.

A first step in that direction is already our simplistic computational implementation as described by continuous nonlinear differential equations which themselves show a characteristic time course. In order to test the validity of the model we designed several experimental paradigms according to the objective of making statements about the inner temporal structure of a negative priming trial. Some of the experiments recorded EEG data which has shown to be a beneficial tool in identifying systematic differences in trial processing, both spatially and temporally. For two experiments we developed techniques to record additional time makers during trial processing, making it possible to temporally localize the emergence of reaction time effects. In order to tackle the problem of

diverse paradigms and the incomparability of theoretical accounts, we designed a computational framework for perception based action selection by means of physiologically justified building blocks which each obey a biologically plausible dynamics.

Despite all concise and generally understandable theories that seem to have identified the causes of an observed phenomenon, it is very important to keep in mind that psychological fundamental research uses statistical properties of experimental data in order to interprete human behavior. On the one hand, behavioral experiments tend to produce largely varying results, caused by the complexity of the human cognitive system. On the other hand, the interpretation of results is usually not unambiguous. Both aspects provide a base for the arduous and controversial discourse that is necessary for clarifying a certain psychological phenomenon.

One possibility to proceed in the discussion is to solidify theories by mapping their assumptions on measurable processes in the brain, thereby eliminating arbitrariness of the respective interpretations. A second way is to computationally implement theories. Clearly, the obtained implementation inherits the freedom of interpretation from the underlying theory. Yet, the implementation adds further degrees of freedom. But the benefits of an implementation are obvious. It eliminates the risk of misinterpretation, as the source code can be made available for other research groups interested in working with the model instead of leaving them with wordy descriptions. A computational model may provide links to biological data, all the more if it is based on naturally observable processes.

Nevertheless, certain aspects have to be remembered when arguing on the level of implementations. In order to reproduce the observed results, most models have to undergo a precise fitting of parameters, which is also a very subjective process. Therefore, great care has to be taken of the distinction between results due to parameter fits and extrapolations by the internal dynamics of the model itself. A different way to benefit from a computational model is to analyze the structural result after fitting, which carries a formalized version of the fitted data. Or, in the words of Hintzman (1991): *The measure of a model's value lies not in its ability to fit data, but in how much we can learn from it.*

We will comply with the necessity of quantification in two ways. First we will take up a single theory of negative priming, i.e. the imago semantic action model described in section 2.4.7, and build a minimal model producing realistic effects on the basis of the postulated mechanisms, see chapter 3. A detailed implementation was performed in close interaction with the originator of the theory. The presence of the cognitive representation of a certain object is modeled by a single variable, by which we obtain a rather clear dynamic system which is able to deal with realistic stimulus sequences and generates artificial reaction times. In chapter 6 we will show how the model can be extended to generate hypotheses in a more complex paradigm. The generalized model enables us to resolve contradictions arising in the initially attempted modeling approach. This is to be considered as a success of the modeling process, as we are able to falsify an essential assumption of the original theory by means of a straightforward implementation.

The second computational approach is more ambitious with respect to the discussion about the applicability of the theories of negative priming in specific situations. We build a computational model comprising most of the mechanisms suspected to play a role in the neural processing in negative priming. The outcome is not only a meta-model for negative priming, termed General Model, but in itself a simplified model of the brain as a framework for action selection based on perception. We addressed the tradeoff between biological realism and understandability by modeling each assumed mechanism separately but keeping the internal dynamics of each of the corresponding layers very simple by taking over the dynamical framework of our first model.

1.3 Thesis Overview

The present thesis will describe our multi-level approach to reveal the temporal structure of the negative priming effect. Accompanied by computational modeling, we run sophisticated psychological experiments and record and analyze EEG data. We will start with an overview of the phenomenon of negative priming in chapter 2.3 by reviewing the literature for behavioral results and theoretical explanations of the effect.

Based on one of the theoretical accounts introduced in chapter 2.3, namely the Imago Semantic Action Model, see section 2.4.7, we will implement our first computational model for negative priming, the ISAM. The basics of the modeling approach and the implementation will be the first part of chapter 3. The second part will be devoted to testing the ISAM by deriving predictions and reproducing several effects related to negative priming.

Adapting the voicekey paradigm described in section 2.2, we will describe an EEG experiment in chapter 4 that replicates findings from one of the few studies on event-related potentials related to negative priming. Beforehand we will give a detailed introduction to EEG recordings and the corresponding data analysis and thoroughly review hitherto existing findings of EEG correlates of negative priming.

During the preoccupation with EEG data analysis, we came upon an inconsistency in averaging event related potentials. Chapter 5 introduces our solution to the problem to reconstruct a very noisy signal that additionally is subject to erratic temporal fluctuations. As such a new technique first has to prove its validity in a broad discussion, we limit it in the current thesis to an interlude independent of the remainder.

Due to the additional source of uncertainty in EEG research, i.e. the interpretation of differing event related potentials in the different experimental conditions, we determined ourselves to behavioral experiments and designed a paradigm which requires a button press between stimulus identification and target selection phase which is recorded as an additional reaction time. Chapter 6 describes the model based generation of hypotheses by the ISAM of chapter 3, the paradigm itself and finally the results that locate negative priming in the later part of a trial and that contradict the ISAM all along the line.

After separating the stimulus identification phase, the remainder of a trial still contains the two stages of processing of target selection and response generation. One theory predicts negative priming to be exclusively produced in the response generation phase. Therefore, we constructed a second trial splitting paradigm which now singles out the response generation phase. In chapter 7 we will describe the paradigm, go into expected side effects of the altered paradigm and finally display the results, the devotion of negative priming to the target selection phase of a trial.

Not only the nontrivial extension of our identity based priming paradigm given in chapter 7 to a comparison task, but also the counterevidence for the ISAM by the experiment in chapter 6 made us head for a less rigid computational model. Chapter 8 pictures the result in form of our General Model for negative priming which provides an implementation of each theory and the ability to respond in various different negative priming paradigms. Due to the complexity of the model chapter 8 can only be seen as the general introduction to a new framework which will possibly shed light on the questions why different paradigms produce such diverse result patterns, and how the theories can be compared on a par.

The previous chapters are concluded in chapter 9 which also collects all results and forms a complete picture of negative priming as we can give it by our research. This chapter contains also an outlook on future directions to finally conclude the main body of the thesis. Appended is a listing of experimental data in tables, which were excluded from the according chapters for the sake of readability.

1.4 Original Contributions

All work presented in the present thesis is carried out by a closely cooperating workgroup in the framework of section C4 of the Bernstein Center for Computational Neuroscience Göttingen. The results presented here would not have been possible without this collaboration. My personal contributions are not restricted to modeling but have had an increasing influence also on experimental design, data analysis, interpretation of results, and design of algorithms.

Our main contributions to (but not limited to) negative priming research are listed in the following.

- We developed a simple model for the transient of the firing rate response of an integrate and fire network to constant input by the means of a nonlinear Langevin equation, section 3.1.
- We employed the resulting dynamics to build a minimalistic computational model, section 3.2, reproducing priming effects based on the mechanisms of the global threshold theory, section 2.4.7.
- With the good performance of the model, section 3.3, we quantitatively validated global threshold theory (Schrobsdorff et al., 2007b).
- We adapted our voicekey paradigm, section 2.2 to an EEG recording environment, section 4.4 and replicated some of the very sparse event related potential correlates for negative priming found in a rather different paradigm, section 4.6.
- We confirmed that processing in ignored repetition trials first benefits from stimulus repetition similar to the attended repetition condition, but only later in the trial both conditions diverge due to different demands on cognitive control, section 4.8, (Behrendt et al., 2009)
- We developed sophisticated signal processing methods, sections 5.4 and 5.7, which enhance the averaging of event related potentials, section 5.7.5, and provide a measure for the temporal variation in the processing between two trials, section 5.5, (Ihrke et al., 2008, 2009b).
- We designed an enhanced algorithm for line-of-synchrony detection in recurrence plots which outperforms established solutions, section 5.7.4, (Ihrke et al., 2009a).
- We introduced time markers in addition to the usual reaction time into negative priming paradigms, making it possible to investigate the temporal structure of the mechanisms causing negative priming by means of behavioral measures, section 6.1 and 7.1.
- By applying our technique of recording intermediate time markers, we have shown that the stimulus identification phase of a trial carries no negative priming, but only facilitation in the presence of repeated objects, section 6.6.
- By deriving predictions from our computational implementation of the global threshold theory to the task switch paradigm, section 6.2, we provided strong counterevidence for that theory as predicts negative priming to happen already in the identification phase, sections 6.3 and 6.6.
- We showed that negative priming happens in the target selection phase of a trial, section 7.5, by again isolating a part from trial processing, in this case the response generation phase, section 7.1.

- Finally we implemented a neurophysiological model, section 8.3, of the parts of the brain that are assumed to be involved in processing a priming trial, section 8.2. The General Model is able to cope with various paradigms, section 8.1.1, and implements the behavior assumed by any of the negative priming theories, section 8.1.2.

Although partially not yet published as articles, all points are documented by a series of conference contributions listed on page 162 ff. and are available at

`www.nld.ds.mpg.de/~hecke/research.html`

2 Negative Priming

Priming is characterized by a sensitivity of reaction times to how stimuli have been encountered recently. A reduction of the reaction time, positive priming, is usually observed with repetitions of stimuli or responses and is well-known and experimentally understood (Scarborough et al., 1977). Our object of investigation, negative priming, a slowdown in the reaction time usually in response to previously ignored stimuli, is experimentally less tangible (Fox, 1995). The negative priming effect is sensitive on even subtle parameter changes, which poses many methodological and conceptual challenges, but bears exactly for this reason great potential for applications in research fields such as memory, selective attention, and aging effects.

The following chapter will thoroughly introduce the negative priming phenomenon. After a classification of negative priming and a description of the terminology used in negative priming studies in section 2.1, we will discuss a showcase study to give a feeling for what a negative priming experiment looks like in section 2.2. The diversity of findings concerning negative priming will be shown in section 2.3. Then we will give a detailed listing of theoretical accounts to the negative priming effect in section 2.4.

2.1 A Paradigm to Access Selective Attention

Selective attention is the process of extracting behaviorally relevant information from the environment. The focusing on particular stimuli brings along an ignoring of irrelevant information. The process of ignoring is investigated by systematic variation of irrelevant stimuli. Interesting effects like change blindness, the failure to perceive even striking changes in a visual scene that are not behaviorally relevant (McConkie and Currie, 1996), or inattentional blindness, the apparent insensitivity of the cognitive system to unattended stimuli (Simons and Chabris, 1999), demonstrate impressively that our feeling of perceptual accuracy is not objective.

It is still unclear how the selection of stimuli is done. Two classes of mechanisms are assumed, top-down and bottom-up processes (Anderson, 2001). The first process actively guides the attentional focus by highlighting particular features of current interest. The latter one describes selection due to perceptual saliency. In everyday tasks, both of them interact.

As selection and ignoring are two sides of the same medal, the nature of ignoring is crucial, as distracting information can easily be varied in experiments, and thus gives access to the act of selection itself. Even if early attempts assumed a passive ignoring, empirical evidence for an active process comes from the inhibition of return paradigm (Milliken and Tipper, 1998). A prolonged reaction time is observed if a location which has been in the focus shortly before is required to be attended to.

A general approach to the processing of distracting stimuli is provided by the negative priming paradigm. Negative priming is often considered the most direct approach to assess the selective aspect of attentional processing, as the ignored, distracting stimuli can be proven to be actively processed (Houghton and Tipper, 1994).

Selective attention has to permanently deal with distracting information. Most paradigms we will discuss in the following show two items in each trial. One is to be attended, called the target, while the other one, the distractor, is behaviorally irrelevant and has to be ignored. One such

selection trial primes the subject for the next trial. Therefore, a pair of two successive trials is labeled prime and probe respectively. Generally speaking, the repetition of a target stimulus in two successive trials leads to a faster response. This effect is called positive priming. In contrast the presentation of a prime distractor as a target in the probe trial may lead to a deteriorated performance compared to a target which has not been presented immediately before. The behavioral slowdown indicates that irrelevant information is not passively ignored, but actively processed, as no effect on reaction times in subsequent trials would have been found if the information about the distractor is not present.. Negative priming has been found in a wide variety of experimental contexts and is therefore thought to be a reproducible and general phenomenon, see section 2.3.

In the present thesis we will rely on the following definition: Negative priming is a slowdown in reaction time in an ignored repetition condition, where a former distractor has become relevant. As we associate the term negative priming with reaction time differences, we can not use it as a label for the ignored repetition condition but rather chose the condition labels according to the configuration of stimuli in a trial, see (Christie and Klein, 2001). The first letter in the sequence contains information about which part of the prime display is repeated in the probe display. A **D** represents the distractor, a **T** the target. The second letter indicates the role the particular object has in the probe display, see table 2.1. For example, the string DT refers to the condition in which the prime distractor (first letter D) is repeated in the probe trial as target (second letter T), hence it is a traditional negative priming trial. In case both objects are repeated there is a second pair of letters appended for the second object. Because a target and a distractor are shown in the prime and the probe display each, and target and distractor are never identical, seven relevant combinations of target-distractor relations are conceivable, see table 2.1.

	prime display		probe display		
	target	distractor	target	distractor	
TT	A	B	A	C	target($n+1$) = target(n)
DT	A	B	B	C	target($n+1$) = distractor(n)
TD	A	B	C	A	distractor($n+1$) = target(n)
DD	A	B	C	B	distractor($n+1$) = distractor(n)
DDTT	A	B	A	B	target and distractor are repeated
DTTD	A	B	B	A	target and distractor switch
CO	A	B	C	D	two new stimuli

Table 2.1: The seven possible priming conditions of a paradigm with one target and one distractor in each of the prime and probe display.

2.2 A Showcase Negative Priming Experiment

To get a first idea of the characteristics of negative priming, we will now discuss a rather straightforward study which can be taken as starting point for all our experiments. The study was part of our publication (Schrobsdorff et al., 2007a). We use a visual identity priming task where the target is selected by means of its color and then responded to according to its identity. The paradigm was introduced by Tipper (1985) and has been used in the Göttingen gerontology group for some years and has been optimized in several ways. Since negative priming tends to disappear with problem

complexity, identity priming with simple stimuli is an option that maximizes negative priming in this respect. For efficiency reasons we present the trials continuously, such that every trial primes the subject for the following trial.

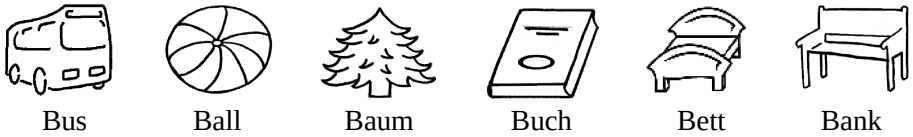


Figure 2.1: Objects used as target and distractor stimuli with their German labels.

Stimuli are six different objects, represented by hand-drawn pictograms, see figure 2.1, that are either shown in green or in red color. We use voice recording together with a sound level threshold to determine the reaction time for every trial. As the experiment is run in German, the names begin with a plosive and consist of a single syllable (*Bus*, *Ball*, *Baum*, *Buch*, *Bett*, *Bank*) for a sharp, thus easily detectable onset of the sound signal. Object presentation is balanced in the different priming conditions as well as their appearance as target and distractor. Priming conditions CO, DT, TT, DDTT and DTTD are a repeated measures factor.

An exemplary sequence of displays of seven trials is shown in figure 2.2. One stimulus display consists of two overlapping line drawings, a green target and a red distractor object. Stimuli appeared entirely in the focal area. The subject is instructed to name the target objects aloud and ignore the superimposed red objects. They were told to answer as quickly and as accurately as possible. Then, after a blank screen period and the presentation of a fixation cross, the next display is presented.

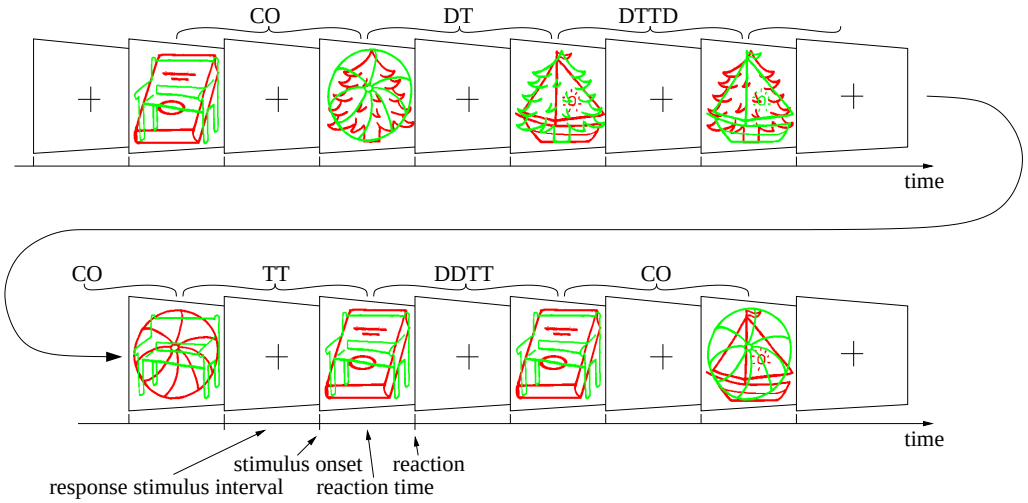


Figure 2.2: Example of a sequence of stimuli. Consecutive screens are shown. Either stimuli or a blank screen followed by a fixation cross is displayed. The meaning of the acronyms is explained in section 2.1. E.g. in the sequence of the second and third stimulus displays, the tree switches from red distractor to green target, but the other two items are unrelated, a DT condition, see table 2.1.

An experimental session starts with a test of color discriminability and memory span, followed by a familiarization with the stimulus objects. Of the entire 420 experimental trials in 10 blocks of 42 trials, 400 trials are analyzed (80 trials of each priming condition), while the first two start trials of every block are excluded from analysis. Each trial consists of the following displays: a fixation cross, centered on the screen for 500 ms, a display containing superimposed pictures shown up to the response, but not longer than 2 seconds and a blank screen for 1000 ms, see figure 2.2. An error is registered when subjects failed to give a correct, clear answer. Participants are 12 adults, 4 male and 8 female, mean age 23.6 years, SD = 4.6.

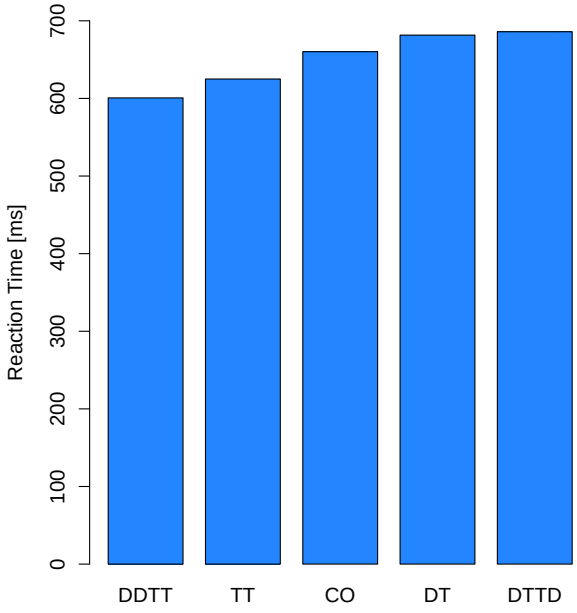


Figure 2.3: Reaction times for the five experimental conditions. Note that the positive priming effects on the left side are larger than the negative priming effects on the right hand side.

Mean reaction time of the different priming conditions, standard deviation, and the effect strength, i.e. the difference to CO trials, are shown in figure 2.3 reported in table A.1 in appendix A.1. Erroneous trials (2.4%) are excluded from analysis. Trials with response latencies less than 250 ms or more than two standard deviations above the individual mean of each priming condition are excluded as well (4.7%). DTTD trials produce the slowest responses, followed by DT and CO trials, whereas the responses to TT trials are faster than control and DDTT trials produce the fastest responses.

For statistical analysis, a one way analysis of variance (ANOVA) was used to explore the effects of CO, DT, DTTD, TT and DDTT. The alpha level for all analyses was set at 0.05. Greenhouse-Geisser corrected degrees of freedom are used as the data violated the assumption of sphericity. Reaction times depend significantly on the priming condition $F(1.45, 15.93) = 23.27$, MSE$= 1938.83$, $p < 0.001$. Planned comparisons show that reaction times in DT and DTTD tri-

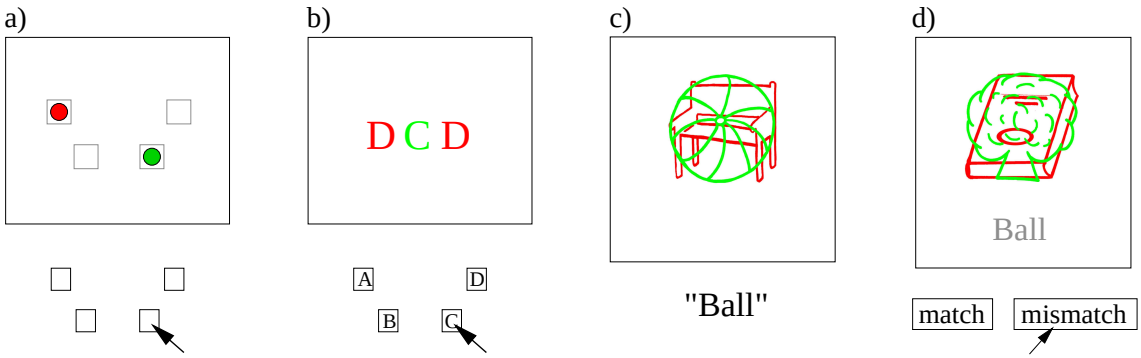


Figure 2.4: Four different paradigms for negative priming: a) location b) flanker task c) voicekey identification d) word-picture comparison task. In all examples green defines the target. The location priming paradigm reveals negative priming in the encoding of space. The flanker task implements a rather difficult stimulus response mapping, whereas vocalization in the voicekey paradigm is a very easy task. The word-picture comparison paradigm has the advantage of a disentanglement of target identity and response.

als were significantly increased, as compared to CO trials (CO vs. DT: $t(11) = -3.57$, $p < 0.01$; CO vs. DTTD: $t(11) = -3.37$, $p < 0.01$). As anticipated, the reaction time for trials in the attended repetition conditions TT and DDTT were significantly decreased (CO vs. TT: $t(11) = 3.11$, $p < 0.01$; CO vs. DDTT: $t(11) = 4.74$, $p < 0.001$). Directed comparison of attended repetition conditions reveal a difference of reaction time (TT vs. DDTT: $t(11) = 6.11$, $p < 0.001$), whereas the reaction time of ignored repetition conditions did not differ (DT vs. DTTD: $t(11) = -0.60$, $p = 0.558$).

The experiment shows how the repetition of stimuli can influence reaction times in a negative priming paradigm. A repetition of relevant stimuli leads to a prominent speedup of the trial processing, whereas a repetition of irrelevant stimuli as target causes a slowdown of the reaction. The results document the important fact: irrelevant stimuli are not filtered out by early perception mechanisms, but rather are subconsciously processed.

2.3 The Negative Priming Effect

Negative priming is subject to fervid discussions among cognitive theorists for several decades. The negative priming effect was discovered first in a Stroop task (Dalrymple-Alford and Budayr, 1966), subjects have to name the ink in which a color-word is written. A usually strong tendency to read the word has to be suppressed. Reaction times are delayed in trials where the color of the ink in the probe is identical to the word in the prime. The semantic meaning of the word serves as the distractor, because it has to be ignored in order to be able to correctly name the color of the ink. The results were replicated in similar settings by Neill (1977).

Negative priming is present in a wide variety of experimental contexts (for reviews see Fox, 1995; May et al., 1995; Tipper, 2001; Mayr and Buchner, 2007). For example, negative priming has been elicited using different stimuli such as line drawings (Tipper and Cranston, 1985), letters (Neill and Valdes, 1992; Neill et al., 1992), words (Grison and Strayer, 2001), auditory stimuli (Buchner and Steffens, 2001; Banks et al., 1995; Mayr and Buchner, 2006) and nonsense shapes (DeSchepper and Treisman, 1996). Negative priming has been obtained with manual (Neill and Valdes, 1992; Tipper et al., 1992) and verbal responses (Allport et al., 1985; Tipper and Cranston, 1985), as well as in situations where the mode of response changed between prime and probe

(Chiappe and MacLeod, 1995). Furthermore, negative priming has also been found in various tasks including naming (Tipper, 1985), same-different matching (DeSchepper and Treisman, 1996), Stroop-like tasks (Neill, 1977) and spatial localization (Milliken et al., 1994; Park and Kanwisher, 1994; May et al., 1995; Kabisch, 2003).

In spite of the obvious universality, the negative priming effect is sensitive to a variety of parameters. Most paradigms show their individual aspect of negative priming, but no global pattern of results exists (Fox, 1995). It has been shown that negative priming can depend on the length of the response stimulus interval (RSI) between prime and probe (Neill et al., 1992; Kabisch, 2003; Frings and Eder, 2009). But there are also studies reporting a constant negative priming effect for varied RSIs (Tipper et al., 1991; Hasher et al., 1991, 1996). Surprisingly for very short RSIs, a DT condition can produce a facilitatory (Lowe, 1985), or hampering effect (Frings and Wühr, 2007a). In the other extreme, an experiment revealed negative priming after a month using nonsense shapes which are very unlikely to be seen in other circumstances (DeSchepper and Treisman, 1996). In continuous presentation of trials, the proportion of preprime RSI and current RSI influences negative priming (Neill and Valdes, 1992; Mayr and Buchner, 2006) but not reliably (Hasher et al., 1996; Conway, 1999).

In the absence of distractors in the probe trial during a DT condition, negative priming vanishes or even reverses to facilitation (Lowe, 1985; Tipper and Cranston, 1985; Allport et al., 1985; Moore, 1994). A more salient prime distractor increases the magnitude of the negative priming effect in DT conditions (Grison and Strayer, 2001; Tipper, 2001). Negative priming is reduced or even reversed to facilitation when the emphasis is put on speed rather than accuracy (Neumann and Deschepper, 1992). Negative priming occasionally depends on age (Spieler and Balota, 1996; Verhaeghen and De Meersman, 1998; Gamboz et al., 2002) and sex (Bermeitinger et al., 2008). Increasing the perceptual load, e.g. by raising the number of distractors presented in a single trial, leads to less negative priming (Lavie et al., 2004). In other settings a higher number of prime distractors causes an increase of negative priming (Neumann and Deschepper, 1992; Fox, 1995). The inclusion of TT trials or single target trials in the presentation sequence enhances negative priming (Neill and Westberry, 1987; Titz et al., 2008).

A short presentation time of prime and probe stimuli attenuates negative priming (Gibbons and Rammsayer, 2004). Negative priming vanishes if the target is presented a bit earlier than the distractor in the prime trial. On the other hand, if the prime distractor is shown simultaneously with the prime target but blanked after a short time, negative priming is observed (Moore, 1994). If the prime display contains a single stimulus that is masked, subjects reporting awareness of the prime object show positive priming, while subjects not aware of the object show a negative priming effect (Wentura and Frings, 2005). In subliminally primed trials the presence of a distractor in the probe leads to negative, the absence of a probe distractor to a positive priming effect (Neill and Kahan, 1999).

It should therefore be noted that negative priming is a universal effect which is sensitive to a variety of factors. Hence, it is not surprising that, although there is a lively theoretical discussion, a consistent explanation of the entire negative priming phenomenon is still lacking. Over the years, various theories have been developed to explain negative priming through a variety of mechanisms. The next section will give a thorough description of the theoretical accounts.

2.4 Theories of Negative Priming

Because of the controversial nature of the negative priming effect, a variety of different theoretical accounts have been developed since its discovery. The negative priming phenomenon has a very

complex nature. None of the theoretical accounts is able to explain all aspects of the negative priming effect found in behavioral studies which have partly been outlined in section 2.3, they all have their strong points as well as their shortcomings. All theories assume different mechanisms to be responsible for negative priming. Their realization is also very diverse, from a purely descriptive formulation to detailed computational modeling. The most influential theories comprise the distractor inhibition account (Houghton and Tipper, 1994), episodic retrieval (Neill and Valdes, 1992) and response retrieval theory (Rothermund et al., 2005).

In the present section we will address each theoretical account by explaining the idea and point to experimental evidence as well as counterevidence. As some theories are derivates of others, we will highlight the differences. Many of the proposed concepts are settled in different areas of cognitive processing such that combinations are conceivable, which actually is one of the theoretical ideas, but some exclude each other.

2.4.1 Distractor Inhibition Theory

The first theoretical account in the context of negative priming was the inhibition hypothesis by Neill (1977) and Neill et al. (1990), before Neill and Valdes (1992) began to promote the episodic retrieval theory, see section 2.4.2. Meanwhile Tipper made himself the spokesperson of the distractor inhibition theory (Tipper, 1985; Tipper and Baylis, 1987; Tipper et al., 1988; Tipper and McLaren, 1990; Tipper et al., 1991; Houghton and Tipper, 1994, 1996; Tipper, 2001; Tipper et al., 2002) accompanied by some early questioning work (Tipper and Cranston, 1985).

The basic idea of distractor inhibition theory is that irrelevant stimuli representations are actively suppressed to support the selection of the relevant target stimulus. The inhibition is assumed to persist for some time. If perceptual input is no longer present, the persisting inhibition drives the distractor representation below a baseline activation. The negative priming effect directly results from the time the probe target representation activation needs to reach baseline from below. There are two complementary processes involved in the attentional selection process: a direct feedforward excitation of the representation of perceived items by the visual pathway and another one that inhibits all irrelevant information. The slowdown of the reaction in the probe trial can be seen as a direct indicator of the amount of activation in the prime display. Distractor inhibition assumes selection to operate on a semantic or postcategorial level (Houghton and Tipper, 1994). It therefore also explains findings that report negative priming in semantic priming tasks (Tipper and Driver, 1988).

From a modeler's perspective, the most important contribution to the domain of distractor inhibition comes from Houghton and Tipper (1994). A computational implementation of an artificial neural network qualitatively explains negative priming by an inhibitory rebound naturally emerging from the network connections between excitatory and inhibitory cells homeostatically balancing the state of a property unit. The initial version of the computational distractor inhibition model is very ambitious, as perception is split into the detection of single features, hardwiredly binding them into objects. The model has a very general connection scheme, to act in a variety of situations. Unfortunately, none of the further projects proposed by Houghton and Tipper (1994) has been realized since then. In order to investigate the time course of negative priming, the role of multiple distractors and different distractor salience, the model is later simplified by looking at only one isolated property unit for the target and one for the distractor, with connections only to their on and off cell (Houghton et al., 1996; Houghton and Tipper, 1998). The aim is to simplify the original model as much as possible while still observing the same dynamics. Regrettably, the generality of the first modeling approach is no longer present. On the one hand, simplifications of complex models are an adequate tool to understand the behavior of the entire system. On the

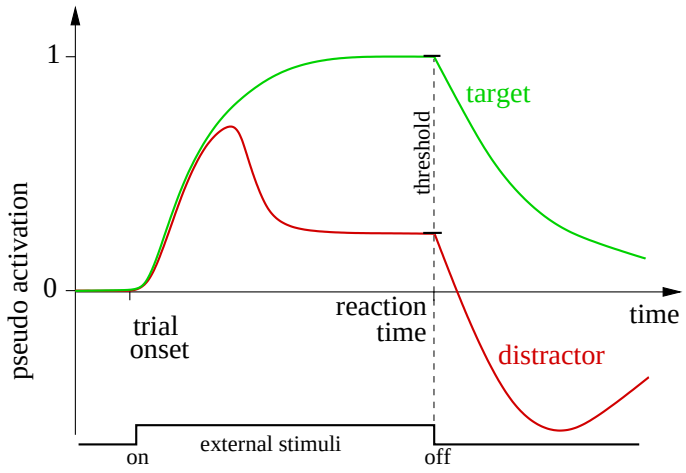


Figure 2.5: Schematic view of target and distractor representation during one trial. At stimulus onset both activations rise driven by the input. The two curves diverge due to the inhibition the distractor receives. As inhibition builds up, it balances the perceptual input to the distractor after some time. If a certain difference between target and distractor representation is reached, the target is assumed to be selected, and an action is taken. Then the input is switched off as the stimuli disappear. The target activation passively decays to zero, whereas the distractor activation is still subject to persisting inhibition, driving the distractor representation below baseline in the so-called inhibitory rebound, being responsible for the negative priming effect in the next trial. Figure adapted from (Houghton and Tipper, 1994).

other hand, the reintegration of the single units into the bigger network always brings along various nonlinear effects that are inherent to the model and can not be neglected when deriving system behavior from results of looking at isolated units. The reintegration does not take place, which might be an indicator for too high complexity of the distractor inhibition model to learn something from it.

A strong point of distractor inhibition theory comes from the study of varied distractor saliency. The negative priming effect increases with growing saliency of the distractor (Lavie and Fox, 2000; Grison and Strayer, 2001; Tipper et al., 2002). This effect can be very well explained in terms of the inhibition model, since a stronger distractor would require more inhibition, causing a stronger inhibitory rebound, and thus leading to a more prolonged reaction time.

Distractor inhibition theory can directly explain the impact of the depth of processing (Craik and Lockhart, 1972; Craik, 2002). Processing on a deep conceptual level produces a bigger negative priming effect. Distractor inhibition theory can explain the results, as deeper processed items have a stronger activation and thus need more inhibition if characterized as distractor. Therefore, more deeply processed stimuli produce larger negative priming.

The original distractor inhibition theory fails to explain the dependency of negative priming on the response stimulus interval. If the representation of a distractor object is inhibited, the impact of inhibition should be strongest immediately after the selection, because the inhibition is assumed to decay to zero with time. Although there is a general trend of negative priming to decay with increasing time between prime and probe (Neill and Valdes, 1992), no negative priming is observed in several studies when the RSI is very short or nonexistent (Lowe, 1985; Houghton et al.,

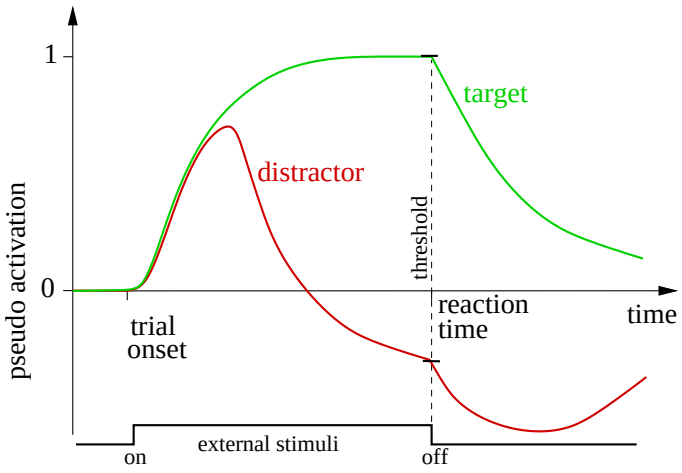


Figure 2.6: Alternative view of the distractor inhibition theory accounting for negative priming effects in the absence of a response stimulus interval. If the situation requires very strong inhibition, the activation of the distractor can drop below baseline already before the end of the trial. Sketch adapted from a talk by Christian Frings, May 30th 2007 in Göttingen, on (Frings and Wühr, 2007a).

1996). In the original model the equilibrium between perceptual input and inhibition is tuned such that the activation of the distractor stays positive. If then a new display is shown directly after the response, a facilitatory effect in the DT condition is expected. Unfortunately, already the study that brought negative priming to light of Dalrymple-Alford and Budayr (1966) shows negative priming without any delay between succeeding stimulus pairs, subjects held cards with several colored words which they processed in order. Therefore, Wentura and Rothermund (2003), Frings and Wentura (2006) and Frings and Wühr (2007a) proposed an extension of distractor inhibition by assuming that the amount of inhibition is proportional to task difficulty. In demanding paradigms like the Stroop task, inhibition may exceed excitatory input thus pushing the distractor activation below baseline even before a reaction, see figure 2.6.

Distractor inhibition is incompatible with target only probe displays. In the absence of a distractor priming constellations that usually produce negative priming effects can show facilitatory priming as reported by Moore (1994). A suitable extension of distractor inhibition theory concerns the notion of what is actually inhibited. Neill (1977) suggests that the semantic representations of the distractors themselves are inhibited, which matches with spreading inhibition through semantical networks (Quillian, 1966). Tipper and Cranston (1985) propose inhibition to act on the link between semantic representation and the response system. More explicitly, they assume a selection state of the response system in which the time-consuming resolving of the inhibition of the link between representation and response produces negative priming. In situations where no such selection is necessary, the response may still be facilitated because of the residual activation of the distractor representation. Unfortunately, the response inhibition account was not integrated in later papers.

Distractor inhibition theory is also challenged by the empirical finding of long-term negative priming effects (DeSchepper and Treisman, 1996; Grison et al., 2005). Tipper (2001) integrates these findings by emphasizing that different mechanisms might underlie the behaviorally similar effect in different settings. It is also stated that a retrieval of an episode (as postulated by the

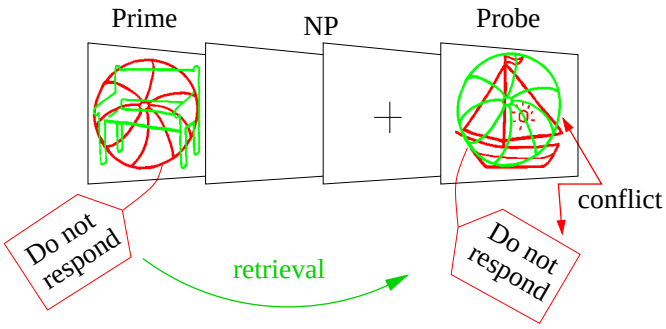


Figure 2.7: Episodic retrieval assumes figuratively a do-not-respond tag that is attached to the prime distractor. If a probe display contains matching information, the former episode is retrieved and with it the tag. Removing this tag in order to respond to the former distractor which has become target in DT trials takes time which is equivalent to the negative priming effect. Disadvantageous for the theoretical discussion, episodic retrieval theory was often reduced to the picture of the tag, which is only introduced as a metaphor in the original work.

episodic retrieval theory, described in the next section) might also retrieve the inhibitory status of the previously ignored distractor.

2.4.2 Episodic Retrieval Theory

In recent years the majority of negative priming studies have interpreted their results according to the episodic retrieval theory. It has originally been introduced by Neill and Valdes (1992). The theory builds on the instance theory of automatization by Logan (1988). If identical tasks have to be fulfilled over and over again, memories of past trials are more and more used to completing the current trial. Negative priming is assumed to be the result of conflicting information caused by automatic retrieval of the prime episode during probe processing. It is argued that the retrieval is triggered by the similarity of prime and probe episodes. Because the object information from the retrieved episode in a DT trial is inconsistent with the current role of the object as a target, retrieved and perceived information are in conflict. Resolving the conflict is time consuming and results in the slowdown of the reaction time. Some of the negative priming phenomena listed in section 2.3 can be more easily explained by episodic-retrieval than inhibition mechanisms, such as effects of prime-probe temporal discriminability (Neill and Valdes, 1992), prime-probe similarity (Fox, 1998) and long-term negative priming from single-trial presentations (DeSchepper and Treisman, 1996).

According to later extensions by Neill (1997), the main determinants of the strength of retrieval are the recency of the memory trace and the strength of the memory representation of the trial. It is assumed that more recent memory traces are more likely to be retrieved than older ones. This assumption allows for the interpretation of experimental settings with many repetitions of highly similar episodes. Recency as a relevant factor receives empirical support from studies that show a negative correlation between response stimulus interval and negative priming effect (Neill and Valdes, 1992). Neill et al. (1992) reports an influence of the interval preceding prime onset on the negative priming effect, which challenges the inhibition based accounts, but is easily explained in

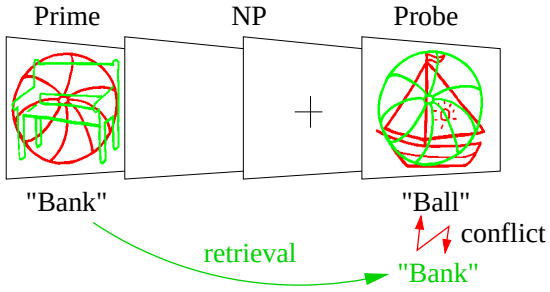


Figure 2.8: Response retrieval assumes that just the response given in the prime trial is retrieved triggered by perceptual similarities of the two displays. If the appropriate response is repeated, a positive priming effect is expected, whereas a response switch results in a negative priming effect. Note that in the voicekey paradigm shown in the current picture, TT trials are always associated with a response repetition, and DT trials always require a response switch. Only more complex paradigms allow for a disentanglement of priming condition and response relation.

terms of episodic retrieval. When the pre-prime response stimulus interval is larger than the one of the current probe, the prime-probe pair is more easily separable from the sequence of stimuli. Hence retrieval of the prime episode in the probe is enhanced. The memory trace produced by the prime episode is more elaborated with deeper processing of the stimuli. Therefore, the depth of processing of the stimuli can influence the strength of retrieval of an episode positively, the data by Yee et al. (2000) can thus be explained by episodic retrieval.

Strong support for episodic retrieval comes from a study by Stolz and Neely (2001). They found an increased negative priming effect when the contextual similarity between prime and probe situations is increased. Prime/probe episodes were more similar in terms of visual characteristics and the required response. On the contrary, a facilitated response at very short response stimulus intervals (Lowe, 1985) is also difficult to explain in terms of the episodic retrieval framework. Another weakness of episodic retrieval is semantic negative priming. The absence of perceptual similarity should eliminate retrieval, and therefore no effects are to be expected.

2.4.3 Response Retrieval Theory

Marczinski et al. (2003) investigate different priming effects for young and old subjects in a two-alternative forced-choice task, which can not be explained by episodic retrieval theory. Therefore she introduces the notions of specific and unspecific encoding or retrieval respectively. The reaction time difference between trials with a response switch and trials with repeated responses are called response repetition effect. The idea was borrowed by Rothermund et al. (2005), who points to the inherent entanglement of priming condition and response relation in most negative priming paradigms. Usually DT trials are accompanied by a response switch, whereas TT trials require the same response. The response retrieval approach postulates that every reaction time difference in priming paradigms is due to the retrieval of a past response depending on perceptual similarity between the two displays.

A letter-matching task initially developed by Neill et al. (1990) is adapted in order to test the hypothesis (Rothermund et al., 2005). Strings of five letters are presented to the subjects and they

have to answer yes or no by an appropriate button press depending on whether the second and the fourth letter are identical. The remaining three letters are identical and function as distractors. The advantage of the paradigm is that the repetition or non-repetition of the response can be varied independently from priming condition. It is possible to require both a yes or a no answer in every priming condition because the answer does not depend on the identity of the target, but on its relation to the second target. In their series of experiments Rothermund et al. (2005) could provide solid justification of response retrieval.

The letter-matching paradigm is problematic in the view of episodic retrieval theory. The similarity of the prime and probe display is dependent on whether both targets match or not. In fact the highest similarity is achieved in a TT condition which requires to answer yes in both prime and probe trial. If the answer no is correct, only one of the targets is repeated, possibly shortening reaction times of yes answers. However, two other experiments described by Rothermund et al. (2005) implemented task-switching paradigms where the described problem does not occur.

2.4.4 Feature Mismatch Theory

In a series of four experiments Park and Kanwisher (1994) developed their feature mismatch theory. If, at a certain location, an object changes identity, a mismatch of features is detected which causes a slowdown of processing (Milliken et al., 1994). They presented two letters, a particular one coding for the target, a specific other one for the distractor, on two out of four possible locations on a screen. Subjects indicated via button press the target location. In order to discriminate between distractor inhibition and feature mismatch, they varied the feature that identifies the target between prime and probe. In the prime trial, the letter **x** had to be responded to, and **o** had to be ignored. In the probe trial, subjects should indicate the position of the **o** ignoring the **x**. According to the distractor inhibition account, TT trials should lead to a positive effect, as the probe target location is already positively primed. But it turned out to produce a longer reaction time on TT trials, and a faster response in DT trials, contradicting distractor inhibition. Unfortunately the aspect of task switching effects (see chapter 6 for a discussion of those) was not addressed in the paper.

Feature mismatch theory received quite some attention, and as it is often listed among explanations for negative priming. Nonetheless, there has been no further development of the theory nor experiments in strong support of feature mismatch.

2.4.5 Temporal Discrimination Theory

Milliken et al. (1998) added temporal discrimination theory to the set of explanations for negative priming. Temporal discrimination assumes a classification of stimuli as old, where a response can be retrieved from memory, or new, where a response has to be generated from scratch. The classification takes time depending on the similarity between the current stimulus and a memory trace. The dependency is not monotonic: the classification as new is fast when prime and probe stimulus are very dissimilar. The classification as old is fast when the displays are identical. But intermediate similarities such as in DT trials where the prime distractor is repeated but not in its former role but as target, the decision whether the display is old or new takes more time. Experimental replications of the effect on which the initial paper by Milliken et al. (1998) is based, i.e. negative priming without selection in the prime trial, do not discuss the theory itself (Neill and Kahan, 1999; Healy and Burt, 2003).

Even if the temporal discrimination account is often cited, it is mostly cut down to the description above. It is rarely theoretically addressed. Only Frings and Wentura (2005) and Frings and Wühr (2007b) develop the model further. They address the question whether the awareness of the

prime distractor, which was masked in most of the experiments in favor of temporal discrimination, plays a role for negative priming or positive priming to occur in single prime distractor DT trials. If the masked prime distractor was supraliminal, a positive effect is expected as participants were not instructed to ignore items eventually contained as a flicker in the mask.

As both temporal discrimination and episodic retrieval theory rely on the question of if the response can be retrieved from memory or if it has to be computed directly, it is hard to delineate them from one another. The difference usually pointed out in the literature is the presence or absence of a *do not respond*-tag, which was introduced metaphorically in episodic retrieval theory, see section 2.4.2, but then falsely made a central part by commentators.

On a closer look, temporal discrimination tacitly involves two different processes. The direct computation of a response is completely different from a retrieval of the answer from memory. No clue exists that these processes take an equal amount of time. Nevertheless, most discussions of temporal discrimination seem to have been misdirected by an illustration Milliken et al. (1998) draws to explain his theory. He equates an experimental trial with the time a PhD student takes to finish his thesis depending on whether he found a project suitable to his interest. If he chose a not matching subject, the degree of match with his interest will determine how long it takes him to realize and switch the project. No match causes an early switch, and he barely loses time in comparison to a perfect match with his interests. But if there is some overlap, he will proceed for some time with the project, until he finally realizes he has chosen the wrong project, then change and again take a full period of PhD research. The analogy raises the following problems: the student always has to write a thesis. It is only one and the same process involved, but response retrieval or direct response computation are very different; and secondly, the old-new classification is always performed, whereas the student would have no switching time if the project is suitable. The acknowledgment of two independent processes renders most of the criticism on temporal discrimination less striking, as most critical arguments rely on the equality of processing time after the old/new-classification.

The weakest point of temporal discrimination theory is the assumption of a serial processing of classification and retrieval or direct generation of a response. Most processes in the brain work simultaneously, and therefore a parallel computation of the old/new signal together with a directly computed answer and the retrieval of past episodes is rather probable. Dropping the assumption of seriality would certainly bury the temporal discrimination account which mostly appeals due to its simplicity.

2.4.6 Dual Mechanism Theory

Obviously there is evidence in support of each of the theoretical accounts and no approach is clearly favored over the others. As distractor inhibition and episodic retrieval are by far dominant in the domain of explaining the negative priming effect, integrating accounts are called dual mechanism theories.

The original dual-mechanism account of negative priming by May et al. (1995) proposes that inhibition as well as memory retrieval can be the source of negative priming and the experimental context specifies which of the two mechanisms is expected to operate. Another attempt to integrate inhibition and retrieval perspectives was made by Tipper (2001). He argued that it is important to note that distractor inhibition and episodic retrieval theories are not mutually exclusive, and both inhibitory and retrieval processes could be involved in the emergence of negative priming. Although retrieval processes can be responsible for producing negative priming effects, inhibitory processes are still required in selecting information for goal directed behavior. Supposedly, in most tasks negative priming will be caused by a mixture of contributions from persisting inhibition

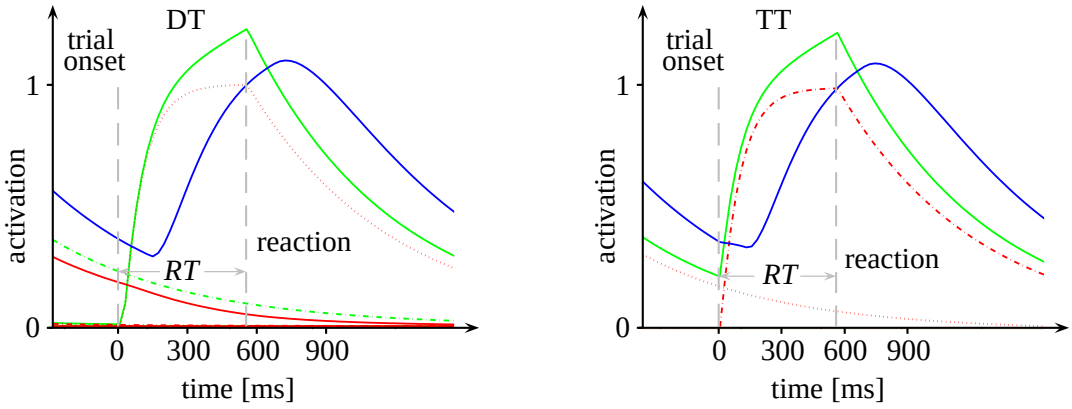


Figure 2.9: Time series of the activation variables during one simulated trial processed by the computational implementation of the ISAM, see chapter 3. Only the representations of objects of the current probe trial or the preceding prime trial have nonzero activation. The threshold level (blue) is driven by the total activation. If only one superthreshold activation is left, a decision is made. The DT condition is shown on the left. The forced decay of activity of the former distractor variable is visible in a subtle kink of the red solid curve. Right is shown the TT condition. The activation of the former target green solid line has not yet decayed to zero when the variable is again subject to input, shortening the reaction time significantly.

and interference from retrieval. Because these processes may sometimes oppose each other, it is difficult to distinguish them by means of behavioral measures like reaction times and error rates (Gibbons, 2006). However, depending on contextual conditions and other experimental factors, the contributions of inhibitory and retrieval processes might vary considerably (Kane et al., 1997; Tipper, 2001). Gamboz et al. (2002) revealed in a meta analysis of age-related negative priming experiments that there is no significant evidence for a paradigm to produce patterns of results favoring either inhibition or retrieval theories.

These integrative concepts seem to provide the better framework to explain the various facets of negative priming, as they raise the level of complexity and the number of possibly contributing parameters. However, a more explicit formulation than the one given by the mentioned accounts would be desirable in many cases. For example, the determination whether encoding or retrieval processes play the main role or to what extent they contribute to the effect is widely left to interpretation and thus largely depends on the paradigm. A comprehensive model that allows for exact, ideally quantitative predictions for a wide variety of paradigms, tasks and other experimental parameters would be desirable.

2.4.7 Global Threshold Theory

In the imago-semantic action model (ISAM), Kabisch (2003) developed the hypothesis of a threshold variable deciding which items from perceptual input can be responded to. The threshold is assumed to adapt according to the current average activation of representations of objects. Additionally, the ISAM proposes a forced decay of activation if residual activity is partly overwritten by perceptual input of a new stimulus. These two mechanisms together can account for positive as well as negative priming. The ISAM differs from the distractor inhibition theory explained

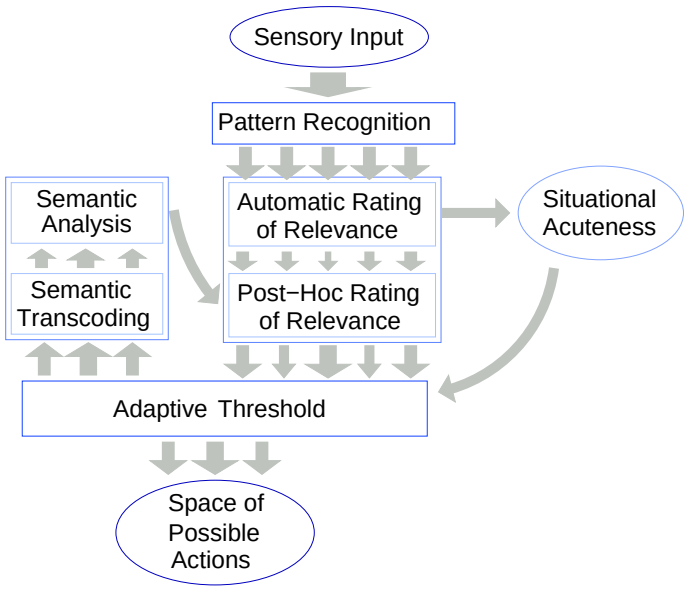


Figure 2.10: Schematic view of the ISAM. The right blue box represents the preliminary rating of relevance for the perceived objects. The direct relevance of certain objects drives the factor of situational acuteness that controls speed and accuracy of the computation. The level of global activity affects a threshold which truncates the list of perceived objects into a space of possible actions. A decision is achieved if in the set of applicable actions only a single one is left. The blue box on the left contains a semantic analysis of the perceived objects. It is able to project back to the posterior rating of relevance in a so-called *semantic feedback loop*. The interaction among the components leads to a decision about a reaction by a dynamical process.

section 2.4.1, by postulating only facilitative input and passive decay in the absence of input.

The assumed processing is sketched in figure 2.9 which actually shows data from our computational simulations of the ISAM described in chapter 3. Perceptual input drives the corresponding representation variables to a high value. The target receives additional support, which drives target and distractor activation apart. When the threshold surpasses the second largest activation, the one of the distractor, the system has selected the target and is able to respond. In DT trials the decay of the former distractor is faster than for an unrepeated object. Therefore the overall activation is lower, resulting in a slower rise of the threshold, which then delays the reaction. In TT trials, target activation has not yet reached zero. Thus more activation is present, which speeds up threshold and subsequently the reaction.

The ISAM forms a comprehensive concept of action selection. The presented objects are assumed to undergo a pre-attentive processing and perception stage, resulting in an abstract cognitive representation of the objects. Formally, the decision between target and distractor is determined by the task instruction, which is made accessible to the model via the semantic feedback loop (left in figure 2.10). In contrast to the abstracted early visual processes, the decision is guided by attention and a conscious application of the task instruction.

The stimuli are assumed to be initially processed automatically according to a relevance rating based on low-level features such as motion or color. The stimuli are sorted hierarchically by their (automatically assigned) relevancy. The relation between stimulus and associated action incentive

is that strong that we will speak of both synonymously. The subject has to decide which one of the current action incentives to follow.

Attention is modeled as trimming the perceived stimulus set by an adaptive threshold which is driven by the overall activation. If more than one option for actions exists, the threshold adapts such that only one option remains. During the adaptation of the threshold, representation variables can still be subject to input, be it top-down or bottom-up.

According to the dual-code hypothesis of Krause et al. (1997), assigning modified relevancy values to the respective objects happens in a semantic space where stimuli are processed jointly with the implied actions. The relative relevance of stimuli can be affected in a posterior rating in the semantic space. The activation corresponding to a target is amplified such that even if low-level perceptual features result in a higher saliency of the distractor, the target representation becomes significantly stronger than that of any distractor.

Kabisch (2003) even found a reversal of priming effects compared to the expected ordering shown in section 2.1 for certain response stimulus intervals. He argued that activation is subject to diffusion within its layer on a faster timescale than the one of the passive decay, which means that global activation is more persistent than specific activation of a certain object variable.

The dependence of negative priming effects on distractor saliency can easily be explained by the ISAM, as we will point out in section 3.3.3. If the distractor becomes more salient, the target and distractor activations split later in the trial processing, delaying a decision substantially.

2.5 Summary

Mechanisms of selective attention are made accessible in the framework of negative priming, a slowed reaction to previously ignored stimuli. In standard negative priming paradigms, an experimental trial consists of a target which requires a response and a distractor object that is irrelevant to the task. Experimental conditions are classified by the stimulus relations between two subsequent trials, called prime and probe, respectively. Whenever a stimulus constellation occurs where a prime distractor becomes target in the probe, the DT condition, a negative priming effect is expected.

The variety of paradigms that showed a negative priming effect revealed a robust nature of the effect. But the overall pattern of results also points to a strong dependency on various parameters, often in an inconsistent manner. Therefore a large body of theoretical accounts have evolved, each based on a certain experimental setting. Due to the contradictory effects, a comprehensive theory is improbable.

3 Imago Semantic Action Model

The first mainstay in our quest to reveal the time course of negative priming is to implement a computational model that describes negative priming by the means of a dynamic system. The evolution of the model variables provides detailed temporal information. Characteristic differences between trials of different experimental conditions can be localized in time and thus give insight in the stage of processing where negative priming is produced.

We will present the computational implementation of global threshold theory, see section 2.4.7. The original work on the imago semantic action model by Kabisch (2003) describes the interaction of concrete representation activations which can be straightforwardly translated into model variables. We only had to specify the actual dynamics. Similar concretizations are necessary whenever a quantification is sought and only a qualitative description is available.

Our simulations nicely reproduced experimentally observed reaction times. Without further fitting several dependencies of negative priming are matched by our model. We thus show that the adaptive threshold mechanism proposed by global threshold theory is sufficient to explain both positive and negative priming effects. Additionally, the implementation provides testable predictions with respect to hitherto untested stimulus combinations, e.g. single object trials, see section 3.3.4.

We will start the following chapter with considerations on the transient behavior of an artificial neural network to the onset and offset of input, section 3.1. The resulting exponential fixpoint dynamics is then used to implement the ISAM. Implementation and resulting dynamics are described in section 3.2. Section 3.3 shows how the explanatory power of the ISAM in several aspects, i.e. artificial reaction times, the sensitivity of negative priming to the length of the response stimulus interval, varying distractor saliency and the presentation of single target trials. The results are discussed in section 3.4. The present chapter is based on (Schrobsdorff et al., 2007b).

3.1 Deriving Simple Activation Dynamics

The simulation of the ISAM, as it will be discussed in the following, acts on a rather abstract level. It contains a single variable for the semantic representation of an object triggered by visual input. For the abstraction we make several assumptions. First, we consider representation of a concept as an increase of firing rate for a specific, densely coupled assembly embedded in a larger network. For simplicity we model the sensory presence of an object or a certain feature as a constant input to the corresponding cluster. The of input either present or absent. The variables of the ISAM subsume the firing rate of such an assembly as it is driven by input. In order to determine a realistic dynamics of this activation of the corresponding concept, we consider an isolated cluster as an all-to-all coupled network of integrate and fire neurons. We then average the firing rate of the network over many presentations and analyze the shape of rise and decay.

3.1.1 Networks of Integrate-and-Fire Neurons

The membrane potential h_i of neuron $i = 1 \ldots N$ is driven by external input $I_i(t)$ and recurrent connections which deliver spikes of adjacent neurons modulated by a synaptic strength $w_{i,j}$.

$$\dot{h}_i = I_i(t) + \sum_{j=1}^{N} w_{i,j}\delta(t - t_{sp}^j) \tag{3.1}$$

If h_i reaches the firing threshold $\theta = 1$, it delivers a spike to its postsynaptic neurons and is reset by the threshold value $h_i^{\text{post-spike}} = h_i^{\text{pre-spike}} - \theta$.

The external input $I_i(t)$ is drawn independently in each time step from a Gaussian distribution, with a mean chosen such that a single neuron receives on average the same input equal to the difference of threshold θ and resting potential h^0. On average and without the recurrent coupling a neuron would fire once during presentation of the stimulus. The firing rate of the network was determined by summing up all spikes occurring in one timestep.

We simulated a network of $N = 1000$ neurons. A stimulus was shown for 50 timesteps, and the interstimulus interval was also 50 timesteps long. The total output of a neuron was fixed to $\alpha = 0.87$. The stochasticity of the input and the sensitivity of the network for fluctuations result in a rather random single trial firing. But on average a coherent behavior becomes visible. For the results shown in figure 3.1 we averaged 10.000 trials in order to obtain a good estimation of the firing rate over time.

3.1.2 Network Response to Input Onset and Offset

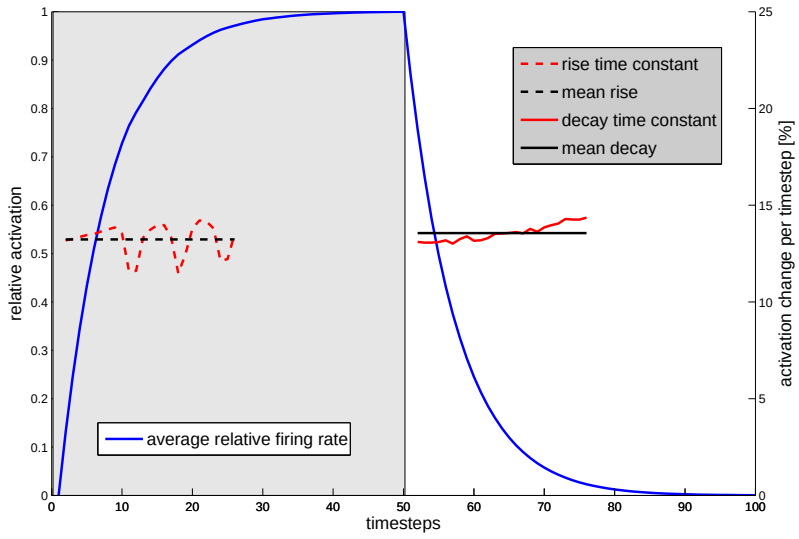


Figure 3.1: Average firing rate of the network during input (gray shaded region) and no input. The fraction of two subsequent values is shown in red. Black lines, the averages of the respective fractions, are indicating that they remain rather constant over time.

In order to derive a computationally simple dynamics for the representation variables of the ISAM, we tried to fit the time course of rise and decay of the firing rate of the network. A good candidate seems to be an exponential fixpoint dynamics, i.e. a variable approaches its current input value by a fixed fraction of the distance in every timestep. This fraction is called time constant of the variable.

In figure 3.1 we show the averaged firing rate and plot the percentage of change from one timestep to the succeeding one in reference to the actual fixpoint, i.e. maximum firing rate or zero. The observed time constants are only marginally constant, but sufficient for a justification of the simplified dynamics we will use for the implementation of the ISAM.

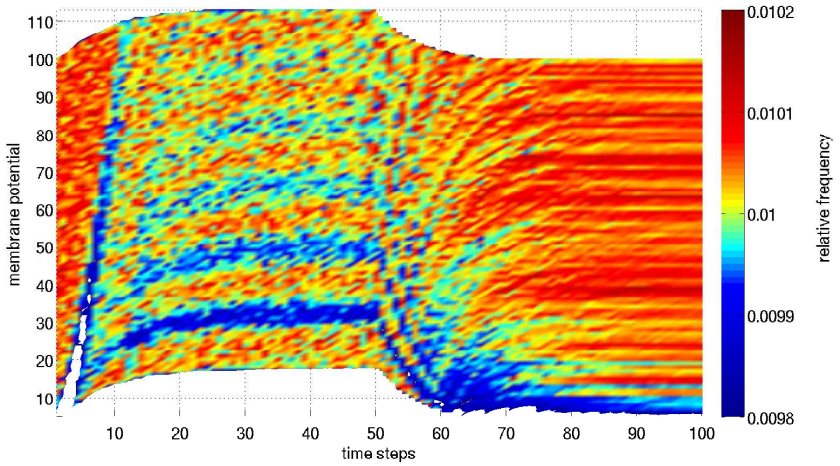


Figure 3.2: Distribution of membrane potentials averaged over 10.000 trials. Note that the potentials are mostly equally distributed, as the colormap only covers values from 0.0098 to 0.0115. Nevertheless, the fine grained plot reveals the processes generating the firing rates analyzed in figure 3.1: Initially all neurons are pushed towards higher membrane potential by the input, leaving a relative gap that is propagated upwards. Then, bands of neurons with a certain membrane potential form as the recurrent input builds up. Finally, the system relaxes and the less regular spikes rebuild a more equally distributed picture until no further spikes are generated.

Besides the result of a dynamics to use in our implementation of the ISAM, the periodicity of the time constant, even after severe averaging, points to interesting behavior of the system. Therefore, as a short excursion, we will spend some time on it. Figure 3.2 shows the distribution of membrane potentials averaged over 10.000 trials as in figure 3.1. During input, all neurons are shifted in their membrane potential such that small potentials become very improbable, to the benefit of superthreshold potentials. Most potentials within the band that is shifted upwards during stimulus presentation have a probability of 0.0098 and 0.0115, which is near an equal distribution. But there is some inner structure that survives the averaging process, which is revealed when having a closer look. In the beginning, all units receive only external input. They are shifted upwards, leaving a relative gap which then propagates through the entire range of potentials. Neurons that spiked are not reset to zero but reset by 1 and still receive recurrent as well as external input, which results in virtually no neurons having membrane potentials between zero and 0.15. As recurrent input tends to settle at a certain value, there is a trend of jumping into the band between 0.18 and 0.28 after spiking. This band is now shifted upwards by the same amount of activation, being smeared out during travel to the threshold. In every timestep a neuron jumps from one band to the next one. After offset of input there is only remaining recurrent excitation which decays rather fast. As the number of spikes decreases smoothly, the bands are washed out. In the beginning of the decay, there are still jump bands visible, but from time step 60 there is just the small trend upwards until settlement in the absence of input.

The approximate exponential dynamics of the firing rate in the network is attributed to the successive stimulation of neurons to spike. As more spikes are generated, the more neurons are superthreshold. The only deviations from the exponential dynamics are driven by the external input, which generates a constant baseline activity and the preference for a certain individual input which results in the band structure.

3.1.3 Exponential Fixpoint Dynamics

The dynamics can thus be formulated by a set of coupled nonlinear Langevin equations (Risken, 1996) with the basic form

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = h(\boldsymbol{x},t) + g(\boldsymbol{x},t)\Gamma(t) \tag{3.2}$$

where $\boldsymbol{x}$ is the overall state of the system, t is time, h is a function that describes drift forces that depend on the actual state and time and $\Gamma(t)$ is a Gaussian diffusion term with zero mean $\langle t\rangle\Gamma(t) = 0$ and δ correlation function $\langle t\rangle\Gamma(t)\Gamma(t') = 2\delta(t-t')$.

We will drop any noise because of our intention to describe the imago semantic action model as concretely as possible. The theory does not make any statements about noise influences to negative priming. Therefore we expect the noise component to just shadow the mechanisms in question. Another problem is the influence of the subject's age on priming effects. In recent work on modeling the influence of age noise was found to play an important role, see e.g. (Li, 2005). In future work we plan to model aging questions by taking up the noise component and relating its level to age.

The first part h of equation (3.2) will contain all driving forces of the system, external input and internal activation exchange via couplings. This results in an exponential fixpoint dynamics. The difference between system state and a given fixpoint multiplied by a certain time constant determines the change of the system.

3.2 Implementation of the ISAM

The simulations essentially follow the paradigm that we introduced in section 2.2. For an overview of the course of a simulated time step see figure 3.3. We consider a set of n stimuli each of which is represented by the activation of a detector function which models a cell assembly that represents the stimulus. In addition to the stimulus-driven activation, the detector is subject to intrinsic dynamics.

3.2.1 Representation Variables

Since the stimuli can occur either as a target or as a distractor, each is coded by two variables. One, x_i^τ indicates the presence of a stimulus $i \in \{1,\ldots,n\}$ as the target object, while x_i^δ codes the activity of the distracting stimulus. x_i^τ and x_i^δ are assumed to represent feature combinations which are considered to precede an object representation, see e.g. (Schrobsdorff et al., 2007a) how such feature combinations can be combined with an object representation. The system is modeled using a set of differential equations which determine the time course of the cluster activations and of the common threshold variable. Stimulus input is presented to the model in the form of jumps in the fixed points of the activation variables, i.e. they are modeled as delta-pulses to be integrated by the internal units (3.5), (3.6). The integration in the neural units is modulated by a temporal convolution with the time constants α and β in the following way: the activity x_i^ν, where $\nu \in \{\tau,\delta\}$

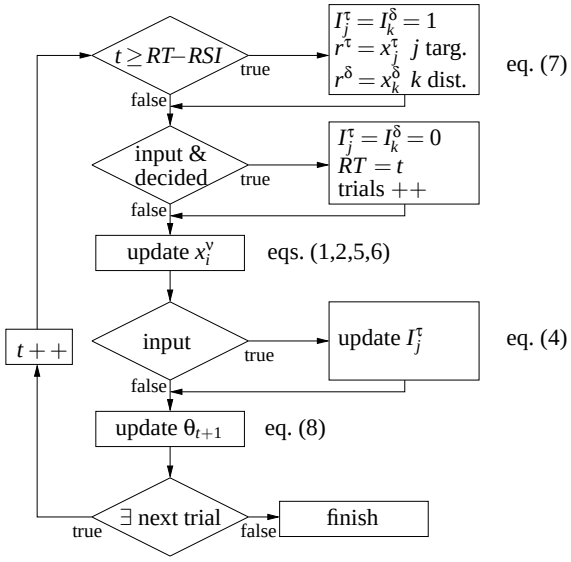


Figure 3.3: Schematic view of the loop structure of ISAM. Discrete time is denoted by t. Reaction time and response stimulus interval in the discrete time units are labeled by *RT* and *RSI* respectively. It is assumed that stimulus index j is the present target and k indicates the distractor of the current trial.

denotes target or distractor respectively, rises exponentially with time constant α to a certain level I_i^ν, which is determined by the input.

$$\frac{1}{\alpha}\frac{dx_i^\nu}{dt} = I_i^\nu - x_i^\nu \quad \text{if } x_i^\nu < I_i^\nu \tag{3.3}$$

If the input is switched off the population activity exponentially decays towards zero. If the present input is lower than the population activity it tends towards this input, equation (3.4). In both cases the decay constant is β.

$$\frac{1}{\beta}\frac{dx_i^\nu}{dt} = I_i^\nu - x_i^\nu \quad \text{if } x_i^\nu > I_i^\nu \tag{3.4}$$

3.2.2 Visual Input

We assume an abstract recognition mechanism that activates stimulus-specific units. Inputs are represented by an activity level of one unity in the presence of an input or otherwise by zero activity (3.5), (3.6). The semantic feedback loop amplifies the activity of the unit representing the target I_i^τ linearly with feedback strength ξ, see equation(3.6).

$$I_i^\delta = \begin{cases} 1 & \text{during presentation of object } i \text{ as distractor} \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

$$I_i^\tau = \begin{cases} 1+\xi\Delta t & \text{during presentation of object } i \text{ as target} \\ & \Delta t \text{ denotes the elapsed time while stimulus onset.} \\ 0 & \text{otherwise} \end{cases} \quad (3.6)$$

3.2.3 Interference of Semantically Identical Objects

As x_i^τ and x_i^δ belong to the same object i, they interfere. This interference is given by a negative interaction. One and the same object cannot be target and distractor at the same time. Therefore, these variables cause conflicts with the strength of the symmetric negative interference ζ during simultaneous activation. Assuming the absence of input, equation 3.4 becomes

$$\frac{1}{\beta}\frac{dx_i^\tau}{dt} = -x_i^\tau - \zeta x_i^\delta x_i^\tau \quad (3.7)$$

for the target and vice versa for the distractor activation

$$\frac{1}{\beta}\frac{dx_i^\delta}{dt} = -x_i^\delta - \zeta x_i^\tau x_i^\delta. \quad (3.8)$$

3.2.4 Adaptivity of the Threshold

The crucial point of the model is the adaptation of the threshold. The threshold θ (3.9) is driven by the average activity $\bar{x}$ (3.10). The parameter γ denotes an adaptation constant and Δ is a delay time. This is a standard version of a winner-takes-all mechanism.

$$\frac{1}{\gamma}\frac{d\theta}{dt} = \bar{x}(t-\Delta) - \theta \quad (3.9)$$

The average activity depends on the object representations x_i and also on two memory variables: r^τ and r^δ, which are residual activities from previous activations in the present target and distractor, respectively.

$$\bar{x} = \frac{1}{2}\left(r^\tau + r^\delta + \sum_{i=1}^{n}(x_i^\tau + x_i^\delta)\right) \quad (3.10)$$

At the moment of input onset, r^τ and r^δ are set to the activity level of the appearing target and distractor. If the new display contains stimulus j as target and k as a distractor, then we set the following at the point of stimulus onset:

$$r^\tau = x_j^\tau, \quad r^\delta = x_k^\delta \quad (3.11)$$

The variables r^τ and r^δ form a separate temporary representation of the current objects, which undergoes a different time course after initialization.

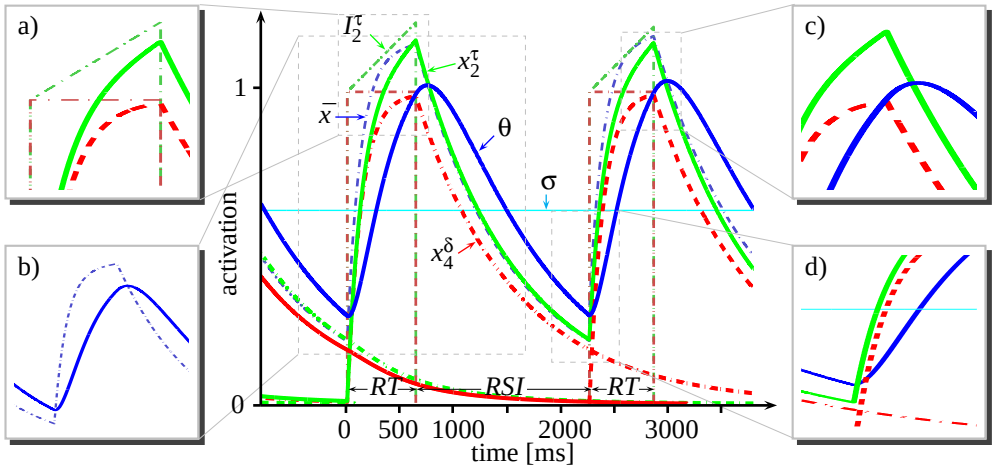


Figure 3.4: Activity vs. time. All traces relevant to the model are shown. Distractor x_i^δ activation is red and target x_i^τ activation green, the threshold level θ is blue. Additionally, the average activity $\bar{x}$ is plotted in dash-dotted blue, the absolute sensitivity level σ in cyan. The inputs together with the semantic amplification are shown as a thin dash-dotted pattern. Their color corresponds to their role as target or distractor input. Insets: **a)** The activation curve of target and distractor approach their input exponentially. The input of the distractor is fixed to one. The input of the target is linearly amplified by the semantic feedback loop. **b)** The exponential approach of θ towards the global average $\bar{x}$ is delayed by a certain time interval. **c)** At the moment when θ crosses the distractor activity only the target activation is above θ, and a decision can be made. **d)** During the fast rise of the activation of the new stimuli there is a short time interval where also only one activation surpasses θ. Due to the global sensitivity level σ, no decision is provoked.

3.2.5 Response Generation

The decision about the target object is performed globally. We test whether exactly one variable is activated above threshold level θ. This test is additionally conditioned due to the fact that the threshold is higher than a fixed sensitivity level σ. We thus avoid decisions without a significant activation, e.g. when the adaptive threshold θ is near zero. All variables that form the model are presented in a joint plot in figure 3.4, for details cf. the insets. Most activity variables x (green for x^τ, red for x^δ, in different textures for the different objects i) are effectively zero. Only the variables that are or had just been subject to input have a significant activity. The threshold θ (blue) is oriented along the average activity level $\bar{x}$ (dash-dotted blue).

Parameter values are chosen following several principles. Freely scalable parameters such as distractor input I_i^δ, are chosen such that they are kept as simple as possible. Certain bounds have to be respected, e.g. the sign is usually predetermined, but also the order of dependant variables is fixed. Most desirably parameters can be derived from experimental data. At the abstract level of the ISAM, we only can rely on behavioral experiments. Thus we decided to fit the model behavior to the priming effects of our negative priming study presented in section 2.2. The fitting is done manually.

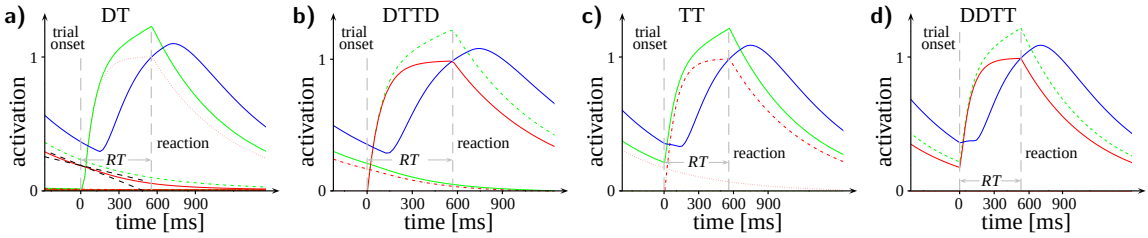


Figure 3.5: Time evolutions of the activation variables during one simulated trial. Only the activation of objects presented in the current trial or the preceding prime trial are visible, the other eight activation variables are virtually zero. **a)** DT condition. The forced decay of activity of the former distractor variable x_i^δ, where x_i^τ is the target in the second trial, is visible in a subtle kink that is pointed up by two black dashed tangential lines. These tangentials are not drawn in subplot b) for comparison. **b)** DTTD condition. Kinks are present in both the former target and distractor, as they are influenced by the rising activation of the new target and distractor. **c)** TT condition. The activation of the target, which is the same object as the previous target, rises from a certain level above baseline, as it has not yet decayed to zero when the new input is switched on. The activation approaches the input from this step. This shortens the reaction time significantly. **d)** DDTT condition. Here, both stimuli stay the same and approach their input from a higher level, shortening the reaction time even further.

3.3 Computational Results

3.3.1 Comparison to the Experimental Data

Figure 3.5 shows the time series of a trial in the simulated model. The four subfigures refer to the four main cases (DT, DTTD, TT, DDTT) described in section 2.1. Since the scale of time is a degree of freedom in the simulation, we performed a temporal gauge to match the average reaction time in the control trials.

In the DT-trial in figure 3.5a, the activation of target and distractor from the prime stimulus decay with a characteristic exponential profile, equation (3.4). The threshold θ follows these activations. At time $t_0 = 0$ms the RSI has elapsed and a new stimulus is shown. This drives the activations x_j^τ and x_i^δ of the new target and distractor (respectively) exponentially towards the input level $I_i^\delta = 1$ and $I_j^\tau = 1 + \xi(t - t_0)$. At time t_0, these inputs are similar. Later the difference between the activations increases. At $t = 565$ ms, the threshold θ follows with a delay t_{delay}. It approaches x_i^δ, which has meanwhile arrived close to I_i^δ. If the activity of the distractor is surpassed by the adaptive threshold, then the distractor is cut out of the space of possible actions. Thus, an unambiguous decision in favor of the target can be made.

Now, the display presents a fixation cross, i.e. the inputs I_i^δ and I_j^τ are switched off and the activations x_i^δ, x_j^τ and θ decay correspondingly. The characteristic time course in a DT trial shows a kink in the activation of the distractor variable x_j^δ when the distractor becomes the target. Via the coupling strength ζ, the rise of x_j^τ causes an acceleration of the decay of the activation and, consequently, a slower rise of θ. The threshold reaches x_i^δ later than in a control trial. In the DTTD trial, see figure 3.5b, the two decaying activations from the previous stimulus are decaying faster. The overall activity is thus even lower and thus the adaptation of θ is further delayed, such that the response is even slower in DTTD trials.

In the positive priming condition on the other hand, see figure 3.5c, the same variable x_j^τ is activated again by input and therefore starts off at t_0 at a higher level of activation compared to a control trial. As discussed in section 3.2, the residual activity of x_j^τ before stimulus onset is stored in the appropriate variable r^τ, which is subject to an intrinsic decay, and adds to the overall activation level $\bar{x}$ that drives θ. The augmented activity level of x_j^τ at the beginning of input causes a higher $\bar{x}$ that shortens the time that θ needs to reach x_i^δ significantly. Therefore, the reaction time is reduced in comparison to the control trials. In the DDTT trial, see figure 3.5d, the reaction is even faster due to the departure of both target and distractor activations from a higher level of remaining activity.

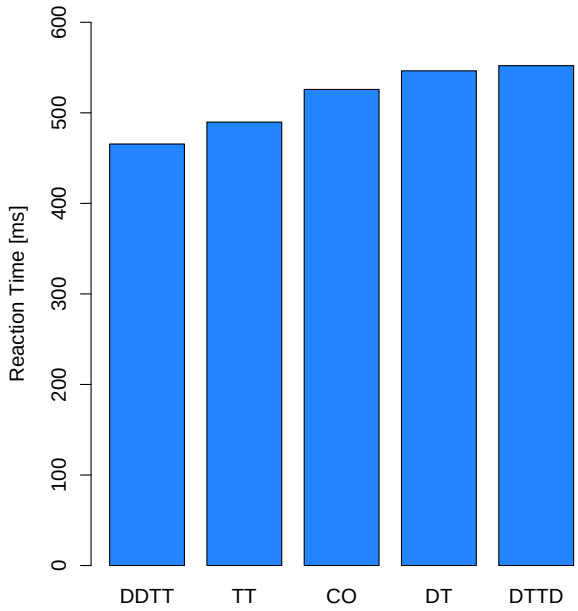


Figure 3.6: Reaction times for a response stimulus interval of 1500ms. To match the experimental data given in section 2.2 an additional delay of ≈ 135 ms for perception and action initiation has to be assumed. These processes will also further raise the variance but not alter the strength of the effects. These reaction times were obtained in an exemplary simulated session of 400 trials. The parameters of this simulated experiment were: $\alpha = 0.028$, $\beta = 0.003$, $\xi = 0.0016$, $\zeta = 0.0053$, $\gamma = 0.013$, $t_{delay} = 15$ms and $\sigma = 0.62$. A simulated time step was $0.3ms$.

We simulate sequences of 400 trials. The results from the simulations show a close correspondence to the results we obtain in our experimental studies. The results of a simulated session are presented in table A.2. We thus added 200 ms to all of the reaction times in order to account for perception and action initiation processes which are not covered by the ISAM.

3.3.2 Dependence on the Response Stimulus Interval

An interesting aspect for the evaluation of the model is the dependency of priming effects on the response-stimulus interval (RSI). The adaptation of the threshold critically depends on the length of the RSI. This is due to an adaptation to the specificity of the stored information rather than the exact average level of activation. In the computational model, this corresponds to a weighting of the averaged activity in the process of the adaptation of the threshold. We simulate the state of the model system during medium RSI from about 1s to 2.5s. Both very short (below one second) and very long RSI (on the order of hours or more) are not covered by the model and are discussed in section 3.4. In the first case, the perception process needs to be specified in more detail, which would lead to a stronger dependency of the model on the particular experimental setup. In the latter case, we would need to extend the model with specific memory mechanisms, which is beyond the scope of the present approach.

It is a known fact that reaction times in human subjects are strongly affected by the RSI (Kabisch, 2003). At very short RSI for example, a TT trial is paradoxically slower than baseline and in medium RSI the priming effect in DTTD trials is weaker than in simple DT trials. The occurrence of these so-called paradoxical effects in experimental studies (Kabisch, 2003) for very short RSI is confirmed by the computational model, see figure 3.7. In the model some of these results are due to a backward threshold crossing by the activation variables.

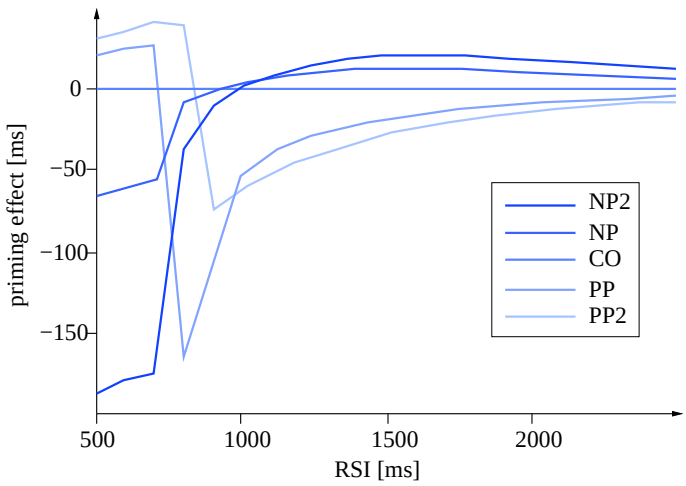


Figure 3.7: Differences in the reaction times for the five experimental conditions in dependence upon the response-stimulus interval. Note that the change of effects for short response stimulus intervals is caused by the situation that the threshold is not yet fallen under the representation activations.

3.3.3 Variation of Distractor Saliency

One of the characteristics of the negative priming phenomenon is the influence of distractor saliency on negative priming. The empirical finding that the negative priming effect increases with growing saliency of the distractors is generally labeled reactive inhibition (e.g. Grison and Strayer, 2001; Houghton et al., 1996; Lavie and Fox, 2000; Tipper et al., 2002).

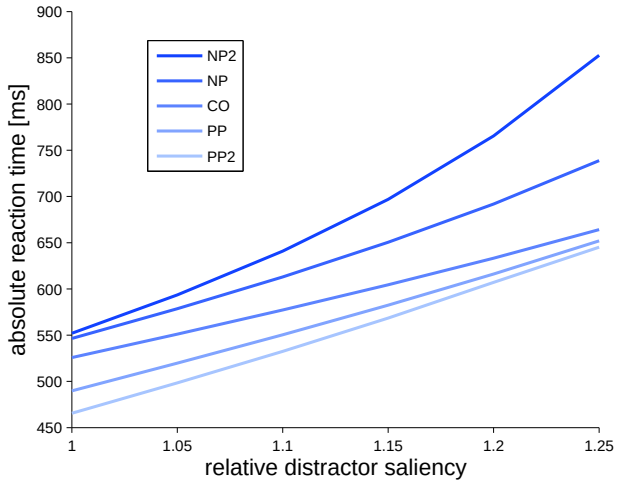


Figure 3.8: Differences in the reaction times for the five experimental conditions in dependence on the relative distractor saliency. For a low distractor saliency positive priming effects are pronounced whereas for higher distractor saliency, positive priming virtually vanishes and negative priming becomes lager.

In two experiments, Grison and Strayer (2001) manipulated the perceptual quality of the target or the distractor. Their data indicates that negative priming is dependent on the activation of postperceptual representations of the distractor. Thus a more salient postperceptual distractor representation leads to a stronger negative priming effect. They found stronger negative priming and weaker positive priming for a degraded target compared to a degraded distractor. This relation was predicted by the inhibition model (Houghton et al., 1996) and could be confirmed by Tipper et al. (2002).

The question how distractor saliency affects reaction times in terms of the ISAM is not easily answered, as reaction times depend on the sensitive interplay of all parts of the model. The saliency of the distractor can be altered by changing equation (3.5), such that in the presence of a distractor stimulus the distractor input I_i^{δ} equals the saliency s. It denotes the saliency relative to the target input. Equation (3.6) describing the target input I_i^{τ} stays unchanged. Figure 3.8 shows the reaction times produced by the ISAM if the distractor saliency s is changed from 1 (equal distractor and target saliency) to 1.25 (distractor saliency 25% stronger than target saliency).

The ISAM shows several effects when exposed to distractors of varying saliency. Generally, a stronger distractor saliency leads to longer reaction times. More time is needed to actively ignore the distractor to be able to respond to the target. Computationally, this is due to the higher level the threshold has to reach in order to cut out the distractor activation from the space of possible actions. Counteracting is the fact that the residual activations from the former trial are still higher, due to the higher level of target and distractor activation at the point of the prime decision. This contributes to a higher threshold level at trial onset, thereby accelerating an answer. But this effect is weaker than the influence of the higher distractor activation level the threshold has to reach for a decision.

The two positive priming effects TT and DDTT become weaker with growing distractor saliency. The reason for this is the exponential dynamics of the activation variable. Positive priming is produced by greater residual activation of the target variable compared to a control

trial. This activation difference between positive priming and control also decays exponentially. Thereby, longer reaction times and thus a higher distractor saliency reduce positive priming effects exponentially in the framework of the ISAM.

Both negative priming effects grow with higher distractor saliency. Negative priming is produced by an interference between the former distractor activation and the new target activation accelerating their decay reciprocal. With higher distractor saliency more activation from the prime trial carries over. The interference in equation (3.7) is dependent on the former distractor activation. Therefore, a stronger activation leads to a stronger interference and thus to a bigger negative priming effect.

3.3.4 Predictions for Single-Object Trials

The present model is critically tested by the application of single stimuli during the course of the experiment. In single-object (SO) trials only a single object is presented. With respect to the experimental setting, it is only conceivable to use single-object trials with the object being a target. If single distractors were to be included, a no-response answer would be necessary, disturbing the flow of the experiment. Nevertheless, the implementation of the ISAM enables us to predict reaction times in these conditions. A non-reaction is possible at the moment, when the distractor activation reaches a strength that would lead to the classification of a target.

We introduce a number of new categories in addition to the ones named in section 2.1. SO-trials occur in three variants, realizing special cases of the standard conditions DT, TT and CO relative to the preceding target/distractor pair (see upper half of table 3.3.4). We are also aware of numerical effects in the reaction times of the ISAM of preceding SO-trials to standard trials. Therefore we further separate these three trial conditions from trials which were preceded by both target and distractor. In our notation, four letter abbreviations denote the six new conditions: If the preceding display contained a single stimulus, the first two letters are SO, and the last two letters indicate the condition of the actual display. If the present trial shows only a single object, the first two letters denote the condition it matches in relation to the preceding display. See table 3.3.4 for the appropriate naming of the resulting additional cases.

	prime display		probe display		
	target	distractor	target	distractor	
DTSO	A	B	B	-	distractor(n) = target($n+1$)
TTSO	A	B	A	-	target(n) = target($n+1$)
COSO	A	B	C	-	One object this trial which was
	A	B	-	C	not presented the last trial.
SODT	-	B	B	C	distractor(n) = target($n+1$)
SOTT	A	-	A	C	target(n) = target($n+1$)
SOCO	A	-	C	D	None of the two stimuli matches
	-	B	C	D	the only preceding stimulus.

Table 3.1: Additional priming conditions resulting from the introduction of single object trials. The upper part of the table shows the single objects trials themselves, the lower part show trials that immediately follow a single object trial. The ISAM predicts that responses directly after single object trials have to be considered separately.

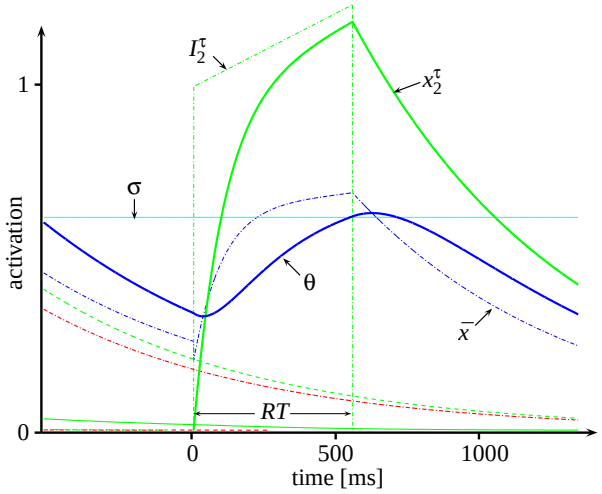


Figure 3.9: Representation variables over time for the analysis of the behavior of the ISAM during SO-trials. There is only one object present. This results in a lower average activity $\bar{x}$. Without the sensitivity cutoff via σ, the reaction time in the case of an SO-trial would not be longer than the time necessary for perception and action initiation. In the trial directly after an SO-trial, the baseline activation is lower in comparison to non-SO trials, as the activation of the second object is lacking.

When SO-trials are presented to the model, a dramatic reduction in reaction time is observed, as shown in table A.3. Columns 2 – 4 contain much smaller values than expected in trials with two simultaneous objects. The priming effects are clearly present without specific tuning of the model to this case, See figure 3.9. However, the global sensitivity level σ plays an important role, as it finally determines the moment of reaction.

Figure 3.10 also displays the reaction times for trials that immediately succeed SO-trials. They are classified in dependence on the reoccurrence of the former object. It is obvious that a significant slowdown of reaction time after SO-trials is inherent to the model due to the lower overall activity if only one object representation is activated by the input. This justifies the separate consideration of trials immediately succeeding SO-trials. If these trials were subsumed under the classical conditions, they would artificially increase the variance. In experiments, this slowdown can occur due to a disturbance of the response routine of subjects by SO-trials, where the switch back to two stimuli displays is also a distracting factor. We are currently running experiments to check whether this effect really is present in human behavior.

The observed effects during the presentation of single-object trials are readily explained in terms of the formal model: The overall activation that drives the threshold is lower in SO-trials, therefore θ takes longer to reach the activation level of the distractor during the next trial. This causes a longer delay until the correct object is singled out in order to allow for an unambiguous decision of the system.

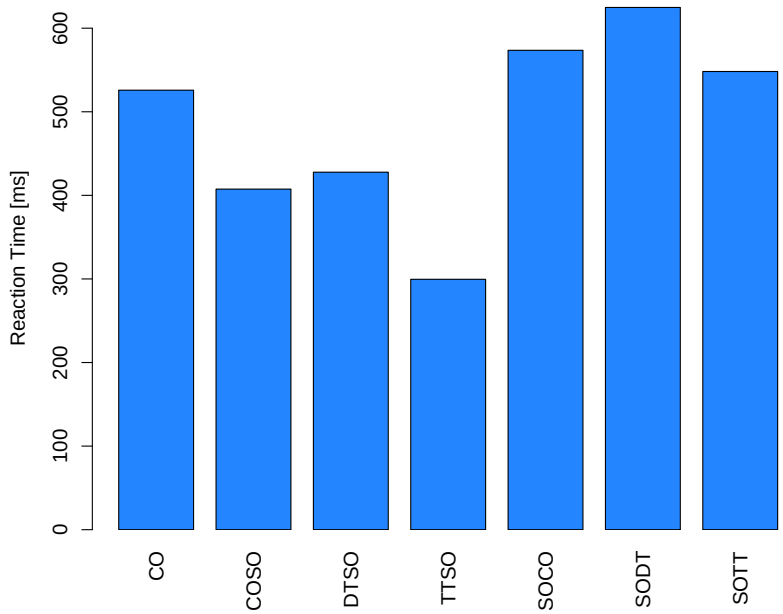


Figure 3.10: Predictions for single object reaction times for a response stimulus interval of 1500ms, in dependence of experimental condition. The usual priming conditions are not shown. The parameters are the same as in figure 3.6.

3.4 Discussion

3.4.1 Modeling Priming

The ISAM impressively reproduces experimental data without contradictions. Additionally it shows reasonable behavior when confronted with the touchstone of RSI-dependency. The implementation of the ISAM presented here is robust enough to provide predictions about stimulus constellations that have barely been considered in priming experiments thus far. This circumstance serves as a good opportunity to test the ISAM experimentally. Among the many theoretical approaches to priming, we select three of the most popular models and compare their implications to empirical results.

The distractor-inhibition model (Tipper and Cranston, 1985; Houghton and Tipper, 1998) is one of the most influential theoretical accounts. Following this line of reasoning, the negative priming effect is supposed to be a cognitive index for the inhibitory component of selective attention. It is assumed that irrelevant stimuli representations are actively suppressed to support selection of the goal-relevant target stimulus, and that this inhibition persists for some time. When the former distractor becomes the behavioral relevant target in the subsequent display, responding is hampered because of the persistence of the inhibition imposed on it during the prime trial. However, shortcomings of this model emerge when explaining certain experimental results: Because negative priming is described as an after-effect of distractor inhibition, a facilitatory effect of ignored

repetition trials without a distractor in the probe trial (e.g. Tipper and Cranston, 1985) cannot be reasonably accounted for in terms of this model.

A second important account, the episodic retrieval model, has originally been proposed by Neill and Valdes (1992). They argue that negative priming is the result of conflicting information caused by a retrieval of the prime episode when exposed to the probe stimulus triggered by similarities of the two situations. The probe target thus causes retrieval of the prime episode due to repetition of the prime distractor. The information from the retrieved episode (i.e. "ignore the stimulus") conflicts with the need to respond to this same stimulus in the current probe episode, resulting in a time-consuming process. This causes an increased reaction time, characteristic of negative priming. A particular advantage, in comparison to the other models, is its explanatory power regarding the influence of temporal discriminability of the prime episode relative to the preceeding episodes. This is more easily accounted for by episodic retrieval than by inhibition-based models. The same is true for experimental evidence which shows that the effect increases quantitatively when the contextual similarity between prime and probe situation is increased (e.g. Stolz and Neely, 2001). However, the model falls short of explaining semantic negative priming effects, where the response to a probe target is only semantically related to the prime distractor (such as dog to cat), which are more easily accounted for by the inhibition model (Tipper and Cranston, 1985). Both the inhibition model and the episodic retrieval model explain general reaction time increases in negative priming conditions in a straight-forward manner.

In contrast to the two previously described models, the temporal discrimination model (e.g. Milliken et al., 1998) does not assume selection processes during the prime task as the basis of negative priming. Instead, it assumes that negative priming is caused at the moment of response formation during the probe processing. Two response modes are postulated in the following context: the response can either be computed or directly retrieved from memory. Which of the two response modes guides behavior in the probe task depends on a categorization of the probe target as either "old" or "new". If the probe target is categorized as new, a response is generated on the basis of perceptual analysis. Otherwise, if the same task situation has already been encountered, the response can rely on a direct retrieval of the former response. For a probe target in an ignored repetition trial, an ambiguity in the categorization process is assumed, resulting in a longer reaction time. But there is also empirical evidence which contradicts the predictions derived from this model: In a repeated distractor condition, the model predicts a slowdown in responding, whereas a speed-up is repeatedly demonstrated in psychological experiments (e.g. Frings, 2005).

These considerations show that modeling of the negative priming effect is still in its developmental stages. Therefore we do not hesitate to implement the ISAM as a new contribution to the ongoing discussion of negative priming explanations. At first glance the ISAM looks similar to the inhibition based model (Tipper, 1985) as negative priming is the result of a negative interference. However considering details the two models differ in crucial aspects. In the ISAM, priming effects are an interplay of remaining activation from the prime display and activation driven by the probe input.

Generally, none of the effects are present in the ISAM until the probe trial starts. For the inhibition based model, this is not the case, as the activation of the distractor is inhibited in the prime trial and negative priming is just an after-effect. The negative interference is not due to a control mechanism ruling the prime decision as postulated in the inhibition based model, but rather an intrinsic interference within the neuronal circuits accounting for the prime distractor representation. The negative priming effect is thus obtained by the added "effort" of the distractor representation, which must switch states from "do not respond to" to "respond to". This interference is more like the postulated conflict between memory traces and the current activation pattern responsible for negative priming in the episodic retrieval model. Another marking-off feature of the ISAM is that

representations cannot be pushed below baseline, as the absence of a representation is modeled by zero activation. The interference only affects time constants of decay or rise towards input equations (3.7) and (3.8). Similarities to Tipper's inhibition model can be found in the assumption of an external control mechanism responsible for the decision. But rather than inhibiting target activation in the probe trial, the control mechanism postulated in the ISAM boosts the target activation due to its semantic classification relative to the task.

Since the ISAM produces the priming effects at a very early stage of the probe trial, it displays no obvious similarities to the mechanisms presupposed in the temporal discrimination theory (Milliken et al., 1998). In terms of the functional effectiveness, it can rather be placed between the two models first discussed, the inhibition and the episodic retrieval theory. The ISAM thus incorporates advantages from both approaches and integrates assumptions from both models into a more complete theory comprising all (positive and negative) priming effects.

3.4.2 Phenomenological and Neural Models

Modeling provides an attractive approach to attentional mechanisms and thus in particular to priming phenomena (Houghton and Tipper, 1994). First, mathematical modeling of a theory requires a precise formulation of the theoretical assumptions and mechanisms. This is of even greater importance as some of the explanatory models for negative priming presented above do not specify how and on what level attentional processes operate. Second, an explicit computational model can help to derive more specific predictions than possible from more abstract models, since not only qualitative but also quantitative predictions can be generated. Also, conducting computational experiments can produce new and even unexpected results that, in turn, can lead to experimental confirmation and extension of the model.

The ISAM presented here is implemented with effective equations. A refinement of neural networks seems desirable but cannot provide deeper insight into the mechanisms generating negative priming, as the effect appears at a high level of brain function where computation is strongly distributed. Therefore, no plausible neural model can be given at this stage of brain research. The reduction to an effective theory, in contrast, allows for direct observations of the functioning of the model itself without exhaustive investigations of the simulated data.

The difficulties of the model with short RSI point out the specificity of a certain approach that can never describe the whole behavior of the brain which, for example, in the context of changing the RSI uses several different strategies, such as short term memory and long term memory when changing the RSI from 500 to 1500 ms. With the present implementation, the ISAM moves up in line with the other explanations of priming effects like the inhibition or episodic retrieval approach. It shows possible simplifications of existing models, having in mind a minimal model that describes the generation of priming effects in the human brain.

3.4.3 The Implementation of the Model

The numerical implementation revealed an insufficiency of the ISAM concerning RSIs that involve short-term memory effects. For very short RSI its consequent implementation shows a reversal of priming effects. The reason for this strange behavior is an empty space of possible actions. At stimulus onset, the threshold level is still greater than any activation of the variables representing object recognition. Therefore, a decision is made when the first variable reaches the threshold. An increase of this limited range of RSI also requires the consideration of low-level neurophysiological details, which will be incorporated in an improved version of the model. Furthermore, our implementation is sensitive to the number of stimuli in one display. The factor $\frac{1}{2}$ in equation 3.10

was chosen for the situation of one target and one distractor. If multiple distractors were present, the model may not come to a decision. In extending the ISAM to more paradigms the factor should be adjustable according to the stimulu configuration in a trial display.

All this together with the restriction of the ISAM to identity based voicekey paradigms built the basis for constructing the General Model for negative priming more or less from scratch. Chapter 8 will be devoted to the development of a computational model that is capable of dealing with numerous stimuli and response schemes while being complex enough to incorporate all mechanisms that are suspect to produce priming effects.

3.5 Summary

The imago semantic action model by Kabisch (2003), which has been described in section 2.4.7, is the attempt to explain negative and positive priming by one single mechanism, the presence of an adaptive threshold which decides on the semantic representations of stimuli which one of them to attend. The original description is sufficiently precise to make a straightforward implementation as a dynamic system possible.

We first derived a very simple dynamics describing the transient of the spike rate response of a neuronal compound to the onset and offset of external input. The resulting dynamics is used to model object representations as activation of a single variable even though this representation is most likely held by a distributed network of interconnected neurons. On the basis of these representation dynamics we implemented the ISAM with minimal additional assumptions and found that the reproduction of frequently encountered priming effects is easily achieved. We thus prove quantitatively that priming effects, both positive and negative, can be explained in terms of the ISAM. Further phenomena that are well described by the ISAM include the speedup of responses to single-object presentations, the preservation of priming effects in these cases, and the reproduction of the dependency of priming effects on varied distractor saliency.

4 EEG Correlates of Negative Priming

The second mainstay in the determination of the time course of negative priming is electrophysiologically recording the processes happening during a negative priming trial. The electroencephalogram (EEG) is well suited for negative priming experiments due to its high temporal resolution. Especially the consideration of event-related potentials (ERPs) is well suited to reveal the differentially time-consuming components of information processing that are responsible the negative priming effect.

In search of a cooperation, we were offered to use a fully equipped state-of-the-art 64 channel EEG-laboratory in the department for medical psychology Göttingen. We seized the opportunity to run a series of priming experiments while recording EEG activity. The results of the adaptation of our voicekey paradigm, presented in section 2.2, to the EEG environment provide the replication of the results of a recent study on EEG correlates of negative priming (Gibbons, 2006).

We will describe the derivation of EEG signals and traditional way of data analysis in section 4.1. In section 4.2, we will give a detailed review of correlates of negative priming. In sections 4.3 to 4.5 we will describe our study on electrophysiological correlates of negative priming, present the results in sections 4.6 and 4.7. The results are integrated into previous ERP correlates and evaluated in sections 4.8 and 4.9. Our publication (Behrendt et al., 2009) is the basis for the current chapter.

4.1 Introduction to Electroencephalography

Electroencephalographic (EEG) data is the electric potential change on the skull surface on a 100μV scale. These potential fluctuations are most likely produced by electrical fields generated by ion flux around axons of firing neurons. Axons that are oriented perpendicular to and not far from the skull surface contribute best to the signal. Around 10.000 neurons are required to fire in synchrony to obtain a good signal. But even contributions from deeper brain regions can not be excluded and a potential reversal can occur at electrodes near cortical folds. The measured signal is the superposition of all signals finally low-pass filtered by the skull.

4.1.1 EEG Recording

For clean recordings, a shielded recording environment is necessary. Electrical equipment should run on low DC current. Very sensitive difference amplifiers are used that feed the signal into the A/D converter at a recording computer. Electrodes are attached to the skull by an electrode cap that ensures a correct positioning. Electrical contact is maximized by degreasing the electrode sites with alcohol and by the application of a conductive gel until the impedance is below 5 kΩ. The sintered ring electrodes are made of highly conductive material (Ag/AgCl) in our recording environment. In our case placing on the scalp follows the standardized extended 10-20 system by Jasper (1958), see figure 4.1. Electrodes are named according to their position from frontopolar via anterior, frontal, central, parietal and temporal to occipital. Odd numbers are situated on the left side, even numbers on the right. Care has to be taken, when looking at head plot topographies, as the EEG-researcher looks at a head from above, whereas fMRI-researchers look

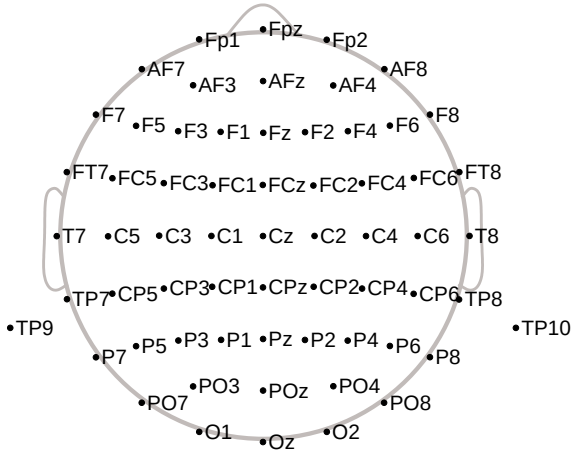


Figure 4.1: The extended Ten-Twenty system of electrode placement as introduced by Jasper (1958).

at their subjects from the feet and thus sketch the head inversely. Additional electrode positions are located beneath the eye, a rather strong dipole, to record an electrooculogram in order to clean the data from collateral artifacts.

The reference for the difference amplifier is usually placed at a position where low brain activity can be expected, such as the mastoids TP9 and TP10 or a position above the great longitudinal fissure like FCz. Recording a difference signal has the advantage that global noise can be expected to be homogeneously present at all sites and thus not enter the data. Unfortunately global EEG signals do not as well. Choosing the reference is crucial as the signal of the reference is present in every time series. But if certain electrodes have been identified to contain a signal of interest, an offline rereferencing is still possible which can enhance the signal. In the raw recording setup we set a high sampling rate of 5000 Hz, applied a band pass filter between 0.1-70 Hz and a notch filter to suppress 50Hz mains hum.

A second computer is dedicated to stimulus presentation and recording of the behavioral data. It transmits time markers to the computer recording the EEG, ca. 4GB of data per session. Markers can be given for the onset of the successive stimulation displays and for the subjects' responses and make a meaningful segmentation of the EEG data possible.

4.1.2 Data Processing

Before considering EEG data, the behavioral data is analyzed and outliers are rejected according to the following procedure. The outliers can be simple behavioral errors which invalidate the current trial and the successor, as the priming may be mixed up. Reactions that were faster than 250 ms or slower than 2000 ms were removed, as they most likely contain signals from processes differing from the ones under investigation. Finally, reaction times where the difference to the mean of the experimental condition exceeded two times the standard deviation were excluded too.

Because most inferential statistics assume a normal distribution of the data, Kolmogorov-Smirnov tests were conducted for the reaction times within the experimental conditions. If the test showed that the assumption of a normal distribution was violated, single values were removed based on their probability given the normal distribution model until the Kolmogorov-Smirnov test yielded insignificant results. Overall, we ensure that not more than 10% of the trials are excluded from the analysis for each participant. If this is the case, the subject is excluded completely from the analysis.

In order to obtain neural correlates of a certain behavioral effect, the effect has to be identified first by analysis of the behavioral data, i.e. by comparing average reaction times of the different priming conditions. Then the EEG is segmented according to the surveyed time markers. The segments should contain some information preceding the trial for a baseline correction as the EEG is no absolute measure but just potential differences. Also the reaction in each trial should be captured, thereby defining the segment length. Only segments belonging to behavioral valid trials enter the analysis.

The data still contains artifacts due to body movements, fluctuations in electrode conductance, electrostatic charges or technical errors. A major problem of recording the average activity of large numbers of neurons (as compared to intracellular measurements of single neurons) is the very strong background activity in the data, apparent as very strong noise. In the EEG, noise is significantly stronger than the ERP itself (Flexer, 2000). Whereas the electrical background activity is in the range of $1 - 200\mu$V, the evoked potentials have an amplitude of only $1 - 30\mu$V (see figure 4.2).

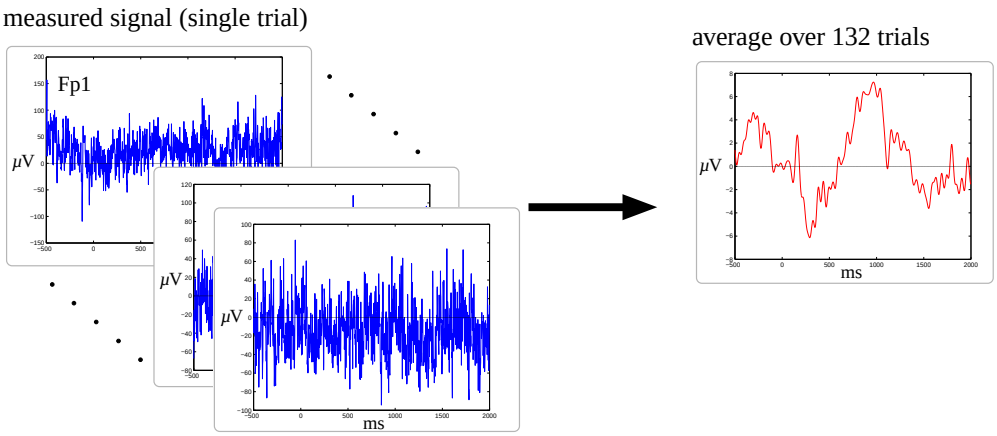


Figure 4.2: Noise and signal in EEG-data. Note that the approximate amplitude of the average ERP of 132 trials is in the order of 10μV, while single trial recordings show amplitudes of $\approx 100\mu$V.

Thus a cascade of data cleaning and trial rejecting procedures is carried out. Basic techniques are simple filters, baseline corrections and visual inspection and manual removal of erroneous electrodes or segments. Automatic data rejection excludes trials with a maximum potential surpassing a certain threshold. Furthermore, eyeblinks cause major disturbances in the data, especially in the fronto-polar regions. Their impact has to be removed either by traditional regression based techniques (Gratton et al., 1983) or by more recently developed approaches using independent component analysis (Joyce et al., 2004; Jung et al., 2000; Delorme et al., 2007).

After these cleaning procedures, segments of one experimental condition are averaged pointwise forming the averaged event-related potential (AERP). The AERP usually shows a stereotypical form, a sequence of minima and maxima. Height and latency of the peaks, which are named according to polarity and sequential number, alternatively approximate latency, as N2, P3 or P300, are assumed to contain information about cognitive processes. Therefore latency and amplitude of the peaks are systematically compared between the priming conditions in order to determine which mechanisms in the trial processing carry the difference between the conditions.

4.2 Review: ERP Correlates of Negative Priming

We will now review electroencephalographic correlates of negative priming. As we have seen in the last section, EEG recordings provide insight into what is happening in the brain, but the interpretation is difficult due to a very bad signal to noise ratio. One way to relate EEG data to behavior is finding discrepancies in the averages of two experimental conditions. Those Differences can be interpreted in favor or in denial of a certain theory. The different theoretical accounts vary by assuming different cognitive processes to cause negative priming. Those Processes are found to be present in defined ERP components. Thus the latency and amplitude of the signal differences can be attributed to such a process and therefore give evidence for a theory for negative priming.

To give a comprehensive overview of ERP Correlates of Negative Priming, we will go through the ERP components in the order of their latency. Particularly the N200 and P300 components are discussed as candidates for correlates of behavioral negative priming. The observed patterns of differences in latency and amplitude are interpreted as evidence for a particular theoretical approach (for a review see Mayr and Buchner, 2007).

4.2.1 N200 Component

In general, the N200 component in the ERP, a negative deflection of the EEG around 200 ms, has been interpreted to reflect early stimulus evaluation and especially active or passive discrimination processes subsequent to the perceptual identification of the stimulus features. Several negative priming studies revealed modulations or additional negative components in the N200 range in frontocentral or occipital areas in DT trials compared to a control condition. For example, Daurignac et al. (2006) analyzed negative priming correlates in a Piaget-like numerical task, in which strategies rather than stimulus features were repeated between prime and probe. They observed an enhanced N200 in DT trials, which they interpreted as evidence for inhibition mechanisms. Similarly, Frings and Groh-Bordin (2007) found more negative-going waveforms in the frontal N200 complex in DT relative to control trials using a classical flanker paradigm. In addition, they reported that this effect was modulated by magnitude of the individual behavioral negative priming effect in reaction times. Participants who did not show behavioral negative priming also did not have a significant negative priming effect on N200 amplitude. The authors concluded that ERP components in the N200 time range represent cognitive processes that cause negative priming and interpreted them as additional effort required in DT probes to select a previously inhibited stimulus against distractors. ERP studies investigating episodic memory processes revealed that episodic recognition of a stimulus is usually associated with ERP components emerging not earlier than 300 ms (Groh-Bordin et al., 2005, 2006). Therefore, the time window of the DT-specific enhanced N200 component (around 200 ms after stimulus onset) is said not to fit a retrieval explanation. But the fact that no correlates of retrieval are found at that early latencies does not imply that retrieval is not yet in progress.

4.2.2 P300 Component

The P300 is often interpreted as reflecting the updating of resources needed for stimulus evaluation within the context of a model of the environment (Donchin and Coles, 1988). In a study investigating location-based and identity-based negative priming in a direct comparison, Kathmann et al. (2006) reported differential ERP effects for the two tasks. They observed increased P300 amplitudes in DT relative to control trials in an identity task with overlapping black and grey digits, whereas this effect was absent in a location negative priming task. Kathmann et al.

(2006) concluded that there was no evidence in their data for inhibitory mechanisms in identity-based negative priming. The authors suggested that their observed P300 modulation would reflect the updating of object representations when they are repeated, and concluded that an increased processing effort was required in the DT condition.

Unlike Kathmann et al. (2006), Gibbons (2006) argued that the P300 enhancement seems to be well in line with the inhibition view of negative priming. According to the context-updating hypothesis of Donchin and Coles (1988), the increased P300 amplitude reflects the completion of a process by which a task-relevant stimulus is integrated into the current context (stimulus evaluation). Inhibition theory assumes that the activation level of the internal representation of a DT trial probe target should be below baseline at display onset (Houghton and Tipper, 1994). Relative to control, more activation is therefore needed to reach a criterion threshold required for stimulus identification, which is reflected in increased P300 amplitude and/or latency. Based on similar considerations, one can expect reduced P300 amplitude and/or latency in the TT condition, due to above-baseline activation of the representation of the target at probe display onset (Stahl and Gibbons, 2007).

Whereas findings by Ceballos et al. (2003) of larger P300 amplitudes and delayed P300 latencies adopting a visual identity negative priming task are well in line with the results of Kathmann et al. (2006), Stahl and Gibbons (2007) found a reduction of the left posterior P300 amplitude for both TT and DT conditions compared to control in a recent study investigating ERP correlates of identity negative priming in a flanker task. Furthermore, the DT-related P300 amplitude reduction was negatively correlated with behavioral negative priming; those participants who showed strong behavioral negative priming showed less DT-related P300 reduction. Thus, it appears likely that DT shifts produce at least one facilitatory effect, being reflected in P300 reduction. This facilitatory effect may reduce the behavioral negative priming effect in weak-negative priming participants. Stahl and Gibbons (2007) further argued that the P300 reduction reflects an initial evaluation of both TT and DT probe displays as similar to the prime and postulated a resulting tendency to repeat the prime response. This tendency is correct in the TT condition but wrong in the DT condition, causing facilitation and response conflict, respectively, like it was postulated by response retrieval theory. The response conflict account of their flanker negative priming effect also receives support from the analysis of the lateralized readiness potential (LRP), differing EEG signals on the two primary motor areas. Onset of the response locked LRP occurred significantly earlier for DT than control, indicating increased duration of motor processing, which fulfills the predictions that can be derived from retrieval theories. Stahl and Gibbons (2007) suggested that DT-related P300 reduction is better in line with episodic retrieval than with inhibition view. Because of the initial equivalence of TT and DT processing and the resulting tendency to repeat the prime response, persisting activation rather than persisting inhibition seems to be present in DT trials. Therefore, to explain the emergence of behavioral negative priming effects, a process operating at later stages counteracting the presumed facilitatory DT effect during stimulus evaluation has to be assumed. However, it should also be taken into account that there is an alternative interpretation of the P300 reduction. Instead of a true modulation of that component, the P300 reduction for DT and TT conditions could reflect superimposed processing negativity in the P300 range (Kok, 2001). Yet, this alternative account of DT related P300 reduction would also not be in line with distractor inhibition theory because it leaves open the question of why DT and TT conditions behave in a similar fashion (Stahl and Gibbons, 2007).

The divergent empirical findings of Kathmann et al. (2006) and Stahl and Gibbons (2007) concerning the P300 component may be accounted for by three methodological differences between the two studies. First, Stahl and Gibbons (2007) employed shorter prime-probe than probe-prime intervals (Neill and Valdes, 1992), whereas Kathmann et al. (2006) realized identical intervals.

The former study may therefore have provided conditions favoring episodic retrieval, due to the better temporal discriminability of the prime episode from prior episodes (see Allport et al., 1985). Second, Stahl and Gibbons (2007) used a significant proportion of TT trials, which potentially increase the probability of episodic retrieval (on TT trials, the retrieval of prime information would indeed be useful), while Kathmann et al. (2006) presented only DT and control trials in their study. Third, compared to the side-by-side arrangement of target and distractor stimuli in the flanker experiment by Stahl and Gibbons (2007), Kathmann et al. (2006) used overlapping digits as target and distractor stimuli, so that successful selection of the target might require particularly strong distractor inhibition. The discussion of the diverging results concerning the P300 component seems to agree with the assumptions of the dual mechanism hypothesis proposed by May et al. (1995). These authors proposed that inhibition as well as memory retrieval may contribute to the emergence of negative priming, but that the experimental context specifies which of the two mechanisms is expected to operate. It is conceivable that Stahl and Gibbons (2007) provided conditions that encouraged episodic retrieval to become effective, while the experimental setup by Kathmann et al. (2006) favored the emergence of persisting inhibition. This may have been responsible for the different ERP correlates of negative priming in the two studies.

4.2.3 Positive Slow Wave Component

Two studies of negative priming in the auditory domain by Mayr et al. (2003) and Mayr and Buchner (2006) reported modulations of ERP amplitude in the time window of the late positive complex (LPC). They investigated reaction time and ERP effects of auditory negative priming in a categorization task, where the participants had to categorize sounds presented to a cued ear as musical instruments or animal voices. In the study by Mayr et al. (2003), ERP analysis revealed an attenuation of a late parietal positivity ranging from about 300 to 600 ms in the DT condition compared to both TD and control. This critical finding was replicated by Mayr and Buchner (2006), showing a parietally located negativity in the DT condition compared to the control conditions between 550 and 730 ms post stimulus onset. The authors interpreted these results as support for an episodic retrieval account, because the late ERP component is known to be sensitive to stimulus recognition and familiarity. Mayr et al. (2003) argued that the only ERP effect that was uniquely related to the DT condition was a late posterior complex which shares polarity, time course, and topography of the so-called old/new ERP effect obtained in studies investigating recognition memory (Rugg and Doyle, 1994). This old/new effect is characterized by more positive parietal ERP amplitude at about 300 ms following the onset of an old compared to a new stimulus. Several studies found a relationship between the enhancement of the LPC and an increase of the familiarity of successive events (Rugg and Allan, 2000; Rugg and Doyle, 1994). Greater familiarity, in turn, has been associated with more fluent processing of an event (Johnston et al., 1985). If the DT condition evoked a smaller LPC, the processing of a previously ignored stimulus is equivalent to the less fluent processing of novel stimuli, hence resulting in negative priming in reaction times.

This line of argumentation should also hold when comparing TD and control condition, because of the greater familiarity between prime and probe events in TD trials, which was not observed in the data. Mayr and Buchner (2006) pointed out that the stimulus familiarity in their experiment was generally high because of the small set of only six different auditory stimuli that were presented throughout the study. Therefore, partial stimulus repetition from prime to probe as in the TD or DT condition was per se not expected to be sufficient to increase the baseline familiarity of the stimulus, and to induce an old/new ERP effect. This expectation was fulfilled in the TD condition, which did not differ from the standard control with respect to the LPC. The apparent contradiction could be resolved by assuming that the old/new effect is restricted to the attended parts of the probe display, i.e. the probe target.

As a further result, Mayr and Buchner (2006) reported that the LPC negative priming correlate was more pronounced for slow trials than for fast trials. The dependency of the effect size on the reaction time and ERP effects of negative priming on the reaction time level is consistent with the episodic-retrieval account of negative priming. However, lower-than-baseline familiarity in the negative priming condition could also be induced by other mechanisms. For example, a comparable effect might result from a persisting inhibition of the prime distractor. Nevertheless, it seems that the concept of a frontal inhibition process is not compatible with the results of Mayr et al. (2003) and Mayr and Buchner (2006).

4.2.4 Summary of ERP Correlates

Further ERP studies on visual identity negative priming have been published, but either showed no ERP effects of negative priming (Gibbons, 2006; Hinojosa et al., 2007), or the interpretation of the ERP correlates of negative priming was difficult because no negative priming effects were found at the behavioral level (Gibbons, 2006). Two other studies using lexical decision tasks found small N400 amplitude reductions in the DT condition relative to control (Heil and Rolke, 2004; Wagner et al., 2006). It has however been argued that this effect was a correlate of a detected prime-probe repetition, rather than a specific correlate of negative priming (Heil and Rolke, 2004).

On the basis of the relatively small number of existing studies and the heterogeneous results it is not possible to get a clear picture of the ERP correlates of visual identity negative priming up to now. Overall, the majority of studies seem to support an episodic-retrieval explanation of negative priming. Yet the inconsistency of previous results and especially the lack of replication of the ERP correlates of negative priming have hitherto prevented significant contributions of ERP research to the theoretical debate on negative priming. Only the findings of left posterior P300 reduction for both DT and TT conditions compared to control and the significantly earlier onset of the response locked LRP for DT than control trials (Stahl and Gibbons, 2007) were replicated by Gibbons (2009) using the same flanker task. Showing that the P300 effect generalizes over different identification tasks would substantially add to the discussion. It should also be emphasized that the findings only indirectly support retrieval explanations; at least, a genuine functional late-range ERP correlate of negative priming which could represent the conflict between retrieved prime information and information extracted from the probe is still missing.

4.3 Hypotheses

The present study of identity based negative priming aimed at putting episodic retrieval theory to a critical test. We derive two testable predictions from episodic retrieval theory. First, retrieval of prime information should only occur if the stimuli from the probe display in a DT trial is perceived as similar to the ones of the prime display. Second, a late processing conflict should emerge, because the prime episode retrieved due to a similarity cue is misleading for the required probe response in DT trials. For both processes, an early similarity signal and a late correlate displaying a conflict, ERP correlates are to be found. The ERP correlate of perceived prime-probe similarity and/or the retrieval process itself should be largely the same for DT and TT trials (for both conditions, the probe target is a repeated prime stimulus). We therefore expect a conceptual replication of the consistently observed ERP correlate of visual negative priming, that is, P300 amplitude reduction in both TT and DT trials relative to CO (Gibbons, 2009; Stahl and Gibbons, 2007). By contrast, the second late-range ERP correlate should distinguish negative priming from both, TT and CO conditions, because only in DT trials conflict should emerge. Anterior predominance can be expected for this late-range ERP correlate of negative priming, given the crucial role of frontal cortex in the processing of various types of conflict (Ridderinkhof et al., 2004).

4.4 Experimental Setup

The relatively small behavioral negative priming effects (10-15 ms) found in previous ERP studies of visual negative priming (Frings and Groh-Bordin, 2007; Gibbons, 2009; Kathmann et al., 2006; Stahl and Gibbons, 2007) may be one reason why no late range, conflict related ERP effect of negative priming has been established until now. Therefore, in the present study a classical picture naming task introduced by Tipper (1985), and already introduced in section 2.2, is employed. No comparable identity-priming paradigm has been used in ERP studies of negative priming so far. Obviously, the larger effects increase the chance to find a functional late range ERP correlate of negative priming. More complex stimuli producing generally larger reaction times than in the simple flanker task may increase the chance to find a late conflict related ERP correlate of negative priming.

Sixteen participants, five of which were male, eleven female, mean age 23.3 years, SD = 5.5 years, were tested individually in sessions that lasted no longer than 70 minutes. Prior to the main experiment, subjects were tested for their color discrimination abilities. Participants were instructed to name the green target object as quickly and correctly as possible while ignoring the superimposed red distractor object. A 30-trials practice session preceded the main session.

In a single trial, subjects saw the following series of events: (a) a fixation cross, centered on the screen for 500 ms (b) a display containing two superimposed object or a single green target object until the subject responded, but no longer than 2 seconds (c) a blank screen for a randomized duration between 0 and 1000 ms. The length of the response-to-stimulus interval (RSI) was thus randomized between 500 and 1500 ms (blank screen plus fixation cross). This approach yielded good results in terms of strong negative priming effects in previous experiments. Behavioral errors were noted when subjects accidentally named the distractor, used a wrong name identifier, stuttered, or failed to answer.

EEG was continuously recorded from 63 head electrodes (see figure 4.4) arranged in an extended 10-20 system, using a 64-channel BrainAmp MR amplifier and an electrode cap (Brain Products Inc., Germany) with sintered Ag/AgCl electrodes. Vertical electro-oculogram (EOG) was monitored from an electrode positioned 1 cm below the right eye. FCz served as active reference electrode. All impedances were kept below 5 kΩ.

From continuous EEG, ERPs were derived separately for priming conditions (CO, DT, and TT) and participants, according to the following steps. First, EEG segmentation was performed, resulting in $[-100, 1500]$ ms epochs relative to trial onset. Epochs were then baseline corrected with respect to the $[-100, 0]$ ms interval, and an initial artifact rejection was performed to identify epochs with technical artifacts, i.e. amplitudes exceeding ± 1 mV. Afterwards, we performed an EOG correction according to Gratton et al. (1983). A final, more sensitive artifact rejection eliminated all epochs containing amplitude values exceeding $\pm 100\mu$V. This resulted in a reduction of trials of less than 10% for each participant and priming condition. Epochs were then averaged separately for priming conditions and participants. In a final step, these individual averaged waveforms were re-referenced against algebraically linked mastoids (TP9, TP10).

4.5 Data Analysis

An analysis of variance (ANOVA) with a repeated-measures factor priming condition (CO, DT, and TT) was computed for reaction times, employing correction of *p*-values according to Geisser and Greenhouse (1958). One-tailed Bonferroni-corrected t-tests were applied to test differences to control for conditions DT and TT (expecting negative and positive priming effects, respectively).

We determined the grand-averaged ERP waveforms computed for each of the three priming conditions (see figure 4.4). Identified potentials included posterior P1 (mean peak latency 102 ms) and N1 (mean peak latency 174 ms). A subsequent P2 potential peaked at around 230 ms, followed by N2 (270 ms). At posterior electrodes, P300 was observed at around 360 ms. In addition, over frontal areas a late positive complex (LPC) occurred at around 700 ms after probe display onset. Second, in an attempt to systematically investigate ERP priming effects, amplitudes and latencies of these ERP peaks were determined separately for priming conditions and individuals. Separately for electrodes, the P1 peak was found as the most positive amplitude value between 80 and 120 ms. The N1 peak was found as the most negative amplitude value between 140 and 200 ms, and the P2 peak as the most positive amplitude value between 220 and 250 ms. N2 was found as the most negative amplitude value between 250 and 300 ms. The P300 peak was determined at the most positive amplitude value between 300 and 500 ms, and the frontal LPC was found as the most positive peak between 400 and 900 ms.

Amplitudes and latencies of P1, N1, P2, N2, P300, and LPC potentials were subjected to separate analyses of variance (ANOVAs), treating priming condition (CO, DT, and TT) and Electrode as repeated-measures factors, with levels of the latter factor depending on the potential of interest: For P1, N1, and P300, electrodes from the most posterior three rows (see figure 4.4) were included in the analysis, whereas for P2 and N2, only electrodes from the most posterior two rows were considered. In the analysis of LPC, the most anterior four rows were included. Note that P300 amplitude was measured as mean voltage between 300 and 500 ms, to account for the broader temporal distribution of P300. Since analysis revealed a strong priming effect on LPC peak latency, LPC amplitude was not determined from a fixed time window. Rather, to assess priming effects on LPC amplitude and latency independently of each other, LPC amplitude was computed separately for participants, priming conditions and electrodes as mean voltage in 200-ms time windows adjusted for LPC peak latency. Intervals ranged from 500 to 700 ms in the TT condition, and from 550 to 750 ms in the DT and control conditions. Geisser and Greenhouse (1958) correction was applied when necessary and significant effects were further explored using post-hoc Scheffe's test.

4.6 Behavioral Results

Mean reaction times in each priming condition, standard deviations, and effects for DT and TT trials compared with CO trials are shown in figure 4.3 table A.4 in appendix A.3. Trials in which an error was committed (1.2 %) and directly subsequent trials are excluded from further analysis. Trials with response latencies below 250 ms or more than two standard deviations above the individual mean for each participant and priming condition are excluded as outliers (4.7 %).

One-way ANOVA is used to analyze priming effects. The effect of priming condition (CO, DT and TT) is significant $F(2,30) = 85.68, p < 0.001; \varepsilon = 0.85$. Planned comparisons show that the mean reaction time for DT trials is significantly increased by 27.5 ms compared to CO trials (CO vs. DT: $t(15) = -5.62, p < 0.001$). As anticipated, reaction time for trials in the TT condition are significantly decreased by 132.9 (CO vs. TT: $t(15) = 8.57, p < 0.001$).

4.7 ERP Results

Figure 4.4 displays the grand-grand average ERP waveforms computed across priming conditions. Components P1 (mean peak latency 102 ms), N1 (174 ms), P2 (230 ms), N2 (270 ms), and P300 (360 ms) can be identified. In addition, a frontal/fronto-central LPC component can be seen at around 700 ms post-stimulus. The data suggests ERP priming effects on N2, P300, and LPC.

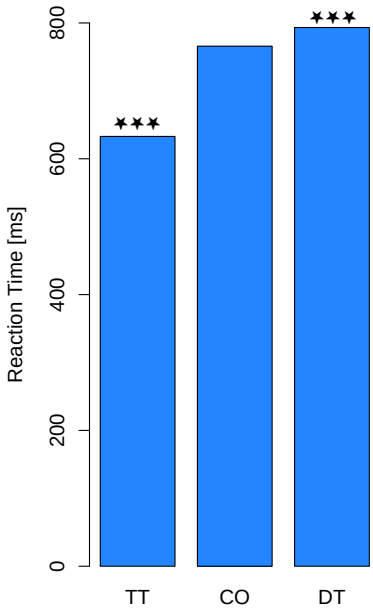


Figure 4.3: Summary of the reaction times. All three conditions differ significantly (CO vs. DT: $p < 0.001$, CO vs. TT: $p < 0.001$).

Posterior N2 is increased specifically in the DT condition. While P300 amplitude is reduced for both conditions TT and DT relative to CO, particularly over left posterior areas, amplitude of the anterior LPC increases in the order DT, CO, and TT. Moreover, LPC peaked earlier for TT compared to both DT and CO, which is not surprising due to the much shorter reaction time in TT trials. In the following repeated-measures ANOVAs are carried out for peak amplitude and latency of components P1, N1, P2, N2, P300, and LPC, employing factors priming condition (CO, DT, and TT) and Electrode, with levels of the latter factor depending on analysis (see section 4.4).

P1 In the analysis of P1 latency, only the main effect of electrode is significant, $F(16,240) = 5.0, p < 0.01; \varepsilon = 0.25$. This effect is not further explored because it was irrelevant to priming. Neither the main effect of priming condition, $F(2,30) = 0.7, p = 0.47; \varepsilon = 0.80$, nor the interaction by Electrode, $F(32,480) = 0.8, p = 0.51; \varepsilon = 0.13$, proves to be reliable. Also, the analysis of P1 amplitude does not yield any significant priming effects; main effect of priming condition, $F(2,30) = 1.7, p = 0.20; \varepsilon = 0.88$; interaction priming condition $\times$ electrode, $F(32,480) = 0.4, p = 0.84; \varepsilon = 0.14$.

N1 There is a significant priming effect on N1 peak latency, $F(2,30) = 3.7, p < 0.05; \varepsilon = 0.97$. According to Scheffe's test, in the DT condition the N1 peak is reliably delayed compared to TT (176 and 172 ms, $p < 0.05$). N1 latency in the CO condition (174 ms) is not significantly different from DT and TT conditions ($p > 0.50$). Neither the main effect of priming condition on N1 amplitude, $F(2,30) = 1.5, p = 0.23; \varepsilon = 0.93$, nor the interaction by electrode, $F(32,480) = 0.8, p = 0.56; \varepsilon = 0.19$, are significant.

P2 Also for P2 latency, a significant effect of priming condition can be established, $F(2,30) = 3.8, p < 0.05; \varepsilon = 0.79$, which is due to P2 peaking slightly earlier for DT than CO (237 and 241

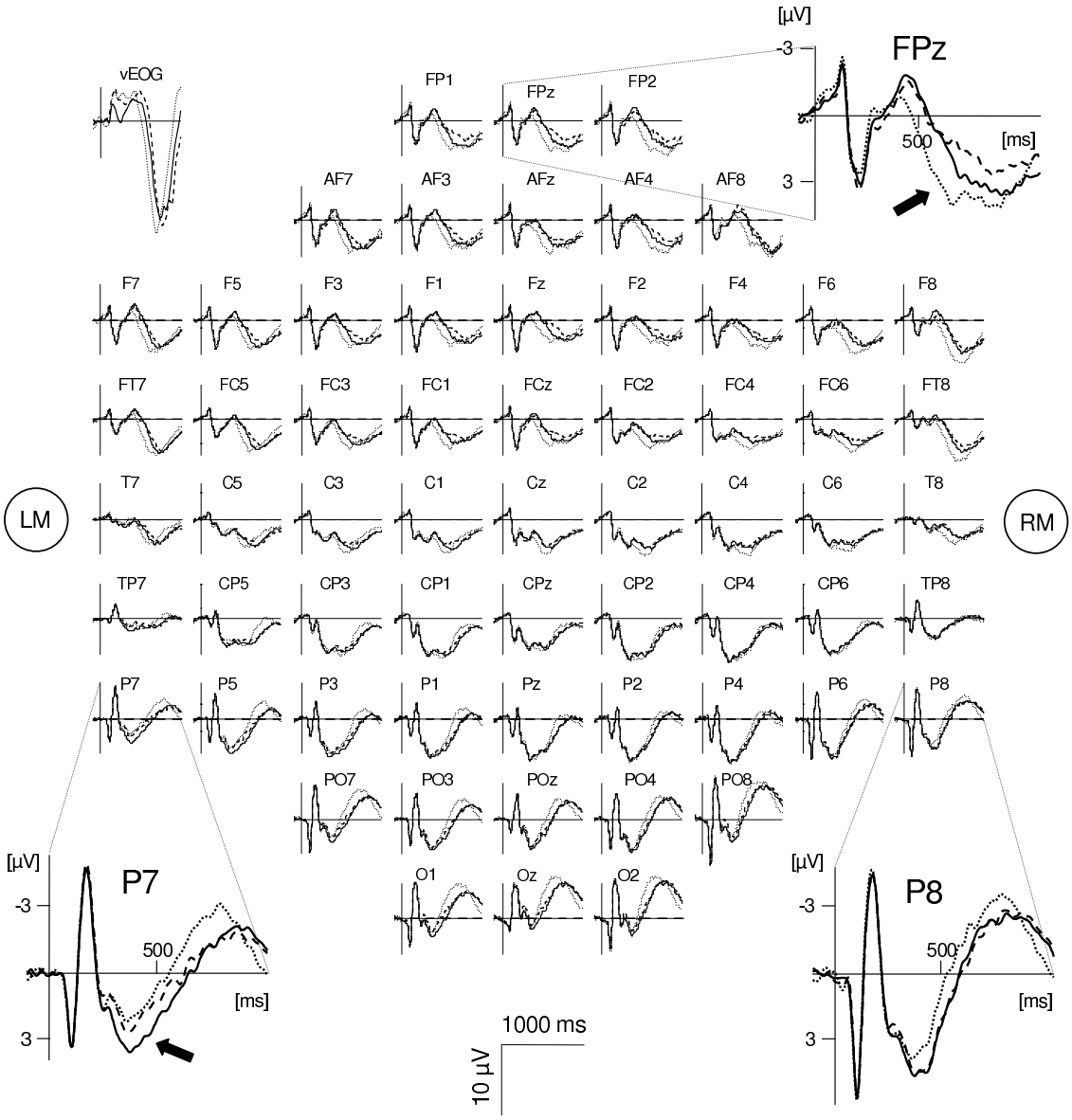


Figure 4.4: Grand-grand averaged ERP waveforms for each priming condition (CO = continuous line, DT = dashed line, TT = dotted line). Negativity is plotted upwards. Note the reduced left-posterior P300 amplitude for the DT and the TT conditions as compared to CO trials (e.g. P7). Amplitude and latency of a frontal late positive complex (LPC) between 500 and 700 ms post-stimulus were found to be sensitive to the priming manipulation. DT trials showed significantly reduced LPC amplitude compared to CO trials, while TT trials produced significantly increased amplitude (see section 4.8). In addition, the LPC peak occurred earlier for TT trials (e.g. FPz).

ms, $p < 0.05$ according to Scheffe's test). P2 latency in the TT condition (239 ms) does not differ significantly from the other two conditions ($p > 0.50$). The interaction priming condition $\times$ electrode is not significant, $F(14,210) = 1.1, p = 0.36; \varepsilon = 0.31$. Neither the main effect of priming condition on P2 amplitude, $F(2,30) = 0.4, p = 0.61; \varepsilon = 0.75$, nor the interaction by electrode, $F(14,210) = 1.4, p = 0.24; \varepsilon = 0.36$, are significant.

N2 No reliable priming effects are observed for N2 latency; main effect of priming condition, F(2, 30) = 0.7, p = 0.50; ε = 0.85; interaction by electrode, F(14, 210) = 0.5, p = 0.67; ε = 0.33. By contrast, a significant main effect of priming condition is established for N2 amplitude, F(2, 30) = 3.7, p < 0.05; ε = 0.91. N2 is significantly larger in the DT condition than in the CO condition (-0.8 μV and 0.0 μV; p < 0.05 according to Scheffe's test). Both, the differences between DT and TT and between TT and CO are not significant (p > 0.30). The interaction by electrode is not significant, F(14, 210) = 1.8, p = 0.14; ε = 0.31.

P300 The analysis of P300 latency yields a significant main effect of priming condition, $F(2,30) = 4.3, p < 0.05; \varepsilon = 0.74$. Scheffe's test revealed that the P300 peak occurs in the TT condition reliably earlier than in the DT condition (351 and 361 ms, $p < 0.05$). The difference of reaction times between DT and CO is not significant ($p = 0.83$), nor is the difference between TT and CO ($p = 0.10$). The interaction by electrode is not significant, $F(32,480) = 1.6, p = 0.16; \varepsilon = 0.20$. Also for P300 amplitude, the main effect of priming condition is significant, $F(2,30) = 4.3, p < 0.05; \varepsilon = 0.78$. As indicated by Scheffe's test, P300 amplitude is reliably larger in the CO condition (3.6 μV) than in both the TT condition (2.9 μV; $p < 0.05$) and the DT condition (3.0 μV; $p < 0.05$). TT and DT condition do not differ from each other ($p = 0.95$).

The main effect is further qualified by a significant interaction by electrode, $F(32,480) = 4.5, p < 0.001; \varepsilon = 0.19$. Since an a-priori hypothesis has been formulated based on the results by Stahl and Gibbons (2007) and Gibbons (2009), regarding left-parietal predominance of priming effects on P300 amplitude (see section 4.2), planned comparisons between priming conditions are performed for clusters of left parietal (P1, P3, P5, P7) and homologous right parietal electrodes (P2, P4, P6, P8). Over left side parietal areas, P300 amplitude in the CO condition (4.3 μV) is significantly larger than in both the DT condition (3.6 μV, t[15] = 4.1, $p = 0.001$), and the TT condition (3.5 μV, t[15] = 2.7, $p < 0.05$). By contrast, right parietal P300 amplitude does not differ significantly for DT and CO (4.6 and 4.8 μV, t[15] = 1.3, $p = 0.22$), and TT and CO (4.4 and 4.8 μV, t[15] = 1.7, $p = 0.11$).

LPC For analysis of the frontal late positive complex at around 700 ms post-stimulus, the four most anterior rows of electrodes are considered. The main effect of priming condition on LPC peak latency is significant, $F(2,30) = 10.1, p < 0.001; \varepsilon = 0.99$. According to Scheffe's test, LPC peaks significantly earlier in the TT condition (711 ms) than in both the CO condition (789 ms; $p < 0.01$) and the DT condition (769 ms; $p < 0.05$). The difference between DT and CO is not significant ($p > 0.5$). There was no reliable interaction between priming condition and electrode, $F(50,750) = 1.3, p = 0.20; \varepsilon = 0.17$. Analysis of LPC amplitude determined as mean voltage in a 200-ms window centered around the condition specific LPC peak (see section 4.5) yields a significant main effect of priming condition, $F(2,30) = 13.7, p < 0.001; \varepsilon = 0.71$. Scheffe's test reveals significantly larger LPC amplitude in the TT condition (3.6 μV) compared to both DT (2.1 μV; $p < 0.001$) and CO conditions (2.8 μV; $p < 0.05$). Also the difference between DT and CO is significant at $p = 0.05$. The interaction by electrode is not significant, $F(50,750) = 0.6, p = 0.66; \varepsilon = 0.09$.

Given substantial blink activity at around (or, immediately after) the overt response (see figure 4.4, vEOG), it is necessary to ensure that the frontal LPC effects are not merely due to differences in vertical EOG activity between priming conditions. Mean EOG activity in the 550-750 ms interval (which is exactly the time window used for LPC amplitude analysis) is compared for

CO and DT conditions. No significant difference was found, t(15) = 1.5, $p = 0.16$. To compare CO and TT conditions, the time window had to be adjusted for the TT condition (500-700 ms) to compensate for LPC latency differences, thus exactly matching the time window chosen for the TT condition in the analysis of LPC amplitude. The t test is also not significant, t(15) = 0.4, $p = 0.68$.

4.8 Discussion

The present experiment was designed to localize the negative priming effect in time and space by means of electrophysiological correlates. Concerning theoretical explanations of negative priming our main goal is to find evidence for two ERP correlates predicted from episodic retrieval. First, the finding of left-posterior P300 amplitude reduction for DT and TT conditions compared to control (Gibbons, 2009; Stahl and Gibbons, 2007) which may reflect the processing of prime-probe similarity should be replicated in a completely different task. A second late conflict-related correlate of negative priming which distinguishes the DT condition from both, TT and CO conditions was expected, and may reflect the negative priming-specific processing of conflict between retrieved prime information and information extracted from the probe. Only one such direct ERP correlate of visual identity-based negative priming has been reported before, in terms of negative priming specific amplitude enhancement of a left posterior, N400-like component (Gibbons, 2009).

Relative to other ERP studies using digit identification (Kathmann et al., 2006) or flanker tasks (Gibbons, 2009; Stahl and Gibbons, 2007), a strong advantage of the present picture-naming task is that it produces strong behavioral priming effects (see Titz et al., 2008). In the current study, mean reaction time was delayed by 28 ms in DT trials and accelerated by 133 ms in TT trials, compared to control. Because both TT and DT effects were sufficiently large, substantial differences in the corresponding ERPs can be expected.

The observed ERP correlates of negative priming mainly concerned two components, P300 and frontal LPC. Two other findings were not followed up further: a small increase in DT of the N1 latency by 4 ms compared to TT, which neither can account for the 28-ms behavioral negative priming effect nor for the 161-ms reaction time difference between DT and TT conditions; and a small N2 amplitude increase for the DT condition relative to control without significant difference of DT and TT. Thus, in contrast to the studies by Daurignac et al. (2006) and Frings and Groh-Bordin (2007), no significant negative priming specific effects on early ERP components were observed which were assumed to support distractor inhibition. In the following, possible explanations for the difference in P300 and LPC amplitudes depending on priming conditions and their theoretical consequences are discussed in more detail.

The present study provides a successful replication of reduced left-posterior P300 in both DT and TT trials by Gibbons and Stahl (2007) and Gibbons (2009). Since their Eriksen flanker task is rather different from the present task employing overlapping objects, the convergent results regarding priming effects on P300, even with respect to the left posterior scalp topography, are all the more remarkable. Note that the present left posterior P300 effect was observed in a (300-500 ms) time window, whereas Gibbons (2009) and Gibbons and Stahl (2007) analyzed an earlier time window (300-400 ms). Given the mean reaction times of the present task (650-800 ms) and the two earlier studies (500-650 ms), it seems likely that functionally the same processes were involved in P300 reduction in both tasks. Therefore, following Gibbons (2009) we interpret TT and DT related reduction in P300 amplitude as reflecting processes related to perceived prime probe similarity corresponding to a retrieval cue according to episodic retrieval (Neill and Valdes, 1992; Rothermund et al., 2005).

Repetition usually triggers larger ERP positivity (see Rugg, 1995, for an early review). However, the repetition effect typically has its maximum at around 600 ms, as opposed to the present joint positive priming/negative priming effect on P300 around 400 ms. There are at least two other striking differences between the present priming task and the repetition paradigm: First, in the present experiment all stimuli are repeated over and over, whereas in repetition studies first and second presentations of a stimulus are compared, with lags often ranging from many seconds to several minutes. Second, unlike the present task repetition experiments do not involve selection; typically, on each trial one single stimulus is presented. The present P300 component therefore seems to be rather different from the later positive ERP components observed in ERP repetition studies.

We did not find differences in P300 peak latency between DT and CO conditions, suggesting similar speed of stimulus evaluation processes in these two priming conditions. When the relatively strong behavioral negative priming effect is considered, in line with Stahl and Gibbons (2007) it can be argued that at least one additional process in the post-P300 time range has to be assumed to explain behavioral negative priming.

The present study provides a basis for a consistent explanation of the reaction time differences between priming conditions. We found a reduced amplitude and (qualitatively) delayed peak latency of the frontal late positive complex (LPC) for the DT condition, as well as increased LPC amplitude and reduced LPC latency for the TT condition, both compared to control. The effects were widely distributed over the anterior scalp, including the first four rows of electrodes (frontopolar to fronto-central). Note that LPC priming effects were analyzed between 500 and 700 ms, but frontal ERP differences-to-control already started at around 380 ms for TT and 520 ms for DT (see figure 4.4). Given that mean reaction times ranged from 632 ms (TT condition) to 793 ms (DT condition), the LPC effects may well reflect processes that were responsible for priming effects on reaction time, and not merely consequences of the mean reaction time differences between conditions.

Frontal brain activity is known to be particularly present in situations where an individual's behavior is not simply controlled by stimulus-response relationships, but requires careful selection from different response options (see Ridderinkhof et al., 2004, for a review). One possibility is to explain the present LPC differences in terms of the necessary amount of cognitive control. Processes related to cognitive control are often observed as late range frontal ERP negativity in situations of conflict processing (e.g. Krigolson and Holroyd, 2007; Lorist et al., 2000; West et al., 2004; Yeung et al., 2004). Control processes should be strongest in the high conflict DT condition, thereby reducing and/or delaying the ERP positivity that normally indicates the completion of trial processing. In contrast, in the low conflict TT condition no strong cognitive control is required, causing earlier and stronger ERP positivity.

Response retrieval theory (see section 2.4.3) can well explain negative priming in terms of late processing conflicts. In case of TT trials, both prime response retrieval and probe processing converge at the same response, which causes behavioral positive priming and, moreover, should be accompanied by a very low conflict. No cognitive control is required. This can explain particularly strong and early frontal ERP positivity in the TT condition, as positive ERP deflections typically occur when mental resources are updated when they are no longer needed for ongoing information processing.

In contrast, in the DT condition retrieval of the prime response and algorithmic probe processing activate different responses. Cognitive control processes become necessary, resulting in late-range frontal brain activation. This type of brain activity typically manifests in ERP negativity. Superimposed on positive components indicating the completion of stimulus analysis, the negativity is observed as an amplitude reduction or delayed latency (see Kok, 2001, for the general argu-

mentation). In sum, the observed differences in frontal LPC amplitude between DT, CO, and TT conditions are well in line with response retrieval theory.

A second explanation for the present LPC priming effects may directly refer to memory processes. Mayr et al. (2003) and Mayr and Buchner (2006) observed negative priming related reduction of posterior LPC amplitude in the auditory domain. The authors interpreted their finding as support for episodic retrieval mechanisms and argued as follows: familiarity of a stimulus is reflected in larger posterior LPC, compared to novel stimuli. If however negative priming is accompanied by reduced LPC amplitude, this may indicate that former distractor stimuli, when repeating as targets in the DT condition, are in some sense less familiar than the novel targets in the CO condition. This below baseline familiarity of negative priming targets may then result in less efficient processing of these stimuli, causing the negative priming effect. In principle, a similar mechanism may also be responsible for our frontal LPC reduction. The different topographies of the two effects might be explained in terms of modality differences (visual vs. auditory) and/or differences in task demands (i.e. naming a visual object vs. classifying a sound by means of a button press). Note that there is evidence in the literature that visual memory tasks indeed may show frontal LPC effects (e.g. Hayama et al., 2008; Wilding and Rugg, 1996). The effects on frontal LPC presented here are therefore well in line with the assumption that familiarity of the probe target in the TT condition is particularly high and that the familiarity of a recently ignored target seems to be below baseline.

To conclude, there seems to be little evidence in the present data for an interpretation in favor of distractor inhibition theory. According to the literature, persisting inhibition can be expected to operate in the N200 time range (Daurignac et al., 2006; Frings and Groh-Bordin, 2007) where we find no significant negative priming correlates. P300 amplitude is often understood as an index of mental effort (e.g. Ullsperger et al., 1988; Kok, 2001). Thus, persisting inhibition as a cause for negative priming would produce larger P300 amplitude for DT compared to TT trials. Reduced P300 amplitude in TT trials would reflect facilitation of processing due to persisting activation of internal representations of the target stimulus. Increased P300 amplitude in DT trials would reflect greater mental effort, as the activation of DT targets is still below baseline, due to persisting inhibition. As noted earlier, this pattern of P300 amplitudes is not present in the current data.

Episodic retrieval therefore seems most appropriate to interpret our ERP correlates of negative priming in a picture-naming task, as it can explain the reduced P300 amplitudes in both conditions DT and TT, and the modulation of LPC amplitude. Prime-probe similarity which is reflected in smaller P300 amplitude, may serve as cue for the retrieval of prime information (Neill and Valdes, 1992). As explained above, the retrieved prime information then has opponent effects on further processing in TT and DT trials. This is consistent with our observation that up to the P300 time range ERPs do not reveal any major differences in the processing of DT and TT trials. Hence, the present negative priming effect seems to originate at a later, post-P300 stage of processing.

A cautionary note concerns the fact that the three priming conditions relevant to the present study (TT, DT, and CO) were randomly presented among other trial types not containing distractor objects. These trials were included to answer a separate research question not relevant to the present study. Their presence, however, may have affected possible strategies used by the participants and, thereby, processes related to negative priming. For example, the frequent inclusion of single-target trials may have increased the saliency of a distractor. We assume the target-only trials to have strengthened the negative priming effect, which is consistent with the aim of the present study, i.e. to increase the chance to find ERP correlates of negative priming by employing a task known to produce strong negative priming.

4.9 Conclusion

One of the two major results of our study, the smaller left-posterior P300 amplitude for DT and TT trials compared to the CO condition, represents a conceptual replication of the results of Stahl and Gibbons (2007) and Gibbons (2009) in a classical picture-naming priming task. Thus, the processing of prime-probe similarity is important for visual negative priming in general and is not restricted to a certain task. In addition, the reduced LPC amplitude in DT trials shows more effortful processes related to cognitive control and memory retrieval. The present negative priming effect on frontal LPC amplitude has been observed for the first time. It points to a specific feature of the picture naming task, which is more difficult than e.g. the flanker task, and produces long reaction times and strong behavioral negative priming effects. Both these aspects may provide good conditions to identify late range ERP correlates of negative priming. Overall, our results favor a retrieval-based explanation of negative priming.

4.10 Summary

Electroencephalographic recordings are a tool well suited to assess the time course of negative priming due to the high temporal resolution. The obtained EEG data is extremely noisy as the signal common to all similar trials is embedded in the prominent and not stimulus locked ongoing activity in the brain. Also geometric, biological and technical issues deteriorate the recorded signal. The common way to deal with the bad signal to noise ratio is the averaging of the EEG signal over all trials of the same experimental condition, after the data has passed some cleaning routines. The differences of the obtained signals, the so-called averaged event related potentials, are subject to interpretation according to their latency, amplitude and localization on the skull. Due to the specialty of negative priming, not many EEG correlates of the effect have been identified so far. Even between those few no clear picture can be drawn.

We presented a study that adapted a traditional picture-naming paradigm to EEG recordings. We found a less positive P300 component in both DT and TT compared to control, which points to a similar processing of trials which contain repeated stimuli at early stages. A distinction between positive and negative priming, separated by the control condition as it is observed in the reaction times, was present in the late positive complex, which indicates little cognitive control for TT trials, intermediate effort for control trials and strong cognitive intervention in DT trials. An interpretation in favor of retrieval theories is obvious, as they claim the resolution of a conflict between current and retrieved information in DT trials. And a benefit of the retrieved information in TT trials.

5 Interlude: Advanced EEG Analysis

During the occupation with EEG experiments, we were unsatisfied with several aspects of traditional EEG analysis. As a side project we developed enhanced techniques to extract information from EEG data. As the resulting algorithms produce much more prominent components than simple point-wise averaging, the benefit of our methods to access the time course of negative priming is obvious. Nevertheless our methods are not yet sufficiently established to make accredited statements about the processing in a negative priming trial. For this reason we call the current chapter *interlude* and refrain from using the methods in the other chapters.

Our main complaint concerning the point-wise averaging of ERPs is the ignoring of the variability of reaction times. Primarily later in a trial the variations cause different cognitive processes to enter he average simultaneously. If we assume a common series of underlying cognitive processes in all trials of one experimental condition, we should be able to adjust the time flow of the individual trials before averaging, such that similar processes are matched temporally when averaging. We find that a natural assumption about processing speed yields a complex but nevertheless robust algorithm for the analysis of electrophysiological data. Results from artificial and real ERP-data show that our algorithm outperforms existing solutions.

We will point out the characteristics of EEG data in section 5.1 and derive realistic models for event-related potentials in section 5.2. The main technique of our approach, dynamic time warping, is reviewed in section 5.3. The recursive application of the technique is presented in section 5.4. The obtained algorithm also provides a metric for event-related potentials which can be used for clustering as described in section 5.5. An integration of additional time markers as described in the next two chapters is also possible as shown in section 5.6. Finally we develop the construction of the time warping function by the means of recurrence plots and show the resulting performance in section 5.7. The present chapter appeared as two successive contributions to the conference IDEAL (Ihrke et al., 2008, 2009a) and was also part of a book on coordinated activity in the brain (Ihrke et al., 2009b). The presented algorithms are available for download from the project page `http://libeegtools.sf.net`.

5.1 EEG Analysis in Cognitive Research

In electroencephalographic (EEG) data, noise levels of -25 dB are standard (Flexer, 2000), and for electromyography (EMG) or functional magnetic resonance imaging (fMRI) the situation is similar. The arising problem of the recovery of relevant information from such data has been dealt with extensively (Whalen, 1971; Castleman, 1996). It seems reasonable to exploit intrinsic structures in the data, i.e. to identify patterns in the data that reoccur under specific conditions, e.g. at the onset of a stimulus or in relation with other events in the course of the experiment.

A straight-forward solution consists of averaging single trial event-related potentials (ERPs) in order to obtain an averaged ERP (AERP) that is expected to be comparable across different experimental setups (Picton et al., 1995), see chapter 4. The reliability of the AERP allows in turn the identification of characteristic features of the time course of the signal such as the latency and amplitude of major minima and maxima. We will discuss a number of algorithms which are

theoretically justified, and have proven useful in an experimental project (Ihrke, 2008). Systematic changes in the AERP components between different experimental conditions are consistent with the hypothesis that ERP components reflect stages of information processing in the brain. The idealized noise-free ERP represents the signal of interest, whereas variability across trials is considered as noise. We assume the EEG signal to comply with the following:

1. It contains relevant aspects of neural activity.
2. Task-specific activations form a significant fraction of the EEG signal.
3. The brain solves similar tasks in a similar way.

The signal can thus be defined as a minimal variance curve within the data set obtained for many repetitions of the same task. The axioms imply that variations due to external conditions should be excluded, and that the external conditions and even the state of the subject should be kept as constant as possible for all trials. Yet, data mining techniques reveal that for comparable data only a fraction of 60% of the pooled epochs contribute to the AERP waveform while the other 40% just increase the variance (Haig et al., 1995).

Thus, it cannot be decided unambiguously whether the variability of the ERPs is caused by the stochastic nature of the underlying neural dynamics or by the application of different strategies to a task. The comparison of two single trial ERPs and their AERP, see figure 5.2, points already to a basic problem: Simple averaging will particularly deteriorate late components of the ERP which makes their interpretation difficult, see figure 5.3.

5.2 Models for Event-Related Potentials

In psychological studies, the signal-to-noise ratio (SNR) of EEG data is typically enhanced by combining data epochs in a pointwise average

$$\langle s_i(t)\rangle_i = \frac{1}{N}\sum_{i=1}^{N} s_i(t) \qquad i = 1,\ldots,N. \tag{5.1}$$

Here $s_i(t)$ is the measured signal in trial $i = 1,\ldots,N$ at time t. The signal-plus-noise (SPN) model (Truccolo et al., 2002) or fixed-latency model (de Weerd, 1981) underlying the point-wise average assumes the following.

1. Signal and noise add linearly.
2. The signal is identical in all trials.
3. Noise is a zero-mean random process drawn independently for each trial.

Assuming additive noise of zero mean, i.e. $\langle \varepsilon(t)\rangle = 0$, we can represent the data by

$$s_i(t) = s_0(t) + \varepsilon_i(t), \tag{5.2}$$

where $s_0(t)$ denotes the signal that is to be recovered from $s(t)$. Under the above conditions the pointwise average is an unbiased and optimal estimate in the mean-square error sense. It has been argued on theoretical grounds, that an improvement beyond pointwise averaging is not possible if no a priori knowledge about the characteristics of signal and noise is given (Nagelkerke and Strackee, 1979).

The validity of the SPN model is challenged by the analysis of the residuals ζ_i^{avg} obtained by subtracting the mean from the raw data

$$\zeta_i^{\text{avg}}(t) = s_i(t) - \langle s_i(t)\rangle_i, \tag{5.3}$$

which is expected to be just the noise $\varepsilon_i(t)$. The fact that the repetition of a task is typically accompanied by coherent ongoing neural activity (Truccolo et al., 2001) can be analyzed as follows. Given the SPN model, ζ_i^{avg} should not contain any event-related modulation because the noise is assumed to be independent and identically distributed. Therefore, statistical coherence measures such as the autocorrelations

$$(\zeta_i^{\text{avg}} \star \zeta_i^{\text{avg}})(\tau) = \int \zeta_i^{\text{avg}}(t)\zeta_i^{\text{avg}}(t+\tau)\,dt \tag{5.4}$$

and the spectral densities

$$\text{PSD}(\zeta_i^{\text{avg}}) = \mathcal{F}\{\zeta_i^{\text{avg}} \star \zeta_i^{\text{avg}}\} \tag{5.5}$$

computed on ζ_i^{avg} should not show any event-related modulation (i.e. a flat spectrum and cross correlations that behave like a δ function are to be expected). Empirical evidence shows that these assumptions are violated for real data (Truccolo et al., 2001, 2002), see an analysis of data from our experiment described in chapter 4 shown in figure 5.1.

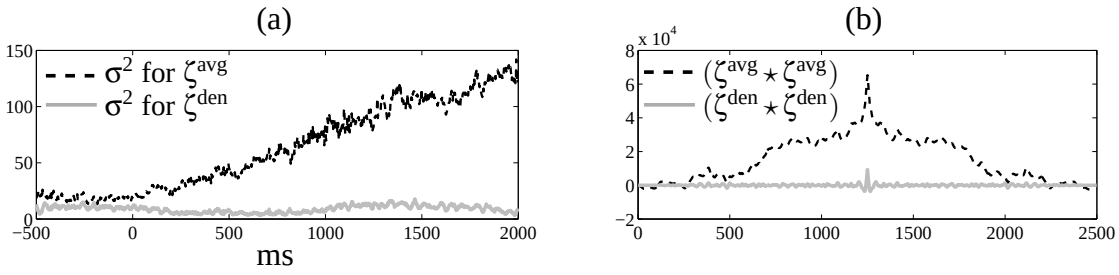


Figure 5.1: Coherence measures computed on residuals ζ^{avg}. (a) The variance σ^2 over trials shows event-related modulation for the residuals after subtracting the average. Given the SPN model, we expect a flat curve as obtained from computing σ^2 on the single-trial denoised residuals ζ_i^{den}. (b) Crosscorrelation computed on the residuals for a sample trial. Again, unexpected (from SPN) correlations show up for ζ_i^{avg} whereas the function approximates a δ-function for the denoised single trial residuals.

Extending the SPN model, the variable latency model (VLM) (Woody, 1967) introduces a trial-dependent scaling factor α_i and a time lag τ_i

$$s_i(t) = \alpha_i s_0(t+\tau_i) + \varepsilon_i(t). \tag{5.6}$$

One possibility to obtain the time lags τ_i is the maximization of the crosscorrelation

$$\tau_i = \arg\max_t (\langle s_i \rangle_i \star s_i)(t) \tag{5.7}$$

between the data and the pointwise average $\langle s_i \rangle_i$. After the shift according to τ_i, the data can be interpreted by the SPN model. However, in an empirical evaluation of analytically derived predictions of this model, patterns that are not consistent with the predictions were found (Truccolo et al., 2002). Therefore it is reasonable to go back to Ciganek (1969), who already showed that the intertrial variability of the evoked potential can go beyond the simple time shift. A more general model for order-preserving time warping of the data is given by

$$s_i(t) = \alpha_i s_0(\phi_i^{-1}(t)) + \varepsilon_i(t), \tag{5.8}$$

where ϕ_i are monotonous functions that map the time scale of the individual trials to that of a template. The functions $\alpha_i > 0$ determine the local scaling of the curve. The advantage of this variable-components-plus-noise model (VCPN) is illustrated by figure 5.2. The VCPN model (5.8)

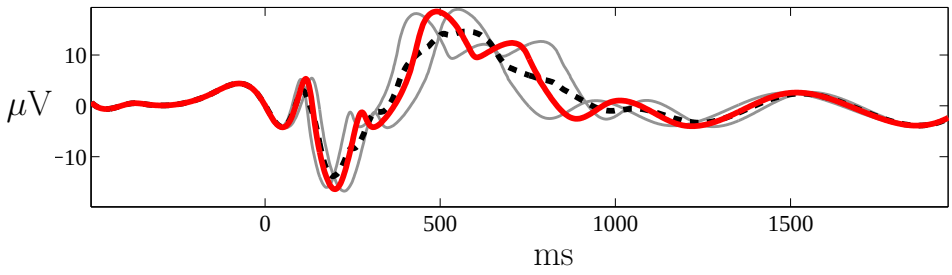


Figure 5.2: Smearing of components in the simple average (black dashed) due to temporal variance. Two input signals (gray curves) were simulated according to the model in equation 5.8. An averaging procedure incorporating temporal variance (red) .

takes into account temporal variations of the individual signals in addition to differences in scale, and is thus able to identify systematic distortions due to single-trial fluctuations that otherwise are averaged out.

We use two datasets, one containing artificial data the other one consisting of real EEG data. Real data reported in this chapter is taken from the study reported in chapter 4. The artificial data was generated according to the VCPN model introduced in equation 5.8.

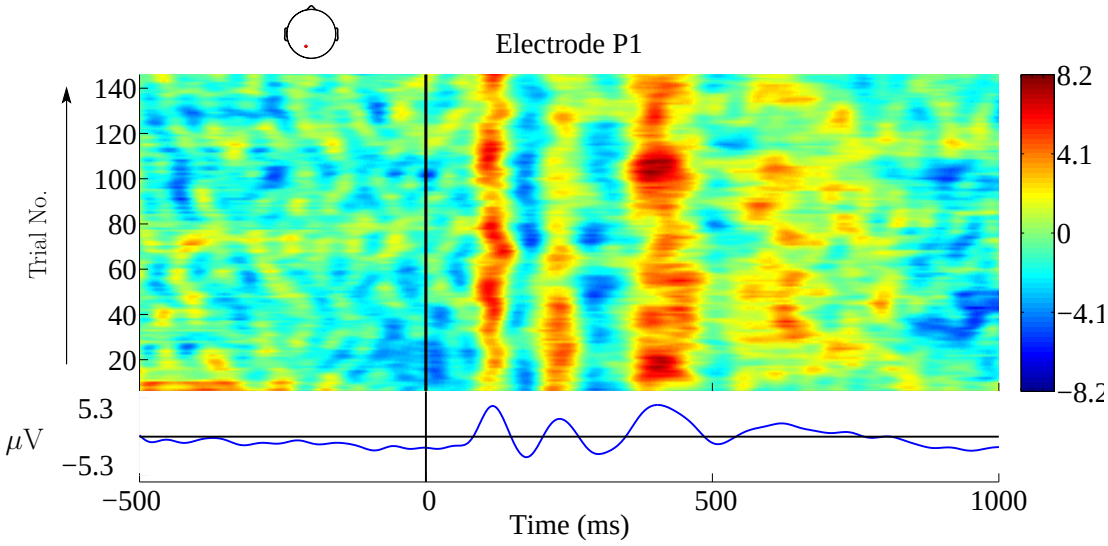


Figure 5.3: Real data ERPs for 150 trials (color coded) and their average (bottom). With growing number of trials (i.e. time spent in the experiment) the shape of the ERP is subject to changes. E.g. while the P2 is very pronounced for the first 50 trials, its amplitude is decreased later. The N2 amplitude decreases over time and is missing completely in some trials. A general shift in the pattern with growing number of trial is observable (P2/N4 amplitude).

For any two trials s_i and s_j, we express the problem of optimizing the individual warping function $\phi_{i\to j}$ in terms of an energy function

$$E(\phi) = \int |s_i(\phi_{i\to j}(t)) - s_j(t)|^2 \, dt. \tag{5.9}$$

Minimizing equation 5.9results in the function $\phi_{i\to j,\,\mathrm{opt}}$ such that the difference between the warped s_i and s_j in Euclidean distances is minimized.

5.3 Dynamic Time Warping

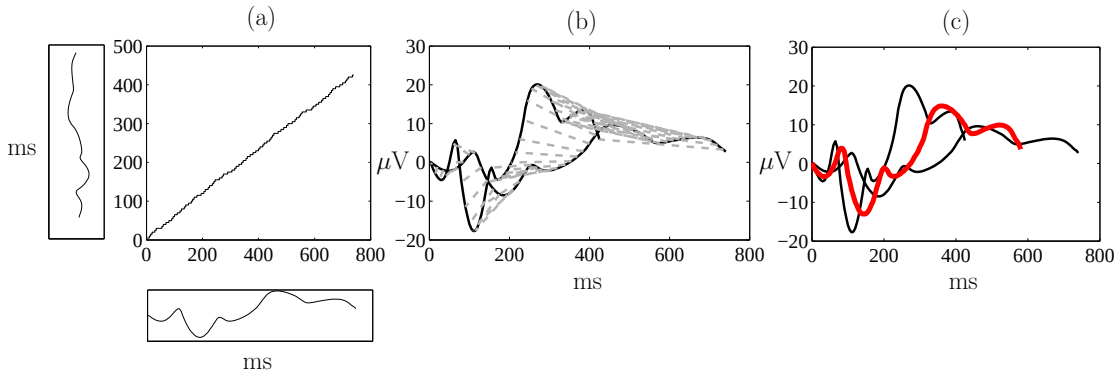


Figure 5.4: Dynamic time warping. (a) Optimal path p_i through the cost-matrix $\mathbf{D}_{jk}$ for two signals (black curves) from (b) and (c). Note the different scales of the x and y axis. (b) Illustration of DTW matching corresponding points in s_i and s_j. (c) An average produced by ADTW (red).

Dynamic time warping (DTW), see figure 5.4, will be shown to provide a distance measure that directly implements our assumptions. The method was first used in speech analysis (Myers and Rabiner, 1981), appears to present a more suitable approach to EEG analysis. DTW aims at matching two trial time courses onto one another, see figure 5.4b). First, a pointwise dissimilarity measure between two signals s_i, s_j is defined

$$d(s_i(t_i), s_j(t_j)) := |\tilde{s_i}(t_i) - \tilde{s_j}(t_j)| + |\tilde{s_i}'(t_i) - \tilde{s_j}'(t_j)|, \tag{5.10}$$

where

$$\tilde{s}(t) := (s(t) - \langle s(t)\rangle_t)\langle s(t)^2\rangle_t^{-1/2} \tag{5.11}$$

is the normalized signal and s' the first derivative of s. The distance (5.10) gives rise to derivative DTW (Keogh and Pazzani, 2001) because it is based on amplitude and slope of the signal.

An optimal map $\phi_{i\to j,\mathrm{opt}}$ is determined by a path p_k that satisfies recursively

$$\text{if } p_k = (l,m) \text{ then } p_{k+1} \in \{(l+1,m),(l,m+1),(l+1,m+1)\} \tag{5.12}$$

and minimizes the sum over the corresponding elements d_{lm} of the dissimilarity matrix (Picton et al., 1988)

$$\mathbf{d}_{lm} = d(s_k(l), s_k(m)). \tag{5.13}$$

This path can be found by backtracking through the cumulated cost matrix

$$\mathbf{D}_{lm} = \mathbf{d}_{lm} + \min\{\mathbf{D}_{l,m-1}, \mathbf{D}_{l-1,m}, \mathbf{D}_{l-1,m-1}\}, \tag{5.14}$$

i.e. via the minimum of the downward, right, and down-right neighbors, from $\mathbf{D}_{L,M}$ to $\mathbf{D}_{1,1}$. The final element $\mathbf{D}_{L,M}$ constitutes a measure for the dissimilarity of the two curves based on their overall shape, due to the cumulation.

In order to recover the original signal, we use the obtained path to simultaneously average two trials in time and amplitude by

$$\begin{pmatrix} \hat{t} \\ \hat{s} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}[t_i + \phi(t_i)] \\ \frac{1}{2}[s_i(t_{(i)}) + s_j(\phi(t_i))] \end{pmatrix} \tag{5.15}$$

and find the required sampling points by linear interpolation. N signals are averaged by subsequent application of (5.15) to trials randomly drawn from $\{1,\ldots,N\}$.

It is further possible to introduce constraints on the DTW method that penalize path deviations (Picton et al., 1995) from the main diagonal, thereby reducing the bias on the cost of an increased variance. Before applying the DTW algorithm, it should be ensured that trials are sufficiently similar to each other, e.g. by applying time warping only within distinct clusters, see section 5.5. In the next section, we propose to apply external time markers that can act as an objective reference for trial matching and advanced averaging.

5.4 Pyramidal Averaging Dynamic Time Warping

For N trials, a straightforward solution proposed by Picton et al. (1988) is to simply combine pairs of single-trial ERPs using ADTW. In a next step, pairs of the results from this combination can be averaged again and the entire process iterated until only one average is left. We proceed recursively, namely

$$\mathrm{D}^{\mathrm{ADTW}}\{s_1,\ldots,s_{2N}\}(t) = \mathrm{D}^{\mathrm{ADTW}}\left\{\mathrm{D}^{\mathrm{ADTW}}\{s_1,\ldots,s_N\}, \mathrm{D}^{\mathrm{ADTW}}\{s_{N+1},\ldots,s_{2N}\}\right\}(t). \tag{5.16}$$

An extension of DTW incorporates information about latency variability by hierarchically choosing pairs of similar trials before averaging. The cumulated path coefficient obtained from DTW is used as a measure for the dissimilarity of two time courses, see figure 5.6. Selecting the minimum element from the resulting difference matrix $\Delta_{ij} = \mathrm{DTW}\{s_i, s_j\}$ by $(i,j) = \arg\min_{(j,k)} \Delta_{jk}$. Now the minimal-dissimilarity trials are combined and row i and column j are removed from Δ_{jk}. This procedure is iterated until the matrix is empty. The complete process is repeated with about half the number of pairwise averaged trials. The entire tree-like process is continued until all trials are merged. We refer to this method as pyramidal ADTW (PADTW) because of the successive subdivision of the set of trials. For realistic data, the procedure performs substantially better than the DTW algorithm described in the last section, see figure 5.5 and (Ihrke, 2007).

5.5 Trial Clustering for Cleaner Averages

Unsupervised classification techniques can help to identify ERPs that were generated by distinct processing mechanism in the brain. The temporal variance introduced in this way, however, is not

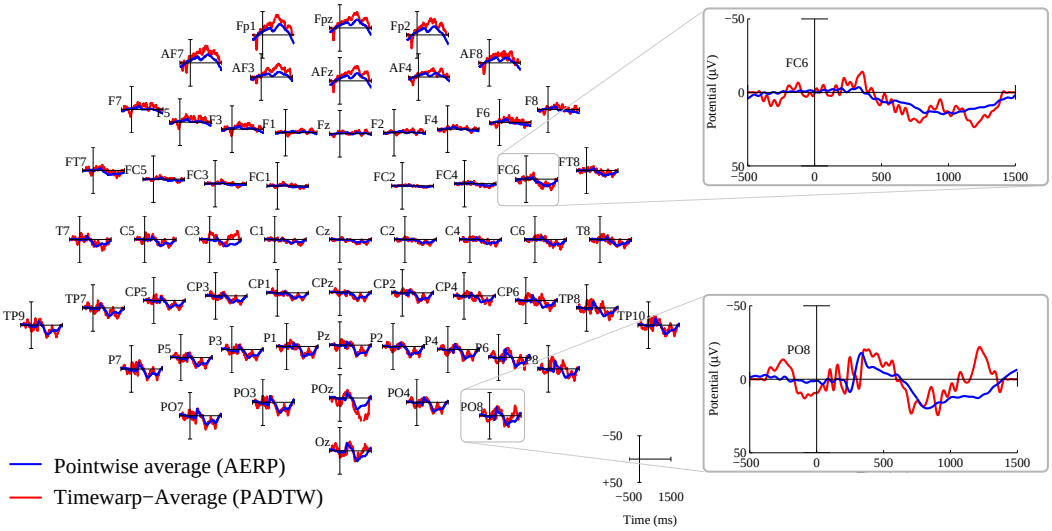


Figure 5.5: Comparison of the traditional point-wise average and the PADTW algorithm for real data. Notably PADTW does not flatten late components, but still shows pronounced peaks. It also shows fast oscillations which are not present. Experimental validation has to show whether the oscillations are present in the data, or if they are produced by the attempt of PADTW to match peaks in the noise of the signal.

resolved by the VCPN model, equation (5.8). Averaging techniques should therefore be applied selectively to trials within distinct clusters.

Selective averaging schemes use only specific episodes for averaging (Basar et al., 1975) in order to exclude artefacts such as muscular activity. In order to reduce visual inspection of the data it is possible to assist the selection process by clustering (Lange et al., 2000). An inherent problem of any clustering algorithm is the decision about a sensible number of clusters. If a specific experimental design allows a theory-driven estimate of that number it is of course to be preferred. Otherwise, strategies based on within-cluster scatter coefficients can be applied (e.g. Tibshirani et al., 2001).

5.6 Enhancing Averaging by Integrating Time Markers

Some experimental setups suggest an alignment of the data with respect to response markers, i.e. instead of stimulus locked (sERP) now response locked (rERP) ERPs are used. If such information is available then a more reliable time order is achieved for the effects, in particular relatively late after the stimulus. Gibbons and Stahl (2007) propose to stretch or compress the single-trial signals in order to match the average reaction time by moving the sampling points in time according to a temporal low-order power law.

From this formulation, it follows naturally that not only two but arbitrarily many time-markers can be integrated to guide the formation of the average. This approach has been used in (Ihrke, 2007), where the onset of an eye movement served as an additional marker.

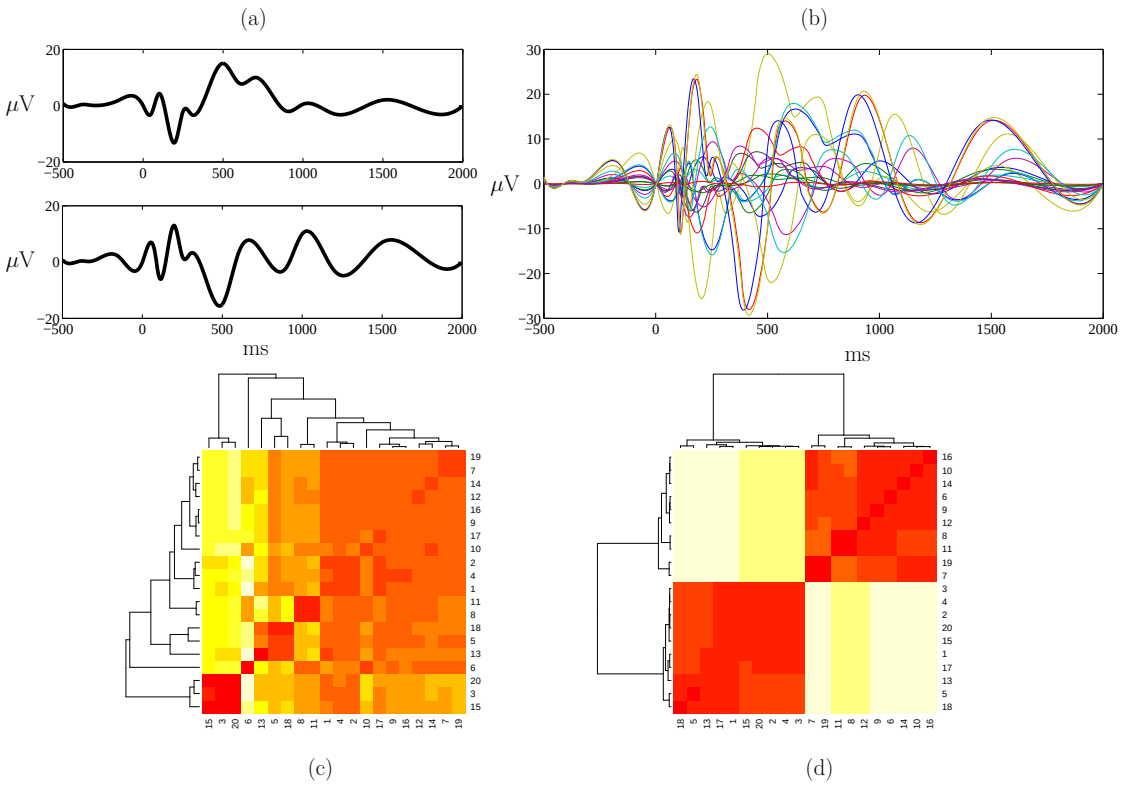


Figure 5.6: Cluster analysis of denoised single-trial ERP data. (a) Two template trials were used to derive the single-trial instances in (b) according to equation 5.8. (c) Heat map based on Euclidean distances, (d) same for DTW. While the DTW metric correctly classifies all trials to be generated by one of the templates in (a), the Euclidean fails to do so in several instances.

5.7 Recurrence Plots to Obtain the Warping Function

We propose to utilize recurrence plots (Eckmann et al., 1987), a tool from nonlinear dynamic system theory, to minimize equation 5.9 and synchronize single-trial ERP data.

5.7.1 Recurrence Plots

Recurrences, i.e. points in state space that are visited more than once, are fundamental characteristics of many dynamic systems, ranging from the financial market to epidemics and brain dynamics (Marwan et al., 2007). In order to visualize and quantify such recurrences, Eckmann et al. (1987) introduced the notion of recurrence plots. Considering a general dynamic system

$$\frac{d\vec{x}(t)}{dt} = \vec{f}(\vec{x}(t))$$

its recurrence plot is defined as

$$R^{\vec{x}}(t_1,t_2) = \Theta(\varepsilon - ||\vec{x}(t_1) - \vec{x}(t_2)||) \tag{5.17}$$

where Θ is the Heaviside step function. That means, $R^{\vec{x}}$ is a binary matrix, that is equal to one whenever the transition in phase-space approaches a previously visited point and falls within an ε-ball of the previous vector. Such a plot (e.g. figure 5.10a) provides the educated observer with a powerful facility to assess important properties of the underlying system (e.g. stationarity, cyclicism, laminar states; see Marwan et al., 2007, for a review).

In previous studies, recurrence plots have already been applied to ERP-data analysis (Marwan et al., 2007; Jeong et al., 1998; Marwan and Meinke, 2004) but mostly in order to make statements about the nature of the underlying dynamic system in the brain. In the current work, we propose to use a detection of the line-of-synchrony based on cross-recurrence plots to synchronize individual trials prior to averaging, thus reducing the bias due to timing variability of the curve.

5.7.2 Phase-Space Reconstruction

In empirical studies, it is almost always impossible to directly measure the different variables of a dynamic system. For example, it is impossible to simultaneously record the activity of all nonlinear devices (i.e. the neurons) that produce a recorded EEG-potential. The only access to the dynamics of the system is therefore the time-series of the EEG. Packard et al. (1980) has shown that reconstruction of the phase space vectors $\vec{x}(t)$ is possible using the time delay method, where

$$\vec{x}(t) = \sum_{j=0}^{m} s(t - j\tau)\vec{e}_j \tag{5.18}$$

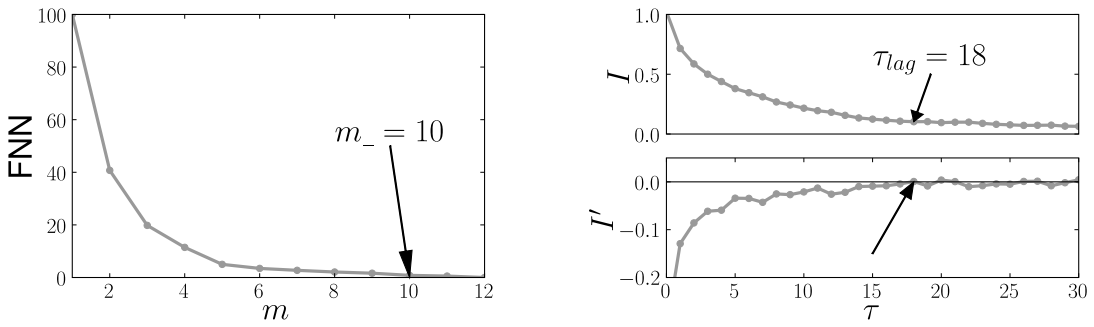
$\vec{e}_j$ being the unit-vector in j-direction. However, this method is strongly dependent on the choice of the embedding dimension m and the time-lag τ (which is often chosen as a multiple of the sampling time Δt). Finding an optimal choice for the two parameters is not trivial and has been subject to intense research (Kennel et al., 1992; Fraser and Swinney, 1986). We use the false nearest-neighbors method proposed in (Kennel et al., 1992) to find the minimal embedding dimension $m\_$ where "false" neighbors (due to a too small embedding dimension) are detected using the fact that they disappear when going from m to $m+1$. We applied this method to our data and estimate $m\_ = 10$ from figure 5.7a, which is similar to earlier results for ERP data (Jeong et al., 1998). To estimate the time lag τ, we use an information theoretic approach (Fraser and Swinney, 1986) where the first local minimum of the mutual information serves as the criterion to determine the time lag, see figure 5.7b.

Finally, the neighborhood criterion ε from equation 5.17 needs to be determined. We apply a fixed amount of neighbors (FJARN) criterion in order to determine ε on a per-point basis for each of the points t on the trajectory (Marwan et al., 2007). That means, that $\varepsilon(t)$ is determined such that each point has the same amount of neighbors. This has the advantage of eliminating the impact of scale difference in the applied curves such that no normalization of the amplitudes is necessary.

5.7.3 Line-of-Synchrony Detection in Cross-Recurrence Plots

An extension to recurrence plots are the so-called cross-recurrence plots (CRPs) in which two signals are simultaneously embedded into the same phase-space (Marwan et al., 2007). Therefore, equation 5.17 is modified as

$$\vec{C}^{\vec{x},\vec{y}}(t_1,t_2) = \Theta(\varepsilon - ||\vec{x}(t_1) - \vec{y}(t_2)||), \tag{5.19}$$



(a) (b)

Figure 5.7: (a) Percentage of false nearest neighbors (ANN) as a function of embedding dimension m. After $m = 10$, no false nearest neighbors are detected. (b) The first local minimum of the mutual information I estimates the time-lag τ (first zero-crossing of the first derivative of I).

where $\vec{x}$ and $\vec{y}$ are the phase-space trajectories of the two signals. It has been shown, that CRPs of similar signals with varying time-scale, show a "distorted main diagonal", the line-of-synchrony (LOS), that can be extracted and used for resynchronization (Marwan et al., 2002).

This approach has been applied to geological data (Marwan et al., 2002) and could be shown to yield results comparable to manual tuning. The algorithm used in this study applies a growing window strategy, where a window is set around a detected recurrence point that grows in x or y direction until no more recurrences are found. The method is parameterized by d_x and d_y, that indicate the maximum size of the window in x and y direction, respectively. It has been shown, that this algorithm works very well in a setting where only minor deviations from the main diagonal are expected. However, applying the method to EEG-data is a somewhat more delicate undertaking, because (i) the initial SNR is extremely low, (ii) the temporal variability can be quite large (Ihrke et al., 2008) and (iii) the underlying signal can be different. In behavioral experiments, huge variations in reaction times are observed (Schrobsdorff et al., 2007a) that are not even one order of magnitude less than the mean reaction time. Given that the reaction times mirror the same state of the brain in two trials, a significant temporal distortion of the signal has to have taken place. We therefore propose a new and more flexible algorithm for determining the line-of-synchrony and evaluate it in comparison to Marwan et al.'s method (Marwan et al., 2002).

For illustrative purposes, consider the example already treated in (Marwan et al., 2002), where the recurrence plot of two sines that differ in their time-scale

$$f(t) = \sin(t) \qquad \text{and} \qquad g(t) = \sin(t + a\sin(t)) \tag{5.20}$$

is investigated. Via parameter a, it is possible to manipulate the amount of distortion of the LOS as illustrated in figure 5.8. While the algorithm from (Marwan et al., 2002) works well to extract the LOS in case of low distortion ($a = 0.3$, figure 5.8a), parameterization is difficult and produces errors for medium distortions ($a = 0.9$, figure 5.8b) and we failed completely to find working parameters for strong distortions ($a = 1$, figure 5.8c). A problem with the parameterization of the algorithm is its dependence on the characteristics of the recurrence plot. Our alternative, an unparameterized algorithm presented in the next paragraphs successfully extracts the LOS in all settings (figure 5.8, red curves).

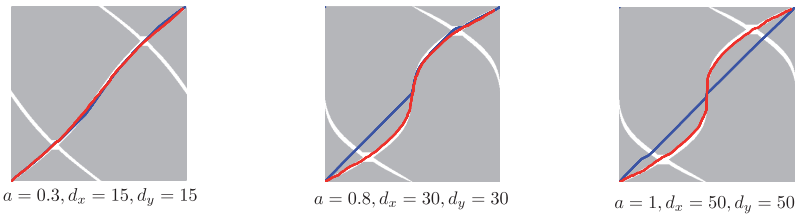


Figure 5.8: The recurrence plot of two sine functions $\sin(t)$ and $\sin(t + a\sin(t))$ is shown for different values of a. The algorithm by Marwan et al. (blue curves) works well for small distortions (a), is hard to parameterize for medium distortions (b) and fails for strong distortions (c). Our parameterless algorithm (red curves) correctly tracks the LOS in all cases.

5.7.4 An Algorithm for Line-of-Synchrony Detection

In the following, we compare the performance of two trial-registration methods based on the CRPs, the method introduced in (Marwan et al., 2002) and a method to estimate the function ϕ with the smallest number of non recurrences and the shortest path-length.

We propose to directly optimize the amount of recurrences while simultaneously preferring the shortest possible path. Let ϕ' be the first derivative of ϕ, then we seek to simultaneously minimize the line integral over the cross-recurrence plot (i.e. the number of non recurrences)

$$\int_{\mathcal{T}} \vec{C}^{\vec{x},\vec{y}}(t,\phi(t))\sqrt{1+\phi'(t)^2}\,dt \tag{5.21}$$

and the length of the path

$$\int_{\mathcal{T}} \sqrt{1+\phi'(t)^2}\,dt. \tag{5.22}$$

When dealing with binary matrices $\vec{C}^{\vec{x},\vec{y}}$, there are many possible solutions for minimizing equation 5.21 because in a neighborhood containing zeros, it is arbitrary which path is taken. We therefore use equation 5.21 to restrict our search space and then look for the shortest of these paths. In practice, we solve these minimizations by a single application of the dynamic time-warping algorithm.

In our case, application of DTW to the binary matrix $1-\vec{C}^{\vec{x},\vec{y}}(t_1,t_2)$, where $t_1,t_2 \in \{i\Delta t | i = 1,\ldots,n; i\Delta t \in \mathcal{T}\}$ results in the warping function minimizing the number of non-recurrences along the way, i.e. minimizing equation 5.21. As stated above, there is more than one solution to this minimization since the cost is not influenced by additional zeros along the path. To impose the path length criterion from equation 5.22, we modify the recurrence matrix that enters the DTW algorithm

$$\mathbf{d}_{t_1,t_2} = (1-\vec{C}^{\vec{x},\vec{y}}(t_1,t_2)) + \varepsilon \tag{5.23}$$

where ε is uniformly distributed, uncorrelated, positive noise of an amplitude much smaller than 1 (in our simulations, we used $\varepsilon \propto \mathcal{U}(0,0.01]$). There are two effects of this added noise component: First, since $\varepsilon > 0$, we add a penalty to ensure that the length of the path is taken into account. Second, we remove the arbitrariness that the algorithm encounters when choosing among three equal

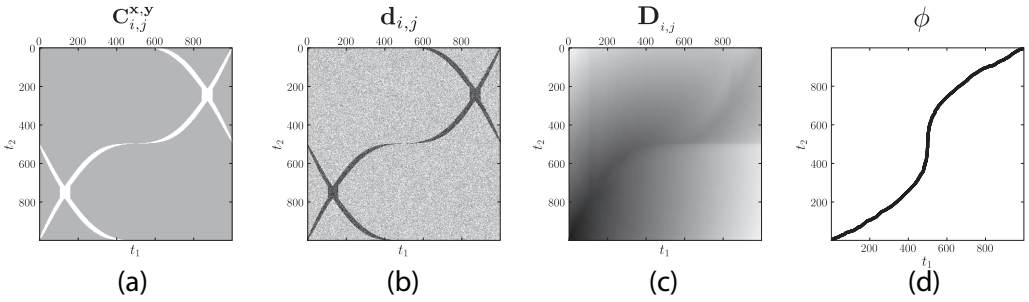


Figure 5.9: Illustration for the proposed algorithm. (a) The CRP for two sine-functions $\sin(t)$ and $\sin(t + a\sin(t))$ and (b) the modified distance matrix with added noise (exaggerated) are shown. (c) The matrix is cumulated (DTW) and used to recover the distortion function (d).

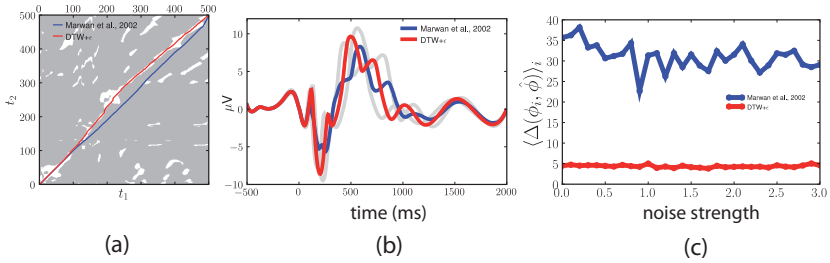


Figure 5.10: (a) Recurrence plot for two trials from the artificial dataset and the LOS calculated according to the two methods from above ($m = 10, \tau = 6, d_x = 25, d_y = 25$). (b) The raw data (gray) and the averaged curves for each algorithm. (c) The error $\Delta(\phi, \hat{\phi})$ (see text) averaged over 20 realizations is plotted as a function of noise-level ($m = 10, \tau = 10, \mathrm{FAN} = 20$).

values in the backtracing stage. For illustration, consider again the two sines from equation 5.20 in figure 5.9, where we plot the subsequent steps of the algorithm.

5.7.5 Results

In order to compare Marwan et al. (2002) and our algorithm, we plot the recurrence matrix, the extracted line of synchrony, see figure 5.10a, and the averages, equation 5.15, using these distortion functions, see figure 5.10b. The function correctly identified by our method improves the shape of the average (i.e. its shape is closer to the underlying signal) while a non-optimal path adds further distortions rather than recovering the original signal. Because we directly construct the distortion function for the artificial data, we can define a distance measure $\Delta(\phi, \hat{\phi})$ that, for each point in the recovered $\hat{\phi}$, returns its distance to the closest point in the real distortion function ϕ. Because we encounter high noise-levels in realistic settings, it is important that the algorithms are stable against noise. We present in figure 5.10c the result of comparing both algorithm with varying noise-level. Our method (red) performs significantly better in this task but both algorithms are relatively stable with respect to the level of noise in the data.

Finally, we show examples for the temporal averaging based on the extracted LOS. In figure 5.11a, the shape of the original signal can be fully recovered even from very noisy data (gray

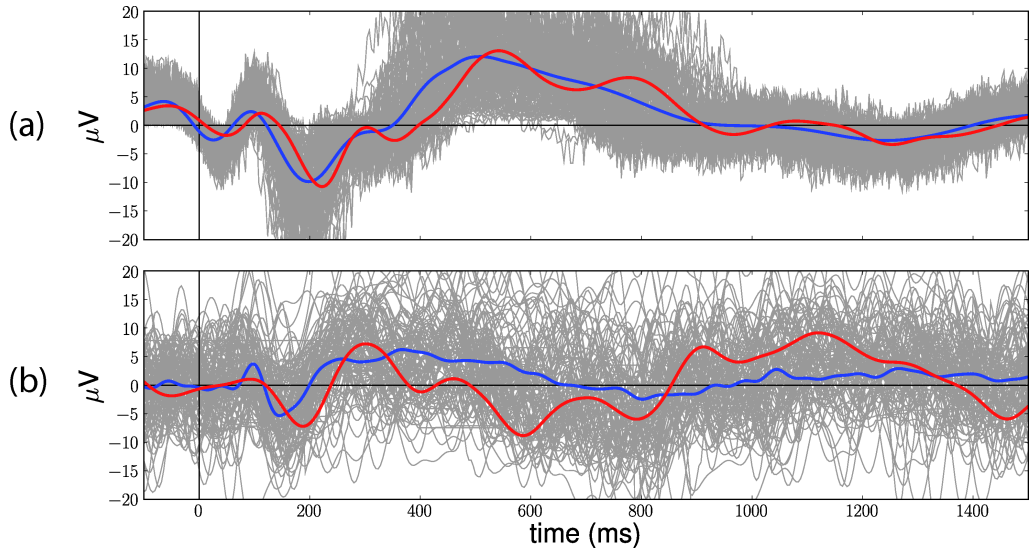


Figure 5.11: (a) The template from which the artificial data was generated is recovered by our method (red) while the conventional average (blue) fails to produce some of the peaks. (b) Real ERP-data from 134 trials at electrode location CPz. Our resynchronization method shows more pronounced and better localized peaks (e.g. P300). Embedding parameters were $m = 10, \tau = 20, \mathrm{FAN} = 100$, time-scale of single trials was transformed for plotting according to the described algorithm.

curves are single trials). In contrast, the conventional pointwise average (blue) washes out some of the peaks.

The data presented here was taken from the priming experiment described in chapter 4. We applied a bandpass filter with cutoff frequencies 0.5 and 20 Hz before data analysis. The results of applying our method to the data is exemplified in figure 5.11b. Our average produces pronounced peaks in the expected time window (e.g. P300) that are better visible and more tightly localized in time than when using conventional averaging.

5.8 Summary

Traditional averaging procedures upon which the entire field of event-related potential research is relying have certain drawbacks, we addressed in the current chapter. We clarified the situation by the formulation a few assumptions on EEG signals: the signal of interest is contained in the recorded time trace, the task-specific activations are a significant amount of the signal and cognitive processes are in some way stereotypical, such that there is a common part in the signal of similar trials. Based on these assumptions we show how more sophisticated analysis methods can be derived.

In order to robustly extract meaningful signals from noisy electrophysiological data, averaging over many similar trials is unavoidable. The nature of these data sets, i.e. correlations between electrodes, clustered time courses across trials and prior knowledge from the design of the experiment, suggests a number of more complex procedures for cleaning data and enhancing the quality of the signal.

We introduced our open source project, an algorithm toolbox `libeegtools` providing mechanisms to enhance averaging of trials of a certain experimental condition. It is possible to reduce

the variablity of the data by allowing for variable internal processing speeds. If two trials produce different reaction times, their processing must also have taken a different timecourse, thus a simple averaging does not comply with the situation. Our methods identify the relative temporal difference in processing and uses the information in the averaging of both trials. The identification of the required time warping function can be done by an adaptation of the concept of recurrence plots. As a side product we develop a metric which can successfully be used to cluster trials according to their timecourse of processing. Our algorithms can successfully be applied to artificial and real ERP-data significantly improving the quality of event related potentials compared to the traditional point-wise average. But the specific application to EEG data does not limit the generality of the approach which may as well be used for other imaging techniques.

6 Perception or Selection Effect

The third mainstay to temporally localize the negative priming effect are experimental paradigms that divide trial processing into several parts which can each be measured separately, as presented in the current and the following chapter. By assuring that the subtasks have to be accomplished in serial manner we can assign temporal differences in the processing of a certain experimental condition to specific parts of the trial.

The aim of the present study is to single out the stimulus identification phase from an experimental trial. The question whether negative priming is produced in the identification phase or in the phase of selecting the target over the distractor is believed to be crucial for the distinction between memory based or representation based theories. The former assume a conflict in the target selection phase and the latter like Distractor Inhibition, see section 2.4.1, or the ISAM, chapter 3, assume interferences already in an early stage of a trial. We realize the trial splitting by showing a color cue indicating which of the two objects to attend after the two objects already disappeared. Both objects have to be considered first, and afterwards the target has to be selected from memory.

We will describe the current experiment in section 6.1 which will also cover an introduction of the side effects to be expected when introducing a task switch as well as the extension of the set of experimental conditions. The hypotheses in section 6.3 are partly generated by a simulation of the ISAM. The necessary adaptations of the implementation are presented in section 6.2. To obtain comparability of the obtained results with previous studies, we run two preparatory experiments which are given in section 6.4. Section 6.5 is devoted to the presentation of the final post-cue experiments which will provide two individual reaction times for the two phases of each trial, stimulus identification and target selection. Finally all experiments of the study are discussed in connection in section 6.6.

6.1 Task Switch Paradigm

We realized the division of the trial processing by introducing a task switch dimension into the voicekey paradigm seen in section 2.2. The basic idea is that if the subject is given a color cue indicating the target after the presentation of the two stimulus objects, no information regarding which of the two shown objects is to be selected is available as long as the stimuli are visually present. Therefore, a complete identification of both objects is required before the trial can proceed by showing the color cue.

Additionally, the design brings along several difficulties. The most prominent one is the dimension of task switching which is known to produce large behavioral effects (Monsell, 2003). We will discuss the impact of task switching to our paradigm in section 6.1.2 after the introduction of our series of experiments in the following section.

6.1.1 Sequence of Experiments

Our final goal in the current series of experiments is the extraction of processing times of the stimulus identification phase of a negative priming trial. The introduction of the target color cue brings along unwanted side effects. To best control and monitor the changes in the paradigm, we prepend

two experiments. The first one surveys the differences between a response to a red target or to a green one by showing one block with red-target trials and a second one, separated by a break, where targets are green. Supposedly responding to a red target is faster, as red is generally more salient to humans. The issue is important, as the relative saliency between target and distractor is suspected to affect priming effects, see section 3.3.3. The second pre-cue experiment assesses the impact of the task switch dimension on priming effects in the usual processing of a negative priming trial. We present a color patch previous to the stimulus objects. Any modulations in the priming effects can then be attributed to task switching.

Finally we make the step to the post-cue experiment, which alters the way of presentation by exchanging the order of cue and stimuli such that while the stimuli are shown, no information is available as to which one of the two will be target. Therefore, subjects have to identify both objects in the stimulus identification phase. Time markers are recorded at trial onset, the subject's button press that changes the display from stimuli to cue and finally the response. We label the reaction time in the stimulus identification phase R^{si} and the one in the target selection phase R^{ts}.

6.1.2 Task Switch and Negative Priming

The current series of experiments is primarily designed to reveal the temporal localization of the different priming effects by segmenting the trials as described above. As such, a differentiation is only possible if the current target is not known to the subject during the stimulus identification phase. Thus we introduced a target color switch to our voicekey paradigm. Even if the task for the subjects ("name as fast and accurately the object of the color you will see after a button press that also eliminates the stimuli from the display") is constant over the entire post-cue experiment, the possibility of a target color switch can trigger to some extent effects that have been investigated under the label of task switching (for a review see Monsell, 2003). Task switching is usually considered by interchanging rather different tasks. Nevertheless, the very similar tasks we implement in the current experiment are also suspected to produce task switching phenomena. Negative priming has already been mixed with a task switch, but the switch happened in the dimension of response (word reading vs. picture naming) rather than in the specific feature as in our case (naming green or red picture, respectively). Waszak et al. (2005) shows overlaid pairs of a word and a picture. A cue tells the subject whether to respond to the word or to the picture every two trials. The switches occurred alternatingly and were therefore predictable. Nevertheless, as Monsell (2003) reports, predictable switches also cause switch costs, i.e. reaction time differences between switch and non-switch trials. Unfortunately, Waszak et al. (2005) focus on the reoccurrence of identical word-picture pairs later in the experiment and compare the reaction times separately for switch trials (directly after a cue) or non-switch trials (before the next cue). No analyses of the interaction between two directly successive trials as we are interested in are possible due to the special stimulus sequence.

When dealing with task switch experiments, two effects of the switch are present. The first one, called global switch costs is a general slow-down of the mean reaction time in a task heterogeneous block compared to a block which always requires the same task. Within one heterogeneous block, where trials directly succeeding a task switch are slower than trials that repeat the task from the preceding trial, the slow-down is labeled specific or local switch costs.

In a study that comes very close to the basic setting in the current chapter, MacLeod et al. (2002) investigate differences in negative priming between words that are repeated and word pairs from the same semantical category. Their design includes blocks with a change of the target color between paired prime and probe trials and blocks where the target color stayed constant. In the former case, a DT condition repeats the object of interest identically, the latter in a different color.

The authors find a difference of approximately 70–90 ms between mismatch and match blocks, which can be seen as global switch costs. Specific switch costs can not be determined from the experiment, as a task switch block does not contain non-switch trials.

6.1.3 Condition Set

We opted for a comprehensive set of stimulus relations in the current experiment, even if the main focus of attention is still on the two conditions DT and TT. If only CO, DT and TT trials are presented, the probe target is repeated from the prime display in 2/3 of all trials, which biases the subjects to make additional effort to keep track of the two items. Adding DD and TD trials to the condition set balances the repeated object between being relevant and irrelevant in the probe trial. A positive side effect is the broadening of the data of the two repeated distractor conditions that are only rarely surveyed. For each of the five priming conditions implemented in the experiment, CO, DT, TT, TD, and DD, a task switch variant exists, see a summary in table 6.1. We label them by appending the abbreviation SW.

	prime display			probe display		
	red	green	cue	red	green	cue
CO	A	B	■	C	D	■
DT	A	B	■	B	C	■
TT	A	B	■	A	C	■
TD	A	B	■	C	A	■
DD	A	B	■	C	B	■
CO_SW	A	B	■	C	D	■
DT_SW	A	B	■	C	B	■
TT_SW	A	B	■	C	A	■
TD_SW	A	B	■	A	C	■
DD_SW	A	B	■	B	C	■

Table 6.1: The priming conditions of our task switch paradigm with one target and one distractor in each of the prime and probe display and a cue for the target color respectively.

Note that the stimulus configuration is shared by a pair of conditions whereof one keeps target color, the other one incorporates a task switch, i.e. a change of the target color. The following pairs have an identical stimulus display: (DT&DD_SW), (TT&TD_SW), (TD&TT_SW) and (DD&DT_SW). In our last experiment of the current chapter, we present the target cue after the subject has pressed a button to announce a full identification of the two stimulus objects. Until the cue is shown, the above condition pairs are indistinguishable. Only after the cue onset can the processing diverge.

6.2 Task Switch and the ISAM

As the current series of experiments is based on the voice-key paradigm, we can easily derive predictions about priming effects by simulations of the ISAM. For all implementational details

see chapter 3. In order to make the simulation of a task switch paradigm with the ISAM possible, we had to extend the model by a task variable. Additionally, the factor in the calculation of the global activity level $\bar{x}$ which drives the threshold has become a parameter as it is paradigm specific.

6.2.1 Extension of the ISAM

In the implementation of the ISAM given in chapter 3, the target input rises with rate ξ. We now introduce two variables that hold the evidence for responding to one of the two colors. Whenever a color cue is presented, the corresponding variable is driven by exponential dynamics to its maximum $\hat{\xi} = \xi$, see equation 6.1 and 6.2. As we have a two alternative forced choice situation, the two tasks, to respond to the red or green object respectively, are complementary. They always add up to the maximum target amplification rate $\hat{\xi}$. Therefore, ξ_{red} and ξ_{green} can be seen as cumulated evidence or probability for one of the two alternative tasks, where $\hat{\xi}$ is the level of certainty.

If a cue is presented, the task variables adapt with time constant τ to their fixpoints $\hat{\xi}$ and 0, respectively. The time constant τ determines the global switch costs, i.e. the difference between the overall mean reaction time in an experimental block without any task switch and the overall reaction time in a task-switch block. For a green color cue, ξ_{green} and ξ_{red} adapt according to

$$\frac{1}{\tau}\frac{d\xi_{\text{green}}}{dt} = \hat{\xi} - \xi_{\text{green}} \tag{6.1}$$

$$\frac{1}{\tau}\frac{d\xi_{\text{red}}}{dt} = -\xi_{\text{red}}. \tag{6.2}$$

In between two trials of a task switch paradigm, the task variables relax with the same time constant towards a baseline evidence $\underline{\xi} \geq \hat{\xi}/2$ and $(\hat{\xi} - \underline{\xi}) \leq \hat{\xi}/2$ that slightly favors the former task to account for specific task switch costs, see section 6.1.2. Assume a green cue in the former task, gives the following dynamics. A preceding red cue would produce the opposite.

$$\frac{1}{\tau}\frac{d\xi_{\text{green}}}{dt} = \underline{\xi} - \xi_{\text{green}} \tag{6.3}$$

$$\frac{1}{\tau}\frac{d\xi_{\text{red}}}{dt} = (\hat{\xi} - \underline{\xi}) - \xi_{\text{red}} \tag{6.4}$$

The resulting dynamics of ξ_{green} and ξ_{red} is shown in figures 6.2 and 6.4 at the bottom. In the presence of a cue, the task of the required target color rises to full evidence, suppressing the opposite task. If the trial is over, both variables approach each other, still having a certain offset representing the tendency to assume the same task cue in the next trial.

The normalization factor of the activation mean $\bar{x}$, equation 3.10, that is the basis for the adaptive threshold, is labeled by ν.

$$\bar{x} = \nu\left(r^{\tau} + r^{\delta} + \sum_{i=1}^{n}(x_i^{\tau} + x_i^{\delta})\right) \tag{6.5}$$

Depending on the number of stimulus objects in a certain paradigm, ν can be adjusted to values that allow for an accurate but fast decision on which object to respond to. Future extensions of the ISAM could include a self-organization mechanism for ν according to the current tasks of the system.

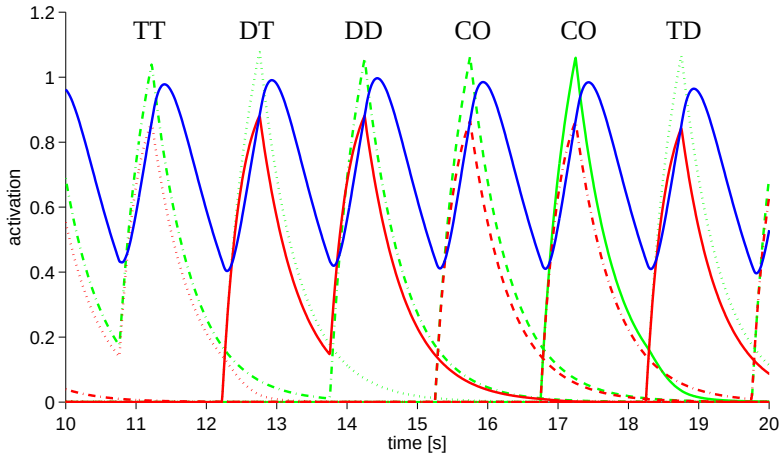


Figure 6.1: Object representation variables over 10 seconds of simulated time. Several different priming conditions are shown. Red lines correspond to red objects, green lines to objects shown in green. Textures label object identity. The threshold θ is shown in blue. Note the slow rise of target and distractor activation as compared to e.g. figure 3.4 which causes the threshold to run only slightly below the distractor activation making the model very sensitive.

6.2.2 Calibration

In order to produce realistic predictions, the ISAM is first subject to a calibration phase. We fit the model such that it matches the priming effects of the study reported in section 2.2 to a satisfying extent. Calibration is done manually by intuition and yielded the following parameter values: rsi = 1000 ms, $\zeta = 0.0025$, $\xi = 0.0007$, $\alpha = 0.004$, $\beta = 0.0018$, $\gamma = 0.003$, $\sigma = 0.845$, $\nu = 0.64$, $\Delta = 25$, a time step was 1 ms. Again as in chapter 3 we do not intend to match the particular reaction times but merely the priming effects, as the simulation of the ISAM disregards perceptual time lags and also delays due to motor commands and their execution. We also took the priming conditions DTTD and DDTT of two repeated objects into account for the calibration phase to obtain a most accurate behavior of the model. In the experiments we exclude those conditions and instead consider the two conditions TD and DD where the repeated object is shown as a distractor in the probe, see section 6.1.3. The simulation produced priming effects similar to those in the mentioned experiment (to-be-fitted target effect in parentheses): DT -24.81 ms (-21.36 ms), TT 34.77 ms (35.20 ms), DTTD -23.44 ms (-25.70 ms) and DDTT 49.10 ms (59.53 ms). A comparison of the obtained reaction times and priming effects is given in table A.5, appendix A.4. The simulation produced a very small TD effect of -4.91 ms and also a weak DD effect of 6.70 ms.

The behavior of the model variables can be seen in figure 6.1. A comparison with figure 3.4 or 3.5 shows the differences in the model behavior obtained by the calibrated parameter set and the parameters we used for our simulations in chapter 3. We adjusted in the first place the sampling rate, i.e. the simulated time step from 0.3 ms to 1 ms for convenience. But we also shifted the point of target selection to an earlier phase of unsaturated distractor activation by a slower rise and decay of activation variables and a short delay in the dynamics of the threshold θ.

6.2.3 Pre-Cue Simulation

We now take the parameter values derived in the last section and give in between the trials a color cue which is subject to unpredictable change. The time course of the corresponding task values as well as the response of the model system are plotted in figure 6.2. Note that the task variables start to adapt to the appropriate values at 500 ms before the next trial starts, at the moment the color cue appears.

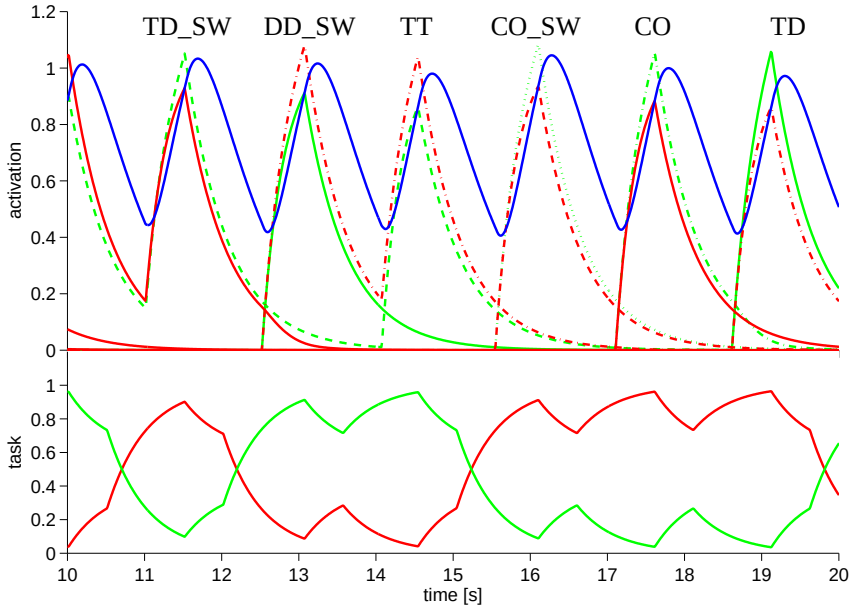


Figure 6.2: Simulation results for the pre-cue paradigm with the parameters determined in section 6.2.2. The top plot shows the dynamics of representation activations and threshold θ in blue. The bottom plot shows the time course of the two task evidence variables in units of $\hat{\xi}$. Note that the six trials are perceptually identical to the six trials in figure 6.1.

The results show a familiar pattern in the non-switch trials, see also table A.6 in appendix A.4. The conditions DT and TD are delayed by 14.67 ms and 15.61 ms, respectively. TT and DD are accelerated by 27.76 ms and 21.92 ms, respectively. The result pattern is very different in switch trials, where DT_SW trials are largely accelerated, 32.82 ms, TT_SW is delayed by 13.87 ms, TD_SW is 22.71 ms faster than CO_SW and DD_SW is again delayed by 14.78 ms. Overall, a change of object color results in a slow-down, whereas the repetition of objects of the same color leads to a faster response.

Even in the switch trials the tendency to respond faster to identically repeated stimuli is conserved. It is much stronger than any priming effect stemming from the role of an object as target or distractor. The ISAM is a model where priming acts on a perceptual level. The global switch costs of 33 ms are given by the difference between the mean of all reaction times (of the relevant conditions: CO, DT, TT, TD, DD) of the baseline simulation of section 6.2.2, 498.50 ms, and the overall mean of the pre-cue simulation of the no-switch trials, 531.54 ms. Remarkably, the specific switch costs taken from the control condition due to the oppositional reaction time differences

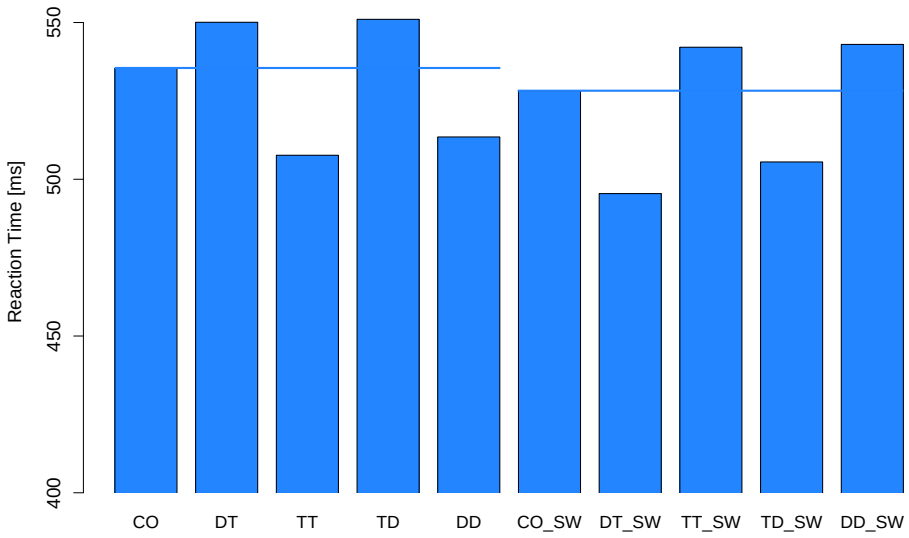


Figure 6.3: Simulated reaction times in the pre-cue paradigm. Note the negative specific task switch costs. A switch trial is faster than non-switch trials. The repetition of an object in the identical color leads to an acceleration, while if the repeated object changes color, the response is slower.

in the other conditions are negative. The CO_SW trials are on average 7.17 ms faster than their non-switch counterparts. If the mean of all no-switch trials, again 531.54 ms, is compared to the average of all switch trials 522.87 ms, a benefit for switch trials of 9.67 ms is still present.

6.2.4 Post-Cue Simulation

The ISAM solves the post-cue paradigm such that it first receives equal input by the two objects of the stimulus display. If both the activations surpass the sensitivity value σ, the objects are said to be recognized and stored, such that their activation is persistent even after the display switches to the color cue.

If we run a simulation with the parameter values derived in the calibration phase described in section 6.2.2 on the post-cue paradigm, we encounter a problem we already know from investigations of the dynamics of the ISAM in a setting with varying response stimulus intervals. The adaptation of the threshold is too quick to accurately solve the task of the post-cue paradigm. The threshold shoots over both the target and the distractor representation activation. To solve the problem, we assume a mechanism that is able to adjust the threshold behavior to the particular task demands. In the present case, the threshold is slowed by manually adjusting the following parameters to $\nu = 0.55$, $\gamma = 0.002$ and $\Delta = 50$. With these parameter values, the threshold stays below both relevant activations until the target color boost adds enough total activation to the system.

Note that both target and distractor activation rise similarly, separated only by their activation values right at stimulus onset, see figure 6.4. Therefore, only perceptual issues of the probe trial

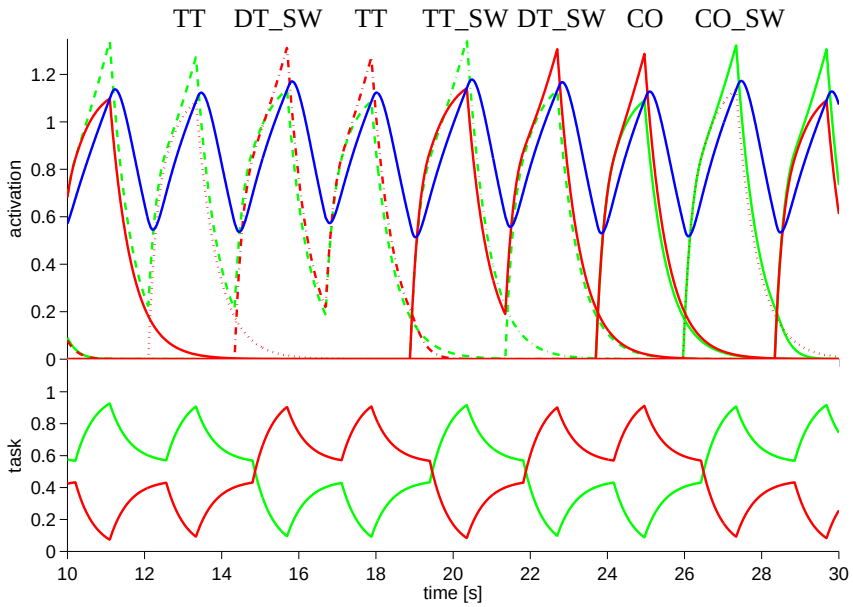


Figure 6.4: Simulation run for the post-cue paradigm with the adjusted threshold parameters $\nu = 0.55$, $\gamma = 0.002$ and $\Delta = 50$. The top plot shows the dynamics of representation activations in red and green and different textures for different objects and threshold θ in blue. The bottom plot shows the time course of the two task evidence variables in units of $\hat{\xi}$. Note that the slower adaptation of θ does not lead to artifacts as the attempt to respond before the cue appears.

can influence the first reaction time R^{si}. A detailed listing of simulated reaction times, see figure 6.5, and effect strengths is given in table A.7. In summary, the ISAM predicts only slowing effects of around 40 ms in R^{si} for the conditions that are subject to the forced decay if a probe stimulus reappears with a different color, i.e. DT, TD, TT_SW and DD_SW. In R^{ts}, the above conditions are again slowed, TD and DD_SW weaker than DT and TT_SW. The other four conditions, TT, DD, DT_SW and TD_SW, are faster than the respective control condition with TT having the strongest benefit.

6.3 Hypotheses

As already mentioned in section 6.1.3 we will mainly focus on the analysis of priming effects in the two conditions DT and TT, the classical negative and positive priming conditions. Therefore, we only give hypotheses for further conditions, if a particular hypothesis also applies to other conditions.

As denoted in section 6.2.1, the ISAM is able to generate detailed hypotheses for the current series of experiments. For the pre-cue experiment, the ISAM predicts a slow-down of about 14 ms for trials where the repeated object changes color and a speed benefit of 20–30 ms for the conditions TT, DD, DT_SW and TD_SW. The post-cue experiment is predicted to produce very large slowing effects of about 70–95 ms distributed equally between the two reaction times in the conditions with a perceptual switch, and a comparably weak positive effect of 18–27 ms which is exclusively present in the target selection phase R^{ts}.

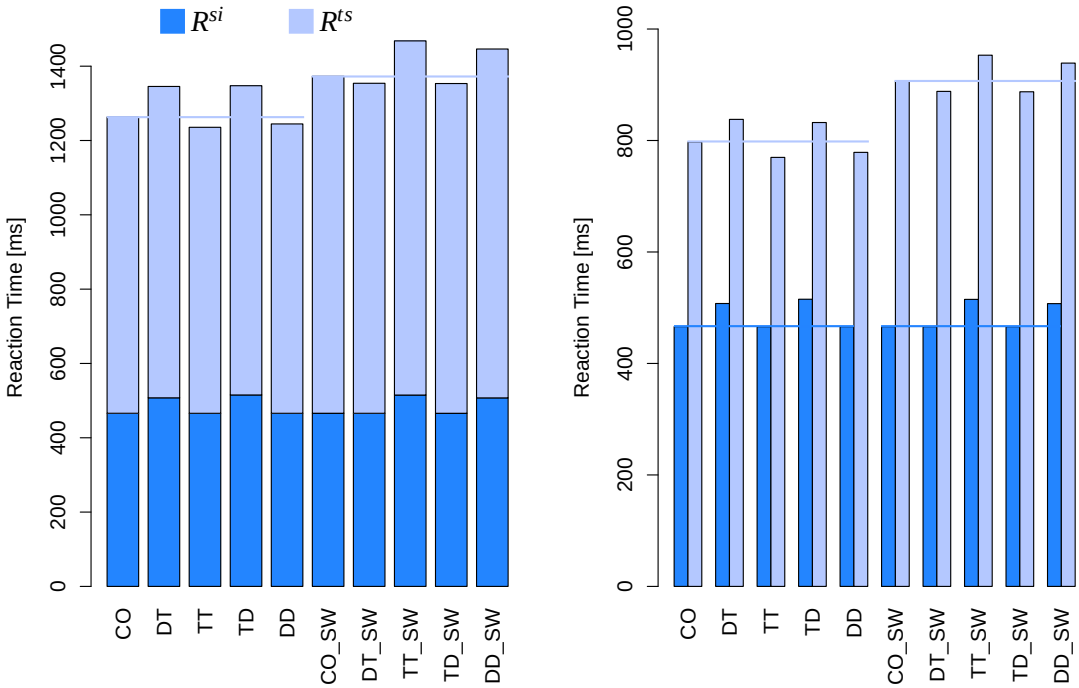


Figure 6.5: Simulation results for the post-cue paradigm. The left plot shows both partial reaction times stacked on top of each other to point up the priming effects in the overall reaction times. The right plot compares both partial reaction times individually. Note the identical $R^{si}$with perceptually identical displays. Perceptually, only a negative effect is present which has its origin in the forced decay of activation the ISAM incorporates. Positive effects are only present in the stimulus identification phase.

Episodic retrieval as well as response retrieval theory predict a prime retrieval in all conditions apart from CO, with a tendency for a stronger memory retrieval for identically repeated objects. As response retrieval postulates only the response to be retrieved, no effects are expected in the first part of the trial. Only the target selection phase should carry all priming effects in a way that only repeated responses benefit from retrieval, i.e. only the TT condition and a bit weaker the TT_SW condition are faster than control. All other priming conditions should trigger by their perceptual similarity a retrieval of the, in this case false, response. Thus all other conditions are expected to produce negative priming. The same effects are to be expected for the first two experiments. Traditional episodic retrieval postulates the retrieval of the entire prime episode depending on stimulus similarity. Therefore, a perceptual benefit is to be expected in every priming condition besides CO. In cases where the to-be-attended object is retrieved together with some conflicting action advice, i.e. DT, DT_SW, a counteracting deceleration in the later part of the trial is conjectured. If the action advice matches, as in TT and TT_SW trials, Episodic Retrieval makes the prediction of additional benefit.

Distractor Inhibition expects perceptual facilitation in conditions where the prime target is repeated, and a perceptual slowing if the prime distractor is repeated. Predictions for the reaction time of the target selection phase $R^{ts}$are hard to obtain, as the dynamic system underlying dis-

tractor inhibition theory is too complex to guess the outcome in the second phase of the post-cue experiment. Late priming effects depend on the mechanism by which objects are classified and identified, be it at a saturation level, where no priming effects could be carried over to the late phase, or at intermediate activation levels like in the simulation of the ISAM, when positive as well as negative priming can still cause behavioral effects in the later stage.

6.4 Preparatory Task Switch Experiments

In order to achieve comparability of results of the post-cue experiment and our earlier studies, we make two preparatory experiments that help to successively add complexity to the original voicekey experiment, see section 2.2. We begin with two baseline blocks, one of which follows the paradigm introduced in section 2.2, the other requiring attention to the red object. The comparison of the two blocks helps us to understand the difference of trials with a red target and those with a green one that will be naturally present in the task switch experiments.

The next step is the introduction of task switching in the form of a color cue indicating the target color in the following trial. Despite the task switch, the paradigm stays identical. Thus, any changes in priming effects are due to the cue and its demand to attend to its color. Besides the impact of updating a task set, physical alterations are also made to a trial. In particular, the response stimulus interval is longer in the pre-cue experiment, as the subject's reckoning of the cue adds to it. Also, the additional percept of a color patch before the actual trial may change processing compared to the baseline experiment.

In the current section we will present the two experiments together. One reason is their nature of providing prerequisites for the experiment of interest, the post-cue paradigm. Another is that both experiments are performed by the same subjects in one session. The baseline experiment also provides a familiarization with the stimuli and the task for the pre-cue experiment.

6.4.1 Design

We first only slightly altered our traditional negative priming experiment. Two blocks of the simple voicekey task are run, one with green as target color, while the other required a response to the red item. The sequence of these blocks was balanced over subjects. The two baseline parts included five priming conditions, CO, DT, TT, TD and DD.

After the two baseline blocks followed a task switch experiment of 420 trials with breaks every 42 trials paced by the subjects themselves. We implemented a standard way of unpredictable task switches by pseudo randomly presenting a color cue indicating the target color each trial anew. The color patch was removed by a button press by the subject which then started the actual trial. The switch dimension is reflected in a doubling of experimental conditions, as a task switch destroys the confounding of whether an object is repeated in the same color or not and the object to respond to.

6.4.2 Participants

Thirty undergraduate students (19 female, 11 male) from the University of Göttingen took part in the study. Their ages ranged from 20 years to 35 years (M = 24.5 years, SD = 1.6 years). The participants were rewarded by course credits or are paid 15 €. All subjects had normal or corrected-to-normal vision and no color discrimination disabilities. They were not informed about the specific purpose of the experiment and had not taken part in a previous study employing similar stimulus material.

6.4.3 Procedure

The experiments took place in two special chambers optimized for low noise and standardized lighting conditions. Participants were tested individually in sessions that lasted no longer than 90 minutes. Before the start of the priming tasks, the line drawings of the experimental stimuli along with their names printed in black were shown to the subjects. The subjects were told that they will see these objects overlapping one another, one drawn in green and the other in red. Realizing a classical picture naming task, participants were instructed to name the target objects as quickly and correctly as possible while ignoring the superimposed distractor object. To familiarize participants with the experimental procedure, a 30-trials practice session preceded the main session.

The first two experimental tasks consisted of 210 trials each, which were divided into 5 blocks of 42 trials. After each block, subjects were allowed to take a short break. In a single trial, subjects encountered the following series of events: a fixation cross, centered on the screen for 500 ms; a display containing two superimposed objects until subject responded, but no longer than 2 seconds and then a blank screen for a randomized duration between 0 and 1000 ms.

In the pre-cue experiment, an additional visual cue indicated if the target color of the trial was green or red. The number of trials was doubled compared to the first tasks to 420 trials, so that the participants had to complete 10 blocks of 42 trials. A single trial consisted of the following series of events: a fixation cross for 500 ms, the color cue until the subject pressed a button, the stimulus display and then a randomized RSI of 0-1000 ms. Behavioral errors were noted when subjects failed to give the correct answer. After the participants completed all trials, they were asked to comment on the experimental procedure.

6.4.4 Data Analysis

First, we want to make sure that the baseline experiments show strong priming effects as anticipated. The two different target colors are then compared. One aspect is a main effect of red/green, which indicates a direct influence of target color to processing speed. Another aspect is an interaction between target color and priming effects. If such an interaction exists, we have to group the task switch trials into conditions with homogeneous target color.

A comparison between overall mean reaction time in the baseline experiment and the pre-cue paradigm yields a measure for global switch costs, i.e. the influence of the presence of task switching on processing speed. Comparing switch and non-switch trials within the pre-cue experiment leads to specific switch costs. We then check for priming effects in the respective condition, especially in the task switch trials, as no data exists for similar settings.

To clean the data, some procedures to reject outliers from the behavioral data are employed. In a first step, all reaction times from trials in which a behavioral error occurred, are excluded from the analysis as well as the directly following trial. Reaction times below 250 ms and above 3000 ms were removed from the analysis as they presumably are based on other processes than the ones under investigation due to their unusual duration. In a next step, reaction times where the difference to the mean of the experimental condition exceeded two times the standard deviation were also excluded. Overall, for each participant not more than 10% of the trials per condition were excluded from the analysis.

6.4.5 Results, Baseline Experiment

The overall $2 \times 5 \times 2$ (target color $\times$ priming $\times$ color block order) ANOVA treating reaction times as dependent variable showed a main effect of target color $F(1,116) = 16.11$, $p < 0.001$, a main effect of priming $F(4,29) = 23.21$, $p < 0.001$, but no interaction of target color and priming

$F(4,116) = 0.77$, $p = 0.54$ and no effects at all of the order in which the two color blocks are tested, see table A.8. Especially the absence of effects of the order in which red or green indicate the target allows a pooling of all subjects, regardless of which target color the subjects encounter first.

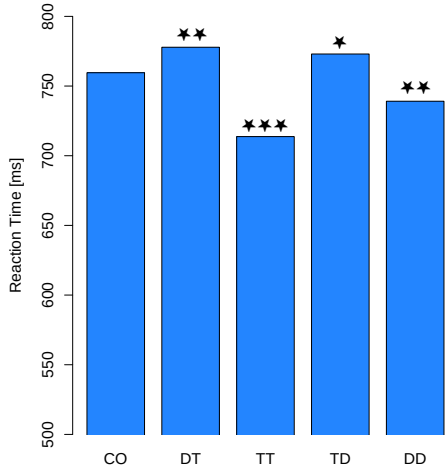


Figure 6.6: Priming effects of the baseline experiment. Significances of comparisons to the control condition are given. Expectedly, trials with a repeated object that changes its role between prime and probe are slower, identical repetitions of objects lead to an acceleration.

The absence of an interaction between target color and priming allows the determination of priming effects in the entire baseline experiment, i.e. the mean reaction time differences over both target color blocks, see figure 6.6 and table A.9 in appendix A.4. Trials with a green target, 776.10 (95.80) ms, are on average 47 ms slower than trials with a red target, 729.20 (104.59) ms. Even in the pooled situation all priming effects are significant, CO−DT= −18.28 ms, $t(29) = -3.09, p < 0.01$, CO−TT=45.83 ms, $t(29) = 4.56, p < 0.001$, CO−TD= −13.43 ms, $t(29) = -2.27, p < 0.05$ and CO−DD=20.45 ms, $t(29) = 3.55, p < 0.01$.

6.4.6 Results, Pre-Cue Experiment

The overall 2×5 (switch $\times$ priming) ANOVA on reaction times shows main effects of switch $F(1,29) = 19.94$, $p < 0.001$ and priming $F(4,116) = 15.12$, $p < 0.001$ as well as an interaction of both $F(4,116) = 8.21$, $p < 0.001$. Specific switch costs, 27 ms, are determined as the difference between overall mean reaction times in switch trials (783.28 (108.52) ms) and no-switch trials (810.25 (123.80) ms) in the pre-cue experiment. We find global switch costs in the two-sided comparison of the mean reaction time in the baseline experiment (752.65 (95.05) ms), and the pre-cue experiment (796.76 (115.23) ms), which differ by 44.11 ms significantly, $t(29) = -3.37, p < 0.01$.

For priming effects we compared the appropriate control with the priming conditions and found effects in CO−DT= −24.93 ms, $t(29) = -2.83, p < 0.01$; CO−TT=54.76 ms, $t(29) = 5.09, p <$

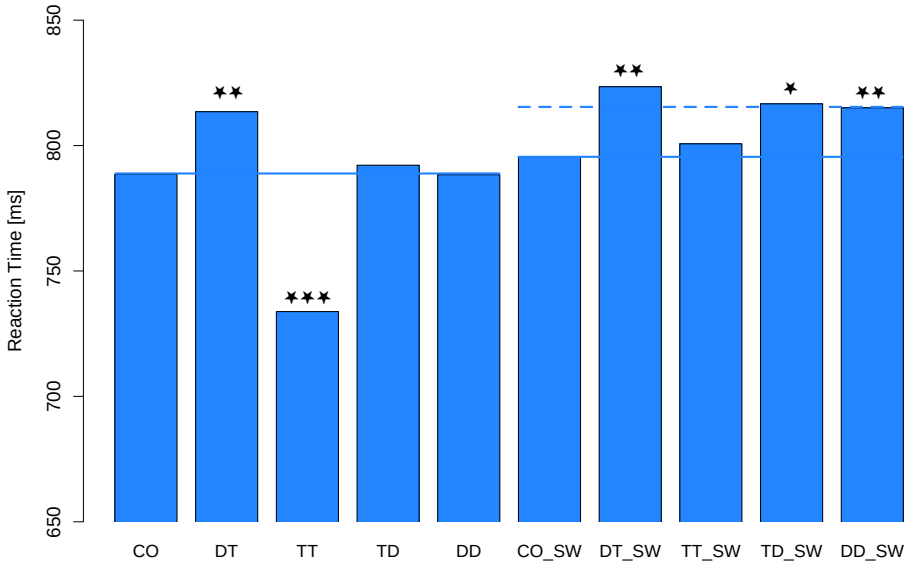


Figure 6.7: Priming effects of the pre-cue experiment. Significances of comparisons to the appropriate control condition (indicated by the horizontal lines) are given. Disregarding the control condition the pattern of the priming conditions looks similar in the switch and non-switch stimuli. For illustrative purposes we indicate the position of such a hypothetical control with a dashed line.

0.001, but no effects in the two conditions CO−TD= −3.60 ms, $t(29) = -0.54, p = 0.29$ and CO−DD=0.20 ms, $t(29) = 0.024, p = 0.48$. For trials that incorporated a target color switch from the prime, the present effects are CO_SW−DT_SW= −28.07 ms, $t(29) = -2.54, p < 0.01$, CO_SW−TD_SW= −21.23 ms, $t(29) = -2.31, p < 0.05$ and CO_SW−DD_SW= −19.66 ms $t(29) = 2.63, p < 0.01$, no effect shows CO_SW−TT_SW= −5.30 ms, $t(29) = -0.66, p = 0.25$. The results are summarized in figure 6.7, details given in table A.10 in appendix A.4.

6.4.7 Discussion

The baseline experiment shows that it is easier to respond to a red target than to a green one, presumably because of the higher saliency of the color red. But despite the faster response in the red target block, priming effects did not differ. Further, the different saliency does not influence priming in the baseline experiment as it does in other settings (Tipper et al., 2002), see also section 3.3.3. We therefore pooled both parts and still found convincing priming effects, i.e. negative priming in the DT and TD condition and positive priming for TT and DD, see figure 6.6. The two condition pairs are each associated by a reoccurrence of an object in the same or in a different color, respectively. The results of the baseline experiment once more prove the suitability of our paradigm to access negative priming.

Introducing the task switch in the pre-cue experiment destroyed the effects in the TD and DD condition, as the influence of the probe distractor apparently is diminished. But the effects in the

most interesting conditions, DT and TT, are even more pronounced than in the baseline experiment without task switch, see figure 6.7. Comparing the pre-cue experiment with the baseline reveals global switch costs visible in the general slowing of reactions after the introduction of the task switch. Comparing non-switch conditions in the pre-cue experiment with the switch conditions reveals specific switch costs of comparable order to the global ones. Trials following a task switch are slower than trials that require the attention to an identically colored target.

Remarkably the priming effects in the switch trials show an overall deceleration in the case of an object repetition. Only TT_SW trials are not distinguishable from CO_SW. The reoccurrence of an object in conjunction with a task switch seems to overload the subjects, such that they react slower than if no object is repeated from the prime trial. A closer look at figure 6.7 reveals that the relations between the priming conditions themselves look very similar to the conceptually matching non-switch trials. Seemingly, the repetition of an object in any color produces a constant offset into the processing of switch trials, but the particular priming effects remain.

6.5 Post-Cue Task Switch Experiment

The presentation of the target color cue after the identification of the stimulus objects is a technically easy manipulation. Nevertheless, the small manipulation has several implications. First of all, the desired effect that the identification of stimuli has to be separated from the target selection puts way higher demands onto the subject. Both objects have to be in the attentional focus. Additionally, the information about their identity bound to the correct color has to be memorized in some way. Secondly, the effective RSI is now augmented by the cue presentation phase which is much longer than the cue phase in the pre-cue experiment.

The high demands led several subjects to use an after-image strategy which shifts part of the stimuli identification to the later phase of the trial. This strategy interferes with our assumption of seriality. In order to enforce a serial trial processing, we introduce a mask between the stimulus presentation and the appearance of the target cue after 15 of the total 30 participants. Such a mask again disturbs the experimental flow and drives us further away from the original voice-key paradigm. Notwithstanding, we assume only a small impact on the trial processing when compared to the subjects in the non-masked group who did not use the after image but completely identified the two objects before proceeding with the trial.

6.5.1 Design

The post-cue experiment only differs from the pre-cue experiment by the order of presenting the cue and the stimulus objects. Now, the subjects see both stimuli in the beginning of a trial, have to press a button when they are ready to proceed, and then the two objects are replaced by a color cue which is present until the subject gives the answer. Fifteen subjects saw a mask for 100 ms between the stimuli and the color cue in order to erase any after image. The mask consists of red and green dots at the location of the stimulus compound in a similar density.

6.5.2 Participants

Overall thirty young adults were tested and received course credit or are paid 10 €. One person was excluded from the sample because of a high rate (38%) of response errors and statistical outliers. The remaining sample of participants consisted of 17 females and 12 males with mean age of 23.7 years, SD = 1.6. All participants had normal or corrected-to-normal vision, no color discrimination disabilities and were naïve about the aims of the experiment.

6.5.3 Data Analysis

Prepended to the analysis of reaction times is the outlier correction we introduced in section 6.4.4 We will consider the two parts of each trial, R^{si} and R^{ts} separately. In both cases we will begin with the analysis of priming effects for the two groups, with and without mask, but later on take the two groups together for an overall analysis of the post-cue experiment. Despite one rather weird exception, the pooling of both groups is possible.

During the analysis of the stimulus identification phase, R^{si}, there is a perceptual congruence between two conditions, see section 6.1.3. We therefore compare the pairs of matching object displays and perform a pooled analysis of priming effects for the four resulting conditions.

6.5.4 Results, Stimulus Identification Phase

When considering the first part of a trial, R^{si}, there are always pairs of conditions that are identical from the percept of objects, until the color-cue is shown and they diverge into a switch condition and a non-switch condition, see section 6.1.3. In the post-cue experiment without a mask, no significant difference of reaction times within the pairs is present. Two-sided tests yield: CO−CO_SW=18.40 ms, $std = 128.43, t(14) = 0.55, p = 0.58$; DT−DD_SW=6.40 ms, $std = 71.98, t(14) = 0.34, p = 0.73$; TT−TD_SW= −3.46 ms, $std = 82.42, t(14) = 0.16, p = 0.87$; DD−DT_SW= −46.06 ms, $std = 108.46, p = 0.12$ and TD−TT_SW=23.60 ms, $std = 93.97, t(14) = 0.97, p = 0.34$.

In most experimental conditions of the masked experiment the situation is similar. A two-sided t-test reveals no difference between the following conditions: CO−CO_SW=26.07 ms, $std = 147.44, t(13) = 0.66, p = 0.52$; TT−TD_SW= −20.07 ms, $std = 110.22, t(13) = -0.68, p = 0.50$; DD−DT_SW= −24.57 ms, $std = 79.56, t(13) = -1.15, p = 0.26$; TD−TT_SW=36.92 ms, $std = 74.73, t(13) = 1.84, p = 0.087$. Only the conditions DT and DD_SW, which are perceptually identical, show a significant difference: DT−DD_SW= −143 ms, $std = 120.56, t(13) = -4.44$. As such, an effect can not be explained by systematic variations between the two perceptually identical trials. The only source of the effect is a disbalancing of trials, e.g. in the number of occurrences of an individual object or the even distribution over all experimental blocks, by data cleaning like error trial or outlier removal.

Disregarding the significant difference between DT and DD_SW trials, we pool the perceptually identical conditions over both experiments, with and without masking of target and distractor object, to obtain priming effects in the stimulus identification phase of a trial, see figure 6.8. A one-sided t-test yields significant priming effects only in the two pooled conditions: (CO&CO_SW)−(TT&TD_SW)=87.51 ms, $std = 118.48, t(28) = 3.97, p < 0.001$; (CO&CO_SW)−(DD&DT_SW)=35.55 ms, $std = 78.37, t(28) = 2.44, p < 0.05$. No effect is present for (CO&CO_SW)−(TD&TT_SW)=11.51 ms, $std = 69.79, t(28) = 0.88, p = 0.19$ and also the invalid subsumption of DT and DD_SW does not show an effect (CO&CO_SW)−(DT&DD_SW)= −6.87 ms, $std = 76.03, t(28) = -0.48, p = 0.31$. Whenever the repeated object is shown identically, i.e. in the same color, the stimulus identification is accelerated. No perceptual slow-down for repeated objects with non-matching colors is visible.

If we now assume an incomparableness of the conditions DT and DD_SW due to their abnormal behavior, which may be caused by some violation of balancing, we can not consider the data set in a global ANOVA, but rather have to exclude the DT and DD_SW conditions in advance. We run a $4 \times 2 \times 2$ (priming×percept×mask) ANOVA which yields the expected main effect priming $F(3,81) = 6.19$, $p < 0.01$, but no interaction of both priming× percept $F(3,81) = 2.51$, $p = 0.082$ and priming×percept×mask $F(3,81) = 0.22$, $p = 0.87$. The above ANOVA takes into account,

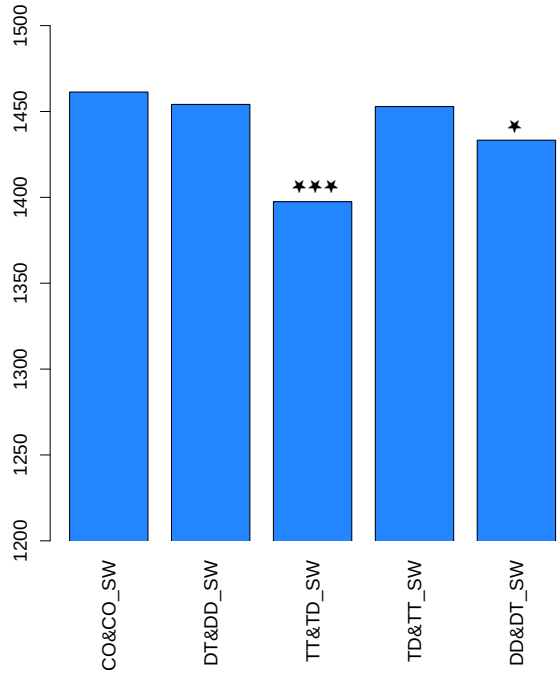


Figure 6.8: Priming effects in the stimulus identification phase of the post-cue experiment in dependence on the pooled conditions. Significances of comparisons to the control condition are given. The stimulus identification phase shows only benefits for the conditions where an object is repeated in the same color. No condition is slower than the two controls, the first part of a trial only shows positive priming.

that not e.g. TT and TT_SW have to be compared but TT and TD_SW as they are identical in the stimulus identification phase,

For better comparison with R^{ts}we pool the masked and non-masked group, ignoring the interaction of mask and priming in the DD_SW condition, see figure 6.9 and table A.12. Priming effects are revealed by 2-sided t-tests of the differences of the appropriate control and the priming condition: CO−TT=104.31 ms, $t(28) = 3.41, p < 0.01$; CO−DD=64.44 ms, $t(28) = 2.50, p < 0.05$; CO_SW−DT_SW=70.72 ms, $t(28) = 3.25, p < 0.001$ and the inappropriate CO_SW−DD_SW= −50.82 ms, $t(28) = -2.32, p < 0.05$. No priming effects are found in the conditions CO−DT=37.06 ms, $t(28) = 1.65, p = 0.11$; CO−TD=7.55 ms, $t(28) = 0.34, p = 0.73$; CO_SW−DT_SW=6.65 ms, $t(28) = 0.43, p = 0.66$ and CO_SW−TT_SW=15.48 ms, $t(28) = 0.96, p = 0.34$.

6.5.5 Results, Target Selection Phase

We now come to the analysis of the second part of the trials, R^{ts}. The detailed R^{ts} and the corresponding priming effects are listed in table A.13 in appendix A.4. A global $5 \times 2 \times \times 2$ (priming×switch×mask) ANOVA reveals a main effect priming $F(4,108) = 10.33, p < 0.001$, a main effect mask $F(1,27) = 4.44, p < 0.05$ and an interaction between priming and switch $F(4,108) = 11.18, p < 0.001$. There is no main effect switch $F(1,27) = 3.26, p = 0.082$ and neither an interaction priming×mask $F(4,108) = 0.63, p = 0.64$ nor an interaction

priming×switch×mask $F(4,108) = 0.66$, $p = 0.61$. Despite trials in the masked experiment being 99 ms slower on average, due to the missing interaction of priming and mask, we are able to combine both experiments and determine priming effects for the post-cue experiment globally. Results are shown in figure 6.9, details given in table A.14.

In the pooled data, we find in a 2-sided t-test priming effects for: CO−DT= −17.86 ms, $t(28) = 2.12, p < 0.05$; CO−DD=31.51 ms, $t(28) = -2.59, p < 0.05$; CO_SW−DT_SW= −32.41 ms, $t(28) = -2.74, p < 0.05$; CO_SW−TT_SW= −60.65 ms, $t(28) = -5.26, p < 0.001$ and CO_SW−TD_SW=29.96 ms, $t(28) = 2.77, p < 0.05$. No effect is present in CO−TT=6.55 ms, $t(28) = 0.56, p = 0.57$; CO−TD=10.75 ms, $t(28) = 0.98, p = 0.33$; CO_SW−DD_SW= −14.31 ms, $t(28) = -1.60, p = 0.12$.

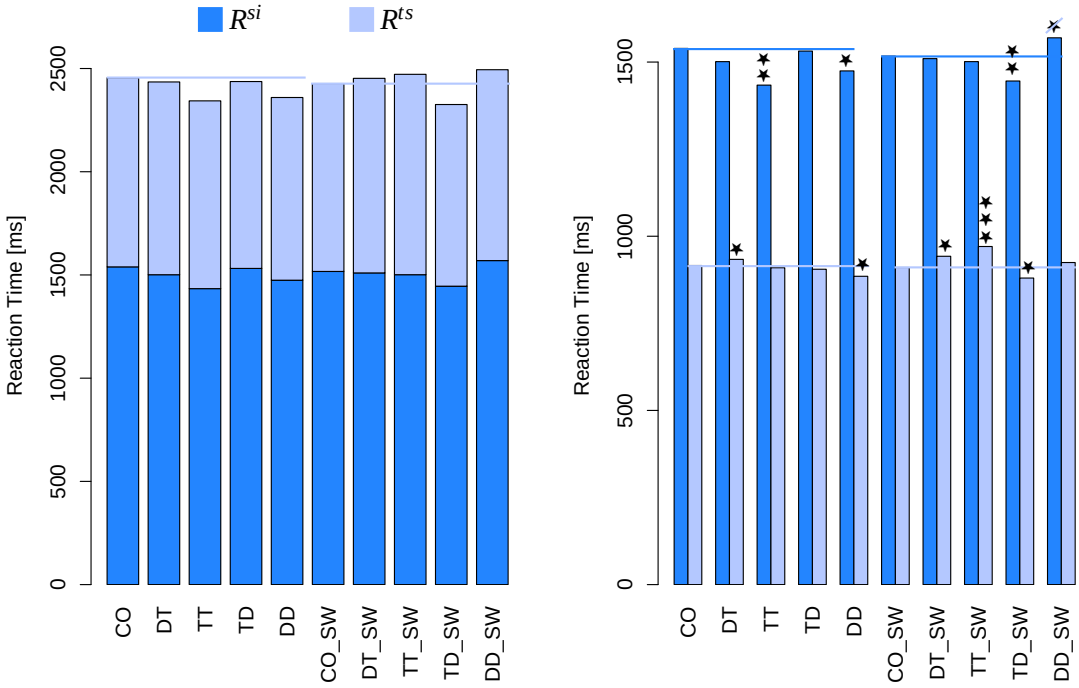


Figure 6.9: Partial reaction times for the post-cue paradigm pooled over masked/nonmasked. Left hand side shows stacked bars, allowing for a look at the overall priming effects, where mostly accelerations are visible. On the right hand side, we show the two partial reaction times beneath each other in order to make priming effects in the later reaction time R^{ts}visible. Condition DD_SW does not allow such a pooling in R^{si}. Therefore, the according bar is not to be taken too seriously.

6.5.6 Results, Comparison of Partial Reaction Times

So far the data suggests that positive priming is a more perceptual phenomenon, i.e. is present in R^{si}, and negative priming is produced while target selection and response generation, i.e. is found in R^{ts}. We limit the analysis to non-switch trials. Due to the significant interaction of switch and priming, a task switch destroys the priming effects we are looking for. Specific 2×2 (control/priming×R^{si}/R^{ts}) ANOVAs are carried out for every priming condition separately.

DT: Main effect for $R^{si}/R^{ts} F(1,28) = 27.39$, $p < 0.001$, no main effect for priming $F(1,28) = 0.56$, $p = 0.45$. But the expected interaction reaches significance, $F(1,28) = 6.07$, $p < 0.05$. DT responses are 38 ms faster than control responses in R^{si}. Conversely, responding to a DT trial is delayed by 18 ms in R^{ts}. TT: Main effect for $R^{si}/R^{ts} F(1,28) = 26.91$, $p < 0.001$, a main effect for priming $F(1,28) = 9.84$, $p < 0.01$ and also the interaction reaches significance, $F(1,28) = 10.79$, $p < 0.01$. TT responses are 104 ms faster than control responses in R^{si}. The acceleration by 7 ms in $R^{ts}$is not significant. TD: Only a main effect for $R^{si}/R^{ts} F(1,28) = 29.75$, $p < 0.001$. No main effect for priming $F(1,28) = 0.50$, $p = 0.48$ and also the interaction does not reach significance, $F(1,28) = 0.02$, $p = 0.88$. TD responses are nonsignificantly 8 ms faster than control responses in R^{si}. The acceleration by 11 ms in $R^{ts}$is not significant. DD: Main effects for $R^{si}/R^{ts} F(1,28) = 29.16$, $p < 0.001$ and priming $F(1,28) = 10.90$, $p < 0.01$, but the interaction does not reach significance, $F(1,28) = 1.40$, $p = 0.24$. DD responses are 65 ms faster than control responses in $R^{si}$and 31 ms in R^{ts}.

Summarizing, the interaction between R^{si}/R^{ts} and priming in the DT and TT conditions supports the hypothesis that negative priming is an effect of attentional processes, and positive priming is produced during perception. The two conditions where the probe distractor is repeated from the prime trial do not produce such a clear pattern.

6.5.7 Discussion

The overall impression when looking at the post-cue experiment is that it implements a very demanding task. Reaction times are nearly three times longer than in the pre-cue experiment. The difficulty, supposedly, led to the development of a facilitating strategy in certain subjects. Subjects applying such a strategy did not process the objects consciously as we wanted them to, in order to achieve a clear separation of stimulus identification and target selection phase. Mostly, the mask we introduced for the second half of the subjects deleted the possibility to use an after image, but did not disturb the obtained priming effects, despite a strange deceleration in DD_SW trials in the masked group. The most puzzling aspect is the difference of R^{si} in DD_SW trials and R^{si} of DT trials. Both conditions are perceptually identical, and up to the point of the button press that determines R^{si} absolutely no information about whether the trial will be a DT trial or a DD_SW trial is available to the subjects. We therefore suppose a technical issue like an imbalance of stimulus objects in the DD_SW condition after outlier removal to be responsible for the seeming precognition.

Ignoring the difference in DD_SW trials, no further interaction of priming and mask is present, and we are able to pool both groups with and without mask in order to determine the location of priming effects in a perceptual stage or the target selection phase of a trial. The overall reaction time shows no negative priming effects in non-switch trials, but positive priming for trials that identically repeat an object, i.e. TT and DD. In switch trials the identical repetition only speeds up trial processing in the TD_SW condition, i.e. where the prime target reappears. The other three switch conditions slow the trials down, see figure 6.9, left side.

The separation of the trials reveals that in the first part, R^{si}, only acceleration effects are present, again disregarding DD_SW, whereas the target selection phase shows a more differentiated picture. The second reaction time R^{ts} shows a negative priming effect for DT and DT_SW but even more prominent for TT_SW. An acceleration is present for DD and TD_SW, see figure 6.9, right side. The data indicates that positive priming is a perceptual phenomenon and negative priming is produced by interferences during selective attention to the target object.

6.6 General Discussion

Overall, the introduction of a task switch, including responses to green as well as red targets seems not to alter the basic negative priming effect in a DT condition. Also, the positive priming effect in the TT condition is conserved. Another pattern shows up in the trials that immediately follow a task switch. In the pre-cue experiment, a repetition of an object only slows down the reaction, but no benefits are visible. As the pattern is different in the post-cue experiment, the postponing of the cue has a strong impact on trial processing.

Predictions of the ISAM for conditions TD and DD are well matched in the baseline experiments, see figure 6.6. TD produces a negative priming effect that is weaker than the one in DT trials, and DD trials are accelerated to a smaller extent than TT trials. Already in the pre-cue experiment the simplicity of the ISAM is no longer able to reproduce the complexity of the data. The behavior of the ISAM relies on perceptual parameters. The situation for the ISAM gets even worse when assuming an additive process occurring during switch trials with a repeated object as discussed in section 6.4.7. The effects obtained this way follow the conceptual nature of the trial and not the perceptual one as the ISAM predicts. Finally the predictions of the ISAM for the post-cue experiments further show the inappropriateness of the model. The ISAM predicts only negative priming but no acceleration for identically repeated objects in the perceptual phase. Only with the onset of the color cue, positive priming emerges and still some decelerations in perceptually nonmatching object repetitions occur, which is completely contrary to the experimental results.

The assumption of a forced decay of activation to be responsible for negative priming via the influence on the adaptation speed of a global threshold variable as the ISAM makes, directly leads to an inevitable slow-down in the stimulus identification phase, as at a later stage in the trial all activation from the prime trial is gone. The experimental data merely shows a tendency to facilitation of DT trials in the R^{si}, due to perceptual similarity. Therefore, our experiment buries the ISAM, at least in the implemented form we introduced in chapter 3.

Depending on how literally the statements of response retrieval theory are taken, the localization of positive priming in the first partial reaction time is problematic or not. On the one hand, response retrieval explicitly links priming effects to the automatic retrieval of a response, which would suspect no priming effects in the perceptual part of the trial. But response retrieval does not exclude other facilitation mechanisms to happen, possibly even in the perceptual phase. Without the assumption of an additional facilitating process, response retrieval theory is not able to explain an acceleration of DD trials as compared to CO.

Besides the slowing in the DD_SW condition, episodic retrieval can well explain the acceleration in all the other priming conditions, as the similarity triggers the retrieval of the former episode. This then positively interferes with the identification of the current objects. In the target selection phase, the predictions of episodic retrieval are compatible with non-switch trials, but do not fit for switch trials. In TT_SW trials, which repeat the prime target in a different color (which should still trigger retrieval) the object is to be responded to again. Contradicting the facilitation presumed by episodic retrieval the experiment shows a prominent slow-down of R^{ts}. Similarly the shorter R^{ts} in TD_SW trials is in contrast to episodic retrieval theory.

Unfortunately, we are not able to exactly reproduce the computational simulations of distractor inhibition theory. Therefore, no predictions to be tested are available for the second reaction time, R^{ts}. Concerning R^{si}, the perceptual facilitation for TT is predicted correctly. But distractor inhibition theory would as well predict a perceptual slowing in DT trials, which could not be verified by our experiment. As distractor inhibition theory hypothesizes the negative priming effect to stem from persistent inhibition carried over from the prime trial, negative priming should be most prominent at the beginning of a trial, which is in opposition to our results.

Summarizing, the current series of experiments shows an interesting way to split trial processing into several parts, each devoted to a certain aspect of processing a trial in a negative priming experiment. The changes that have to be made to the paradigm in order to assure serial processing obviously introduce multiple effects that are not present in traditional negative priming experiments. Apparently negative priming does only show a similar pattern after task switches if the relation to CO is disregarded but only relative reaction times between the priming conditions are considered. In any case, if we restrict our considerations to the two conditions of most interest, i.e. DT and TT, the experiments confirm the statement that positive priming is a perceptual phenomenon, whereas negative priming is produced at later stages of a trial, target selection or response generation. The pattern of results is best explained by episodic retrieval theory and strongly challenges distractor inhibition and the ISAM.

6.7 Summary

We introduced a way to assess the time course of negative priming on a behavioral level. We altered the voicekey paradigm such that a color cue indicates each trial anew which one of the objects is the target. With a series of three experiments, we built the bridge from our previous studies to the post-cue paradigm. This gives the target cue after the presentation of the two objects, enforcing the subjects to first identify both objects and to select the target and generate the response in the second part of a trial.

The extension of the paradigm brought along interferences from task switches which were shown to interact with priming effects. The consideration of trials immediately following a task switch also shows interesting phenomena. Apparently task switches introduce a general slowdown of reaction time if an object is repeated, but relative reaction times in the task-switch priming conditions show a pattern similar to the conditions that do not require a task switch. Seemingly the coincidence of task switch and object repetition triggers a resolving mechanism independent of the concrete priming condition.

The splitting of trials into two phases shows first that whenever the repeated object is shown in the same color, the stimulus identification is accelerated. Also, if the repeated object changes color, identification is unaltered. We show that positive priming is a perceptual effect, and negative priming happens during target selection or response generation.

7 Selection or Response Effect

The results of the last chapter left open whether negative priming is produced during target selection or response generation, due to the lack of a time marker in between those two phases. The present chapter will now solve the question with a second experiment that provides multiple reaction times per trial, i.e. one for perception and target selection, another one for the response generation phase.

The design is tailored for a critical test of response retrieval theory, see Section 2.4.3, which postulates priming to happen during response generation (Rothermund et al., 2005). In order to prevent a decision about the appropriate response until the target is selected, we extend our paradigm by a comparison word which is spatially separated from the stimulus objects, such that a simultaneous evaluation is impossible. If the target object and the word match semantically, a button assigned *yes* is to be pressed, if not the alternative *no* button is the correct response. Due to the spatial separation, a gaze shift from stimulus objects to the comparison word is required which we detect by recording the electrooculogram. As the ISAM structurally depends on a one-to-one mapping of stimuli to actions, we derive our hypotheses mainly from response retrieval theory.

We will begin the chapter with a description of the gaze-shift paradigm in section 7.1 and present the derived hypotheses in section 7.2. The experimental procedure and the analysis are explained in section 7.3. Results are shown in section 7.4 and discussed in section 7.5. The study reported in the current chapter was part of the diploma project of Matthias Ihrke, and has already been presented in the corresponding thesis (Ihrke, 2008).

7.1 Gaze Shift Paradigm

Response retrieval postulates that the systematic variations in reaction time are due to retrieval of the prime response, which then can either facilitate or lead to a conflict, see Section 2.4.3. The fundamental assumption can be tested in an experimental setting that records both the time until target selection is finished and the actual reaction time until the response is given. Such an experimental setup must ensure that no information about the correct response is available before finishing target selection. Admittedly, such a manipulation introduces a severe change in trial processing compared to traditional negative priming experiments that allow for a parallel handling of target selection and response generation.

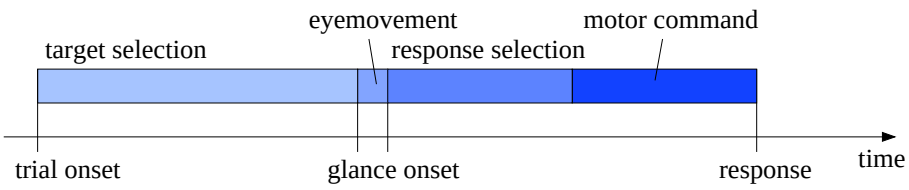


Figure 7.1: Hypothetical time course of processing during one trial in the gaze shift paradigm. First the target is identified and selected, followed by an eyemovement to focus the comparison word. Then the response is generated and performed.

Acquiring information about the time taken until the target selection process finished should not interfere with the natural course of processing of the trial. In order to achieve minimal disturbance, a natural marker of the completion of the target selection has to be measured. In the current study, the movement of the eyes from the target/distractor compound in the direction of the comparison word is used as time marker. The overlapping alignment of target and distractor, as it has already been described in chapter 3, allows for a foveal processing of target and distractor at the same time. To elicit an eyemovement that is strong and reliable enough to be measured validly from the electrooculogram (EOG), the distance between stimulus objects and the comparison word is chosen large enough so that no simultaneous processing is possible. Therefore, we ensure a serial processing of the two parts of the trial as visualized in figure 7.1. Then the reaction time in each trial is composed of several parts: the time taken by the target selection R^{ts}, the response selection part R^{rs} and a constant time for the eyemovement and the motor command ζ

$$R = R^{ts} + R^{rs} + \zeta. \tag{7.1}$$

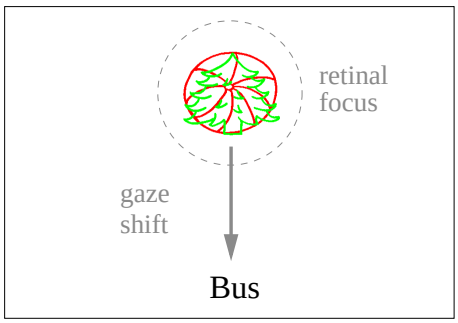


Figure 7.2: Stimulus display of one trial: the comparison word is shown at the bottom of the presentation screen. Participants have to shift their point of fixation in order to complete the task.

Averaging the interval from trial onset to the eyemovement gives an estimation for R^{ts}, while the time taken for response selection cannot be separated from the remainder. Since ζ can be assumed not to vary systematically with priming or response condition, any effects in the interval between eyemovement and reaction can be assigned to the response selection mechanism.

During the experiment we record 60 channels of EEG and four channels EOG for gaze shift detection. As we want to focus on the behavioral part of the experiment and the division of reaction times into two parts, we will only shortly refer to the EEG related results.

7.2 Hypotheses

Rothermund et al. (2005) examined only DT, DD and control conditions. However, the underlying mechanism of response retrieval can be used to derive predictions for further priming conditions. Response retrieval theory postulates an automatic retrieval of the response of the prime trial triggered by any similarity between the prime and the probe display to be responsible for priming effects. Whenever the appropriate response is repeated in the probe trial and some perceptual feature from the prime trial repeats, a facilitatory effect on probe response should occur. Accordingly, a negative priming effect is expected whenever the prime and the probe displays share common

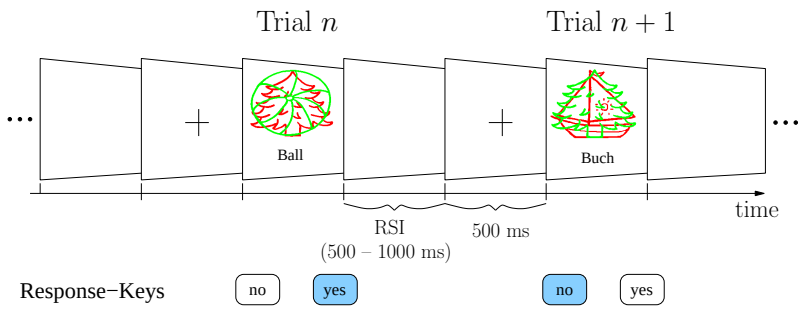


Figure 7.3: A sample trial pair. Trial n primes trial $n+1$, being the probe for trial $n-1$. Shown is a DT condition with a switch of the appropriate response. Note the possibility to vary response and priming condition orthogonally.

features and the required responses do not match. The question remains, whether the assumption of Response Retrieval holds true for other priming conditions like TD as well, i.e. any of the conditions DT, TT, DD and TD leads to an acceleration whenever the appropriate response is repeated, but to a slowdown if the responses changes.

Distractor Inhibition theory predicts a slowdown in reaction time for TD trials, as remaining activation of the prime target has to be inhibited additionally. Episodic Retrieval would expect similar modulations as the retrieved tendency to respond to the probe distractor has to be resolved. Unfortunately, the ISAM as described in chapter 3 is not able to deal with comparison tasks and therefore cannot predict effects of a situation with an orthogonal variation of priming condition and response. Only half of the eight conceivable conditions priming $\times$ response relation can be predicted, see table A.15. Regarding the TD condition which is entangled with a response switch, the ISAM predicts a negative priming effect.

Concerning the localization of effects in one of the two stages of processing, response retrieval theory would expect priming effects in the response selection part of the reaction time R^{rs}. But if negative priming is produced on the level of semantic representations as assumed by the other theories, the effect would be present in the target selection time R^{ts}.

As discussed in chapter 4, the research on the electrophysiological correlates of negative priming is very sparse. A potential candidate for a negative priming sensitive component is the P300. Since the current study is relatively similar to the design described by Behrendt et al. (2009), similar results are to be expected. However, there are some differences in the design that could affect the occurrence of the expected P300 effect. The eyemovement changes the percept and could thus trigger a second P300 for evaluation of the new stimulus, the comparison word. It is thus unclear whether the expected reduction in P300 amplitude would occur for the first or the second part of the display. Also, the modulation in the late PSW could be hard to observe due to the expectation of generally longer reaction times in the current experiment. Regarding the location of negative priming sensitive components on the scalp, we could expect the prefrontal areas to yield the best results since these have been shown to be sensitive to negative priming in fMRI studies (Wright et al., 2006).

7.3 Gaze Shift Experiment

7.3.1 Design

In order to produce strong priming effects, we use a conceptual object comparison task. The response is given via a forced choice equal vs. different decision which is recorded via button presses. In each trial the subjects has to compare a green target object accompanied by a red distractor with a grey word and to decide whether they are semantically identical or not, see figure 7.2.

The main part of the experiment consists of 924 trials. The priming condition in each trial and the repetition of the required response serve as independent variables that are varied within subjects. 5 priming conditions are realized: TT, DT, DD, TD and control. The sequence of responses to trial n and to trial $n+1$ serves as a second independent variable, which can take 4 different values: *yy*, *nn*, *yn* and *ny* depending on whether yes or no is the correct response in the prime and the probe trial respectively. Due to a balancing of Yes and No responses we summarize repeated responses, i.e. yy and nn, under the label sr for same response and altered responses, i.e. yn and ny, under dr for different response. Altogether, a 5 (TT, DT, DD, TD, CO) $\times$ 2 (sr, dr) design with two within-subject variables (priming condition and response relation) is realized.

Other factors known to influence the priming effects are balanced or randomized in the design. The RSI is randomized between 500 and 1500 ms, yielding good results in terms of strong negative priming effects in our previous experiments (see e.g. Behrendt et al., 2009). The assignment of left or right hand to one of the responses is balanced over subjects. The number of trials are equal for each priming condition and for each of the four response relations. The different stimulus objects used in the study are equally distributed over the priming and response conditions as well. The stimulus sequence is thoroughly designed such that predictable patterns are avoided, while the optimal distribution of the conditions over the trials is conserved.

When designing a comparison paradigm, a further effect occurs which might interact with the effect of interest, i.e. in our case negative priming. Responses that require a confirmative Yes answer can be given faster than negative No answers, what Singer (1984) explains by additional cognitive effort is required to generate a negative answer compared to confirmative answers. In order to eliminate a possible confirm vs. negate effect in our results, Yes and No responses are equally distributed over the same response vs. different response trials. Nevertheless, we check for an interaction in Section 7.4.1.

7.3.2 Participants

16 (8 male, 8 female) participants took part in the study with mean age $M = 24.8$ years, and standard deviation $SD = 3.2$ years. Participants were rewarded either with money (10 €), or by course credits. All participants have normal or corrected-to-normal vision and are right-handed. They have no red-green deficiency. All subjects were naïve to the purpose of the experiment and had not taken part in a previous study employing similar stimulus material. The participants are requested to fill out the Activation-Deactivation Adjective Checklist (ADACL) (Thayer, 1986), a vocabulary test (Schmidt and Metzler, 1992) and the number-symbol test from the *Nürnberger Altersinventar* (Oswald and Fleischman, 1982), to ensure a homogeneous group of subjects.

7.3.3 Procedure

The subjects were instructed to make just one eyemovement per trial. They should first identify the target object beyond doubt, then fixate and identify the word and finally, press the correct button.

The experiment started and ended with a baseline phase of 36 trials each, where only green objects were presented and have to be named. After 30 introductory training trials, 924 experimental trials had to be completed, with breaks every 84 trials.

Each of the experimental trials was preceded by the presentation of a fixation cross for 500 ms. When the response was given, the display was cleared. The screen remained completely empty for 500 to 1000 ms randomized over all trials.

7.3.4 Extraction of Partial Reaction Times

In order to find the gaze shift latency, 4 EOG electrodes (left and right hEOG and vEOG respectively) were attached to the ocular muscles. Electroocular data is well suited to record eye movements (Joyce et al., 2002) by measuring the potential fluctuations due to the movement of the eye which is a dipole. The angle of the eye movement is approximately proportional to the resulting change in the recorded potential. Here we are only interested in a valid measure of the latency of the saccade downwards. Therefore we just consider the left and the right vEOG electrodes.

The saccade manifests as a fast positive shift of the potential at both of the two electrodes simultaneously. In a first step, the data from the EOG channels is low pass filtered with a cutoff at 20 Hz. The data is then segmented relative to markers for the trial onset in intervals $[-500, 2000]$ ms. We seek for the global maximum in the first discrete derivative, see figure 7.4. If it is present in both left and right EOG within a time window of 100 ms, we consider its peak latency as R^{ts}. Otherwise the trial is marked as invalid and not further processed.

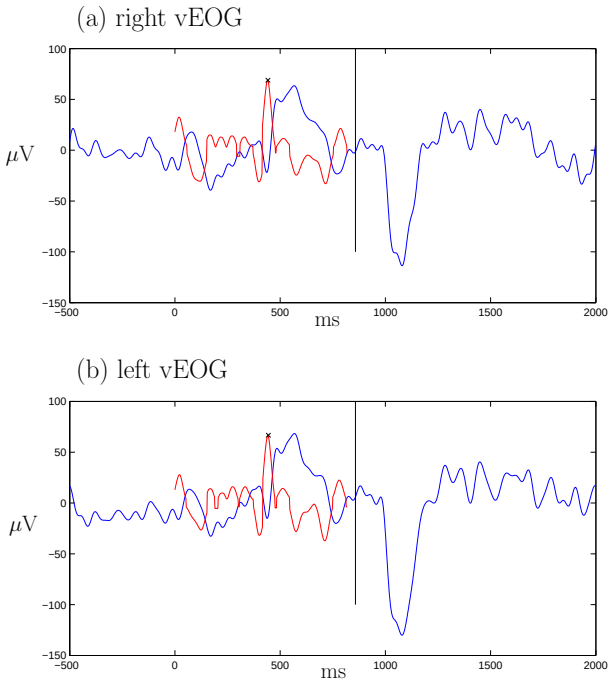


Figure 7.4: Extracting the largest vertical eyemovement. The plot shows the potential fluctuations in the (a) right and (b) left vEOG electrodes (blue line) and the first discrete derivative (red line) in a sample trial. The maximum indicating the gaze is marked by a cross.

7.3.5 Analysis of Behavioral Data

A global analysis of variance is applied to the complete dataset of reaction times. In case of significant main effects or interactions within the factors, single ANOVAs for the factors are computed until individual differences can be contrasted with single t-tests. If more than one t-test is applied to the same dataset, the Bonferroni-correction is applied to the p-values. Generally, a result is treated as significant, when the p-value fell below 0.05. The codes of significance used are as follows: $\star$ for $p < 0.05$, $\star\star$ for $p < 0.01$ and $\star\star\star$ for $p < 0.001$.

7.3.6 EEG Data Analysis

In a first step, visual inspection of ERPs is used as an approximation for the location of possible effects. Once an approximate time window is determined, an analysis of variance is carried out on the mean amplitude value, treating electrode and experimental condition as repeated measures. The main effect of electrodes is considered trivial. In case of significant interactions, pairwise planned contrasts are computed to find out which electrodes or conditions differed significantly. The contrasts are again guided by visual inspection. In case of several simultaneously applied contrasts, the Bonferroni correction is used to avoid a bias in the family-wise error rate.

For analysis, data is downsampled to 500 Hz, bandpass filtered with a low cutoff at 0.5 Hz and a high cutoff at 20 Hz. The data is then split into segments of 2500 ms, i.e. the time window $[-500, 2000]$ ms relative to the markers for stimulus onset. For further analysis, we exclude all trials that were classified as invalid in the outlier analysis of behavioral data. Next, a baseline correction in the interval $[-100, 0]$ ms relative to display onset is carried out, followed by a crude artifact rejection that rejected all data that crossed 1000 μV in absolute values. Then, we rereference algebraically to the mean of TP9 and TP10 and again again perform the baseline correction in the interval $[-100, 0]$ ms. Then, a finer artifact rejection with threshold ± 100 μV is employed. Finally the data is sorted and averaged according to experimental condition.

7.4 Results

We will first have a look at the overall reaction times and look for a response repetition effect. Then we split the reaction times according to the gaze time marker and again determine priming effects in the two segments. In the end we will shortly consider EEG correlates of the observed effects.

7.4.1 Response-Repetition Effect

Concerning the possibility for side effects of the confirm vs. negate effect, the 5×2 (priming $\times$ probe response) ANOVA shows the expected main effect for probe response, $F(1,15) = 39.741$, $p < 0.001$. No responses are consistently slower for all priming conditions. But although the interaction probe response $\times$ priming becomes significant, $F(4,60) = 2.88$, $p < 0.03$. Separate ANOVAs showed that only in the comparison of DD and control is the interaction between probe response and priming significant, $F(1,15) = 5.50$, $p < 0.05$. But on a level further the comparison of priming conditions for each of the responses does not reach significance (Yes answers, $t(15) = 0.9$, $p = 0.38$; No answers $t(15) = -1.77$, $p = 0.10$), finally justifying a pooling of trials according to the response repetition or switch, disregarding the concrete answers.

A 5×2 ANOVA for the reaction times yields main effects for priming condition $F(4,60) = 6.63$, $p < 0.001$ and response repetition $F(1,15) = 19.48$, $p < 0.001$ as well as a significant interaction priming $\times$ response repetition, $F(4,60) = 2.85$, $p < 0.05$. The response is significantly

delayed over all priming conditions when the same response has to be given compared to when the required response switched. To specify the reaction time effects, 4 separate 2×2 (priming $\times$ response repetition) ANOVAs are carried out for all priming conditions against control.

DT Main effects for priming $F(1,15) = 4.62$, $p < 0.05$ and response repetition, $F(1,15) = 9.94$, $p < 0.01$. The expected interaction does not reach significance, $F(1,15) = 0.54$, $p = 0.47$. DT responses are slower than control responses, independently of response repetition. Conversely, responding is delayed only for control trials when the response is repeated, $t(15) = -2.67$, $p < 0.02$ but not so for DT trials, $t(15) = -0.67$, $p = 0.51$.

TT Main effects for priming $F(1,15) = 9.28$, $p < 0.01$ and response repetition, $F(1,15) = 15.78$, $p < 0.01$. The expected interaction does not reach significance, $F(1,15) = 0.91$, $p = 0.36$. Responding is delayed in response repetition trials for both control and TT trials, and the reaction times are shorter in the TT condition, regardless of the response.

TD Only the main effect for the response repetition reaches significance, $F(1,15) = 18.34$, $p < 0.001$. Again, responding is delayed in case of a response repetition.

DD The interaction priming $\times$ response reaches significance, $F(1,15) = 6.41$, $p < 0.03$. There are no main effects. Separate t-tests reveal that the responses are significantly slower in DD trials when the response shifts, $t(15) = -2.02$, $p < 0.05$ but somewhat faster or at least not delayed, when the response has to be repeated, $t(15) = 1.25$, $p = 0.23$.

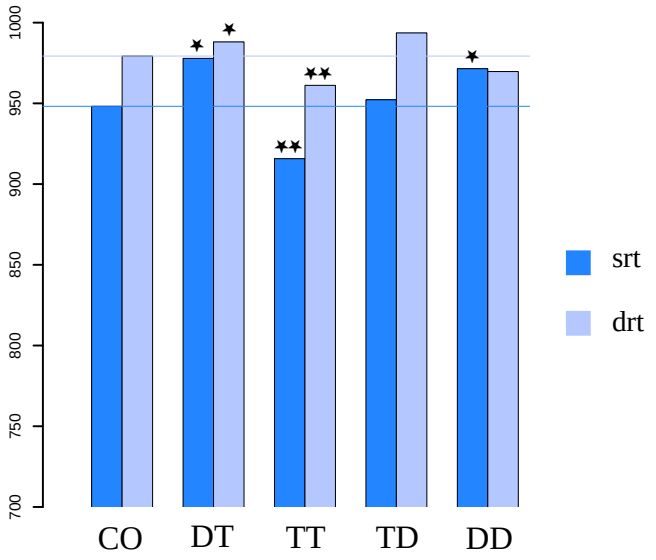


Figure 7.5: Summary of the results clustered by response repetition. Significances in the comparison to the corresponding control condition are indicated.

Figure 7.5 and table A.16 in Appendix A give a summary of the significant results. The results are largely inconsistent with response retrieval. A general slowdown in responding is present for the DT condition, and a general facilitation for TT, independent of response relation. Only the DD condition shows a significant interaction.

7.4.2 Partial Reaction Times

To answer the question whether priming effects are produced during target identification or response selection, we computed the partial reaction times as outlined in Subsection 7.3.4. R^{ts} and

R^{rs} then entered as dependent variables into a similar analysis as described above. Responses were again clustered by response repetition.

The overall 5×2 (priming $\times$ response repetition) ANOVA using R^{ts}as dependent variable reveals only a main effect of priming, $F(4,60) = 5.62$, $p < 0.001$. Because of the absence of an interaction between response repetition and priming, we now pool over both response relations. Separate analyses for the priming conditions show that the effect is present in the DT condition, $F(1,15) = 9.32$, $p < 0.01$ and the TD condition, $F(1,15) = 6.92$, $p < 0.05$, but not in the DD and TT condition ($F(1,15) = 1.79$, $p = 0.2$ and $F(1,15) = 1.78$, $p = 0.2$). No other main or interaction effects are present. The target selection R^{ts} is thus delayed in both response repetition and response shift cases in DT and TD trials ($t(15) = -2.72$, $p < 0.05$ and $t(15) = -2.84$, $p < 0.05$).

The equivalent ANOVA for the response selection time R^{rs} results in a main effect for priming condition, $F(4,60) = 5.17$, $p < 0.001$ and a main effect for response repetition, $F(1,15) = 9.99$, $p < 0.01$. The priming $\times$ response-repetition interaction is not significant, $F(4,60) = 1.34$, $p = 0.26$. Again, we combine response repetition and response switch trials of one priming condition. Separate ANOVAs for the single priming conditions contrasted against control trials show that the main effect for the response repetition is present in all priming conditions (DT, $F(1,15) = 13.89$, $p < 0.01$; TT, $F(1,15) = 6.29$, $p < 0.05$; TD, $F(1,15) = 5.43$, $p < 0.05$; DD, $F(1,15) = 5.72$, $p < 0.05$). The response selection is thus significantly slower when the response has to be repeated in all priming conditions. The main effect priming stems exclusively from the TT condition, $F(1,15) = 11.67$, $p < 0.01$.

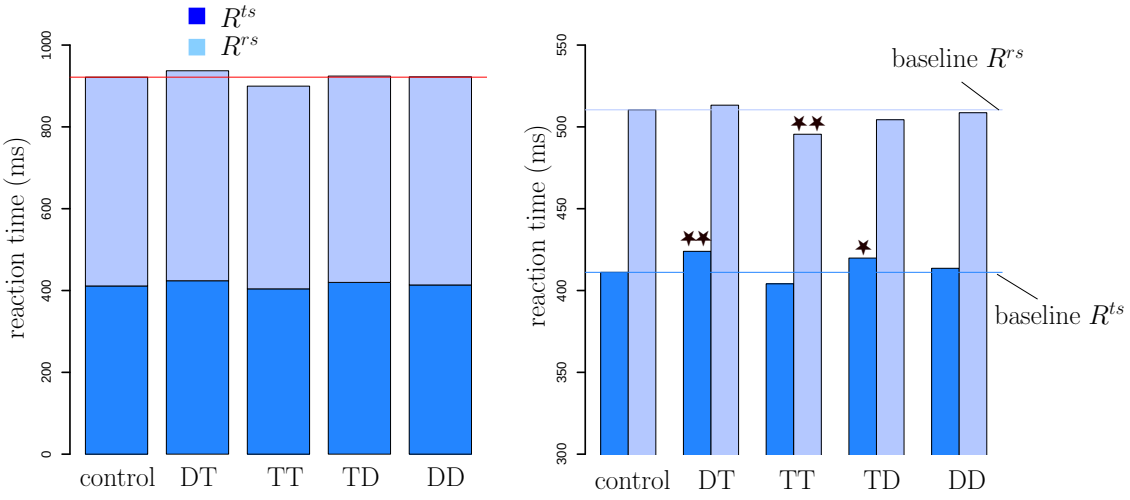


Figure 7.6: Summary of the results for the partial reaction times pooled over response relation. The left plot shows R^{ts} and R^{rs} stacked on top of each other, the right shows the priming effects for both partial reaction times.

7.4.3 EEG Correlates

For completeness, we shortly report the results from EEG analysis. In order to find correlates for the observed behavioral effects, we contrast TD and TT against control. Because no interaction of response relation with the partial reaction times R^{ts} and R^{rs} is present, the ERPs are pooled over response relation. Because the target selection was usually completed after ≈ 500 ms, early ERP components are scanned for correlates of the R^{ts}effect.

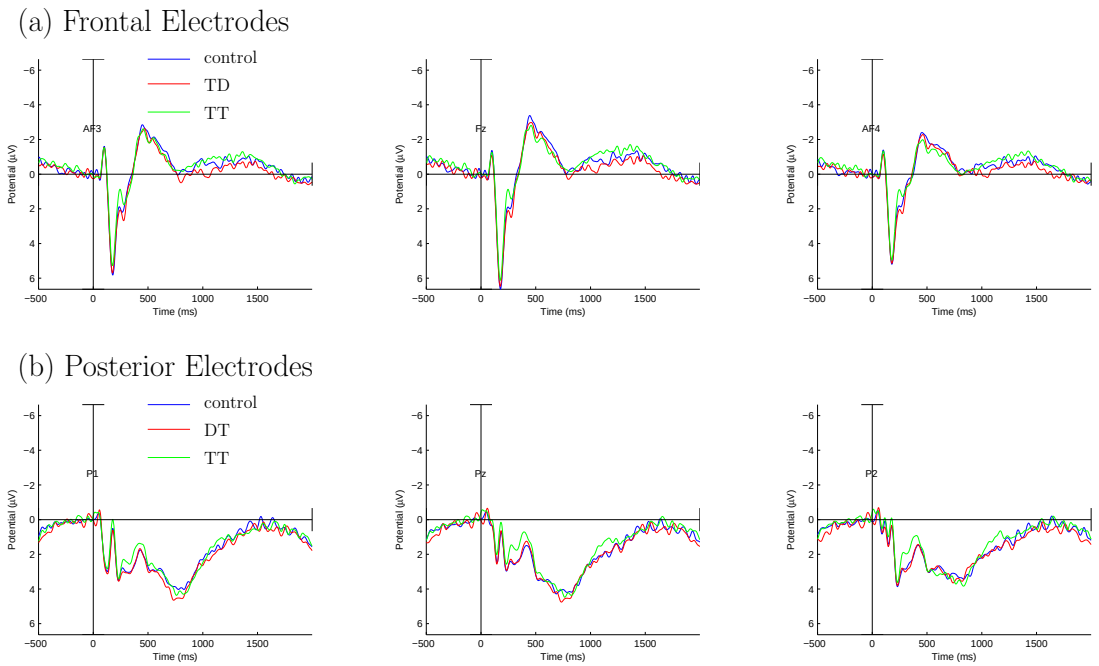


Figure 7.7: Electrodes AF3, Fz, AF4, P1, Pz and P2 for control, DT and TT trials.

Visual inspection of the data reveals differences of the conditions in an early time interval $[200,300]$ ms in the frontal electrodes. There seems to be priming sensitive modulation such that the N1 is enhanced or the P2 reduced for the TT condition relative to control and vice versa for the TD condition, see figure 7.7a. To confirm the visual impression, a 3×62 ANOVA (priming conditions $\times$ electrodes) is applied to the mean amplitude in the interval $[200,300]$ ms post stimulus. There is a main effect for priming, $F(2,30) = 7.45$, $p < 0.001$ and a significant interaction priming $\times$ electrode, $F(122,1830) = 2.92$, $p < 0.001$. Pairwise contrasts of the conditions in the frontal electrodes {FP1, FPz, FP2, AF3, AFz, AF4} with Bonferroni correction show that the priming conditions are indeed different from control, $t(95) = -2.53$, $p < 0.05$ (TD) and $t(95) = 4.01$, $p < 0.001$ (TT) and also that the priming conditions are substantially different from each other, $t(95) = 6.39$, $p < 0.001$. This pattern is only found in the frontal electrodes. In the posterior cluster, only the P300 effect for the TT condition, $t(111) = 3.09$, $p < 0.01$ is present and no difference between control and TD is observed, $t(111) = -0.81$, $p = 0.42$.

7.5 Discussion

The current chapter investigates the validity of the theoretical assumptions underlying response retrieval theory by separating the critical response generation phase from stimulus evaluation and target selection. We use a comprising set of priming conditions and equip each trial with a time marker that divides the overall reaction into a part of target selection and a part attributable to response generation. We are for the first time able to temporally localize priming effects according to the two phases.

We find a main effect of response repetition. There was a consistent delay in response repetition trials compared to response switch for all priming conditions except DD. It is common in forced choice tasks with only a few response alternatives that responses are delayed when the same response has to be given to a new stimulus display (Marczinski et al., 2003). The effect is strong

enough that repetition of the prime response is even slowing TT trials, which is counterintuitive in the framework of priming.

As expected from the literature on yes-no response effects (Singer, 1984), a very strong main effect of probe response is present in the data. Negating responses are significantly delayed compared to confirmations in all trials. However no differential impact of the probe response on priming conditions becomes significant.

We derive predictions for all four priming conditions, predominantly from response retrieval. Surprisingly the overall result for condition TD was reversed compared to what response retrieval suggests. TD trials are significantly faster for response switches compared to control trials instead of response repetitions, a fact that is not only incongruent with response retrieval but also with all other theoretical accounts to negative priming. Distractor Inhibition predicts a decelerated rather than an accelerated response due to persisting activation of the current distractor object stemming from the previous display where it appeared as target. Episodic Retrieval also predicts an increase of reaction time since the tendency to respond to the retrieved object has to be resolved to ignore the particular object in the probe. Also predictions from the ISAM tend in the opposite direction, see table A.15. Since none of the theories can explain the observed TD effect, future research should investigate the TD condition more thoroughly.

Whether the prime response is retrieved or not does not only depend on the overall similarity of the displays and additional processes beyond response retrieval must be at work. The assumption taken by Rothermund et al. (2005) that the prime response is incidentally associated with all aspects of the prime episode might be moderated by a process as proposed in the framework of temporal discrimination theory (Milliken et al., 1998). The full match in a TT trial is detected by the scanning process resulting in an automatic retrieval of the prime response leading to faster or slower reactions depending on whether the retrieved response is correct or not, respectively. For the TD or DT cases the similarities are not strong enough, and the display is classified as new thus leading to a suppression of response retrieval.

The extraction of the partial reaction times produce a separation of priming effects. Apparently, the DT and TD priming effects are produced in the target selection stage, while the TT effect occurs exclusively during response selection. For the TD condition, the division of the reaction time into target and response selection revealed a hidden effect which did not reach significance in the overall reaction time but is present in the target selection part. The temporal separation of priming effects implies that more than one mechanism is responsible for the different priming effects. At first glance, the distinctive feature is whether an object is repeated in a different color (DT and TD) or in the same color (TT). Unfortunately, no significant priming effects in the DD condition are visible in any of the reaction times. Response retrieval postulates the priming effect to stem from a conflict of responses. This is in clear contrast to the fact that the DT effect is observed in the target selection phase.

One reason for our result pattern may be the alterations we made in the paradigm. The forced seriality of processing may produce unexpected side effects. Even if the stimuli are identical to former studies that show strong negative priming, the manipulation of writing the comparison word out of the focal area leads to a split of the stimulus display in time, such that subjectively two different stimuli are visible in each trial. As negative priming is shown to be sensitive to the length of the response stimulus interval, we might use a very long time interval between the two presentations of target and distractor in the prime and the probe trial, as the time to read and process the comparison word adds to the interstimulus interval. Also, the visual input of the grey comparison word can act as a mask, albeit a very weak one as it is very dissimilar to the target and distractor objects, thereby interfering with negative priming. As response retrieval assumes the archiving of the last trial at the time of the response, another reason for our result pattern

may lay in the association of the response with the percept of the comparison word. Therefore, similarities of the two stimulus compounds are not the important factor for triggering a memory retrieval, but perceptual similarities between the comparison words. Unfortunately, we took great care in designing a clean stimulus sequence, thereby eliminated as much additional repetitions as possible, thus permitting repeated comparison words exclusively in TT-yes trials. That is why we are not able to rerun the analysis based on the factor of word similarity.

7.6 Summary

After the last chapter identified retrieval based theories to be better applicable than representation based theories, we then wanted to test the two prominent examples against each other. We introduced a time marker between the target selection phase in the beginning and the response generation phase, where the crucial difference between episodic retrieval and response retrieval is supposed to become apparent. In order to least disturb the subject's processing of trials, we forced a detectable eye movement from the stimulus objects on the top of the screen to the comparison word at the bottom. The obtained results showed that different priming conditions behave differently in the two partial reaction times. Our experiment showed that DT and TD effects occur during target selection while TT produced a facilitation at response generation, thereby favoring episodic retrieval theory over response retrieval.

8 The General Model for Negative Priming

We showed in chapter 6 how computational modeling is able to advance the discussion about the applicability of negative priming theories. Thus it is desirable to have a computational implementation for each theory. A direct reproduction of the only computational model besides the ISAM, the distractor inhibition model by Houghton and Tipper (1994), is not possible. In personal communications, the authors stated to have lost any source code over the years. All other theories are described in a very abstract way, often using keywords in a nonstandard way, like episodic memory does not match the functions postulated by episodic retrieval theory, see section 8.2.3. Such unspecific formulations of theories, each postulating different mechanisms to cause negative priming, are hardly comparable.

We therefore developed a computational model that represents a framework capable of acting on any experimental paradigm and representing the mechanisms proposed by each of the theoretical accounts to negative priming. Starting from the different processes interacting while an action has to be selected on the basis of perceptual input, we constructed several building blocks for the model each having a physiological counterpart. Therefore, several research fields have to be integrated despite their diverging vocabulary. Each of these blocks is implemented using the same realistic activation dynamics as the ISAM. Priming effects emerge by the interplay of all aspects during a simulation of the General Model.

The following chapter will motivate the General Model in section 8.1 and describe its components and their biological counterpart in section 8.2. Section 8.3 presents the dynamic equations on which the General Model is based, and section 8.4 explains how the different theories can be tested. We then show some exemplary simulations in sections 8.5 and 8.6. The present introduction of the General Model will be concluded by a critical discussion about the strong point and the weaknesses of the General Model in section 8.7.

8.1 A Framework to Test all Negative Priming Theories

The starting point for the General Model was our attempt to extend our implementation of the ISAM described in chapter 3 in order to simulate a word-picture comparison task like the one introduced in chapter 7. First of all we sought a way to incorporate different paradigms. The word-comparison task is a good example for what problems arise when aiming for a generalized model. Compared to the voicekey task described in section 2.2, the word-picture comparison provides two different combinations of features in the stimuli: shape-color and word-color. To be accurate, there is also a space dimension in both as they appear on different locations on the screen. Another striking difference between paradigms concerns response modalities. In the voicekey paradigm a direct association of semantic representation and spoken word is possible. Whereas the word-picture paradigm needs an additional stage of processing as the two semantic representations of word and picture have to be compared and the match or mismatch has to be mapped on a yes or no response, still disregarding that those responses are to be mapped onto a left or right hand button press. Thus, we were not able to conserve the simplicity of the implementation of the ISAM.

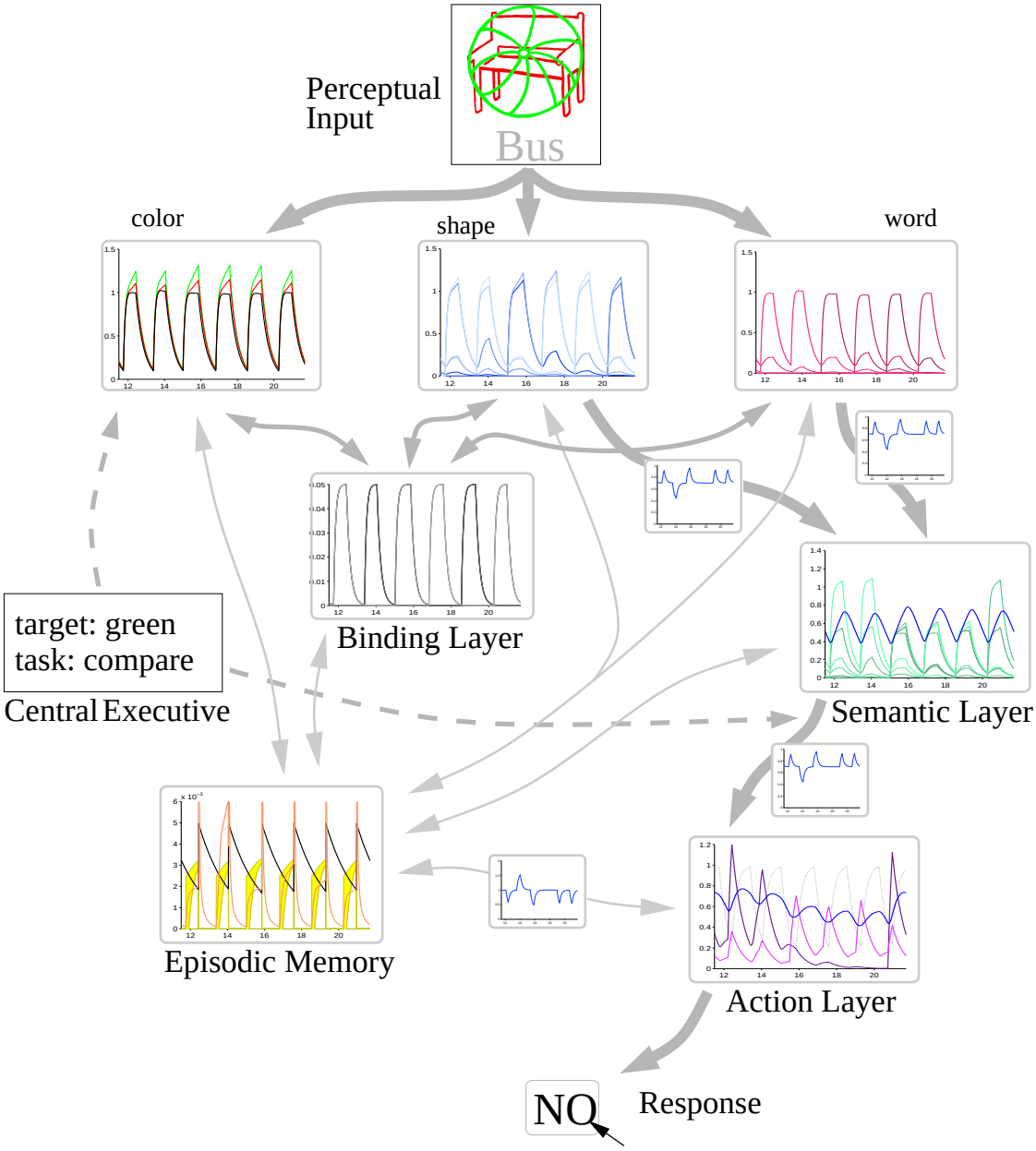


Figure 8.1: Interaction scheme of the different components of the General Model. Perceived stimuli are decomposed into single features, each of which is represented in a single variable in the according layer. Object identity is kept track of by activations in the binding layer, associating the different features of a stimulus object. Most paradigms require a semantic evaluation of the stimuli in order to generate a response. Therefore the semantic layer gates information flow from the relevant features to the action layer which decides on the action to perform. Parallel to the information flow from perception to action a so-called central executive steers the model behavior with regard to the current task, i.e. providing information about the target and the mapping of semantic variables to actions. According to the similarity of the percept and a memorized stimulus configuration, the memory layer feeds back information of the former trial. The similarity signal also affects the effectivity of transmission between features, semantic layer and actions as well as between memory itself and actions, the latter inversely to the first.

8.1.1 Different Paradigms

The description of the initial implementation of the distractor inhibition theory (Houghton and Tipper, 1994) is able to deal with different perceptual input. Stimuli are recognized by specialized feature units. The decomposition of the percept into the relevant feature dimensions offers good flexibility of the model to receive very different stimuli, see section 8.2.1. Such a mechanism nicely determines on a more concrete level the difference of paradigms. This is supposed to be responsible for the diverging results of negative priming studies. After decomposition, it is necessary to keep track of object entities, which can be realized by a flexible feature binding mechanism (Treisman, 1996). For the General Model we decided to implement a simplification of an earlier attempt to bind feature representations by means of persistent spiking activity (Schrobsdorff et al., 2007a), similar to the simplification leading from population activity of a spiking network to the intrinsic dynamics of the ISAM in section 3.1.

In order to account for different response modalities, we assume two separate layers – one containing activations of semantic representations, the other one activations of certain action schemes. Between these two layers, an abstract mechanism we call central executive implements a possibly nontrivial mapping. The central executive also provides information about which feature instance codes for the target and distractor, and which feature dimension is relevant for responding, see section 8.2.5.

8.1.2 Inclusion of Theories

So far we have explained the points that go beyond the ISAM that are necessary to deal with several experimental paradigms. Additionally, the comparison of diverging theoretical concepts needs clear formulations of the different theories. We were thus looking for ways to include the mechanisms proposed by other theories into the General Model and thus make it worthy of its name.

The inclusion of distractor inhibition into our framework is straightforward. The target is no longer boosted, but the distractor is subject to inhibition. Simultaneously, activations below baseline are necessary for the inhibitory rebound. We therefore simply shift the arbitrary baseline activation from zero in the ISAM to $1/2$, as the exponential dynamics of the ISAM which we want to reuse in the General Model does not allow for negative activations.

Episodic retrieval theory requires the explicit modeling of memory and retrieval processes. Therefore, we add a layer which contains (mostly short-term) memory that is able to store a snapshot of the state of the dynamic system subject to decay over time. The memory layer is also able to compute the strength of retrieval determined by the similarity of the current percept and the one stored in memory. Retrieval is modeled by partially restoring former system variables. Memory gets an update at the most prominent point in a trial, i.e. when the decision takes place.

Response retrieval manifests in the General Model as a simplification of episodic retrieval. Only system variables of the action layer are restored during retrieval. The retrieval strength is still determined by similarity of current and stored percept.

Feature mismatch theory states a conflict whenever an object at a certain location changes shape or identity, but not necessarily its role as target or distractor. Abstracting from the very specific location paradigm on which feature mismatch is based, we assume a generalized form to be in line with feature mismatch theory. The binding of one feature into a new stimulus entity is in concurrence with existing bindings of this feature. Such a concurrence is inherent in the General Model, as it only has a single variable for a specific feature instance. If e.g. two red objects are presented, both object representing bindings will point to the same color variable. Any activation

that is exchanged via one of the bindings is then further propagated through the second binding. In the case of still existing bindings from earlier trials, the situation is similar.

Temporal discrimination theory acts on the same episodic memory layer as the episodic retrieval account. The probability if a stimulus display was just seen can be computed by looking at the similarity between current and memorized percept we already have. This value clearly is highest when both configurations entirely match. As a percept is memorized when a decision is made, all stored values of variables that played a role in the last trial already acquired prominent activation and the new percept first has to be built up. The similarity is thus slowly rising from zero to its final value. We now compute the difference of the similarity variable with a prototype similarity signal that separates perfectly similar displays from stimuli that vary in only one feature. In order to account for initial fluctuations, the difference has to surpass a certain threshold that is large at trial onset but shrinks with time in order to generate an old-new classification. If a display is rather similar to the memorized one, the similarity value will stay longest within the uncertainty interval preventing an old-new classification.

As the classification has been done, temporal discrimination theory assumes the affection of information flow. In the presence of new stimuli retrieval is blocked and direct computation is facilitated. And for old stimuli the direct computation is dropped and retrieval will be performed. This is incorporated into the General Model in terms of modulation of transmission strengths between the corresponding layers: from semantic to action for direct computation and from episodic memory to action for retrieval.

The spirit of the dual mechanism hypothesis is inherent to the General Model, as it will cover all the mentioned theories simultaneously. According to the above brief descriptions it becomes evident that the mechanisms postulated by inhibition and threshold theory are located in the more sensory part of the system, whereas the retrieval accounts act in later parts. Thus, coexistence is trivial. The localization of theories in the concrete formulation of the General Model also unveils the differing nature of the two big parties in the discussion about negative priming. As we already pointed out in chapter 6, distractor inhibition and the ISAM are perceptual theories, whereas the retrieval accounts propose influence of selection processes.

8.2 Characterizing System Components

The last section briefly described the minimal setup of the General Model in order to still cope with a variety of paradigms and to incorporate assumptions from all theoretical accounts. The following section will review facts concerning the corresponding building blocks of the General Model, see figure 8.1. There are two sources of information about the respective components. One is the classification by behavioral psychologists. The other one is delivered by neuroimaging studies that allow for a spatial and temporal localization. The latter procedure is usually based on the first one, as neurophysiological results are interpreted on the basis of knowledge from behavioral experiments.

We already introduced the possibilities to obtain information about mechanisms involved in task accomplishment from EEG data in chapters 4 and 5. Another technique, functional magnetic resonance imaging (fMRI) provides very accurate spatial information. But the nature of the measured blood oxygen level dependency (BOLD) is still unclear even on a cellular level, not to speak of its relation to behavior. Nevertheless, fMRI recordings are a widely used tool to noninvasively obtain spatial information about brain processes. If stimuli are repeated, the BOLD response is weaker than for new stimuli, which can be explained by either of the three hypotheses: fatigue, in the corresponding subnetworks less energy is present due to recent activity; facilitation, less

oxygen is consumed, as the units are still in an activated state; and sharpening, neurons coding for features irrelevant for identification show less activity (Grill-Spector et al., 2006). Support for the latter hypothesis comes from a study considering the memorization of novel stimuli with simultaneous recording of EEG and fMRI (Fiebach et al., 2005). However, there is also evidence that the often observed lower BOLD signal is independent of stimulus repetition, as it is possible to vary them orthogonally (Ganel et al., 2006). The hypothesis that a lower BOLD signal reflects accumulation of neural activity (James and Gauthier, 2006) fits well with the findings of increased activation in priming situations (Fiebach et al., 2005). In higher cognitive areas, faster processing is related with less BOLD (Maccotta and Buckner, 2004). Only the presence of distractors reduces activation in the inferiotemporal lobe, whereas targets enhance activation in the prefrontal cortex (Jiang et al., 2000). Concluding, even if the meaning of the measured fMRI signal is still under debate, we can learn from it about the brain areas that exhibit discriminative signals in the different priming conditions. Those areas are good candidates for a source of the systematic reaction time variations. In the current section, we will discuss evidence for the individual components of the General Model.

8.2.1 Feature Layers

In the visual pathway the percept from the retina is decomposed into low level features, each being represented by different subsets of neurons (Van Essen et al., 1992). Later, the low level representations are recombined to form higher order features of objects from visual input (Prinzmetal, 1995). Feature decomposition bears excellent flexibility for a sensory system, but comes with the disadvantage that the distributed information about an object has to be bound together for the recognition of objects as entities (Treisman, 1996). Synchronization might play a role in this context (Singer, 1995), but the particular implementation is yet to be specified. In the realization of feature binding in the General Model, we refer to feature binding in terms of our earlier feature binding model on the basis of localized excitations in a spiking network (Schrobsdorff et al., 2007a).

In order to cover most paradigms with visual stimuli, we equip the General Model in the presented form with feature layers to detect color, shape, location and word(-shape). See Schacter et al. (2004) for the neural distinction between Gestalt-shape and the shape of written words. Feature bindings are realized as a finite list of objects, each a set of features, and a binding strength which specifies activation exchange between the features of the according object.

Biologically, the binding seems to be located in the parietal cortex as some neurological patients show a dramatic breakdown in the ability to see objects (Treisman, 1998). In the formation of binding, attention seems to form a crucial role, as neuromodulators associated with attention are essential for the formation but not for the maintenance of bindings (Botly and De Rosa, 2007). In our terms this means that objects from currently perceived stimuli are bound, and the binding can survive the vanishing of the perceptual input. Bindings are stable against stimulus changes up to the point where the limited resources are in use, i.e. the maximum number of bindings is reached. Then new stimuli will take over the least active binding slots.

8.2.2 Semantic Representations

Many negative priming paradigms require stimulus evaluation on a semantic level, e.g. the word-picture comparison task introduced in chapter 7, the specialized Stroop cards which are the origin of negative priming research (Dalrymple-Alford and Budayr, 1966), or the naming of pictograms

in the voicekey paradigm from chapter 2.3. Semantic representations are closely related to language processing (Demb et al., 1995), which is rather distributed over the entire cortex.

Semantic tasks correlate with higher activity in the left inferior frontal lobe, categories and concepts can be attributed to the temporal lobe, whereas context and figurative demands distribute over the right hemisphere (Bookheimer, 2002). Semantic processing tasks with either pictures or written words are found to activate a network of regions in the left inferior frontal lobe, left anterior temporal lobe, left middle and inferior temporal gyri, left top junction and the right cerebellum (Devlin et al., 2002). Despite the distributed nature of semantic processing, the General Model includes only one layer holding strengths of the semantic representation of a certain stimulus in a similar way to the ISAM. The General Model also inherited the attention mechanism from the ISAM, i.e. an adaptive threshold relying on activations in the semantic layer. The threshold controls information propagation to response generation.

8.2.3 Episodic Memory

Episodic retrieval theory, see section 2.4.2, explicitly refers to episodic memory for storing information about one experimental trial which is retrieved later on. Unfortunately, there is a mismatch with what is called episodic memory in the field of memory research. Episodic memory in that context is generally seen as the capacity to reference personal experiences in the context of both time and space (Tulving, 2002). Experienced sequences are memorized together with temporal information. The basic paradigm to consider such memories is regarding place cells in rats and the encoding of the experienced sequences of visited places first in hippocampus, and then the transfer to long term memory during sleep (Tsodyks, 1999; Sejnowski and Destexhe, 2000; Buzsáki, 2005; Suzuki, 2006).

On the contrary, a trial is more or less static in most negative priming paradigms, and also the temporal sequence of trials is not considered beyond the directly preceeding one by any of the theories. The interference of a memory with behavior is assumed to depend only on the time elapsed and stimuli encountered in the meantime. In spite of the unusual usage, we will stay with the term episodic memory for the memory process involved in negative priming experiments, as the General Model is dedicated to advance the discussion in negative priming research and thus should use the language common in that field. Hence, to give a clear idea of the episodic memory of negative priming research, we will now discuss the process in question in detail.

Following the theory on working memory by Baddeley and Hitch (1974), who assume a conscious awareness to retrieve information from episodic memory (Baddeley, 1998), the episodic memory layer of the General Model is not located in the the episodic buffer but rather in the other parts of working memory (see Baddeley, 2000, 2001). Our memory layer goes beyond the notion of short term memory as the maintenance of activation by attention (Cowan, 1988). During typical priming experiments, no attention is devoted to distractors and past episodes. Nevertheless, the corresponding representation is said to be retrieved, and thus had to be memorized before.

Physiologically, memory encoding is related to activity in the left prefrontal cortex, whereas retrieval is more associated with right prefrontal cortex (Tulving et al., 1994; Fletcher et al., 1997). This is conjectured to be due to different control mechanisms on the two tasks (Craik, 2002). We would rather like to speak of a memory trace ”whatever this turns out to be” (Craik, 2002) as a neural blackboard. We solve the stability-plasticity problem that memories have to be formed reliably and instantly but have to persist for some time even in the presence of interfering input (Norman et al., 2005; Suzuki, 2006) by implementing a limited number of memory slots that hold the entire state of the system at a certain point in time. Such a memory is assigned a strength which decays with time. In our scenario, individual instances are the only forms of experience that are represented, as Logan (1988) postulates for neurological representations of memories.

8.2.4 Memory Retrieval

Memory research distinguishes between involuntary retrieval and willing recollection (Yonelinas, 2002). Physiologically measurable is the so-called familiarity signal, which becomes visible in the EEG past 300ms after stimulus onset. Familiarity is assumed to trigger further retrieval, as a spontaneous recognition can lead to willing recollection (Zimmer et al., 2006; Ecker et al., 2007b,a). There is a functional dissociation of dorsal prefrontal cortex for context monitoring and ventral right prefrontal cortex for retrieval itself (Henson et al., 1999). Context monitoring means the evaluation of the appropriateness of a retrieved episode (Egner and Hirsch, 2005). Topography, latency, and polarity of the familiarity signal in EEG-data bears resemblance to the old/new effect related to episodic memory retrieval (Rugg and Nagy, 1989).

The two approaches of episodic retrieval and temporal discrimination theory predict differing mechanisms controlling the strength of memory retrieval. The first theory assumes that involuntary retrieval is positively correlated with perceptual similarity of the two trials. The latter postulates another perception based classification of the encountered episode being old or new. When significant evidence for an old stimulus display is cumulated, full retrieval is triggered, simultaneously stopping the direct response generation.

The General Model performs in any case the computation of a familiarity signal by comparing the current percept with the memorized one. In one case this familiarity directly determines the strength of retrieval, i.e. very familiar situations cause a strong retrieval, whereas unfamiliar stimuli cause very little but still positive retrieval. In the other case, the system holds a template time course of a familiarity signal separating the time courses of the familiarity signal while encountering a perfect match of stimulus displays and a pair of subsequent displays that vary in one single feature. Greater familiarity than the template indicates the identical stimulus configuration, while lower familiarity is considered as being produced by a new display. In order to account for the uncertainty of the signal early in the trial, the General Model implements a shrinking corridor around the template curve in which the evidence of the display being old or new is not yet significant enough.

8.2.5 Central Executive

Though having a high degree of complexity, the General Model is still a very simplified system compared to biology. The General Model does not comprise a mechanism to autonomously adapt to the diverse demands of the different paradigms. We rather assume a mechanism to be in charge of the adaptation and to adjust several parameters or mappings. The discussion about how the brain solves the adaptation to new situations and tasks brought up the famous, but most likely inexistent homunculus in the brain, steering the system on a high abstraction level and keeping track of long-term goals.

The corresponding building block of the General Model is called central executive (Cowan, 1988), not in the meaning of an “reclusive autocrat” (Craik, 2002), but as an emergent property of interacting subsystems (Barnard, 1985; Teasdale and Barnard, 1993; Bressler and Kelso, 2001). Even if there is no consensus on the necessity of a central executive in memory functions (Baddeley, 1998; Johnson, 2007), we will use the term in order to describe the sudden change in system behavior if it is presented a new task. Concretely, we provide the General Model with the task demands by specifying the top-down input modulating target or distractor activation and a set of mappings each describing the input to the action layer for a certain paradigm.

8.3 Implementation of the General Model

After the thorough examination in the previous section of mechanisms contributing to the processing of a negative priming trial, and thus in question of producing negative priming, we will now come to the concrete description of how we implemented the General Model. Again, we are faced with the problem of translating abstract concepts to concrete mathematical dependencies. The way of implementation determines the behavior of the system. Thus, the correct interpretation of theories is crucial for the success of the model to weigh theories against each other. The following implementation of the the General Model is written in an abstract mathematical way, which is necessary to be explicit enough to allow for a reproduction of the simulations, an aspect we consider to be very important in computational research.

Formally, we will use an exponential fixpoint dynamics like we already successfully introduced in chapter 3. The difference between a state variable and a given fixpoint multiplied by a certain time constant determines the change of that variable. Due to the high degree of interconnections of the system variables, we give an overview on the notation by summarizing all variables, parameters and mappings in tables A.20, A.18 and A.19, appendix A.6, respectively. In the following we will go through all dynamic equations in the order of information flow, beginning with the feature layers, passing the semantic layer and the action layer and finally describe the behavior of the memory variables.

8.3.1 Feature Variables

All objects from input space are represented and fully determined by tuples of features. The number of relevant features can vary according to the task. Information about a perceived object is decomposed and then passed to several layers of the General Model: features drive feature detection variables of the system, whereas the information about the combination of all features to one object entity is held by the binding layer. This also defines the synaptic interconnection of feature variables belonging to the same object.

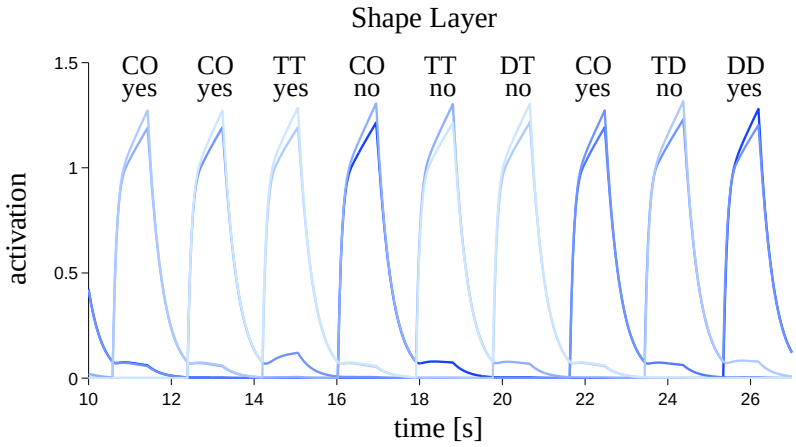


Figure 8.2: Time series of the representation variables in the shape layer f^j_{shape} of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. Different color saturations label different objects. The model is set to episodic retrieval mode, see section 8.4. The plot shows the retrieval of past episodes in dependence of perceptual similarity, see figure 8.9, and also the activation propagation from the color layer, see figure 8.3, via the respective bindings, figure 8.5 as target and distractor activation diverge.

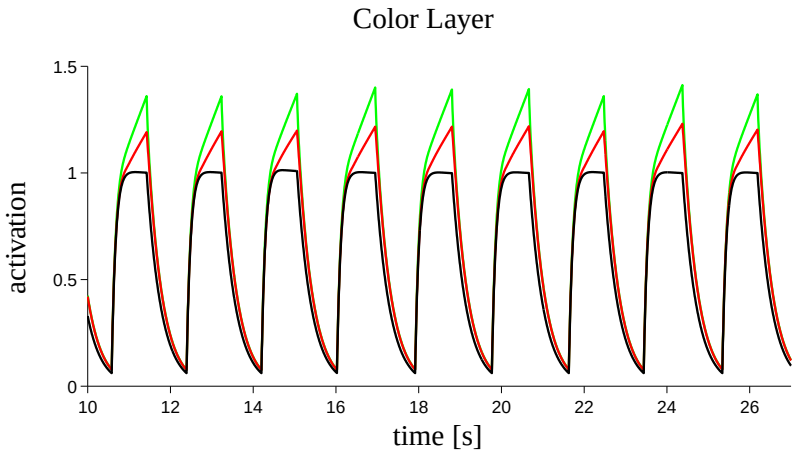


Figure 8.3: Time series of the representation variables $f_{\text{color}}^{\text{green}}$, $f_{\text{color}}^{\text{red}}$ and $f_{\text{color}}^{\text{gray}}$ in the color layer of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. The plot shows the target amplification, equation 8.3, in the faster rise of green.

Feature variables f_i^j are defined by the indices i, denoting the respective feature, e.g. color, see figure 8.3, shape, see figure 8.2 or word shape, see figure 8.4, and j which indicates the concrete instance of the feature, e.g. green. The dynamics of a feature variable is determined by several terms that act simultaneously. The first one is a drift with time constant τ_f towards the external input F_i^j of the feature instance, whereas τ_f takes on two different values ρ_f if the feature variable is lower than the input and rises by an active drive, or δ_f if the input variable is lower and the feature variable passively decays. F_i^j is defined by the presence of the respective feature in the display configuration. If the particular feature instance defines the respective object to be target or distractor in the current task, an additional input, excitatory or inhibitory, is applied to the feature variable. In case of feature presence, F_i^j is set to maximum external input strength $\hat{F}$ plus the current value of the variable accounting for the reception of input by only a subset of neurons in one assembly, similar to the argumentation in section 3.2.4 introducing the residual activities to the ISAM. The residual overshoot of the input decays to the maximum input in the same way feature activation would. In case of feature absence, the input is set to the activation baseline value of $\underline{F}$ which is not necessarily zero.

$$F_i^j = \begin{cases} \hat{F} + f_i^j & \text{at display onset, if instance } j \text{ of feature } i \text{ is present} \\ \delta_f(\hat{F} - F_i^j) & \text{during stimulus perception, as long as } F_i^j - \hat{F} \\ \underline{F} & \text{at display offset} \end{cases} \tag{8.1}$$

Both target selection mechanisms, target amplification and distractor inhibition simply add to the corresponding feature input F_i^j to the overall input $\mathbf{F}_i^j$ in the dynamics of the feature variables. Target amplification A is linearly increasing until a response is given and set to zero afterwards. Distractor inhibition I is said to persist for some time, as it has to be retrenched after a response was given. Therefore, inhibition I increases linearly with slope ι during perception and fades linearly after the decision was made.

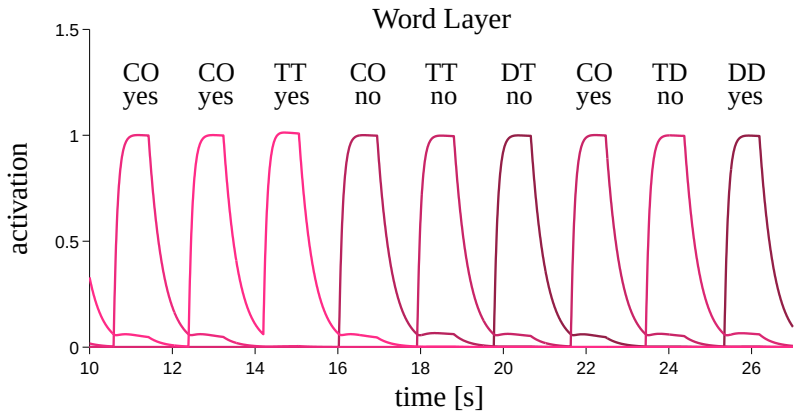


Figure 8.4: Time series of the representation variables f^j_{word} in the word layer of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. Besides a very monotonic behavior, the plot reveals one of the major problems in the design of the word picture comparison paradigm by pointing to the fact that we only allow for a repetition of the comparison word in TT trials that require a confirmative response, which may cause systematic reaction time effects as well.

$$\mathbf{F}_i^j = \begin{cases} F_i^j + A & \text{if } \{i,j\} \text{ defines the target} \\ F_i^j + \mathrm{I} & \text{if } \{i,j\} \text{ defines the distractor} \\ F_i^j & \text{otherwise} \end{cases} \tag{8.2}$$

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \alpha \quad \text{during stimulus presentation} \tag{8.3}$$
$$A = 0 \quad \text{no stimulus present}$$

$$\frac{\mathrm{dI}}{\mathrm{d}t} = \begin{cases} \iota & \text{during external input} \\ -\iota & \text{after the offset of input until } \mathrm{I} = 0 \end{cases} \tag{8.4}$$

The second term governing the dynamics of features is losing of feature specificity in the absence of input defined by a broadening of activation with time constant β, within one feature towards the feature mean $\langle i \rangle f_i^j$, without lowering the total activation of the respective feature layer. Thirdly, feature activation is passed via existing bindings to the other feature instances belonging to the same object. If e.g. the feature tuple {color,green}{shape,ball}{location,bottom} defining a green ball shown at the bottom of the visual scene is held by the binding variable $b_{\{\text{color,green}\}\{\text{shape,ball}\}\{\text{location,bottom}\}}$. Its value defines the amount of activation interchange between the variables $f^{\text{green}}_{\text{color}}$, $f^{\text{ball}}_{\text{shape}}$ and $f^{\text{bottom}}_{\text{location}}$ such that they approach their mean. There exists only one feature variable for green. Therefore multiple green objects experience a natural connection, as they share this variable. The last term that drives feature variables is the back projection of memorized episodes into the feature layer. Weighted by the matching value r_k of the actual percept and the kth last memorized episode and the strength e_k of the respective memory trace the value of the feature variable at the respective response moment $e_k^{f_i^j}$ is fed back to the variable.

$$
\begin{aligned}
\frac{\mathrm{d}f_i^j}{\mathrm{d}t} &= \tau_f(\mathbf{F}_i^j - f_i^j) \\
&\quad +\beta\left(\langle i\rangle f_i^j - f_i^j\right) \\
&\quad +\sum_{b_{\{i_k,j_k\}}:\exists k: i_k=i, j_k=j} b_{\{i_k,j_k\}}(\langle \exists k : i_k = l, j_k = m, \{i_k, j_k\} \neq \{i, j\}\rangle f_l^m) \\
&\quad +\sum_k r_k e_k \left(e_k^{f_i^j} - f_i^j\right) \\
\text{where} \quad \tau_f &= \begin{cases} \rho_f & \text{if } F_i^j > f_i^j \\ \delta_f & \text{if } F_i^j < f_i^j \end{cases}
\end{aligned} \tag{8.5}
$$

8.3.2 Feature Binding Mechanism

The bindings themselves are dynamic variables, as they code the strength of feature interconnection within one object, see figure 8.5. But as the underlying structure, see (Schrobsdorff et al., 2007a), is a flexible but resource constrained layer, there is only a limited number of such binding variables where each can code for any feature combination. If an object appears in stimulus space, it is checked whether the object is already represented, which would correspond to an immediate recognition of the object as the same that has just been represented. If the object is not yet represented by a binding, the weakest binding variable that is not subject to actual input is overwritten, deleting the respective object from working memory. If an object is shown, the respective binding variable is driven with time constant ρ_b towards a maximum interconnection strength $\hat{b}$. If the percept of an object is gone, the respective binding variable passively decays with time constant δ_b to zero.

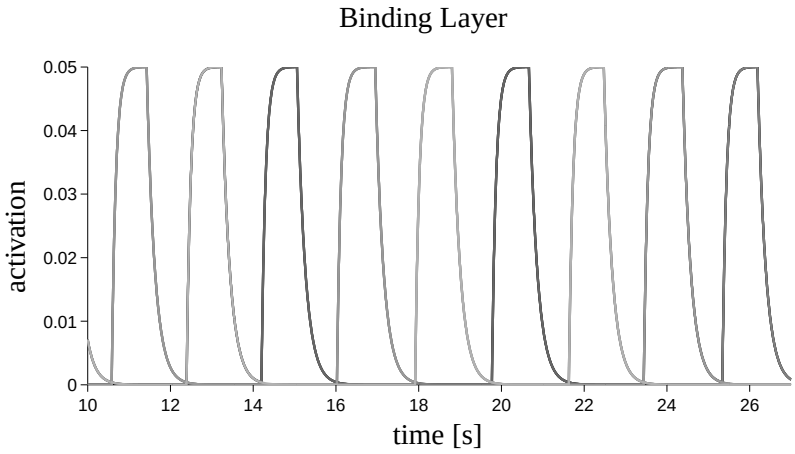


Figure 8.5: Time series of binding strengths $b_{\{i_k,j_k\}_k}$ of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. The plot shows a very stereotypical behavior for all bindings as the only interaction in the binding layer is the taking over of a binding slot if the resources are exhausted and a new object is perceived. Always the weakest binding is replaced by a new one, therefore the effect by the taking over is minimal.

$$\frac{\mathrm{d}b_{\{i_k,j_k\}_k}}{\mathrm{d}t} = \begin{cases} \rho_b(\hat{b} - b_{\{i_k,j_k\}_k}) & \text{if an object with the respective feature is perceived} \\ -\delta_b b_{\{i_k,j_k\}_k} & \text{if the percept is switched off} \end{cases} \tag{8.6}$$

$$b_{\{i_k,j_k\}_k} = 0 \quad \text{if the binding slot is overwritten}$$

i.e. object $\{i_k, j_k\}_k$ was not shown and held by the weakest binding when a new display was uncovered containing a non-bound object $\{i_l, j_l\}_l$.

8.3.3 Semantic Variables

The role of the variables in the semantic layer, see figure 8.6, is assigned by the central executive, depending on task demands. Therefore, a fixed description of the dynamics of semantic variables is not possible. After a hypothetical training phase that introduces a new task, we assume that the central executive has elaborated a reasonable gating function $S(f)$ of feature activations to the semantic layer. This function defines the fixpoint the semantic activation is approaching to with time constant ρ_s or δ_s for an actively driven rise or a passive decay, respectively. Again the variables are subject to retrieval of former episodes analogous to feature variables.

$$\frac{\mathrm{d}s^j}{\mathrm{d}t} = \sigma_{f\to s}\tau_s(S^j(f) - s^j) + \sum_k r_k e_k \left(e_k^{s^j} - s^j\right) \tag{8.7}$$

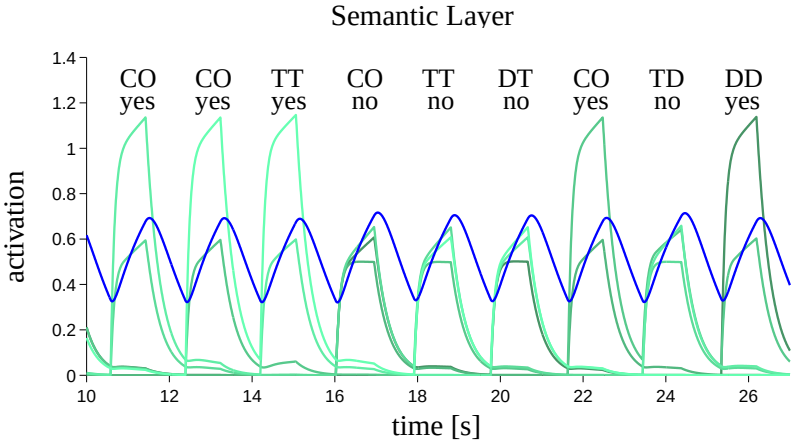


Figure 8.6: Time series of the representation variables s^j in certain shades of green and the threshold s^θ in blue in the semantic layer of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. During yes-trials, when target and comparison word match, the input from chape layer and word layer converge onto the same semantic concept, causing a much higher activation in those cases. The threshold now detects the presence of relatively high activations, as there is always superthreshold activation. In cases of no match, the threshold will surpass all activations which is then seen as evidence for a no response.

$$\text{where} \quad \tau_s = \begin{cases} \rho_s & \text{if } S^j > s^j \\ \delta_s & \text{if } S^j < s^j \end{cases}$$

8.3.4 Short-Term Modulation of Connectivity

The General Model is able to direct information flow such that it can decide whether a response should be computed from scratch by considering the perceptual input or by retrieving information from episodic memory. Therefore, synaptic connections between the layers can be blocked or facilitated, depending on the old-new signal o_k generated by comparing the kth last episode to the actual percept. A blocking variable σ_{block} adapts with time constant τ_{block} to o_k which is either set to 1, $\frac{1}{2}$ or 0 depending on whether the signal is old, unclassified or new, respectively. The synaptic strength is scaled according to σ_{block} between a minimum synaptic strength $\underline{\sigma}_{f\to s}$ and an open channel of $\sigma_{f\to s} = 1$.

$$\sigma_{f\to s} = (1-\underline{\sigma}_{f\to s}) + \underline{\sigma}_{f\to s}\sigma_{\text{block}} \tag{8.8}$$

$$\text{with } \frac{d\sigma_{\text{block}}}{dt} = \tau_{\text{block}}(o_k - \sigma_{\text{block}}) \tag{8.9}$$

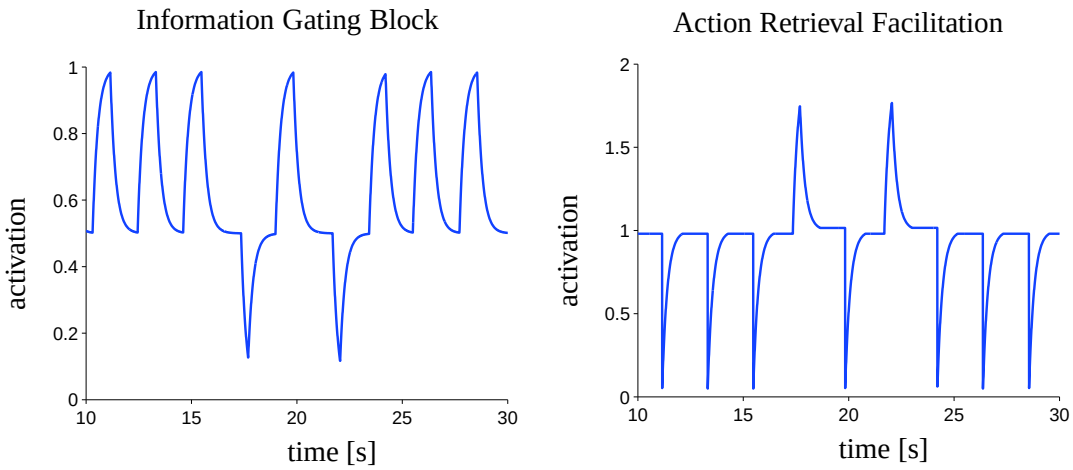


Figure 8.7: Connectivity modulation according to episodic retrieval and temporal discrimination theory. Whenever a perceptual match is high, the direct computation, i.e. the information transmission between feature layers and semantic layer and also between semantic layer and action layer, is closed. If the two percepts are dissimilar, the transmission is enhanced as visible on the left plot. The right plot shows the retrieval strength for the action layer. Whenever a match is detected, the response can be retrieved from memory, retrieval is enhanced, if no match is present, the retrieval is useless, and therefore prevented, see section 8.3.8.

8.3.5 The Adaptive Threshold in the Semantic Layer

As a decision mechanism for comparison tasks, the semantic layer possesses an adaptive threshold s^θ which is computed by a weighted average of activation in the semantic layer, see figure 8.6. The weighting of the average is such that it adapts between the highest two semantic activations.

Therefore the baseline activation $\underline{F}$ that is directly passed from feature layers to the semantic layer has to enter the averaging process.

$$\frac{1}{\tau_{s^\theta}}\frac{\mathrm{d}s^\theta}{\mathrm{d}t} = \nu_{s^\theta}\sum_j \left(s^j - \underline{F}\right) + \underline{F} - s^\theta \tag{8.10}$$

8.3.6 Action Variables

The action layer behaves similarly to the semantic layer, see figure 8.8. Action activation variables are driven towards external input $A(s,f)$ that is computed from semantic and feature representations according to the task, i.e. given by a mapping function incorporated in the central executive. Depending on whether the adaptation is an actively driven rise or a passive decay, two time constants ρ_a, δ_a apply, respectively. An aspect that is easily overseen is the possibility of a person not to respond. For example in cases where no target object is shown. Therefore, action zero a^0 has the special role of representing no reaction. $A(s,f,\sigma_{f,s\to a})^j$ is designed such that whenever there is no target stimulus shown, e.g. between two trials, $A^0(s,f,\sigma_{f,s\to a}) = 1$. In case of stimuli triggering a response $A^0(s,f,\sigma_{f,s\to a}) = 0$.

$$\begin{aligned} \frac{\mathrm{d}a^j}{\mathrm{d}t} &= \tau_a(A^j(s,f,\sigma_{f,s\to a}) - a^j) + r_a\sum_k r_k e_k\left(e_k^{a^j} - a^j\right) \\ \text{where} \quad \tau_a &= \begin{cases} \rho_a & \text{if } A(s,f)^j > a^j \\ \delta_a & \text{if } A(s,f)^j < a^j \end{cases} \end{aligned} \tag{8.11}$$

The relative retrieval of action representations r_a is modulated contrary to the synaptic transmission to the action layer $\sigma_{f,s\to a}$ reflecting the facilitation of action retrieval by an old-new signal indicating an old episode which can be answered by retrieving a former response. Also, the modulation of information flow can decrease the retrieval of a response if a new episode is classified.

$$\sigma_{f,s\to a} = (1 - \underline{\sigma}_{f,s\to a}) + \underline{\sigma}_{f,s\to a}\sigma_{\text{block}} \tag{8.12}$$

$$r_a = \left(1 + \max\left(\underline{\sigma}_{f,s\to a}, \underline{\sigma}_{f\to s}\right)\right) - 2\max\left(\underline{\sigma}_{f,s\to a}, \underline{\sigma}_{f\to s}\right)\sigma_{\text{block}} \tag{8.13}$$

In order to model the decision making process in the action layer we introduce a threshold level analogous to the semantic layer. As input to the action layer ranges from zero to one, we do not have to care about any baseline activation in the present case.

$$\frac{1}{\tau_{a^\theta}}\frac{\mathrm{d}a^\theta}{\mathrm{d}t} = \nu_{a^\theta}\sum_j a^j - a^\theta \tag{8.14}$$

Superthreshold activations $a^j > a^\theta$ define the space of possible actions the system can take. If there is only one action superthreshold, the corresponding action is executed. In case of $a^0 > a^\theta$, the system does not do anything.

8.3.7 Memory Processes

Memory Processes are modeled in a very simplified way. At points in time that mark the closure of an episode, in this case at reaction times, the entire state of the model is written down as one episode. The stored values can be retrieved on the one hand in order to compute similarities between passed episodes and an actual percept, see figure 8.9. On the other hand, this similarity

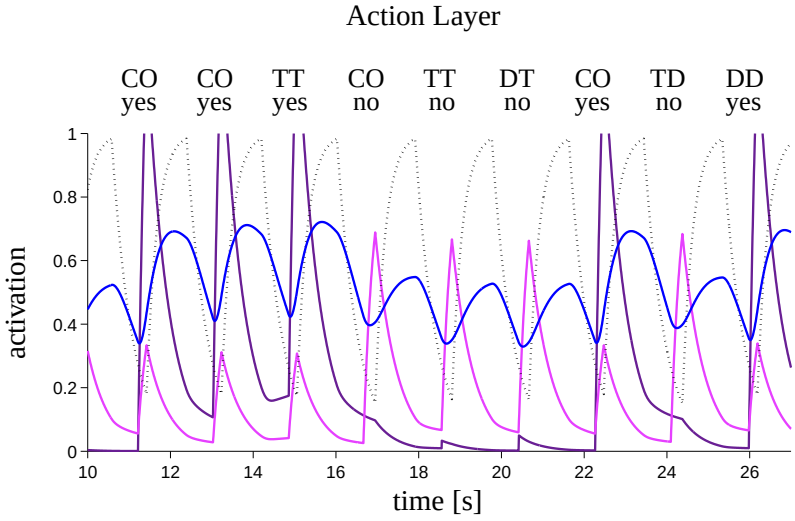


Figure 8.8: Time series of the action variables a^j in shades of purple, the representation of no action a^0 in black dotted and the threshold a^θ in blue in the action layer of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. If the semantic threshold decides for a certain response , its representation receives input. In the absence of stimuli the no-response is activated. The threshold variable adapts in the same way as the semantic threshold. Whenever only one response is above threshold, the corresponding action is executed.

triggers an automatic retrieval of the former episodes. The greater the similarity, the stronger memorized values are piped back into the respective variables. The presence of memorized episodes is set to the initial value $\hat{e}$ when the episode is written down, and then freely decays to zero with time constant δ_e.

$$\begin{aligned} e_k &= \hat{e} && \text{if episode } k \text{ is memorized} \\ \frac{\mathrm{d}e_k}{\mathrm{d}t} &= -\delta_e e_k && \text{otherwise} \end{aligned} \tag{8.15}$$

If a new episode is memorized, the kth last episode naturally becomes the $(k+1)$th last one.

$$\left.\begin{aligned} e^v_{k+1} &= e^v_k \\ e^v_1 &= v \qquad v \in \{f^j_i, b_{\{j_k,i_k\}_k}, s^j, a^j\} \end{aligned}\right\} \text{ when an action is taken} \tag{8.16}$$

Generating a reliable old-new signal from only internal values, i.e. information that is accessible by the system itself, is rather tricky. The intention is to have some value that is higher for a higher degree of similarity between the actual percept and a memorized one. The system is only able to monitor internal variables. Therefore, at first the idea of the actual percept can only be determined by looking at feature variables. As the system is trained for the present task, it has some knowledge about the expected number of objects n in the display. But the actual objects can only be guessed by looking at the n strongest bindings. Besides the summed distances of all feature variables, there is a contribution of the sum of distances between the n largest binding variables and their memorized value if the binding already existed rescaled by the maximum binding strength $\hat{b}$. The

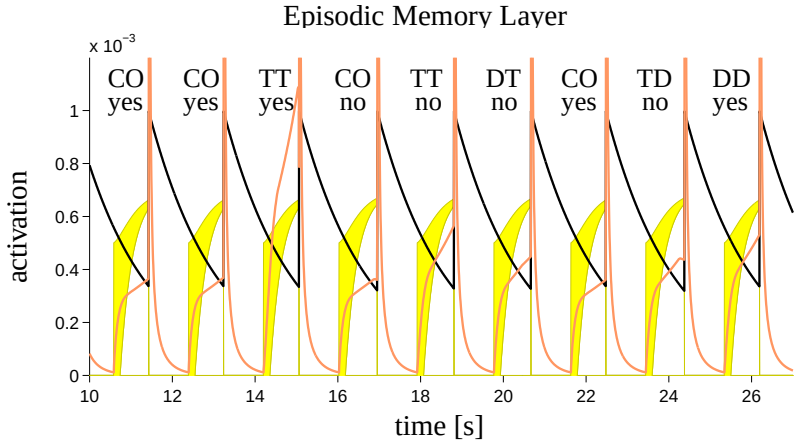


Figure 8.9: Time series of the representation variables in the episodic memory layer of the General Model during accomplishing the word-picture comparison paradigm, see section 8.6. The model is set to episodic retrieval mode, see section 8.4. Additional to the orange similarity signal and the black activation course of memory strength, see section 8.3.7, which are used in the implementation of episodic retrieval, the plot shows the corridor that decides about the classification of a trial as old or new, see section 8.3.8. An old stimulus configuration is correctly only detected in the TT-yes trial, as it repeats the target and the comparison word. In the TT-no trial the comparison word changes between prime and probe, therefore not enough similarity for an old display is detected.

old-new signal computed in this way is biggest for nonmatching percept and memory. Another disadvantage is the bad resolution if percept and memory match to a high degree. Therefore, we take the inverse of this model before applying a normalization by the significance of a percept given by the sum over all actual feature variables, divided by the number of features relevant to the task.

$$r_k = \frac{\sum_{i,j} f_i^j}{\#_f} \left(\sum_{\{i_l,j_l\}_l} \left(\left| f_i^j - e_k^{f_i^j} \right| + \frac{1}{\hat{b}} \left| b_{\{i_l,j_l\}_l} - e_k^{b_{\{i_l,j_l\}_l}} \right| \right) \right)^{-1} \tag{8.17}$$

where $\{i_l, j_l\}_l \in \{\text{subjective percept}\}$

$= \{\text{objects being held by the } n \text{ largest bindings}\}$

with $n = \#\text{objects in one display}$

8.3.8 Connectivity Modulation

Information gating mechanisms are modeled by the dynamic opening or shutdown of synaptic transmissions between the different layers as well as the retrieval channel to the action layer. This modulation is governed by an old-new signal o_k comparing the kth last episode to the actual percept. The comparison process is modeled by locating the kth retrieval signal r_k below, in between or above an uncertainty deviation u, shrinking exponentially with time constant τ_u, from a prototype time course for an intermediate resemblance of displays given by an exponential adaptation from an initial value $\underline{d}$ with time constant τ_d towards a retrieval level $\hat{d}$ dividing old from new displays, see figure 8.9.

$$o_k = \begin{cases} 0 & \text{if } r_k > d+u \\ 1 & \text{if } r_k < d-u \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{8.18}$$

where

$$\left.\begin{array}{l} d = \underline{d} \\ u = \underline{u} \end{array}\right\} \text{ at display onset}$$

$$\left.\begin{array}{l} d = 0 \\ u = 0 \end{array}\right\} \text{ at display offset}$$

$$\left.\begin{array}{l} \frac{\mathrm{d}d}{\mathrm{d}t} = \tau_d \left(\hat{d} - d\right) \\ \frac{\mathrm{d}u}{\mathrm{d}t} = -\tau_u u \end{array}\right\} \text{otherwise} \tag{8.19}$$

8.3.9 Generating Real World Reaction Times

Up to now we presented all dynamic equations governing the behavior of the General Model. For a complete picture of all implementation details, some further points have to be mentioned. The modeling comprises just the internal processes. While deriving biologically realistic reaction times, several additive components have to be considered. Feature input is not set on immediately if a new display is shown, but rather after some recognition latency needed to transfer the information from the screen via the retina and the entire visual pathway to feature layers. Also, there is persisting activation throughout the visual pathway known as the after-image, which drives feature input longer than the stimulus display is shown. Finally, the reaction time is recorded at a button press or similar actions, not at the point in time where one action alternative became selected in the action layer. Therefore some fixed time elapses between the decision in the action layer and the offset of the stimuli. These three time intervals may also be subject to statistical variation and additionally can depend on the subjects preferences or habits. Nevertheless, we take these values as fixed over subjects, stimulus conditions and time. The actual lengths of these time intervals can be taken from the respective research fields.

8.4 Defining Setscrews for the Theories

In order to analyze the consequences of a certain theoretical account, we define single setscrews Ξ that modulate the impact of the assumptions of one theory on the behavior of the General Model. We label the setscrews with abbreviations according to their theoretical scope. A list is given in table 8.1.

Retrieval itself is easiest controlled by adjusting the initial strength of a memory trace as it linearly determines the impact of retrieval. The modulation factor Ξ_{er} just scales the maximum memory strength $\hat{e}$. If Ξ_{er} is zero, no memory is written down, and therefore retrieval has no effect on the system behavior. If $\Xi_{\mathrm{er}} = 1$, memories are stored initially with the maximal strength $\hat{e}$ and retrieval provides the input to the system described in section 8.3.7.

The question whether the entire system state at the last decision is retrieved or only the prime response separates episodic retrieval from response retrieval theory. These two assumptions are contradictory. Therefore the setscrew Ξ_{rr} gradually shuts down the retrieval or activation in layers other than the action layer. If $\Xi_{\mathrm{rr}} = 1$ the entire episode is retrieved, whereas, if $\Xi_{\mathrm{rr}} = 0$, only the action layer receives memory input.

	Model Behavior for $\Xi = 0$	Model Behavior for $\Xi = 1$
Ξ_{er}	no retrieval at all	maximum retrieval
Ξ_{rr}	only retrieval of response	total retrieval
Ξ_{ib}	distractor inhibition	target boost
Ξ_{gt}	no activation interference	forced decay and activation broadening
Ξ_{fsb}	full propagation	retrieval blocks features sematic synapses
Ξ_{sab}	full propagation	retrieval blocks semantic action synapses
Ξ_{td}	classical episodic retrieval	old/new evaluation

Table 8.1: Setscrews controlling the strength of the implementation of a certain theoretical account into the General Model. Their range is continuously between zero and one.

Distractor inhibition theory and the global threshold theory contradict each other by either assuming an inhibition of the distractor or a target boost, respectively. The setscrew Ξ_{ib} modulates input to the feature instance that identifies target and distractor. If $\Xi_{ib} = 0$, only the distractor receives inhibiting input, i.e. $\alpha = 0$. If $\Xi_{ib} = 1$ only the target feature receives excitation, i.e. $\iota = 0$. Ξ_{ib} additionally adjust the baseline activation level from $1/2$ in the distractor inhibition case to 0 with target boost, where no sub-baseline activation is assumed. At this point a major gap in the retrieval accounts becomes visible. They do not make any statements on what the direct computation of a trial may look like. But the general model needs some decision making mechanism. In order to have the least effect of the decision making mechanism on priming effects in the case where we consider retrieval based mechanisms, we chose to have a pure target boost in the feature layers. Forced decay as well as activation broadening as inherent features of the global threshold theory will thus be controlled independently. Ξ_{gt} linearly controls the broadening of activation β and the strength of the forced decay if two concepts are in concurrence of a feature instance.

Both temporal discrimination and episodic retrieval theory postulate a decision of the system on whether the current response should be generated directly from the input, or retrieved from memory. The corresponding modulation in the General Model is done via the setscrew Ξ_{fsb}. If $\Xi_{fsb} = 0$, there is a concurrence between direct computation and retrieval in the system. If $\Xi_{fsb} = 1$, the strength of retrieval, i.e. the similarity signal, triggers a shutdown of the synapses between features and semantic layer, modeling a decision of the system to only retrieve the response and drop the direct determination of the right answer.

In an excursion by Tipper and Cranston (1985) into the episodic retrieval domain, he argued in favor of a blocking of the information flow in the episodic retrieval context right before the action selection state. This manifests in the General Model as a blocking similar to Ξ_{fsb} described in the last paragraph, but the shutdown rather happens between semantic and action layer. The corresponding setscrew is Ξ_{sab}.

A final setscrew is given by Ξ_{td} which controls the evaluation of a stimulus being old or new before retrieval is initiated. In the case $\Xi_{td} = 0$, the similarity signal determines the retrieval strength from the beginning of a trial, whereas if $\Xi_{td} = 1$ there is no retrieval unless the similarity signal surmounts the uncertainty corridor around the prototype similarity signal, as explained in section 8.3.7.

Table 8.2 summarizes the values of the setscrews if the impact of a single theoretical account is to be evaluated. Note that some mechanisms are inherent to the General Model such as activation

	Ξ_{er}	Ξ_{rr}	Ξ_{ib}	Ξ_{gt}	Ξ_{fsb}	Ξ_{sab}	Ξ_{td}
Distractor Inhibition	0	0	0	0	0	0	0
Global Threshold	0	0	1	1	0	0	0
Episodic Retrieval	1	1	1	0	0	0	0
Response Retrieval	1	0	1	0	0	0	0
Temporal Discrimination	1	1	1	0	1	1	1

Table 8.2: Settings of the setscrews in order to produce predictions by just one theory.

propagation via the feature bindings. Therefore the pure settings do not give a minimal computational model of the respective theory, but we rather keep the mechanisms none of the accounts specifies constant in all cases.

8.5 Voicekey Paradigm

The current and the following section will show four examples of the General Model during two different paradigms and while emulating different theory variants. The aim is to show the capabilities of the model concerning different paradigms and the comparison of theories.

The first simulation shows the General Model in a voicekey paradigm, see section 2.2. all relevant variables are plotted over nine trials including all five conditions in figure 8.10. The setscrews are tuned to episodic retrieval, i.e. we have no activation interferences in the feature layers, but perceptual input, a boost of the target color green, the activation exchange via the bindings and activation retrieval from memory.

The presentation of a red and a green pictogram drives the two color and the two shape representations in the respective layers. The central executive delivers additional input to green which augments the activity of the target object's shape via the bindings. The semantic representations are given by a one-to-one mapping of shapes, i.e. $S(f) = \mathbb{1}$, thus the activation is more or less identical in both layers. The plot of the episodic memory layer shows the memory strength in black which decays with time from a fixed value at memory initialization which takes place at the point a response is given. In orange, the plot shows the similarity signal which linearly modulates the retrieval of a former trial. The signal is highest for the TT trial, intermediate for DT, TD and DD in ascending order. In the action layer, the trace for the no-action, see section 8.3.6, is dotted. The selection of the target in the semantic layer again is directly passed to the action layer. Every semantic representation has its corresponding action, i.e. to name it.

Already visible is a prediction of the configuration for TD trials. In this case the intermediate similarity drives the retrieval of the high former target activation which is now applied to the distractor. The threshold needs longer to surpass the distractor, therefore a strong slow-down is observable, see also table A.21.

The present simulation was run with the following values of the relevant parameters: $\Xi_{er} = 1$, $\Xi_{rr} = 1$, $\Xi_{ib} = 1$, $\Xi_{gt} = 0$, $\Xi_{fsb} = 0$, $\Xi_{sab} = 0$, $\Xi_{td} = 0$, $\alpha = 0.0005$, $\underline{F} = 1$, $t_{recognition} = 50$, $t_{afterimage} = 30$, $t_{motor} = 80$, $\rho_f = 0.01$, $\delta_f = 0.003$, $\hat{b} = 0.05$, $\#_b = 7$, $\rho_b = 0.008$, $\delta_b = 0.005$, $\tau_{s^\theta} = 0.002$, $\nu_{s^\theta} = 0.51$, $\rho_a = 0.004$, $\delta_a = 0.002$, $\tau_{a^\theta} = 0.002$, $\nu_{a^\theta} = 0.5$, $\hat{e} = 0.002$, $\delta_e = 0.003$.

Negative priming in DT trials and positive priming in TT trials are with 21 ms and 56 ms at rather realistic regimes, but the deceleration of -157 ms in TD trials points to the fact that

the mechanisms of negative priming are not well reproduced with the current parameter settings. Detailed results are given in table A.21 in appendix A.6.

The second simulation, figure 8.11 is run with a setting to temporal discrimination, i.e. $\Xi_{\text{fsb}} = 1$, $\Xi_{\text{sab}} = 1$, $\Xi_{\text{td}} = 1$. This means that the retrieval of activation is only triggered, if the episode is categorized as old. This is only the case for TT trials, as in all other cases the relevant object changes.

Reaction times are about 200 ms slower than in the setting above, which is mainly due to the missing standard retrieval in every trial which is just modulated in episodic retrieval interpretation of negative priming. Temporal discrimination assumes retrieval only to happen if an episode is classified as old, which is correctly only done in TT trials in the simulation shown in figure 8.11. Also priming effects are not well matched: every condition is decelerated, TT trials are slowest, followed by DD, TD and finally DT trials. The simulation results are summarized in table A.22.

8.6 Word Picture Comparison Task

After the sketch of a comparison of episodic retrieval and temporal discrimination theory we will now show how a relatively small change in the form of a weak modulation of information transfer will enhance priming. We expose the General Model to a word-picture comparison task, as it was introduced in section 7.1 despite the long distance between objects and comparison word, such that a parallel evaluation of shapes and words is possible. In this paradigm both shape and word layers project into the layer of semantic representations. Whenever both converge, much stronger input is delivered to the corresponding semantic concept and a yes response is triggered.

We again start off with a straight episodic retrieval setting. The present simulation was run with the following values of the relevant parameters: $\Xi_{\text{er}} = 1$, $\Xi_{\text{rr}} = 1$, $\Xi_{\text{ib}} = 1$, $\Xi_{\text{gt}} = 0$, $\Xi_{\text{fsb}} = 0$, $\Xi_{\text{sab}} = 0$, $\Xi_{\text{td}} = 0$, $\alpha = 0.001$, $\underline{F} = 1$, $t_{\text{recognition}} = 50$, $t_{\text{afterimage}} = 30$, $t_{\text{motor}} = 80$, $\rho_f = 0.01$, $\delta_f = 0.003$, $\hat{b} = 0.05$, $\#_b = 7$, $\rho_b = 0.01$, $\delta_b = 0.005$, $\tau_{s^\theta} = 0.0015$, $\nu_{s^\theta} = 0.5$, $\rho_a = 0.004$, $\delta_a = 0.002$, $\tau_{a^\theta} = 0.002$, $\nu_{a^\theta} = 0.5$, $\hat{e} = 0.002$, $\delta_e = 0.003$.

Results are given in table A.23 in appendix A.6. No strong artifacts are visible, but the desired facilitation in TT trials does not exist.

Parameter changes in the second simulation introducing a slight modulation of information transmission, shown in figure 8.13, are given by: $\Xi_{\text{fsb}} = 0.3$, $\Xi_{\text{sab}} = 0.3$ $\tau_{\text{block}} = 0.5$. Whenever the model encounters a similarity signal strong enough to classify the percept as old, bottom up information flow is hampered by a relative shutdown of the connection between feature layers and semantic layer and in parallel the retrieval of the former response is facilitated. On the contrary a classification as new eases bottom up information flow and hampers retrieval.

Reaction time results are given in table A.24 in appendix A.6. Even if the overall reaction time is rather fast, the pattern of priming effects shows realistic values, as compared to the vanishing facilitation in TT trials the simulation with pure episodic retrieval linearly based on the similarity signal shows.

8.7 Discussion

The simulated reaction times in the tables in section A.6 show that the behavior of the General Model is far from being robust against even small parameter changes. This may be a hint that the complexity level is chosen about right in order to account for the multitude of different findings in connection with negative priming. But we have to face the question whether the model is built such that it can basically fit any data pattern with just the right parameter settings. Due to the

high dimensionality of the parameter space and the sensitivity of the General Model, this question can not be answered conclusively by the means of parameter scanning techniques. In fact, an important next step for the General Model is the parameter reduction by determining as many values as possible by comparisons with trusted experimental results, e.g. for the availability of after-images, decay times of feature bindings, etc. The detailedness of the General Model is also easily capable of showing partial reaction times. Therefore a good way to limit the range of the parameter space would be to have a series of time-marker experiments specially designed to reveal processing stages that are also measurable in the General Model. Up to now the General Model can only be a basis for discussion on the concrete nature of negative priming theories and paradigms.

In order to really decide which of the theories explains what part of the negative priming effect, a thorough dialogue with the psychologists who invented the theories is necessary. As the discussion can best be triggered by a prominent introduction of the General Model to the priming community we are in a vicious cycle. No acknowledged simulation results are possible without an exchange between theoreticians and modelers, but also no discussion in the community is possible without recognized results. The first attempt to break the cycle by advertising the General Model at several conferences did not succeed yet. The second attempt, trying to reproduce a large portion of empirical data, is still work in progress.

8.8 Summary

We presented the proof of concept for our neurophysiological model of perception based action selection. Based on the cognitive demands of a negative priming paradigm we reviewed experimental findings as well as theoretical concepts that characterize the mechanisms suspected to contribute to trial processing. We then implemented several model layers for the different stages, each devoted to a specialized purpose and with certain characteristics. But all layers are working with the same realistic rate dynamics we introduced with the implementation of the ISAM.

The result is a comprehensive model able to recognize perceptual objects by feature decomposition and a binding mechanism that keeps track of the object entities. These objects are translated into a semantic representation where the attentional mechanism selects the most important item and propagates the information to an action selection layer. This chooses the appropriate responses and triggers its execution. In parallel a memory component observes the repetition of perceptual stimuli and triggers the retrieval of previously encountered stimuli together with the results of trial processing in order to facilitate responding.

The General Model gives a unified framework to quantify each of the theories for negative priming. The identification of setscrews for the different accounts makes it convenient to compare the different predictions in a certain setting. But the application to clarify the explanations of negative priming is still confronted with several hurdles to take. The high dimensionality of parameter space makes it impossible to exclude possibilities of behavior. And the implementation of the different theories still should undergo a debate with the priming community in order to best match the concepts of the theories, which are a matter of common sense.

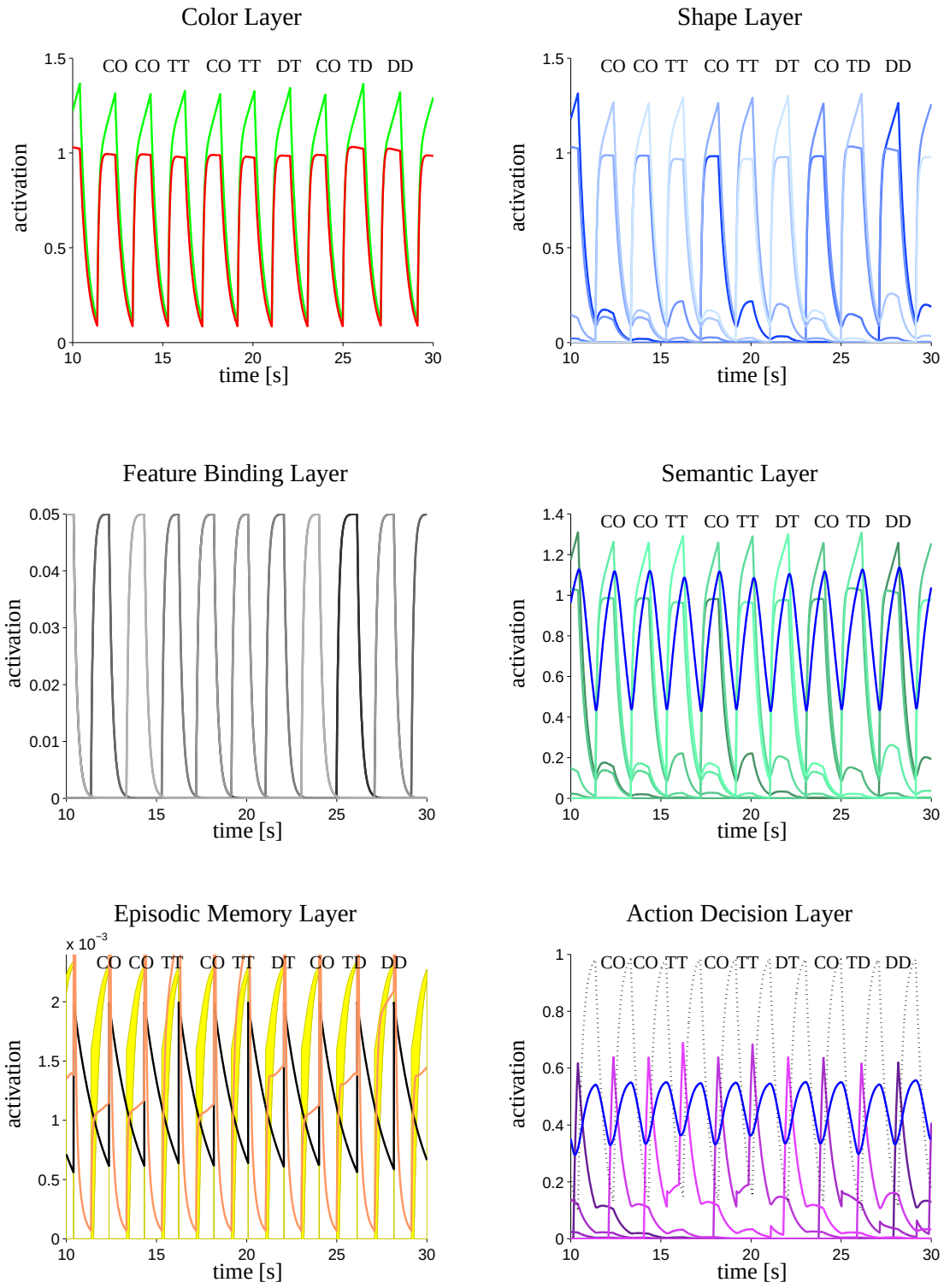


Figure 8.10: Activation traces in the different layers of the General Model in the voicekey paradigm described in section 2.2. The model is in classical episodic retrieval mode, i.e. $\Xi_{er} = 1$, $\Xi_{rr} = 1$, $\Xi_{ib} = 1$, $\Xi_{gt} = 0$, $\Xi_{fsb} = 0$, $\Xi_{sab} = 0$ and $\Xi_{td} = 0$. That means we have a target boost and retrieval of the entire episode, but no forced decay of activation, no old/new classification. Retrieval is visible by the relatively small rise of formerly active variables. Resulting reaction times are summarized in table A.21.

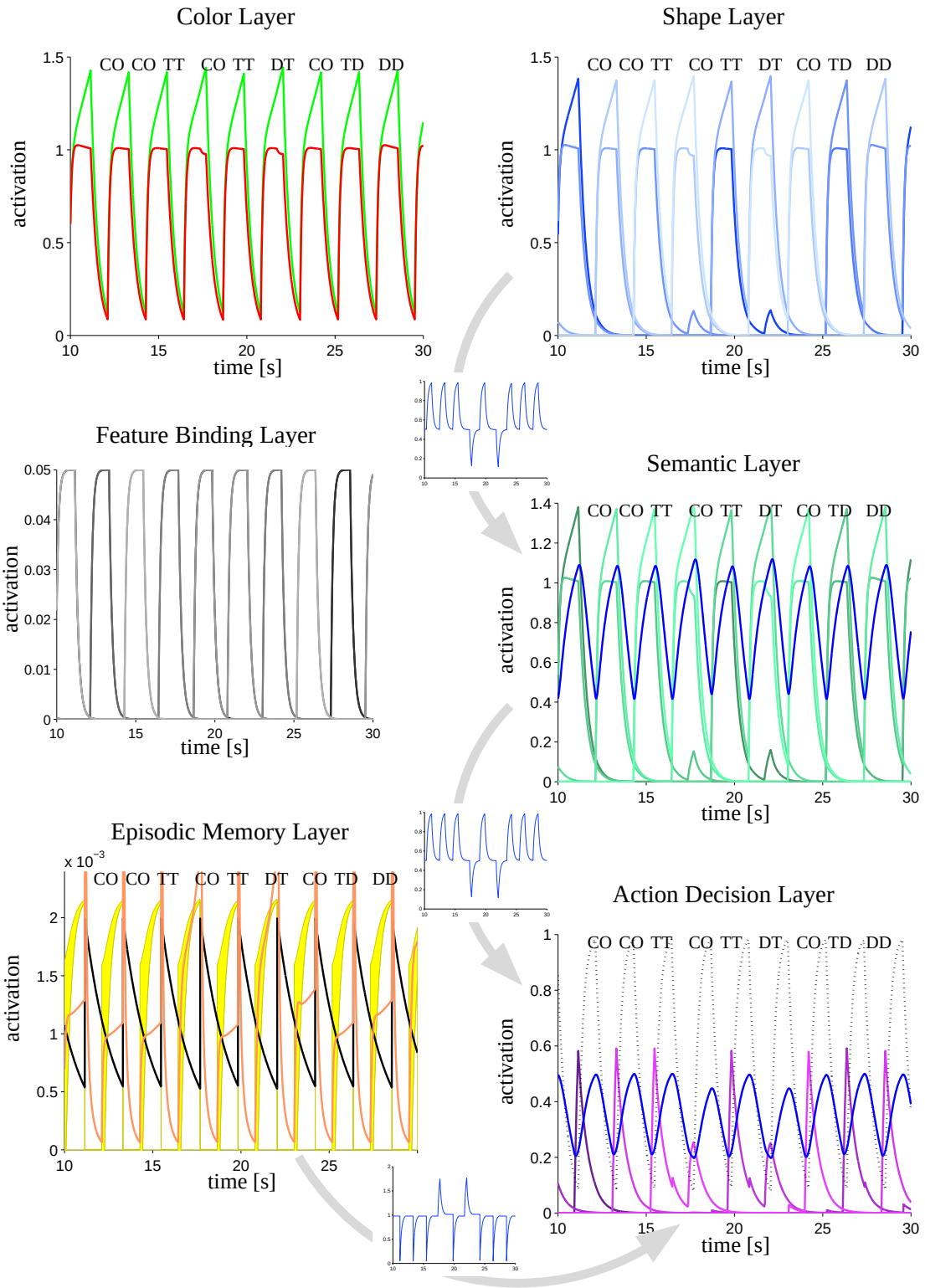


Figure 8.11: Activation traces in the different layers of the General Model in the voicekey paradigm described in section 2.2. The model is in temporal discrimination mode, i.e. $\Xi_{er} = 1$, $\Xi_{rr} = 1$, $\Xi_{ib} = 1$, $\Xi_{gt} = 0$, $\Xi_{fsb} = 1$, $\Xi_{sab} = 1$ and $\Xi_{td} = 1$. That means we have a target boost and retrieval of the entire episode, but only if an episode is classified as old, what happens if the retrieval variable (orange) leaves the uncertainty region (yellow). Resulting reaction times are summarized in table A.22.

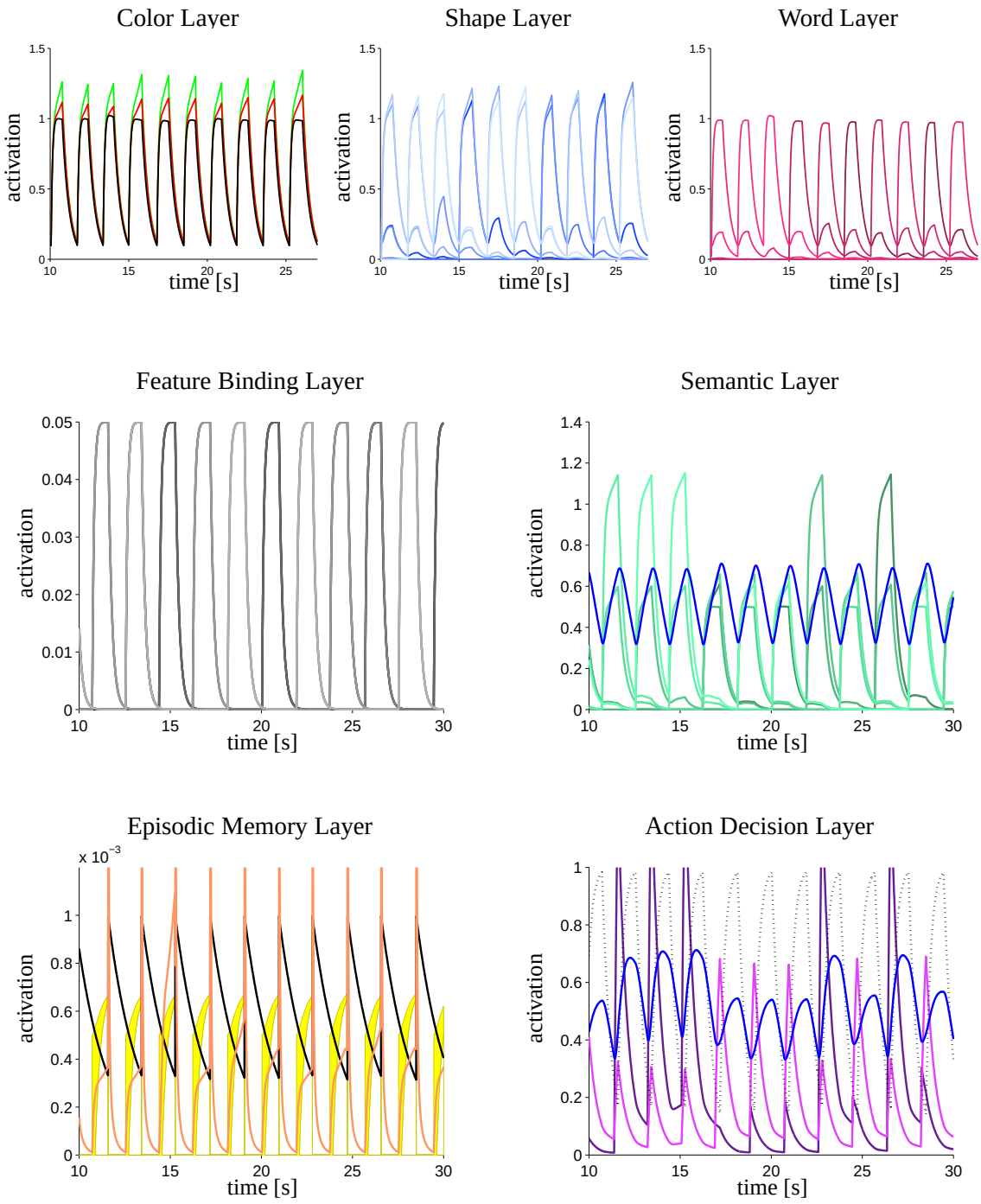


Figure 8.12: Activation traces in the different layers of the General Model in the voicekey paradigm described in section 2.2. The model is in classical episodic retrieval mode, i.e. $\Xi_{er} = 1$, $\Xi_{rr} = 1$, $\Xi_{ib} = 1$, $\Xi_{gt} = 0$, $\Xi_{fsb} = 0$, $\Xi_{sab} = 0$ and $\Xi_{td} = 0$. That means we have a target boost and retrieval of the entire episode, but no forced decay of activation, no old/new classification. Retrieval is visible by the relatively small rise of formerly active variables. Resulting reaction times are summarized in table A.23.

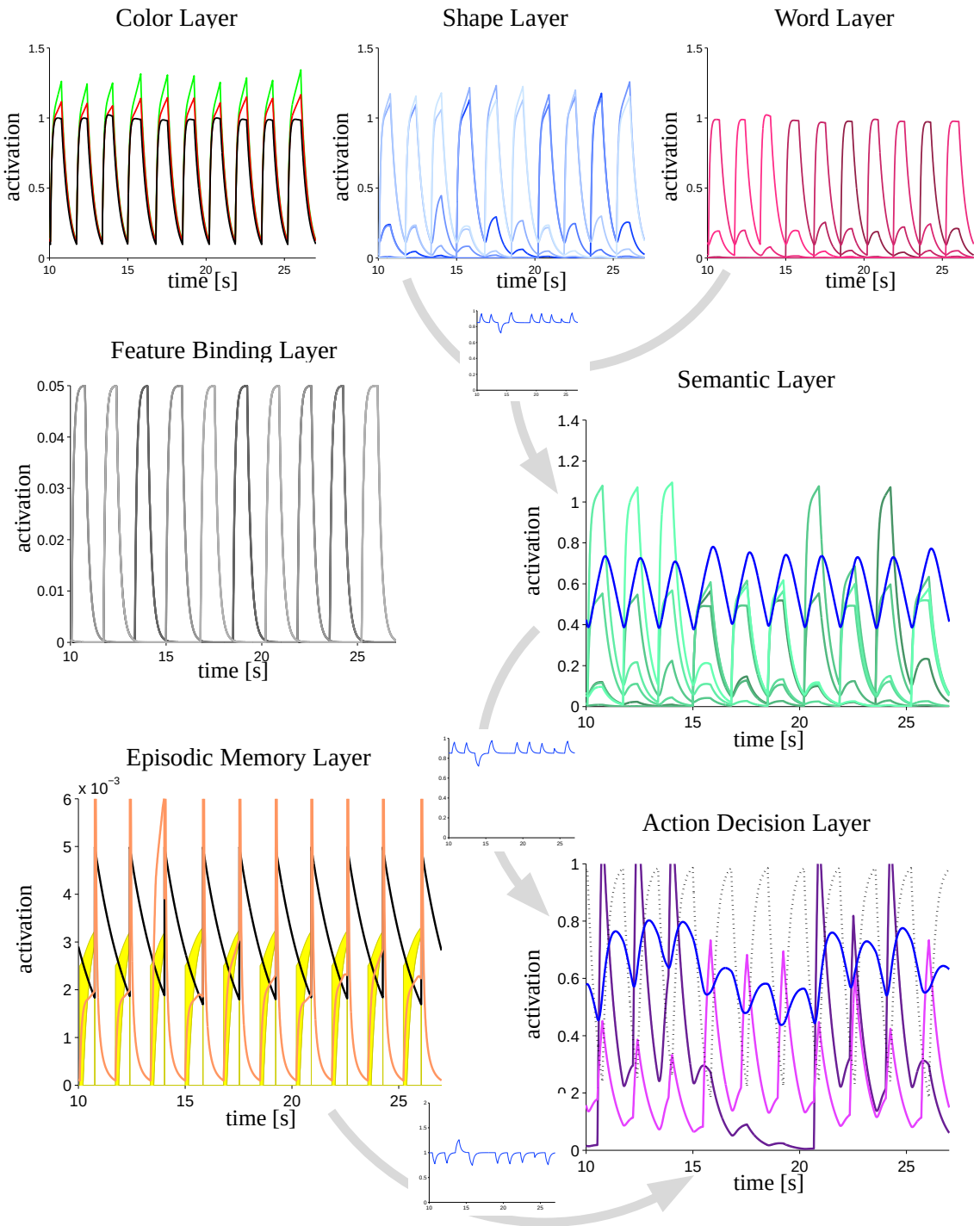


Figure 8.13: Activation traces in the different layers of the General Model in the voicekey paradigm described in section 2.2. The model is in episodic retrieval mode additionally applying information gating via a modulations of information flow according to similarity, i.e. $\Xi_{er} = 1$, $\Xi_{rr} = 1$, $\Xi_{ib} = 1$, $\Xi_{gt} = 0$, $\Xi_{fsb} = 1$, $\Xi_{sab} = 1$ and $\Xi_{td} = 1$. That means we have a target boost and retrieval of the entire episode, but only if an episode is classified as old, what happens if the retrieval variable (orange) leaves the uncertainty region (yellow). Resulting reaction times are summarized in table A.24.

9 Conclusion

The present thesis described how we applied different techniques to reveal the temporal structure of the negative priming effect, i.e. the slow down of reactions to previously ignored stimuli. We have approached negative priming by modeling, data analysis and by the design of novel experimental paradigms, each of which led to specific results. We implemented two computational models describing the origins of negative priming. Detailed insight into the temporal and spatial characteristics of priming was gained by EEG recordings. Finally we described two attempts how the temporal structure of negative priming can be assessed by introducing paradigms with subtasks for which we record reaction times separately. We will conclude on each of the three points in the following sections 9.1, 9.2 and 9.3, respectively. After a review of the results obtained from each of the approaches to negative priming, section 9.4 is dedicated to collect all results in order to get insight in the time course of negative priming. Section 9.5 will give some final remarks on our project and future directions.

9.1 Computational Modeling in Psychology

For the prediction of behavior, psychological theories are a good tool. But they neither provide a clear link to physiology, nor make precise predictions. It is rather the current interpretation of the models characteristics that gives rise to educated guesses on how the mechanisms postulated by the particular theory interact and what consequences this has. If a negative priming experiment seems to provide negative evidence for a certain theory, most likely this evidence would be less conclusive if the involved theories were considered from a different perspective. There are only a few agreements on behavioral effects that can be described exclusively by one theory.

With a background in the exact sciences we found it natural to concretize the theoretical assumptions by implementing computational models. The first model follows the global threshold theory, see section 2.4.7, as closely as possible. The resulting simulation easily produced realistic priming effects. The power of the implementation is shown in section 6.2 where we derived hypotheses by simulating the model for a newly designed paradigm. The straightforward adaptation of the implementation to the new paradigm showed the inappropriateness of the assumptions underlying global threshold theory, i.e. the forced decay of residual activity in a clash with external input. This assumption is responsible for the localization of negative priming in an early stage of a trial, as all residual activation will have faded later on.

Apparently the level of complexity of the ISAM was set too low to correctly describe the manifold aspects of negative priming. Instead of implementing each theory separately, we designed a comprehensive framework of action selection mechanisms based on perception. In this way a sufficient level of detail was achieved in order to incorporate several theories of negative priming and to compare them quantitatively. Due to the complexity of the model implied by the diverse theories and the unclear data patterns, we focused on a proof of concept of the General Model. It is clearly of importance to discuss the implications of the model with the founders of the respective theories, a process which was very beneficial during the implementation of the ISAM.

To get some ordering into the different experimental paradigms with their partly very different result patterns, it would be helpful to implement several classical experiments in the framework

of the General Model. Not only would the standardization of the formulation of the paradigms bring up the crucial differences, but also the behavior of the model itself would reveal the different behavioral consequences the paradigms provoke. We should therefore obtain reasons for the diversity of findings on negative priming.

9.2 EEG Correlates

We put a lot of hope into the analysis of EEG data for the specification of the timecourse of negative priming. EEG experiments are difficult because of technical problems, the high variability in the data, or the available commonly acknowledged use of analysis techniques that assume an underlying process which is not subject to temporal variations which is already disproved by behavioral data. Therefore, we did not only advance our skills in recording EEG data with appropriate paradigms and sharpen our view to recognize possible correlates in resounding grand average plots, but we also challenged the established methods and developed advanced analysis techniques.

On the level of correlates our experiments reveal the similar processing of trials with repeated objects in an early, perceptual phase. Later in the trial, the EEG shows a discrimination of both conditions in opposite directions from the control condition, indicating low cognitive control during target selection if the target is repeated, and high control if the target has been ignored previously. These correlates replicate the results from another negative priming study that was based on a flanker task.

For comparability reasons we did not consequently use our own advanced averaging technique, but stayed with the traditional pointwise average when reporting our results. The advanced averaging technique is up to now only promoted separately.

9.3 Behavioral Paradigms Beyond Response Latencies

In order to gain insight into the temporal localization of the mechanisms causing negative priming, we developed two paradigms that single out a certain phase of trial processing while recording reaction times separately for each of the parts. The first paradigm, described in chapter 6 considers the perceptual phase individually by leaving open the question which of the two colors indicates the target in the current trial, until the recognition of both objects has been acknowledged by the participant. Clearly the change of the paradigm compared to its origin, the voicekey paradigm introduced in section 2.2, is quite drastic. We had to deal with interfering effects from the difference between trials, requiring attention to the same color as in the prime trial and trials where the other color coded for the target.

The situation is similar in the paradigm considering response generation separately, as described in chapter 7. In order to disentangle target selection and response generation, we adapted a word-picture comparison task, which is also based on the same stimuli as the voicekey paradigm. The crucial differences are the presence of a comparison word and the response modality, a left or right button press according to a match or mismatch of target object and comparison word. We introduced our time marker that separates the trial by a long distance between the stimulus objects and the comparison word, while asking the subjects to first identify the target object and then to make one single eye-movement towards the comparison word. In this case, we encountered interferences from the different processing of confirmation vs. negation. Another desired particularity of the paradigm is the equal frequency of response switches and response repetition, an aspect that

is crucial for response retrieval theory. This localizes priming to occur during response generation based on change or repetition of the appropriate response.

We obtained rather diverse result patterns over the entire range of experimental conditions. The additional complexity of the paradigms made the introduction of further conditions necessary. Nevertheless, the results concerning the traditionally interesting experimental conditions DT and TT are rather clear. Our experiment unveiled the fact that stimulus repetitions lead only to facilitation, and that negative priming is generated before the response generation phase.

9.4 The Time Course of Negative Priming

Remembering the promise of the title of the current thesis, we will now describe the time course of a negative priming trial. An objective description is attempted here by collecting the major results of our experiments, without referring to theories or simulations, as we will do in the following paragraph. We will also interpret the results in light of the different theories, and then describe a negative priming trial on the level of mechanisms in the framework of the General Model.

A negative priming trial consists of the following sequence. First, during the perceptual phase a comparison with preceding stimuli occurs which results in a similarity signal. It does not need to rely on purely perceptual features, but takes into account also semantic similarities. Thus DT and TT trials show a similar characteristic in the P300 modulation, and are both accelerated in the stimulus identification phase, albeit objects in TT trials are still more quickly identified than in DT trials. As the target is identified by a perceptual feature, the target selection may run in parallel with perception if not artificially delayed. In this perceptual phase, the processing of a DT trial changes from the initial benefit of easier recognition, to the delay that is characteristic for negative priming due to a high demand on cognitive control if a formerly ignored stimulus has to be attended to. The onset of response generation may as well begin before the target is fully identified as such by accumulating evidence. This final section of the trial processing is accelerated if the response follows the same stimulus again, regardless of whether this response is actually the same or not. However, no decelerations were found in our separated response generation phase.

This picture of a negative priming trial is clear enough to draw some conclusions about the validity of the different theoretical accounts for negative priming. The characteristics of episodic retrieval match our results best. Both distractor inhibition and global threshold theory fail to explain why no negative priming effects are present during stimulus identification. Temporal discrimination does not comply with the equal processing in early TT and DT trials. Response retrieval theory is disproved by the fact that no negative priming is found during response generation. Feature mismatch theory is not really applicable in our paradigms, as we show all our stimuli at the same location. Thus, each trial should show the slow-down caused by the present feature mismatch, as we never repeat both objects at once. We therefore chose a description by means of the General Model in episodic retrieval mode with moderate activation gating.

Following the interpretation of the General Model, a negative priming trial begins with the onset of input to the feature layers and the formation of bindings for the respective stimuli. Simultaneously, the feature instance that defines a target receives additional input by the central executive. The bindings exchange activation between the different feature layers, such that the feature which triggers the response to the target object receives also more input than the corresponding one of the distractor. A similarity signal is computed immediately after the stimulus appeared. The higher the similarity, as in TT and DT trials compared to CO, the stronger the feedback of the former trial's activations. This supporting retrieval leads to a perceptual acceleration. Later in the semantic layer the differential retrieval becomes crucial as all activations, the externally driven and the

retrieved ones, enter the calculation of the threshold that will determine the point in time where the target is selected against the distractor. During DT trials, only the low value of a distractor is fed into the probe target variable making it slower than TT, but we also have less overall activation as in a CO trial where both former objects are retrieved due to the inherent similarity of trials of the same experiment. If the selection of the target has been done, a rather stereotypical mechanism of action initiation is triggered, which is exclusively faster in TT trials, a phenomenon which can not be explained by any of the existing theories, and is thus not yet implemented in the General Model. One possible source may be a higher arousal in cases where the same object is to be attended again, accelerating response execution.

9.5 Summary and Outlook

By the means of innovative experimental paradigms, EEG-recordings, and computational modeling, we obtained a comprehensive picture of negative priming. We found that negative priming is neither produced during perception, nor at response generation. A fact only a single theory, episodic retrieval, can explain.

Besides the global aim of unveiling the temporal structure of a negative priming trial, our project yielded numerous additional results. We enriched the set of negative priming paradigms by introducing intermediate behavioral time markers. We investigated the interaction of negative priming with phenomena that occur naturally in several paradigms and that are caused e.g. by task switching or response repetition. We consolidated several EEG-correlates of negative priming and brought up an subtle way to account for reaction time variations during averaging. With our modeling we advanced the discussion about the most likely theory to explain negative priming, by consistently falsifying one account and by constructing a framework in which all theoretical accounts are comparable.

All in all, the question about negative priming is far from being answered, as our description of a generic negative priming trial can not account for all the oddities found in connection with negative priming. We will finally present several ideas to proceed in order to come closer to the origin of negative priming.

Negative priming is sensitive to age, and therefore an interesting phenomenon for aging research. As we aimed at a complete picture of negative priming itself before we introduce aging as an additional factor, a theory of aging still remains to be included into our approach. Experimentally, however, we already acquired a broad base of experimental data comparing old and young subjects. The evaluation of the data and the integration into our theoretical framework is one of the major points in the remaining eight months of our project.

Another direct continuation of our time marker paradigms is the introduction of multiple makers per trial for a future experiment. According to our advanced averaging techniques, the consideration of individual event related potentials for each part of the trial is straightforward. Therefore we are already planning another EEG study.

In order to extend the approach also to other experimental paradigms, we want to introduce self-organization mechanisms which should enable the General Model to act more flexibly and robustly in different situations. The accelerating effect in the response generation phase has to be tested more thoroughly in order to identify the corresponding mechanism to include it in the General Model.

Bibliography

Allport, D. A., Tipper, S. P., and Chmiel, N. R. J. (1985). Perceptual integration and postcategorical filtering. *Attention and performance*, 11:107 – 132.

Anderson, J. R. (2001). *Kognitive Psychologie*. Spektrum Akademischer Verlag, Berlin, 3. edition.

Baddeley (2000). The episodic buffer: a new component of working memory? *Trends Cogn Sci*, 4(11):417–423.

Baddeley, A. (1998). Working memory. *C R Acad Sci III*, 321(2-3):167–173.

Baddeley, A. (2001). The concept of episodic memory. *Philos Trans R Soc Lond B Biol Sci*, 356(1413):1345–1350.

Baddeley, A. and Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, pages 47–89.

Banks, W. P., Roberts, D., and Ciranni, M. (1995). Negative priming in auditory attention. *Journal of Experimental Psychology: Human Perception & Performance*, 21(6):1354–1361.

Barnard, P. (1985). *Progress in the Psychology of Language*, volume 2, chapter interactive cognitive subsystems: a psycholinguistic approach to short-term memory, pages 197 – 258. Lawrence Erbaum Associates Ltd., Hove UK.

Basar, E., Gonder, A., Ozesmi, C., and Ungan, P. (1975). Dynamics of brain rhythmic and evoked potentials. I. Some computational methods for the analysis of electrical signals from the brain. *Biological Cybernetics*, 20(3-4):137–43.

Behrendt, J., Gibbons, H., Schrobsdorff, H., Ihrke, M., Herrmann, J. M., and Hasselhorn, M. (2009). Event-related brain potential correlates of idendity negative priming from overlapping pictures. *Psychophysiology (accepted)*.

Bermeitinger, C., Wentura, D., and Frings, C. (2008). Nature and facts about natural and artifactual categories: Sex differences in the semantic priming paradigm. *Brain and Language*, 106(2):153–163.

Bookheimer, S. (2002). Functional MRI of Language: New Approaches to Understanding the Cortical Organization of Semantic Processing. *Annual Review of Neuroscience*, 25(1):151–188.

Botly, L. and De Rosa, E. (2007). Cholinergic influences on feature binding. *Behav Neurosci*, 121(2):264–76.

Bressler, S. and Kelso, J. (2001). Cortical coordination dynamics and cognition. *Trends Cogn Sci*, 5(1):26–36.

Buchner, A. and Steffens, M. C. (2001). Auditory negative priming in speeded reactions and temporal order judgements. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 54(4):1125–1142.

Buzsáki, G. (2005). Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus*, 15(7):827–840.

Castleman, K. (1996). *Digital image processing*. Prentice Hall Press Upper Saddle River, NJ, USA.

Ceballos, N. A., Nixon, S. J., and Tivis, R. (2003). Substance abuse-related p300 differences in response to an implicit memory task. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 27(1):157–164.

Chiappe, D. L. and MacLeod, C. M. (1995). Negative priming is not task bound: A consistent pattern across naming and categorization tasks. *Psychonomic Bulletin & Review*, 2(3):364–369.

Christie, J. and Klein, R. M. (2001). Negative priming for spatial location? *Canadian Journal of Experimental Psychology*, 55(1):24–38.

Ciganek, L. (1969). Variability of the human visual evoked potential: normative data. *Electroencephalogr Clin Neurophysiol*, 27(1):35–42.

Conway, A. R. A. (1999). The time-course of negative priming: Little evidence for episodic trace retrieval. *Memory & Cognition*, 27(4):575–583.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol Bull*, 104(2):163–191.

Craik, F. and Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6):671–684.

Craik, F. I. (2002). Levels of processing: Past, present ... and future? *Memory*, 10(5/6):p305 – 318.

Dalrymple-Alford, E. and Budayr, B. (1966). Examination of some aspects of the stroop color-word test. *Perceptual and motor skills*, 23(3):1211 – 1214.

Daurignac, E., Houdè, O., and Jouvent, R. (2006). Negative priming in a numerical piaget-like task as evidenced by erp. *Journal of Cognitive Neuroscience*, 18:730 – 736.

de Weerd, J. (1981). A posteriori time-varying filtering of averaged evoked potentials. I. Introduction and conceptual basis. *Biological Cybernetics*, 41(3):211–22.

Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis. *Neuroimage*, 34(4):1443–1449.

Demb, J., Desmond, J., Wagner, A., Vaidya, C., Glover, G., and Gabrieli, J. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional mri study of task difficulty and process specificity. *Journal of Neuroscience*, 15(9):5870.

DeSchepper, B. and Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22(1):27–47.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, M. J., and Tyler, L. K. (2002). Is there an anatomical basis for category-specificity? semantic memory studies in pet and fmri. *Neuropsychologia*, 40(1):54–75.

Donchin, E. and Coles, M. G. H. (1988). Is the p300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11:357 – 374.

Ecker, U. K. H., Zimmer, H. D., and Groh-Bordin, C. (2007a). Color and context: an erp study on intrinsic and extrinsic feature binding in episodic memory. *Mem Cognit*, 35(6):1483–1501.

Ecker, U. K. H., Zimmer, H. D., Groh-Bordin, C., and Mecklinger, A. (2007b). Context effects on familiarity are familiarity effects of context - an electrophysiological study. *Int J Psychophysiol*, 64(2):146–156.

Eckmann, J. P., Kamphorst, S. O., and Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters*, 5:973–977.

Egner, T. and Hirsch, J. (2005). Where memory meets attention: neural substrates of negative priming. *J Cogn Neurosci*, 17(11):1774–1784.

Fiebach, C. J., Gruber, T., and Supp, G. G. (2005). Neuronal mechanisms of repetition priming in occipitotemporal cortex: spatiotemporal evidence from functional magnetic resonance imaging and electroencephalography. *J Neurosci*, 25(13):3414–3422.

Fletcher, P. C., Frith, C. D., and Rugg, M. D. (1997). The functional neuroanatomy of episodic memory. *Trends Neurosci*, 20(5):213–218.

Flexer, A. (2000). Data mining and electroencephalography. *Statistical Methods in Medical Research*, 9(4):395.

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin & Review*, 2(2):145–173.

Fox, E. (1998). Perceptual grouping and visual selective attention. *Perception & Psychophysics*, 60(6):1004–1021.

Fraser, A. and Swinney, H. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2):1134–1140.

Frings, C. (2005). Verstärkte Reaktionszeitvorteile von Distraktor-zu-Distraktor-Wiederholung im Negativen Priming – Klare Gegenevidenz zur Temporal Discrimination Theory. In Kramer, A. F., Coles, M., and Logan, G. D., editors, *47. Tagung experimentell arbeitender Psychologen (TeaP)*, pages 77–106. Universität Regensburg.

Frings, C. and Eder, A. B. (2009). The time-course of masked negative priming. *Experimental psychology*, 56(5):301 – 306.

Frings, C. and Groh-Bordin, C. (2007). Electrophysiological correlates of visual identity negative priming. *Brain Research*, 1176:82–91.

Frings, C. and Wentura, D. (2005). Negative priming with masked distractor-only prime trials: awareness moderates negative priming. *Exp Psychol*, 52(2):131–139.

Frings, C. and Wentura, D. (2006). Negative priming is stronger for task-relevant dimensions: Evidence of flexibility in the selective ignoring of distractor information. *The Quarterly Journal Of Experimental Psychology*, 59(4):683 – 693.

Frings, C. and Wühr, P. (2007a). Prime-display offset modulates negative priming only for easy-selection tasks. *Memory & Cognition*, 35:504 – 513.

Frings, C. and Wühr, P. (2007b). Prime display offset modulates negative priming only for easy-selection tasks. *Mem Cognit*, 35(3):504–513.

Gamboz, N., Russo, R., and Fox, E. (2002). Age differences and the identity negative priming effect: an updated meta-analysis. *Psychol Aging*, 17(3):525–531.

Ganel, T., Gonzalez, C. L. R., Valyear, K. F., Culham, J. C., Goodale, M. A., and Köhler, S. (2006). The relationship between fmri adaptation and repetition priming. *Neuroimage*, 32(3):1432–1440.

Geisser, S. and Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of Mathematical Statistics*, 29:885–891.

Gibbons, H. (2006). An Event-Related Potential Investigation of Varieties of Negative Priming. *Journal of Psychophysiology*, 20(3):170–185.

Gibbons, H. (2009). Functional brain-electrical correlates of negative priming in the flanker task; evidence for episodic retrieval. *Psychophysiology*, 46.

Gibbons, H. and Rammsayer, T. H. (2004). Differential effects of prime-probe duration on positive and negative location priming: Evidence for opponent facilitatory and inhibitory influences in priming tasks. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 57(1):61–86.

Gibbons, H. and Stahl, J. (2007). Response-time corrected averaging of event-related potentials. *Clinical Neurophysiology*, pages 197–208.

Gratton, G., Coles, M. G. H., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4):468–484.

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci*, 10(1):14–23.

Grison, S. and Strayer, D. L. (2001). Negative priming and perceptual fluency: More than what meets the eye. *Perception & Psychophysics*, 63(6):1063–1071.

Grison, S., Tipper, S. P., and Hewitt, O. (2005). Long-term negative priming: Support for retrieval of prior attentional processes. *Quarterly Journal of Experimental Psychology*, 58A(7):1199 – 1224.

Groh-Bordin, C., Zimmer, H., and Ecker, U. (2006). Has the butcher on the bus dyed his hair? when color changes modulate erp correlates of familiarity and recollection. *NeuroImage*, 32:1879 – 1890.

Groh-Bordin, C., Zimmer, H., and Mecklinger, A. (2005). Feature binding in perceptual priming and in episodic object recognition: evidence from event-related brain potentials. *Cognitive Brain Research*, 24:556–567.

Haig, A., Gordon, E., Rogers, G., and Anderson, J. (1995). Classification of single-trial ERP sub-types: application of globally optimal vector quantization using simulated annealing. *Electroencephalogry Clinical Neurophysiology*, 94(4):288–97.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., and Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17(1):163–169.

Hasher, L., Zacks, R. T., Stoltzfus, E. R., Kane, M. J., and et al. (1996). On the time course of negative priming: Another look. *Psychonomic Bulletin & Review*, 3(2):231–237.

Hayama, H., Johnson, J., and Rugg, M. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*, 46(5):1211–1223.

Healy, D. and Burt, J. S. (2003). Attending to the distractor and old/new discriminations in negative priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 56(3):421–443.

Heil, M. and Rolke, B. (2004). Unattended distractor-induced priming in a visual selective attention task: N400 effects in the absence of rt effects. *Journal of Psychophysiology*, 18:164 – 169.

Henson, R. N., Shallice, T., and Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: a functional mri test of the monitoring hypothesis. *Brain*, 122 (Pt 7):1367–1381.

Hinojosa, J., Villarino, A., Pozo, M., Elosua, R., Merino, J., Moreno, E., and Luna, D. (2007). An event-related potential study of identity positive and negative priming. *International Journal of Psychophysiology*, 66:48 – 55.

Hintzman, D. L. (1991). *Relating Theory to Data: Essays on Human Memory in Honor of Bennet B. Murdock*, chapter Why are formal models useful in psychology?, pages 39 – 56. Lawrence Erlbaum Associates.

Houghton, G. and Tipper, S. P. (1994). A model of inhibitory mechanisms in selective attention. In Dagenbach, D. and Carr, T. H., editors, *Inhibitory processes in attention, memory, and language*, pages 53–112. Academic Press, Inc, San Diego, CA.

Houghton, G. and Tipper, S. P. (1996). Inhibitory Mechanisms of Neural and Cognitive Control: Applications to Selective Attention and Sequential Action. *Brain and Cognition*, 30(1):20–43.

Houghton, G. and Tipper, S. P. (1998). A model of selective attention as a mechanism of cognitive control. In Grainger, J. and Jacobs, A. M., editors, *Localist connectionist approaches to human cognition*, Scientific psychology series, pages 39–74, Mahwah, NJ. Lawrence Erlbaum Associates, Inc.

Houghton, G., Tipper, S. P., Weaver, B., and Shore, D. I. (1996). Inhibition and interference in selective attention: Some tests of a neural network model. *Visual Cognition*, 3(2):119–164.

Ihrke, M. (2007). Negative priming and response-relation: Behavioural and electroencephalographic correlates. Master's thesis, University of Göttingen. Available from: `http://www.psych.uni-goettingen.de/home/ihrke`.

Ihrke, M. (2008). Single trial estimation and timewarped averaging of event-related potentials. B.sc. thesis, Bernstein Center for Computational Neuroscience. Available from: `http://www.psych.uni-goettingen.de/home/ihrke`.

Ihrke, M., Schrobsdorff, H., and Herrmann, J. M. (2008). Compensation for speed-of-processing effects in eeg-data analysis. In *Lecture Notes on Computer Science*, volume 5326, pages 354–361. Springer. Proceedings of IDEAL 2008.

Ihrke, M., Schrobsdorff, H., and Herrmann, J. M. (2009a). Recurrence-Based Synchronization of Single Trials for EEG-Data Analysis. In *Lecture Notes on Computer Science*, volume 5788, pages 118 – 125. Springer. Proceedings of IDEAL 2009.

Ihrke, M., Schrobsdorff, H., and Herrmann, M. J. (2009b). Denoising and Averaging Techniques for Electrophysiological Data. In Wennberg, R. and Perez-Velazquez, J. L., editors, *Coordinated Activity in the Brain: measurements and relevance to brain function and behaviour*. Springer.

James, T. W. and Gauthier, I. (2006). Repetition-induced changes in bold response reflect accumulation of neural activity. *Hum Brain Mapp*, 27(1):37–46.

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 20:371–5. Document Type: Article.Citation.

Jeong, J., Kim, S., and Han, S. (1998). Non-linear dynamical analysis of the eeg in alzheimer's disease with optimal embedding dimension. *Electroenc Clin Neurophys*, 106:220–228.

Jiang, Y., Haxby, J. V., Martin, A., Ungerleider, L. G., and Parasuraman, R. (2000). Complementary neural mechanisms for tracking items in human working memory. *Science*, 287(5453):643–646.

Johnson, M. (2007). *Science of Memory: Concepts*, chapter Memory systems: A cognitive construct for analysis and synthesis, pages 353–357. New York: Oxford University Press.

Johnston, W. A., Dark, V. J., and Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11:3–11.

Joyce, C., Gorodnitsky, I., King, J., and Kutas, M. (2002). Tracking eye fixations with electrooccular and electroencephalographic recordings. *Psychophysiology*, 39(05):607–618.

Joyce, C. A., Gorodnitsky, I. F., and Kutas, M. (2004). Automatic removal of eye movement and blink artifacts from eeg data using blind component separation. *Psychophysiology*, 41(2):313–325.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178.

Kabisch, B. (2003). *Negatives Priming und Schizophrenie - Formulierung und Empirische Untersuchung eines Neuen Theoretischen Ansatzes*. PhD thesis, Friedrich-Schiller-Universität, Jena, Germany.

Kane, M. J., May, C. P., Hasher, L., Rahhal, T., and et al. (1997). Dual mechanisms of negative priming. *Journal of Experimental Psychology: Human Perception & Performance*, 23(3):632–650.

Kathmann, N., Bogdahn, B., and Endrass, T. (2006). Event-related brain potential variations during location and identity negative priming. *Neuroscience Letters*, 394:53–56.

Kennel, M., Brown, R., and Abarbanel, H. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6):3403–3411.

Keogh, E. and Pazzani, M. (2001). Derivative Dynamic Time Warping. In *First SIAM International Conference on Data Mining (SDM 2001)*.

Kok, A. (2001). On the utility of p3 amplitude as a measure of processing capacity. *Psychophysiology*, 38:557–577.

Krause, W., B., S., Gibbons, H., and Kriese, B. (1997). On the distinguishability of conceptual and imaginal representations in elementary thinking. *Zeitschrift für Psychologie*, 205(2):169–204.

Krigolson, O. E. and Holroyd, C. B. (2007). Predictive information and error processing: The role of medial-frontal cortex during motor control. *Psychophysiology*, 44:586–595.

Lange, D., Siegelmann, H., Pratt, H., and Inbar, G. (2000). Overcoming selective ensemble averaging: unsupervised identification of event-related brain potentials. *Biomedical Engineering, IEEE Transactions on*, 47(6):822–826.

Lavie, N. and Fox, E. (2000). The role of perceptual load in negative priming. *Journal of Experimental Psychology: Human Perception & Performance*, 26(3):1038–1052.

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (2004). Load theory of selective attention and cognitive control. *J Exp Psychol Gen*, 133(3):339–354.

Li, S.-C. (2005). *Cognitive Neuroscience of Aging*, chapter Neurocomputational Perspectives Linking Neuromodulation, pages 354 – 379. Oxford University Press.

Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review*, 95:492 – 527.

Lorist, M. M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., and Meijman, T. F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology*, 37:614–625.

Lowe, D. (1985). Further investigations of inhibitory mechanisms in attention. *Memory & Cognition*, 13(1):74–80.

Maccotta, L. and Buckner, R. L. (2004). Evidence for neural effects of repetition that directly correlate with behavioral priming. *J Cogn Neurosci*, 16(9):1625–1632.

MacLeod, C. M., Chiappe, D. L., and Fox, E. (2002). The crucial roles of stimulus matching and stimulus identity in negative priming. *Psychonomic Bulletin & Review*, 9(3):521–528.

Marczinski, C. A., Milliken, B., and Nelson, S. (2003). Aging and repetition effects: Separate specific and nonspecific influences. *Psychology & Aging*, 18(4):780–790.

Marwan, N. and Meinke, A. (2004). Extended recurrence plot analysis and its application to erp data. *International Journal of Bifurcation and Chaos*, 14(2):761–771.

Marwan, N., Romano, M. C., T., M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438:237–329.

Marwan, N., Thiel, M., and Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear processes in Geophysics*, 9(3-4):325–331.

May, C. P., Kane, M. J., and Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin*, 118(1):35–54.

Mayr, S. and Buchner, A. (2006). Evidence for episodic retrieval of inadequate prime responses in auditory negative priming. *Journal of Experimental Psychology: Human Perception and Performance*, 32(4):932–943.

Mayr, S. and Buchner, A. (2007). Negative Priming as a Memory Phenomenon A Review of 20 Years of Negative Priming Research. *Zeitschrift für Psychologie*, 215(1):35–51.

Mayr, S., Niedeggen, M., Buchner, A., and Pietrowsky, R. (2003). Erp correlates of auditory negative priming. *Cognition*, 90(2):11–21.

McConkie, G. and Currie, C. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, 22(3):563–81.

Milliken, B., Joordens, S., Merikle, P. M., and Seiffert, A. E. (1998). Selective attention: A reevaluation of the implications of negative priming. *Psychological Review*, 105(2):203–229.

Milliken, B. and Tipper, S. (1998). *Attention and Inhibition*, chapter 5, pages 191–221. Attention. Psychology Press.

Milliken, B., Tipper, S. P., and Weaver, B. (1994). Negative priming in a spatial localization task: Feature mismatching and distractor inhibition. *Journal of Experimental Psychology: Human Perception & Performance*, 20(3):624–646.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3):134–140.

Moore, C. M. (1994). Negative priming depends on probe-trial conflict: Where has all the inhibition gone? *Perception & Psychophysics*, 56(2):133–147.

Myers, C. and Rabiner, L. (1981). A level building dynamic time warping algorithm for connected word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(2):284–297.

Nagelkerke, N. and Strackee, J. (1979). Some notes on the statistical properties of a posteriori Wiener filtering. *Biological Cybernetics*, 33(2):121–123.

Neill, W. and Westberry, R. (1987). Selective attention and the suppression of cognitive noise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13:327 – 334.

Neill, W. T. (1977). Inhibitory and facilitatory processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 3:444–450.

Neill, W. T. (1997). Episodic retrieval in negative priming and repetition priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23(6):1291–3105.

Neill, W. T. and Kahan, T. A. (1999). Response conflict reverses priming: A replication. *Psychonomic Bulletin & Review*, 6(2):304–308.

Neill, W. T., Lissner, L. S., and Beck, J. L. (1990). Negative priming in same−different matching: Further evidence for a central locus of inhibition. *Perception & Psychophysics*, 48(4):398–400.

Neill, W. T. and Valdes, L. A. (1992). Persistence of negative priming: Steady state or decay? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18(3):565–576.

Neill, W. T., Valdes, L. A., Terry, K. M., and Gorfein, D. S. (1992). Persistence of negative priming: Ii evidence for episodic trace retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18(5):993–1000.

Neumann, E. and Deschepper, B. G. (1992). An inhibition-based fan effect: Evidence for an active suppression mechanism in selective attention. *Canadian Journal of Psychology*, 46(1):1–40.

Newell, A. (1994). *Unified theories of cognition*. Harvard Univ Pr.

Norman, K. A., Newman, E. L., and Perotte, A. J. (2005). Methods for reducing interference in the complementary learning systems model: oscillating inhibition and autonomous memory rehearsal. *Neural Netw*, 18(9):1212–1228.

Oswald, W. and Fleischman, U. M. (1982). *Nürnberger Altersinventar (NAI). Testinventar & NAI Testmanual und Textband*. Göttingen: Hogrefe.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time-series. *Physical Review Letters*, 45(9):712–716.

Park, J. and Kanwisher, N. (1994). Negative priming for spatial locations: Identity mismatching, not distractor inhibition. *Journal of Experimental Psychology: Human Perception & Performance*, 20(3):613–623.

Picton, T., Hunt, M., Mowrey, R., Rodriguez, R., and Maru, J. (1988). Evaluation of brain-stem auditory evoked potentials using dynamic time warping. *Electroencephalography and Clinical Neurophysiology*, 71(3):212–25.

Picton, T. W., Lins, O. G., and Scherg, M. (1995). The recording and analysis of event-related potentials. In Boller, F. and Grafman, J., editors, *Handbook of Neuropsychology*, pages 3 – 73. Elsevier Science B.V.

Prinzmetal, W. (1995). Visual feature integration in a world of objects. *Current Directions in Psychological Science*, 4(3):90–94.

Quillian, M. (1966). *Semantic Memory*. Cambridge, MA: Bolt, Beranak and Newman.

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, 56:129–140.

Risken, H. (1996). *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer.

Rothermund, K., Wentura, D., and Houwer, J. D. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3):482–495.

Rugg, M. (1995). ERP studies of memory. In Rugg, M. and Coles, M. G. H., editors, *Electrophysiology of Mind*, pages 132 – 170. Oxford University Press.

Rugg, M. and Allan, K. (2000). Event-related potential studies of memory. In Tulving, E. and Craik, F., editors, *The Oxford Handbook of Memory*, pages 521–537. Oxford Univ. Press, New York.

Rugg, M. and Doyle, M. (1994). Event-related potentials and stimulus repetition in direct and indirect tests of memory. In Heinze, H. J., Muente, T. F., and Mangu, G. R., editors, *Cognitive Electrophysiology*, pages 124–148. Birkhaeuser, Boston.

Rugg, M. and Nagy, M. (1989). Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology*, 72(5):395.

Scarborough, D. L., Cortese, C., and Scarborough, H. (1977). Frequency and repetition effects. *Journal of Expermental Psycholology-Human Perception and Performance*, 3:1–17.

Schacter, D. L., Dobbins, I. G., and Schnyer, D. M. (2004). Specificity of priming: a cognitive neuroscience perspective. *Nat Rev Neurosci*, 5(11):853–862.

Schmidt, K. and Metzler, P. (1992). *Wortschatztest (WST)*. Weinheim: Beltz.

Schrobsdorff, H., Herrmann, J. M., and Geisel, T. (2007a). A feature-binding model with localized excitations. *Neurocomputing*, 70(10-20):1706–1710.

Schrobsdorff, H., Ihrke, M., Kabisch, B., Behrendt, J., Hasselhorn, M., and Herrmann, J. M. (2007b). A Computational Approach to Negative Priming. *Connection Science*, 19(3):203–221.

Sejnowski, T. J. and Destexhe, A. (2000). Why do we sleep? *Brain Res*, 886(1-2):208–223.

Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28:1059–1074.

Singer, M. (1984). Toward a model of question answering: Yes-no questions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2):285 – 297.

Singer, W. (1995). Synchronization of neuronal responses as a putative binding mechanism. In Arbib, M. A., editor, *The Handbook of Brain Theory and Neural Networks*, pages 960–964, Cambridge. MIT Press.

Spieler, D. H. and Balota, D. A. (1996). Characteristics of associative learning in younger and older adults: evidence from an episodic priming paradigm. *Psychol Aging*, 11(4):607–620.

Stahl, J. and Gibbons, H. (2007). Event-related brain potentials support episodic-retrieval explanations of flanker negative priming. *Experimental Brain Research*, 181(4):595–606.

Stolz, J. A. and Neely, J. H. (2001). Taking a bright view of negative priming in the light of dim stimuli: Further evidence for memory confusion during episodic retrieval. *Canadian Journal of Experimental Psychology*, 55(3):219–230.

Suzuki, W. A. (2006). Encoding new episodes and making them stick. *Neuron*, 50(1):19–21.

Teasdale, J. and Barnard, P. (1993). *Affect, Cognition, and Change: re-modelling depressive thought*. Psychology Press (UK).

Thayer, R. E. (1986). Activation-deactivation adjective check list (ad acl): Current overview and structural analysis. *Psychological Reports*, 58:607–614.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423.

Tipper, S. and Cranston, M. (1985). Selective attention and priming: inhibitory and facilitatory effects of ignored primes. *Quarterly Journal of Experimental Psychology A*, 37(4):591–611.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 37(4):571–590.

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? a review and integration of conflicting views. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 54(2):321–343.

Tipper, S. P. and Baylis, G. C. (1987). Individual differences in selective attention: The relation of priming and interference to cognitive failure. *Personality & Individual Differences*, 8(5):667–675.

Tipper, S. P. and Driver, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition*, 16(1):64–70.

Tipper, S. P., Eissenberg, T., and Weaver, B. (1992). The effects of practice on mechanisms of attention. *Bulletin of the Psychonomic Society*, 30(1):77–80.

Tipper, S. P., MacQueen, G. M., and Brehaut, J. C. (1988). Negative priming between response modalities: Evidence for the central locus of inhibition in selective attention. *Perception & Psychophysics*, 43(1):45–52.

Tipper, S. P. and McLaren, J. (1990). Evidence for efficient visual selectivity in children. In Enns, J. T., editor, *Development of attention: Research and theory*, pages 197–210. North-Holland, Oxford, England.

Tipper, S. P., Meegan, D., and Howard, L. A. (2002). Action-centred negative priming: Evidence for reactive inhibition. *Visual Cognition*, 9(4/5):591–614.

Tipper, S. P., Weaver, B., Cameron, S., Brehaut, J. C., and Bastedo, J. (1991). Inhibitory mechanisms of attention in identification and localization tasks: time course and disruption. *J Exp Psychol Learn Mem Cogn*, 17(4):681–692.

Titz, C., Behrendt, J., Menge, U., and Hasselhorn, M. (2008). A reassessment of negative priming within the inhibition framework of cognitive aging: There is more in it than previously believed. *Experimental Aging Research*, 34:340 – 366.

Treisman, A. (1996). The binding problem. *Curr Opin Neurobiol*, 6(2):171–178.

Treisman, A. (1998). Feature binding, attention and object perception. *Philos Trans R Soc Lond B Biol Sci*, 353(1373):1295–1306.

Truccolo, W. A., Ding, M., and Bressler, S. L. (2001). Variability and interdependence of local field potentials: Effects of gain modulation and nonstationarity. *Neurocomputing*, 38-40:983–992.

Truccolo, W. A., Ding, M., Knuth, K. H., Nakamura, R., and Bressler, S. L. (2002). Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity. *Clinical Neurophysiology*, 113(2):206–226.

Tsodyks, M. (1999). Attractor neural network models of spatial maps in hippocampus. *Hippocampus*, 9(4).

Tulving, E. (2002). Episodic memory: from mind to brain. *Annu Rev Psychol*, 53:1–25.

Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., and Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc Natl Acad Sci U S A*, 91(6):2016–2020.

Ullsperger, P., Metz, A.-M., and Gille, H.-G. (1988). The p300 component of the event-related brain potential and mental effort. *Ergonomics*, 31:1127–1137.

Van Essen, D., Anderson, C., and Felleman, D. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science*, 255(5043):419–423.

Verhaeghen, P. and De Meersman, L. (1998). Aging and the negative priming effect: A meta-analysis. *Psychology & Aging*, 13(3):435–444.

Wagner, M.and Baving, L., Berg, P., Cohen, R., and Rockstroh, B. (2006). An erp investigation of semantic priming, repetition priming, and negative priming in schizophrenic patients. *Journal of Psychophysiology*, 20:195 – 211.

Waszak, F., Hommel, B., and Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: Negative priming and competitor priming. *Memory and Cognition*, 33(4):595.

Wentura, D. and Frings, C. (2005). Repeated masked category primes interfere with related exemplars: new evidence for negative semantic priming. *J Exp Psychol Learn Mem Cogn*, 31(1):108–120.

Wentura, D. and Rothermund, K. (2003). The ”meddling-in” of affective information: A general model of automatic evaluation effects. In Musch, J. and Klauer, K. C., editors, *The psychology of evaluation: Affective processes in cognition and emotion*, pages 51–86. Erlbaum.

West, R., Dowry, R., and McConville, C. (2004). Sensitivity of medial frontal cortex to response and nonresponse conflict. *Psychophysiology*, 41:739 – 748.

Whalen, A. (1971). *Detection of Signals in Noise*. Academic Press.

Wilding, E. and Rugg, M. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, 119(3):889.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering and Computing*, 5:539–53.

Wright, C. I., Keuthen, N. J., Savage, C. R., Martis, B., Williams, D., Wedig, M., McMullin, K., and Rauch, S. L. (2006). Brain correlates of negative and positive visuospatial priming in adults. *NeuroImage*, 30(3):983–991.

Yee, P. L., Santoro, K. E., Grey, A. L., and Woog, V. (2000). Deeper processing at target selection increases the magnitude of negative priming. *Memory & Cognition*, 28(8):1429–1436.

Yeung, N., Cohen, J. D., and Botvinick, M. M. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, 111:931–959.

Yonelinas, A. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3):441–517.

Zimmer, H. D., Mecklinger, A., and Lindenberger, U. (2006). *Handbook of binding and memory: Perspectives from cognitive neuroscience*. Oxford University Press, Oxford.

A Tables

A.1 Tables of Chapter 2

	⟨RT⟩ [ms]	(sd)	effect [ms]
CO	660.22	(62.85)	—
DT	681.57	(69.65)	-21.36
DTTD	685.92	(78.04)	-25.70
TT	625.02	(65.29)	35.20
DDTT	600.69	(70.56)	59.53

Table A.1: Reaction times, standard deviation and priming effects, i.e. the differences of control reaction time and reaction time of the according condition.

A.2 Tables of Chapter 3

	⟨RT⟩ [ms]	(sd)	effect
CO	525.84	(14.78)	—
DT	546.40	(19.62)	-20.55
DTTD	552.12	(22.19)	-26.28
TT	489.76	(18.38)	36.09
DDTT	465.56	(17.31)	60.29

Table A.2: Reaction times for a response stimulus interval of 1500ms. To match the experimental data given in section 2.2 an additional delay of ≈ 135 ms for perception and action initiation has to be assumed. These processes will also further raise the variance but not alter the strength of the effects. Given are also the variance and the effect in dependence on the condition. These reaction times were obtained in an exemplary simulated session of 400 trials. The parameters of this simulated experiment were: $\alpha = 0.028$, $\beta = 0.003$, $\xi = 0.0016$, $\zeta = 0.0053$, $\gamma = 0.013$, $t_{delay} = 15$ms and $\sigma = 0.62$. A simulated time step was $0.3ms$.

	⟨ RT ⟩	(sd)	effect
CO	525.84	(14.78)	—
COSO	407.48	(22.91)	118.36
DTSO	427.72	(24.77)	98.12
TTSO	299.58	(36.51)	226.26
SOCO	573.55	(19.98)	-47.70
SODT	624.80	(34.30)	-98.96
SOTT	548.20	(27.70)	-22.36

Table A.3: Predictions of reaction times for a response stimulus interval of 1500ms, together with variances and absolute effects in dependence of the condition of an exemplary simulated session. The usual priming conditions are not shown. The parameters are the same as in table A.2.

A.3 Tables of Chapter 4

	⟨RT⟩ [ms]	(sd)	effects [ms]
control	765.8	(198.8)	—
DT	793.3	(214.2)	⋆⋆⋆ -27.5
TT	632.9	(139.0)	⋆⋆⋆ 132.9

Table A.4: Reaction times for the different priming condition (CO, DT, TT) together with their standard deviation and priming effects as the difference between CO and DT, resp. TT.

A.4 Tables of Chapter 6

	⟨RT⟩ [ms]	Effect [ms]	Target Effect [ms]
CO	500.85	—	—
DT	525.66	−24.81	−21.36
TT	466.07	34.77	35.20
TD	505.75	−4.91	—
DD	494.15	6.70	—
DTTD	524.29	−23.44	−25.70
DDTT	451.75	49.10	59.53

Table A.5: Reaction times for the different priming conditions TT, DT, DDTT and DTTD of the calibrated simulation in comparison to their target effect strengths, see section 2.2. Note that the simulation already provides predictive data for the conditions DD and TD which are not subject to fitting.

	Reaction Time [ms]	Priming Effect [ms]
CO	535.42	—
DT	550.08	−14.67
TT	507.66	27.76
TD	551.02	−15.61
DD	513.50	21.92
CO_SW	528.25	—
DT_SW	495.43	32.82
TT_SW	542.12	−13.87
TD_SW	505.53	22.71
DD_SW	543.02	−14.78

Table A.6: Predictions for the priming conditions TT, DT, DD and TD and their task-switch counterparts in the pre-cue paradigm. Priming effects are determined by the difference of the corresponding control and the respective priming condition. The repetition of an object in the same color leads to a benefit, the reappearance in a different color disturbs processing.

	Reaction Times [ms]			Priming Effects [ms]		
	Total	R^{si}	R^{ts}	Total	R^{si}	R^{ts}
Co	1263.21	465.98	797.23	—	—	—
DT	1345.47	507.57	837.90	−82.26	−41.59	−40.67
TT	1235.54	465.85	769.69	27.68	0.13	27.55
TD	1347.49	515.23	832.26	−84.28	−49.25	−35.02
DD	1244.61	465.89	778.72	18.60	0.08	18.52
CO_SW	1372.73	465.95	906.78	—	—	—
DT_SW	1354.00	465.87	888.13	18.73	0.09	18.64
TT_SW	1468.04	514.97	953.07	−95.31	−49.01	−46.29
TD_SW	1353.33	465.90	887.43	19.40	0.06	19.34
DD_SW	1446.22	507.36	938.86	−73.49	−41.41	−32.08

Table A.7: Predictions for the priming conditions TT, DT, DD and TD in dependency of a task switch or task repetition derived by our simulation of the ISAM. Additional to the total reaction time the ISAM produces predictions for the divided reaction times into the parts $R^{si}$and R^{ts}. Effects are differences of the appropriate control and the respective priming condition. Note that perceptually identical displays like DT and DD_SW produce the same R^{si}.

	target red		target green	
	$\langle$RT$\rangle$ [ms] (sd)	Effect [ms]	$\langle$RT$\rangle$ [ms] (sd)	Effect [ms]
CO	735,36 (110.92)	—	783.76 (103.86)	—
DT	753.40 (113.75)	−18	802.30 (101.37)	−18
TT	686.63 (100.33)	48	740.83 (103.95)	43
TD	757.36 (135.24)	−22	788.63 (94.09)	−5
DD	713.26 (98.83)	22	764.96 (100.26)	19

Table A.8: Mean reaction times, standard deviation in brackets and effect strength for the five priming conditions CO, DT, TT, TD and DD in the two baseline experiments. The left side shows the data for red targets, and the right side shows green targets. Note that trials with green targets are about 48 ms slower than trials with a red target due to the stronger saliency of the color red. But priming effects, the difference of CO and the respective condition, are mostly identical in the two target color cases.

	⟨RT⟩ [ms] (sd)	Effect [ms]
CO	759.56 (104.32)	—
DT	777.85 (100.71)	⋆⋆ −18.28
TT	713.73 (87.50)	⋆⋆⋆ 45.83
TD	773.00 (108.15)	⋆ −13.43
DD	739.11 (91.68)	⋆⋆ 20.45

Table A.9: Mean reaction time, standard deviation in brackets and effect strength for the five priming conditions CO, DT, TT, TD and DD pooled over the two baseline experiments. All conditions produce priming effects. DT and TD have a negative impact, whereas TT and DD lead to facilitation.

	⟨RT⟩ [ms] (sd)	Effect [ms]
CO	788.567 (118.39)	—
DT	813.500 (107.17)	⋆⋆ −24.93
TT	733.800 (102.35)	⋆⋆⋆ 54.76
TD	792.167 (119.42)	−3.60
DD	788.367 (111.97)	0.20
CO_SW	795.400 (120.73)	—
DT_SW	823.467 (145.76)	⋆⋆ −28.07
TT_SW	800.700 (126.88)	−5.30
TD_SW	816.633 (124.90)	⋆ −21.23
DD_SW	815.067 (123.31)	⋆⋆ −19.66

Table A.10: Mean reaction time, standard deviation and priming effect for the priming conditions TT, DT, DD and TD and their task-switch counterparts in the pre-cue paradigm. Priming effects are determined by the difference of the corresponding control and the respective priming condition. DT and TT behave similarly to the baseline experiment. In the task-switch conditions, an overall deceleration is present which may be due to a negative shift of conditions that repeat an object.

	without mask		with mask	
	$\langle R^{si} \rangle$ [ms] (sd)	Effect [ms]	$\langle R^{si} \rangle$ [ms] (sd)	Effect [ms]
CO	1470.53 (633.33)	—	1607.85 (431.49)	—
DT	1457.33 (603.48)	13	1545.21 (418.74)	63
TT	1395.73 (593.51)	75	1471.92 (414.66)	136
TD	1464.66 (626.80)	6	1598.50 (423.73)	9
DD	1410.26 (581.47)	60	1538.92 (407.10)	69
CO_SW	1452.13 (622.73)	—	1581.78 (435.97)	—
DT_SW	1456.33 (613.08)	−4	1563.50 (412.72)	18
TT_SW	1441.06 (609.74)	11	1561.57 (419.11)	20
TD_SW	1399.20 (564.47)	53	1492.00 (380.62)	96
DD_SW	1450.93 (607.58)	1	1688.35 (471.45)	−106

Table A.11: Mean R^{si}, standard deviation in brackets and effect strength for the five priming conditions CO, DT, TT, TD and DD in the two post-cue experiments. The left side shows the data for the non-masked experiment, where the third of the subjects that use an after-image strategy causes a faster mean reaction time and a much higher variance than on the right side, where data from the masked experiment is shown.

	$\langle R^{si} \rangle$ [ms] (sd)	Effect [ms]
CO	1536.82 (540.25)	—
DT	1499.75 (515.27)	37.06
TT	1432.51 (507.40)	⋆⋆ 104.31
TD	1529.27 (533.32)	7.55
DD	1472.37 (500.27)	⋆ 64.44
CO_SW	1514.72 (535.24)	—
DT_SW	1508.06 (519.60)	6.65
TT_SW	1499.24 (520.76)	15.48
TD_SW	1444.00 (478.32)	⋆⋆ 70.72
DD_SW	1565.55 (549.86)	⋆ −50.82

Table A.12: Stimulus identification reaction times R^{si}, standard deviation and priming effect for the priming conditions TT, DT, DD and TD and their task-switch counterparts in the post-cue paradigm pooled over both experiments with and without mask. Priming effects are determined by the difference of the corresponding control and the respective priming condition. Note the inappropriateness to pool DD_SW over the mask and non-mask group.

	without mask		with mask	
	$\langle R^{ts} \rangle$ [ms] (sd)	Effect [ms]	$\langle R^{ts} \rangle$ [ms] (sd)	Effect [ms]
CO	966.46 (125.85)	—	865.50 (158.62)	—
DT	989.66 (114.47)	−23	877.64 (170.83)	−12
TT	957.06 (118.14)	9	862.00 (168.23)	4
TD	953.20 (135.59)	13	857.42 (157.47)	8
DD	918.60 (110.99)	48	851.50 (192.85)	14
CO_SW	957.86 (95.55)	—	862.64 (130.98)	—
DT_SW	994.20 (82.43)	−36	890.85 (145.57)	−28
TT_SW	1028.20 (110.20)	−70	912.92 (144.69)	−40
TD_SW	936.06 (107.76)	22	823.92 (146.35)	39
DD_SW	973.06 (101.21)	−15	876.00 (141.36)	−13

Table A.13: Mean R^{ts}, standard deviation in brackets and effect strength for the five priming conditions CO, DT, TT, TD and DD in the two post-cue experiments. The left side shows the data for the non-masked experiment, the left side the masked part.

	$\langle R^{ts} \rangle$ [ms] (sd)	Effect [ms]
CO	915.98 (26.49)	—
DT	933.65 (26.82)	* −17.86
TT	909.53 (26.83)	6.55
TD	905.31 (27.22)	10.75
DD	885.05 (28.96)	* 31.51
CO_SW	910.25 (21.18)	—
DT_SW	942.52 (21.76)	* −32.41
TT_SW	970.56 (23.77)	*** −60.65
TD_SW	879.99 (23.74)	* 29.96
DD_SW	924.53 (22.70)	−14.31

Table A.14: Target selection reaction times R^{ts}, standard deviation and priming effect for the priming conditions TT, DT, DD and TD and their task-switch counterparts in the post-cue paradigm pooled over both experiments with and without mask. Priming effects are determined by the difference of the corresponding control and the respective priming condition.

A.5 Tables of Chapter 7

	Priming Effects	
	same response	different response
TT	30.98	n.a.
DT	n.a.	−24.80
DD	n.a.	49.44
TD	n.a.	−17.74

Table A.15: Predictions for the difference and reaction times of control and the four priming conditions TT, DT, DD and TD derived by our simulation of the ISAM.

	same response	different response
	Mean Reaction [ms] (sd)	
control	976.0 (301.3)	945.2 (258.3)
DT	984.0 (281.0)	974.7 (287.6)
TT	959.5 (265.2)	912.8 (244.1)
TD	989.8 (323.8)	949.5 (267.3)
DD	965.0 (294.5)	968.0 (300.7)
	Priming Effects [ms]	
DT	⋆ −8.0	⋆ −29.5
TT	⋆⋆ 16.5	⋆⋆ 32.4
TD	−13.8	−4.3
DD	11.0	⋆ −22.8

Table A.16: Reaction times and priming effects in dependence on priming condition and response relation. Priming effects are computed as the difference of the reaction time in the control and the prime condition, respectively.

	R^{ts}	R^{rs}
	Mean Reaction [ms] (sd)	
CO	417.2 (107.3)	515.8 (63.9)
DT	429.4 (107.9)	521.7 (73.1)
TT	414.7 (96.0)	496.7 (60.7)
TD	428.8 (119.0)	510.4 (68.9)
DD	422.2 (109.5)	518.7 (75.5)
	Priming Effects [ms]	
DT	$\star\star$ −12.2	−5.9
TT	2.5	$\star\star$ 19.2
TD	$\star$ −11.5	5.5
DD	−5.0	−2.9

Table A.17: Partial reaction times and priming effects for each priming condition. The color switch of repeated objects gives a slowdown in the perceptual phase, whereas an acceleration of TT trials is located in the second part of trial processing.

A.6 Tables of Chapter 8

	free parameters
ρ_f ρ_b ρ_s ρ_a	rise time constants
δ_f δ_b δ_s δ_a	decay time constants
τ_{s^θ} τ_{a^θ}	adaptation time constant for thresholds
ν_{s^θ} ν_{a^θ}	normalization factors for thresholds
$\hat{b}$	maximum binding strength
$\#_b$	maximum number of simultaneous bindings
	theory specific parameters
α	target amplification
$\underline{F}$	baseline activity in case of feature absence
ι	distractor inhibition
β	time constant of activation broadening within one feature
$\underline{\sigma}_{f\to s}$, $\underline{\sigma}_{f,s\to a}$	maximum synaptic depression while full blocking
τ_{block}	adaptation time constant for σ_{block}
$\hat{e}$	initial memory strength
δ_e	decay time constant for memories
$\underline{d}$	initial value of d at display onset
$\hat{d}$	fixpoint for the adaptation of d
τ_d	adaptation time constant for d
$\underline{u}$	initial value of u at display onset
τ_u	adaptation time constant for u

Table A.18: Parameters of the General Model. The upper part shows parameters determining the overall behavior of the General Model. The lower part describes parameters that control the processes that are postulated by a particular theory for negative priming.

$S(f)$	maps feature activations on semantic input
$A(s, f, \sigma_{f,s\to a})$	maps feature and semantic activations on action input

Table A.19: Mappings to be defined by the central executive.

	Eq.	Description	Comment
f_i^j	(8.5)	activation of the jth instance of feature i	$f_{\text{color}}^{\text{green}}$ is the subjective presence of green
F_i^j	(8.1)	external input to the j-th instance of feature i	$F_{\text{color}}^{\text{green}} = \hat{F}$ while green is shown
I	(8.4)	inhibition of the distractor object	I= 0 no inhibition
A	(8.3)	excitation of the target	A= 0 no excitation
$\mathbf{F}_i^j$	(8.2)	total input including excitation and inhibition	$\mathbf{F}_{\text{color}}^{\text{green}} = \hat{F} + A$ while green is target color
$b_{\{i_k,j_k\}_k}$	(8.6)	binding strength of an object possessing j_kth instance of feature i_k	$b_{\{\text{color,green}\},\{\text{shape,ball}\}} = \hat{b}$ in subjective presence of a green ball
s^j	(8.7)	activation of the jth semantic representation	$s^{\text{ball}} = 0$ subjective absence of the semantic concept *ball*
s^θ	(8.10)	adaptive threshold level in the semantic layer	The relation between the s^j and s^θ determines the input to the action layer.
$\sigma_{f\to s}$	(8.8)	synaptic strength between features and semantic layer	$\sigma_{f\to s} = 0$ all information flow blocked
σ_{block}	(8.9)	synaptic block caused by matching retrieval	$\sigma_{\text{block}} = 1$ full block of bottom up information flow
a^j	(8.11)	activation of the jth action alternative according to the task	$a^{\text{yes}} = 0$ means the impossibility to answer *yes*
$\sigma_{f,s\to a}$	(8.12)	synaptic strength between features and semantic layer projecting onto the action layer	$\sigma_{f.s\to a} = 0$ all information flow blocked
r_a	(8.13)	relative retrieval strength for actions	$r_a = 1$ retrieval equal to other layers
a^θ	(8.14)	adaptive threshold level in the action layer	Action alternatives a^j above a^θ determine possible actions of the model.
e_k	(8.15)	memory strength of the kth last episode	$e_1 = 0$ the last episode is totally forgotten
e_k^v	(8.16)	memorized value of variable v	$e_1^{f_{\text{color}}^{green}} = 1$ green was active at last reaction
r_k	(8.17)	retrieval strength for the kth last episode	$r_k = 0$ no stimulus match, therefore no retrieval
o_k	(8.18)	old-new signal comparing the kth last trial with the actual one	$o_1 = 0$ the actual percept has just been shown
d	(8.19)	threshold for old-new decisions	
u	(8.19)	uncertainty width for old-new decisions	$r_1 > d + u$ the last episode is recognized as old

Table A.20: Variables describing the state of the General Model ordered by their location in the different layers. See the corresponding equations for details on the particular dynamics.

	⟨RT⟩ [ms](sd)	Effect [ms]
CO	959.34	—
DT	981.25	−21.91
TT	902.50	56.84
TD	1116.94	−157.60
DD	1032.87	−73.53

Table A.21: Mean reaction time and effect strength for the five priming conditions CO, DT, TT, TD and DD produced by the General Model in episodic retrieval mode as described in section 8.5.

	⟨RT⟩ [ms](sd)	Effect [ms]
CO	1154.19	—
DT	1169.85	−15.66
TT	1224.12	−69.93
TD	1184.10	−29.91
DD	1187.35	−33.16

Table A.22: Mean reaction time and effect strength for the five priming conditions CO, DT, TT, TD and DD produced by the General Model in episodic retrieval mode as described in section 8.5.

	⟨RT⟩ [ms](sd)	Effect [ms]
CO	839.49	—
DT	852.87	−13.38
TT	839.04	0.44
TD	852.69	−13.20
DD	849.26	−9.78

Table A.23: Mean reaction time and effect strength for the five priming conditions CO, DT, TT, TD and DD produced by the General Model in episodic retrieval mode during a word-picture comparison task as described in section 8.6.

	⟨RT⟩ [ms](sd)	Effect [ms]
CO	698.82	—
DT	715.22	−16.40
TT	676.47	22.35
TD	709.10	−10.28
DD	725.74	−26.91

Table A.24: Mean reaction time and effect strength for the five priming conditions CO, DT, TT, TD and DD produced by the General Model in episodic retrieval mode with a weak modulation of activation transfer during a word-picture comparison task as described in section 8.6.

Publication List

All publications, posters, talk slides and abstracts, as long as available online, can be downloaded from my webpage at `www.nld.ds.mpg.de/~hecke/research.html`

Publications

J. Michael Herrmann, **Hecke Schrobsdorff**, Theo Geisel: *Localized Activations in a simple neural field model*. Neurocomputing 65-66, 2005

Hecke Schrobsdorff, J. Michael Herrmann, Theo Geisel: *A feature binding model with localized excitations*. Neurocomputing 70(10-20):1706-1710 (2007). doi:10.1016/j.neucom.2006.10.049

Hecke Schrobsdorff, Matthias Ihrke, Björn Kabisch, Jörg Behrendt, Marcus Hasselhorn, J. Michael Herrmann: *A Computational Approach to Negative Priming*. Connection Science, Vol. 19, No. 3, September 2007, 203-221, doi:10.1080/09540090701507823

Matthias Ihrke, **Hecke Schrobsdorff** and J. Michael Herrmann: *Compensation for Speed-of-Processing Effects in EEG-Data Analysis*. Lecture Notes on Computer Science, Proceedings of IDEAL 2008. 5326, pp. 354-361.

Matthias Ihrke, **Hecke Schrobsdorff** and J. Michael Herrmann: *Denoising and Averaging Techniques for Electrophysiological Data*. In: Velazquez, Jose Luis Perez; Wennberg, Richard (Eds.). Coordinated Activity in the Brain Measurements and Relevance to Brain Function and Behavior. Springer Series in Computational Neuroscience, 2009. ISBN: 978-0-387-93796-0 165 – 189

Matthias Ihrke, **Hecke Schrobsdorff** and J. Michael Herrmann: *Recurrence-Based Synchronization of Single Trials for EEG-Data Analysis*. Lecture Notes on Computer Science, Proceedings of IDEAL 2009. 5788, pp. 118 – 125.

Jörg Behrendt, Henning Gibbons, **Hecke Schrobsdorff**, Matthias Ihrke, J. Michael Herrmann and Marcus Hasselhorn: *Event-Related Brain Potential Correlates of Idendity Negative Priming From Overlapping Pictures*. Psychophysiology (accepted).

Matthias Ihrke, Jörg Behrendt, **Hecke Schrobsdorff**, J. Michael Herrmann and Marcus Hasselhorn: *Response Retrieval and Negative Priming: Encoding and Retrieval Specific Effects*. Experimental Psychology (submitted).

Invited Talks

24.01.2006	HU Berlin. **H Schrobsdorff**: *Localized Excitations in Networks of Spiking Neurons*
21.03.2006	Institute for Theoretical Neurophysics, University of Bremen. **H Schrobsdorff**: *Negatively Primed Psychophysics*
04.02.2008	German Primate Center. **H Schrobsdorff**: *A Computational Approach to Negative Priming*
19.07.2009	CNS*2009 Workshop, Modeling neural mass action in brain networks using delay differential equations: **H Schrobsdorff**, K Doubrovinski, J M Herrmann. *Localized Solutions in Two Dimensional Neural Fields*

Talks at Conferences

10.03.2004 Frühjahrstagung der DPG in Regensburg. **H Schrobsdorff**, J M Herrmann and T Geisel: *Localized Solutions in Neural Fields*.

23.03.2006 Workshop der Abteilung für Pädagogische Psychologie und Entwicklungspsychologie der Universität Göttingen in Hermannsburg. J Behrendt, **H Schrobsdorff**, M Ihrke, J M Herrmann and M Hasselhorn: *Mathematische Modellierung als Zugang zum Negative Priming*.

20.07.2006 CNS*2006: CANN-Workshop. **H Schrobsdorff**: *Symmetry Breaking in a Two Dimensional Continuous Attractor Neural Network*.

19.09.2006 45. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) in Nürnberg. J Behrendt, **H Schrobsdorff**, J M Herrmann, and M Hasselhorn: *Mathematische Modellierung als Zugang zum Negative Priming - empirische Prüfung der Implementierung von Alterseinflüssen*.

27.03.2007 Frühjahrstagung der DPG in Regensburg. **H Schrobsdorff**, M Ihrke, J Behrendt, B Kabisch, M Hasselhorn and J M Herrmann: *Modeling Neural Correlates of Selective Attention*.

09.06.2007 Psychologie & Gehirn 2007 in Dortmund. J Behrendt, H Gibbons, **H Schrobsdorff**, M Ihrke, J M Herrmann and M Hasselhorn: *Event-Related Brain Potential Correlates of Identity Negative Priming* .

12.07.2007 CNS*2007 Workshop, Synchronization of brain signals: What is real, what is not. **H Schrobsdorff**, Matthias Ihrke: *Response-Time Corrected Averaging of Event-Related Potentials*.

18.07.2007 BCCN Minisymposium. **H Schrobsdorff**: Modeling Aging Effects in Selective Attention *A General Model for Negative Priming*.

25.09.2007 BCCN Symposium Göttingen. **H Schrobsdorff**, M Ihrke, J Behrendt, M Hasselhorn, J M Herrmann: *Unifying Negative Priming Theories*.

26.09.2007 Tagung der Fachgruppe Entwicklungspsychologie. J Behrendt, H Gibbons, **H Schrobsdorff** , M Ihrke, J M Herrmann, M Hasselhorn: *EKP-Korrelate des negativen Priming bei jungen und alten Erwachsenen*.

05.03.2008 TeaP 2008. M Ihrke, J Behrendt, **H Schrobsdorff**, J M Herrmann and M Hasselhorn: *Investigating the Temporal Course of Attentional Processing - A Test of the Response-Retrieval Account of Negative Priming*.

25.07.2008 ICP 2008. J Behrendt, M Ihrke, **H Schrobsdorff**, J M Herrmann H Gibbons and M Hasselhorn: *Negative priming in younger and older adults: Selection or response effect?*.

08.10.2008 BCCN Symposium 2008 in Munich. **H Schrobsdorff**, M Ihrke, J Behrendt, H Gibbons, Marcus Hasselhorn, J M Herrmann: *Aging Deficits as a Side-Effect of Optimization* .

Posters at Conferences

20.07.2004 CNS*2004. **H Schrobsdorff**, J M Herrmann and T Geisel: *Localized Activations in a Simple Neural Field Model*.

10.03.2005 Frühjahrstagung der DPG in Berlin. **H Schrobsdorff**, V David, J Hass, J M Herrmann and T Geisel: *Localized Breathers*.

19.03.2005 Computational and Systems Neuroscience (2005) in Salt Lake City. **H Schrobsdorff**, V David, J Hass, J M Herrmann, and T Geisel: *Localization of neural activity*.

10.10.2005 1. Bernstein Symposium on Computational Neuroscience in Freiburg (Brg.). **H Schrobsdorff**, M Ihrke, B Kabisch, J Behrendt, M Hasselhorn, J M Herrmann: *A Computational Approach to Negative Priming*.

02.02.2006 Japan Germany Symposium on Computational Neuroscience at RIKEN Brain Science Institute. **H Schrobsdorff**, M Ihrke, B Kabisch, J Behrendt, M Hasselhorn, J M Herrmann: *A Computational Account of Negative Priming*.

27.03.2006 48. Tagung experimentell arbeitender Psychologen (TeaP) in Mainz. J Behrendt, **H Schrobsdorff**, M Ihrke, J M Herrmann and M Hasselhorn: *Mathematische Modellierung als Zugang zum Negative Priming - Konzeption und empirische Prüfung*.

09.06.2006 Tagung Psychologie und Gehirn der Fachgruppe Biologische Psychologie und Neuropsychologie der Deutschen Gesellschaft für Psychologie (DGPs) in Dresden. J Behrendt, **H Schrobsdorff**, M Ihrke, J M Herrmann and M Hasselhorn: *Computational modelling as an approach to neural mechanisms of Negative Priming*.

18.07.2006 CNS*2006. **H Schrobsdorff**, M Ihrke, B Kabisch, J Behrendt, M Hasselhorn, J M Herrmann: *A Computational Account of the Negative Priming Effect*.

18.07.2006 CNS*2006. **H Schrobsdorff**, J Herrmann, T Geisel: *A Feature-Binding Model with Localized Excitations*.

09.07.2007 CNS*2007. **H Schrobsdorff**, M Ihrke, J Behrendt, M Hasselhorn and J M Herrmann: *Modeling Negative Priming Using EEG Data Abstract*.

04.12.2007 BCCN Zwischenevaluation. **H Schrobsdorff**, M Ihrke, J Behrendt, M Hasselhorn and J M Herrmann: *The General Model for Negative Priming* .

21.07.2008 ICP 2008. **H Schrobsdorff**, M Ihrke, J Behrendt, J M Herrmann and M Hasselhorn: *The General Model for Negative Priming* .

09.10.2008 BCCN Symposium 2008 in Munich. M Ihrke, **H Schrobsdorff**, J M Herrmann: *Integration of Temporal Information Across Trials in EEG Data Analysis* .

09.10.2008 BCCN Symposium 2008 in Munich. J Behrendt, M Ihrke, **H Schrobsdorff**, J M Herrmann, H Gibbons, Marcus Hasselhorn: *Is the Negative Priming Effect a Consequence of Target Selection or Response Generation? A Study with Younger and Older Adults* .

26.03.2009 Frühjahrstagung der DPG 2009 in Dresden. **H Schrobsdorff**, J Behrendt, M Ihrke, J M Herrmann, H Gibbons, M Hasselhorn: *Cognitive Aging as a Loss of Critcality*.

26.03.2009 Neurobiology Conference in Göttingen. **H Schrobsdorff**, M Ihrke, J Behrendt, H Gibbons, M Hasselhorn, J M Herrmann: *Learning as a Cause for Aging Impairments* .

26.03.2009 Neurobiology Conference in Göttingen. M Ihrke, **H Schrobsdorff**, J M Herrmann: *Compensating for Temporal Variation in Event-Related Potential Analysis* .

28.03.2009 Neurobiology Conference in Göttingen. J M Herrmann, **H Schrobsdorff**, M Ihrke, J Behrendt, H Gibbons, M Hasselhorn *A Computational Framework for Theories of Negative Priming*.

19.07.2009 CNS*2009. **H Schrobsdorff**, M Ihrke, J Behrendt, J M Herrmann, Th Geisel, A Levina: *Are Aging Effects Caused by Optimization?*

Theses

Diploma Thesis Physics (2005): *Localization of Neural Activity*. Published by Shaker, DOI:10.2370/234_197

Diploma Thesis Mathematics (2006): *Quantum Computing: Anwendungen in Kryptographie und der Darstellung Boolescher Funktionen*. Published by Shaker, DOI:10.2370/234_196

Awards

09.07.2007 *Best Poster* Award and *Biggest Poster* Award at CNS*2007: **H Schrobsdorff**, M Ihrke, J Behrendt, M Hasselhorn and J M Herrmann: *Modeling Negative Priming Using EEG Data*.

19.07.2009 *Best Poster* Award at CNS*2009: **H Schrobsdorff**, M Ihrke, J Behrendt, J M Herrmann, T Geisel, A Levina: *Are Aging Effects Caused by Optimization?*

Acknowledgements

The genesis of the current thesis is influenced by many people. I want to express my unlimited gratitude to:

- Theo Geisel for drudging day and night in order to provide the best working conditions ever possible, for helpful advice in crucial points, for freedom to flower out, for a keen pace on the slopes and for the supervision of the current thesis;
- Marcus Hasselhorn for motivation and esteem for our work in critical situations, for gentle but decisive and farsighted project lead and also for putting his inexistent time into the evaluation of my thesis;
- J. Michael Herrmann for nearly ten years of pointing me to the most interesting topics and teaching me how science works, for discursive but rewarding discussions, for loose guidance and the supervision of several theses;
- Matthias Mittner fka. Ihrke and Jörg Behrendt for not simple collegial cowork, but real friendship that made work in project C4 great, for a (sometimes in the very crashy sense) crash course on psychological research, for the statistical analyses presented in this thesis, for thorough corrections and remarks to sometimes unreadable preversions of the different chapters and for innumerable other things;
- Regina Wunderlich, Katharina Jeremias and Tobias Niemann for years of administrative support and help on various levels;
- Yorck Fabian Beensen, Denny Fliegner and Barbara Guichemer for gadget support, open ears and professional training for the time science kicks me out;
- the lunch club: Anna, Georg, Tatjana, Min, Wolfgang, Peter, Michael, Olav and Jakob for a communicative diversification of my eating habits;
- Dominik, Sven, Oliver, Raoul and Christoph for a lot of time, fun but also serious moments we shared, be it in a common office or during conference stays;
- Henning Gibbons and Jutta Stahl for a first swimming free in selling our psychological results;
- Shu-Chen Li and Timo von Oertzen for intense discussions not only on computational modeling of aging;
- Torsten Wüstenberg for a dedicated introduction to the technical issues of EEG recording and enthusiasm for freaky ideas how to abuse the equipment;
- Ralph Meier for deep interest, long discussions and hands-on help when we were clueless about our whole mess of data;
- Cora Titz and Uwe Menge for sharing expertise on Negative Priming;
- Björn Kabisch for the insight in how far scientific thoroughness can go;
- Christian Frings for targeted hints on our work and preliminary information about his research;
- Steven Tipper and George Houghton for open-mindedness in personal communications;
- Florian Waszak for a review of his data in order to help with our task switching paradigm;
- Anna, Caro, Axel, Melanie, Svenja, Ira and Toni for tedious work in insular rooms, tiring filling of endless error lists and toleration of Matthias', Jörg's and mine fervid discussions on the correct use of statistics, the optimal way to design a new experiment and frequent meandering of those;
- Michelle Monteforte for proofreading the entire thesis;
- Astrid and Heiner for finally giving me the opportunity to become adult;
- Ilka, Tom and Ronja for always reminding me of what is really important.

My work was funded by the BMBF in the framework of the Bernstein Center for Computational Neuroscience Göttingen project C4. Grant number 01GQ0432.

Personal Note

During the last years my perception changed fundamentally, not only in private where I learned (and still do) to be responsible for more than myself, but also my view of science and particularly basic research has changed due to various experiences. I hold psychologists in high regard not to surrender in the face of such ambiguous, barely reproducible and highly fluctuating data. I also gained a lot of respect for statistics. During my studies of mathematics I hardly could take a method that subjective seriously. Now I have realized that the mathematical branch of statistics is the most accurate way to extract information from the vague data nature provides us with. In numerous discussions on negative priming theories I learned a lot about how scientific objectivity is produced by the weighting of statements according to reputation of the originator or personal feasibility. All in all I am very happy with my project which not only taught me a lot and stayed interesting over the years, but also was a lot of fun.

Lebenslauf

Dipl. Phys, Dipl. Math. Hendrik Degering, geb. Schrobsdorff
Künstlername: Hecke Schrobsdorff
geboren am 29.07.1978 in Göttingen
Staatsangehörigkeit: deutsch

1985 - 1989	Grundschule *Mascheroder Holz*, Braunschweig
1989 - 1991	Orientierungsstufe *Lindenberg*, Braunschweig
1991 - 1998	Gymnasium *Raabeschule*, Braunschweig
1998	Abitur
1998 - 1999	Zivildienst im Verein *ganzheitlich Leben*, Prinz Höfte
1999 - 2005	Diplomstudium der Physik, Nebenfach Informatik
1999 - 2006	und Diplomstudium Mathematik, Nebenfach Informatik
	an der Georg-August-Universität Göttingen
2000 - 2006	Magisterstudium der Philosophie, Sport und Pädagogik, MA
2005	Diplom in Physik, Titel der Arbeit: *Localization of Neural Activity*
2006	Diplom in Mathematik, Titel der Arbeit: *Quantum Computing:*
	Anwendungen in Kryptographie und der Darstellung Boolescher Funktionen
seit 2005	Wissenschaftlicher Mitarbeiter der Georg-August-Universität Göttingen