Jörg Bahlmann: Neural Correlates of the Processing of Linear and Hierarchical Artificial Grammar Rules: Electrophysiological and Neuroimaging Studies. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2007 (MPI Series in Human Cognitive and Brain Sciences; 85)

Neural correlates of the processing of linear and hierarchical artificial grammar rules: Electrophysiological and neuroimaging studies

Von der Fakultät für Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
genehmigte
DISSERTATION
zur Erlangung des akademischen Grades
doctor rerum naturalium
Dr. rer. nat.,
vorgelegt
von
Diplom-Psychologe Jörg Bahlmann
geboren am 18. Juli 1975 in Frankfurt/Oder

Dekan: Prof. Dr. Annette Beck-Sickinger

Gutachter: Prof. Dr. Angela D. Friederici
Prof. Dr. Erich Schröger
Prof. Dr. Thomas F. Münte

Tag der Verteidigung: 07. Dezember 2006

Acknowledgement

The work on which this thesis is based was only possible because of the contribution of many people in the last three years.
I want to express my special thanks to the following supporters and contributors:

Prof. Angela D. Friederici for her knowledgeable supervision, as well as for providing the opportunity to conduct the experiments contained in this dissertation and for providing a lot of support and the scientific freedom that was necessary for this project.

PD Dr. Ricarda I. Schubotz, Dr. Thomas C. Gunter, Dr. Stefan Heim, Dr. Stefan Zysset, Dr. Jutta L. Müller, Korinna Eckstein, and Henning Holle for helpful comments and discussions along the way.

Cornelia Schmidt, Mandy Naumann, and Domenica Wilfiling for the help with data acquisition. Kerstin Flake and Andrea Gast-Sandmann provided many helpful advices for the production of graphical figures.

Caterina Remde and my parents for their continuous and unconditional support.

Contents

Introduction

It is an ongoing debate whether the processing of hierarchical structures and recursion is the crucial trait that makes human language specific (Hauser, Chomsky, & Fitch, 2002), or if it is one of several necessary aspects of language processing (Pinker & Jackendoff, 2005). To date, the conflicting perspectives about evolution of language can not be disentangled on the basis of empirical studies. Thus, the present work aims to contribute to this debate presenting experiments dealing with recursion on a more abstract level than natural language (i.e. artificial grammar). Moreover, additional evidence for the biological basis of language will be provided by specifying the neural correlates which subserve the processing of recursive structures.

During natural language processing syntactic, semantic, and phonological features usually interact. Artificial grammar processing in psychological experiments represents an efficient tool for the investigation of syntactic processing and rule-based structures in general, since other aspects of language can be held constant. Nonetheless it should be kept in mind that the reduction from a natural language to an artificial language may go hand in hand with a reduction of the external validity. In this respect, results of artificial grammar learning experiments have to be discussed in relation to different types of artificial grammars. Moreover, different methods (behavioral data, electrophysiological measurements, or functional imaging) on the one hand, and different types of samples (healthy participants or patient studies) on the other hand should be considered in order to identify the role of specific brain structures during the processing of artificial grammar tasks.

The present contribution is concerned with the neural basis of the processing of artificial grammar rules. It focuses on the neural generators involved in the computation of rules that varied in complexity, while many other cognitive aspects were held constant. In three different experiments the structures of the rules were systematically changed and electrophysiological and metabolic measurements were applied. The aim of this thesis is therefore to substantiate neural correlates of the processing of two different rule types.

In Chapter 1 the theoretical background will be given which covers an overview of recursion as a possible human specific trait, a description of the artificial grammar rules applied in this work, an outline of applications of these rules in psychological and linguistic research, and finally an introduction to a model of sentence comprehension. The ERP and fMRI methodologies are introduced in Chapter 2. Chapter 3 provides a summary of ERP and fMRI studies in the area of artificial grammar learning/processing and in the field of natural language processing. Chapter 4 specifies the main goals of the present work. Next, one ERP experiment (Chapter 5) and two fMRI experiments (Chapter 6 and 7) will be described. In Chapter 8 the results of the three experiments will be summarized and integratively discussed.

Chapter 1

Theoretical background

The present work is derived (inter alia) from a study by Fitch & Hauser (2004), comparing humans and non-human primates (tamarins) in two artificial grammar tasks. The behavioral experiment by Fitch & Hauser (2004) was anchored in the landmark paper by Hauser, Chomsky, & Fitch (2002). Who put forward some hypotheses about the nature and evolution of language and offered considerations about the interdisciplinary exploration of language evolution. In order to provide a background for the present work, an explanation of hypotheses provided by Hauser et al. (2002) will be given in the first section of this chapter. The controversial discussions about that topic (Fitch, Hauser, & Chomsky, 2005; Jackendoff & Pinker, 2005; Pinker & Jackendoff, 2005), and finally the experiment by Fitch & Hauser (2004) will be briefly described. In the second paragraph a formal description of syntactic rules used in the present work will be provided. Since the Chomsky Hierarchy is the theoretical base of the work by Fitch & Hauser (2004) and hence of the present studies the four types of grammatical rules will be introduced. Additionally, regular grammar and context-free grammars will be compared to each other. Third, behavioral studies which made use of artificial grammars of the regular or context-free type will be described. These studies differ in their theoretical foundations (implicit learning versus statistical learning) and in the type of rule used (regular grammar versus context-free grammar). Grammatical processing in artificial as well as in natural language involves maintenance and manipulation of stored information. Hence, the third paragraph describes a psychological model of sentence processing namely, Syntactic Prediction Locality Theory proposed by Gibson (1998), as this model illustrates the distribution of resources during sentence comprehension.

1.1. Recursion – A uniquely human feature?

The work by Hauser et al. (2002) provides an outline of the interdisciplinary study of the faculty of language. The authors suggested a differentiation between shared versus unique abilities, gradual versus saltational evolution, and continuity versus change of evolutionary functions. The first aspect distinguishes faculties that are uniquely to humans from abilities that are shared with other species. The second topic addresses the question whether the evolution of language was gradual versus saltational. The qualitative distinction among existing species into those which possess language and those which do not could have evolved gradually. Species that would fill the gap between faculties could have become extinct. The third issue deals with the question of whether certain aspects of language represent adaptations of other, previously adapted functions (e.g. spatial or numerical reasoning, tool-making, and navigation). Despite these general issues concerning language evolution, the authors provided a segregation of those aspects of language faculty that are uniquely human versus those abilities that are necessary for language processing but shared with other species. Hauser et al. (2002) labeled the former as faculty of language in a narrow sense (FLN) and the latter as faculty of language in a broad sense (FLB). At least two cognitive systems, namely sensory-motor system and conceptual-intentional system subserve the FLB. What is not included in the FLB is memory, respiration, or digestion. FLN, on the other hand, only consists of the cognitive ability to compute recursive structures[1], which is assumed to be uniquely human and could have evolved from other faculties such as communication, social behavior, or navigation. The authors suggested that "… FLN comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interface." p. 1573. The interface contains mechanisms of speech perception, speech production, conceptual knowledge, and intentions. While FLB is based on mechanism shared with other species and with other non-linguistic cognitive domains in humans, FLN is proposed to be a uniquely human aspect of language.

These hypotheses are highly discussed in the research community. Within the first quarter of 2005, Pinker & Jackendoff (2005), Fitch, Hauser, & Chomsky (2005), and Jackendoff & Pinker (2005) replayed arguments in favor and against these theories. The first article by Pinker & Jackendoff (2005) revolved around the definition of FLN and its implications. The authors argued, presuming only recursion as the core aspect of human language ignores features of grammar that are not recursive, but (in their point of view) human specific (e.g. phonology, morphology, or case). They pointed out that speech perception and speech production could also be part of FLN. Behavioral experiments in favor of these

---

[1] In mathematical terms, recursion refers to a procedure that calls itself. In computer programs a recursive sub-routine involves an embedded calculation that must resume to the original procedure from where it left off. Recursion requires a device with a stack of pointers to trace where to return after an embedded sub-routine has been executed. In the context of language recursion takes place if a constituent invokes an identical type of constituent, e.g. an embedded sentence like "The boy [that the girl saw] was tall.".

assumptions were provided and methodological critiques against studies that would falsify the hypothesis of language as a complex adaptation for communication were mentioned. Additionally, a close link between the definition of FLN and Chomsky's minimalist program (1995) was drawn. Chomsky's thoughts would implicate (among others) that language did not evolved as an adaptation for communication (which reflects the position of Pinker & Jackendoff), but rather for inner speech. The replay by Fitch, Hauser, & Chomsky (2005) endorsed their arguments, delivered theoretical considerations and experiments against the hypothesis that speech perception and speech production are uniquely human, and thus part of FLN, and denied that their concept of the evolution of language was exclusively based on the minimalist program. The return by Jackendoff & Pinker (2005) contained again the argumentation that FLN is based on the minimalist program. The authors additionally uttered concerns about certain definitions (e.g. definition of cognitive capacity or of trait), and suggested once more that recursion is not the only trait that makes human language specific.

Despite of that discussion another branch of theories concerning the evolution and the acquisition of language favors general models based on cognitive functions (e.g. Tomasello, 2006). These theories are also called usage-based linguistics due to the assumption that language structure emerges from language use. Tomasello stated: "Usage-based theories hold that the essence of language is its symbolic dimension, with grammar being derivative. The ability to communicate with conspecifics symbolically (conventionally, intersubjectively) is a species-specific biological adaptation." p. 258. Hence, one should keep in mind that the debate about FLB and FLN is only one aspect of language evolution that is of current interest in the research community.

Empirical data testing the hypothesis of recursion as the core aspect of human language was delivered by Fitch & Hauser (2004). Based on the distinction between FLB and FLN, the authors conducted a behavioral study with humans (college students) and non-human primates (cotton-top tamarins). Both, students and tamarins were exposed to two different types of artificial grammars. The authors raised the question whether humans and non-human primates differ in their ability to process linear and hierarchical syntactic structures. Both rule types were instantiated in two different grammars which were made of two categories of meaningless consonant-vowel syllable (A and B). The categories were coded by pitch (i.e. A = high pitch and B = low pitch). The linear structure was called finite-state grammar (FSG) and followed the rule $(AB)^n$. In this case, simple adjacent transitions between the two categories were generated (e.g. n = 2: ABAB). The hierarchical structure was called phrase-structure grammar (PSG) and followed the rule A^nB^n. This rule generated center-embedded strings of syllables (e.g. n=2: AABB). A familiarization-discrimination task was used to learn the grammar rules. As a result, humans easily learned both types of rules. In contrast, monkeys learned only the FSG, but not the PSG. The authors suggested that the acquisition of the ability to process hierarchical structures could be a possible critical stage in evolution of the human language faculty.

1.2. The Chomsky Hierarchy

Noam Chomsky developed the concept of a hierarchy of formal grammars (Chomsky, 1959). The author suggested that natural language processing is a paradigmatic example of infinite use of finite means. The Chomsky Hierarchy is a containment hierarchy of classes of formal grammars that generate formal languages.[2] The application and verification of formal grammars to natural languages is limited (e.g. Shieber, 1985). Nonetheless, still many psycholinguistic, psychophysiological, and even evolutionary (e.g. Fitch & Hauser, 2004) implications can be drawn from the Chomsky Hierarchy.

A formal grammar represents a mathematical description of rules underlying natural languages. It is characterized as a set of specifications of mechanisms that generate different types of structural regularities (Chomsky, 1957). Generative grammars are a subcategory of formal grammars. Each generative grammar consists of a set of rules for the transformation of strings. Usually, a generative grammar is characterized as a quad-tuple:

$G = (N, \Sigma, P, S)$

The generative grammar (G) comprises a finite set of nonterminal symbols (N), a finite set of terminal symbols (Σ, the letters of words), a finite set of production rules (P), and a start symbol (S). Consider for example the grammar G with N = [S], Σ, = [a, b], and P consisting of the following production rules:

1. $S \rightarrow aSb$
2. $S \rightarrow ab$

Examples of the generation of strings in the grammar are:

- $S \rightarrow$ (1)ab
- $S \rightarrow$ (2)a**ab**b
- $S \rightarrow$ (2)aa**ab**bb

---

[2] Note that the viewpoint of the author of syntax theory may have changed in the recent years (cf. The Minimalist Program, Chomsky, 1995).

This grammar defines the language [$a^n b^n$ | n > 0] where a^n denotes a string on *n* a's. Hence, the entire language consists of any positive number of a's, followed by the same number of b's. This grammar represents in the Chomsky Hierarchy a context-free grammar.

The Chomsky hierarchy classifies generative grammars into four types. Categories differ in the precision of production rules. The two most important types of rule systems are context-free grammars and regular grammars. The context-free grammar was called Type-2 grammar and also Phrase Structure Grammar (PSG). The regular grammar was labeled Type-3 grammar and also Finite State Grammar (FSG). Due to the focus of this dissertation on the processing of phrase structure versus finite state grammar rules, I will emphasize here only Type 2 and Type 3 grammars of the Chomsky Hierarchy.

The Type-0 grammar is the less powerful and unrestricted grammar. This rule system was equated with a Turing machine, a pushdown automaton simulating the logic of a computer.

Type-1 grammars generate context-sensitive languages. In difference to the Type-0 grammar, a Type-1 language is restricted with regards to the length of the input.

The Type-2 grammar (PSG) composes the theoretical basis of the syntax of most programming languages. The example mentioned above, the rule [$a^n b^n$ | n > 0], henceforth called $A^n B^n$, generates *n* sequential "A" terminal symbols, always followed by a single "B" symbol, and such pairs are repeated *n* times. This rule generates center-embeddings of two categories. Moreover, a hierarchical dependency between single members can be produced. The center-embedded structure of this grammar is, however, less common in human languages than other structures (e.g. right-branching or left-branching structures).

The Type-3 grammar (FSG) is the simplest formal model of a generative grammar. This type is also called regular grammar, since it generates a mechanism for the specification of a certain type of structural regularities in a linear way. Moreover, FSG generates simple transitions between a finite set of terminal symbols. Chomsky (1959 p. 150, definition 9) characterized FSG as follows:

The grammar G with N = [S, A, B], Σ, = [a, b], and P consisting of the following production rules:

1. $S \rightarrow aA$
2. $A \rightarrow aA$
3. $A \rightarrow bB$
4. $B \rightarrow bB$
5. $B \rightarrow \varepsilon$ (empty string)

Examples of the generation of strings in the G are:

- S → (1)aA → (2)a**aA** → (3)a**abB** → (4)aab**bB** → (4)aabb**bB** → (5)aabbb

This language comprises of the rule $[a^nb^m \mid m, n > 0]$. Any positive number of a's are followed by any positive number of b's, where numbers may be different.

Another type of Finite State Grammar can be described as follows:

The grammar G with N = [S], Σ, = [a, b], and P consisting of the following production rules:

1. S → abS
2. S → ab

Examples of the generation of strings in the G are:

- S → (1)abS → (2)ab**ab** → (1)abab**S** → (2)abab**ab**

This language was generated by the rule $(AB)^n$, which also produces a Finite State Grammar via linear transitions between two types of categories (A and B).

1.2.1. FSG and PSG in human language

One difference between FSG and PSG in the above mentioned example is the specification that the number of a's and the number of b's must be equal in the context-free language. Moreover, "… the extra power of phrase structure grammars over finite automata as language generators lies in the fact that phrase structure grammars may be self-embedding." (Chomsky, 1959, p. 154). Hence, PSG represents a more complex grammar in comparison to FSG. Furthermore, a FSG is insufficient in generation of all features of any human language (Chomsky, 1957; Miller, 1967). The description of a natural language minimally requires procedures at the next level of complexity (namely PSG). A FSG generates specific transitional probabilities between single adjacent elements. Sentences of natural languages, however, consist of intermediate phrases, or constituents. These constituents contain words, associated to specific word categories. These categories are not uniform across sentences. Instead, certain word categories are more highly independent than others. For example, in English the transitional probability between an article and a noun is very high. In contrast, a noun can be followed by a number of other word classes (e.g. adverbs, auxiliary verbs, main verbs). Hence, within a noun phrase a high conditional probability can be found. Whereas relations between different phrases are denoted by dependencies between constituents themselves. The unity of word classes within a phrase can be illustrated by the fact that a phrase can be replaced by a simple element (e.g. a noun phrase by a pronoun). Only units of words (phrases) can be moved, words from separate phrases cannot be moved together (Chomsky, 1965). Hence, Chomsky (1965) concluded, that human languages are simultaneously organized into linear and hierarchical structures.

1.3. Artificial grammar tasks in psychological and linguistic research

Why should one use an artificial grammar task in psychological research if the full-blown natural language is available? Like many complex systems, natural language is characterized by redundant and confounded features. Hence, it is often impossible to determine which of the linguistic properties are crucial to the processing of syntax. In artificial grammar tasks, those variables which are not to be investigated can be eliminated or held constant. An artificial grammar is semantically empty and morphological and phonological features can be held constant. The reduction of language to the underlying grammatical rules makes the artificial grammar task an efficient tool to investigate syntactic processing in general. Nonetheless, it should be kept in mind that the reduction from a natural language to an artificial language may go hand in hand with a reduction of external validity.

1.3.1. Psycholinguistic implementations: implicit and statistical learning

Several psycholinguistic studies applied a FSG or a PSG in behavioral experiments using an artificial grammar task. These studies differed in their theoretical motivations. Two major theoretical backgrounds can be distinguished namely implicit learning and statistical learning. While the former emphasizes syntax acquisition, lexicon formation is highlighted in the latter. More precisely, implicit learning deals with the phenomenon that many human abilities reflect some kind of adaptation to regularities. These adaptations concern for instance language, perception, motor skills, or social behavior. Behavioral adaptations evolve without intention to learn and without a clear awareness of what was learned. Statistical learning, on the other hand has been used to refer to the human ability to use distributional information in order to segment and organize a continuous stimulus stream. Hence, many statistical learning experiments tested the ability of infants to generalize distributional cues of rule-based structures, while during implicit learning experiments mostly adult participants were recruited. In implicit learning experiments different types of stimuli (e.g. digits, letters, symbols) and different modalities (e.g. visual, auditory) were studied with a specific focus on the question whether the acquired knowledge is consciously accessible by the participant. In contrast, statistical learning experiments often emphasize perceptual or distributional characteristics of the input in order to tackle the problem over which units statistics are computed (Bonatti, Pena, Nespor, & Mehler, 2006). Recently Perruchet & Pacton (2006) suggested certain similarities between statistical learning and implicit learning framework. The authors noted that both approaches use structured material without the instruction to learn and without the instruction to use any analytical processes or hypothesis-testing strategies. Furthermore, both types of design make use of exposure to positive instances during learning. In general, the authors suggested that statistical learning and implicit learning paradigms cover the same aim, namely the study of domain-independent learning acting on attended information in incidental, unsupervised learning situations. In the following I will review results of important implicit and statistical learning studies using an artificial grammar task with a specific focus on the type of the underlying rule which had to be acquired.

1.3.2. FSG rules

Most behavioral studies on implicit learning applied a FSG rule. The first implementation of a FSG in an experimental design was conducted by Reber (1967, see Figure 1.1). In this study participants were presented with a series of items consisting of 6-8 letters that followed a FSG rule. Participants should memorize the items. A control group learned randomly formed sequences in the same way. As a result the number of errors committed by participants in the concurrent grammaticality judgment task decreased successively from trial to trial and the decrease was significant stronger in the group that learned rule-based sequences than in the control group. In a second experiment participants were initially exposed to the same letter strings as in the first experiment and were instructed to memorize these sequences. Thereafter participants were notified that a rule underlies these sequences of letter strings, but they were not explicitly informed about the nature of the rule. Next, participants were presented to novel letter sequences which followed the finite state grammar rule or violated the rule (50%). Participants were instructed to judge the strings for grammaticality. An average of 69% correctly identified new sequences was interpreted as an indication of implicit learning of the finite state grammar. The results of this experiment were replicated several times by Reber and co-workers (see e.g. Reber, 1989 for an overview). The author suggested that participants were able to abstract automatically the artificial grammar rule of the displayed material. Moreover, it was proposed that participants could use the implicit knowledge about the structure independently of the specific stimuli. In other words, a transfer to superficially different, but structurally similar stimuli was implicated. Reber (1989) argued that implicit learning gives access to the deep, abstract structure of the stimuli in addition to the knowledge about the surface structure.

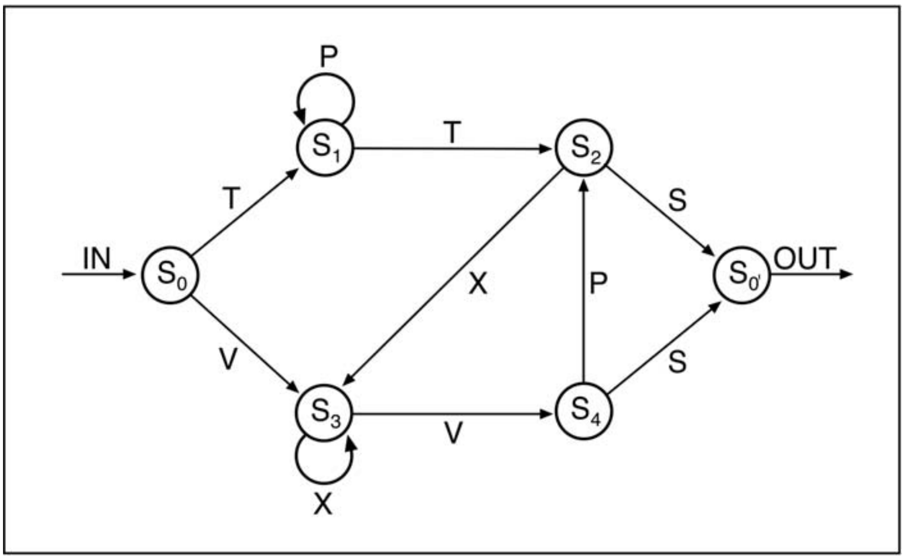

Figure 1.1: Schematic description of the artificial grammar used by Reber (1967)

The implicit nature of learning of artificial grammar rules and/or the type of rule abstraction was repeatedly challenged by other authors (e.g. Knowlton & Squire, 1996; Perruchet & Pacteau, 1990; Vokey & Brooks, 1992). Perruchet & Pacteau (1990) conducted a behavioral experiment in order to preclude the implicit abstraction of a complex rule during learning of rule-based letter strings. For this purpose the Reber grammar was adapted. Two groups of participants learned the artificial grammar from usual letter strings (e.g. M, R, T, V, X). The first group obtained an “incidental” instruction the second group attended to an “intentional” instruction. During learning, participants were exposed to letter strings following the artificial grammar rule. In the following test phase 50 new items (25 grammatical) were presented and participant should judge the grammaticality of letter strings. A third group learned bigrams made up the letters M, R, T, V, and X. The bigrams were derived from letter strings used in the two groups that were exposed to usual items. The frequency of occurrence was balanced over bigrams. Thereafter participants were told that half of the following test items contained new combinations and half of them stemmed from the same pool of letters. During the test phase the same grammatical and ungrammatical items were shown as in the other two groups. As a result, the percentage of correct judgment in the group exposed to the incidental instruction did not differ from the group listening to the intentional instruction.

Moreover, also the group that learned bigram stimuli scored significantly over chance. The value was lower than in the two other groups, though. The type of instruction did not influence the performance of participants. The authors argued that “the performance observed in typical artificial grammar-learning studies does not demonstrate participants’ purported ability to abstract a complex grammar.” (p. 268). Moreover, due to the performance over chance of the bigram group it was suggests that the explicit knowledge of bigrams assessed during a recognition task is sufficient for the above chance performance of participants in an artificial grammar task. Hence, the authors proposed that fragments of strings provide the most efficient coding of information underlying the sequences.

Some authors propose that participants first memorize the displayed strings and then perform the judgment of sequences on the basis of similarity between test items and learn items (e.g. Vokey & Brooks, 1992). On the other hand it was argued, that participants memorize the formation of small chunks (fragments) instead of memorizing the whole string of possible sequences. It would be more efficient to learn and recall the transitional probability within small chunks of items, instead of memorizing the complete sequence (Perruchet, Vinter, Pacteau, & Gallego, 2002). To summarize, the critical learning process assuming during an artificial grammar task was hypothized to reflect the learning of abstract rules (Reber, 1989), exemplar-specific learning (Vokey & Brooks, 1992), and learning of chunk information (Perruchet et al., 2002).

Also the implicit nature of artificial grammar learning was assessed by different authors. For instance, Knowlton & Squire (1994) demonstrated that amnesic patients with several impairments of declarative memory reveal the same performance during artificial grammar judgment as normal controls. Amnesic patients show deficits in storing information in declarative memory. On the other hand, implicit learning and nondeclarative memory are not impaired. The authors suggested that their results demonstrated that explicit knowledge about the structure of the grammar does not contribute to learning of the rule. Note, that another experiment of this work revealed that (similar to the study by Perruchet & Pactau (1990) mentioned above) normal participants could still use the explicit knowledge about training items when they where instructed to do so. A further interesting study by Knowlton & Squire (1996) suggested a dissociation of abstract rule acquisition and similarity based learning. In order to manipulate the similarity between test items and learning items the parameter chunk strength was systematically varied. Chunk strength is the frequency with which bigrams and trigrams in test items had already appeared in the training sets. High chunk strength correlates with a high similarity between test and training sequences and low chunk strength goes in line with a low similarity. A 2 x 2 design with the factors grammaticality (grammatical versus ungrammatical) and chunk strength (high versus low) was conducted. As a result chunk strength influenced the grammaticality judgments to a similar degree as did “adherence” to grammatical rules. It was concluded that learning of artificial grammar rules depends on both, abstract information and also exemplar specific information.

A subtype of implicit learning represents the serial reaction-time (SRT) task. This paradigm has become a preferred task for the investigation of sequence learning in recent years. In a SRT task, participants respond as quickly as possible to successively visually presented stimuli (e.g. letters or digits), by pressing corresponding response keys. Each response of participants initiates the presentation of the next stimulus on the screen. Learning of the underlying rule is typically associated with a significant stronger decrease of reaction times over time for rule-based sequences than for random orders (see Berry, 1994 for a review). Most SRT studies applied a linearly organized rule to the structure of the stimuli (e.g. Cleeremans & McClelland, 1991).

Statistical learning

The rules applied in the majority of statistical learning experiments were simpler than in implicit learning tasks. The aim of these experiments was mostly to assess the ability of infants to extract and generalize information from auditory strings following a simple artificial grammar rule (FSG). One way of looking into the learning mechanisms is to test infants in their ability to generalize the structure of syllable sequences.

In a study by Gomez & Gerken (1999) 12 months old children were exposed to auditory presented pseudo-words. These auditory strings followed a FSG. After 2 minutes of acquisition, new grammatical or ungrammatical items (50% each) were presented in a test session. The detection of ungrammatical items during the test session was assessed via the infant’s looking behavior. The time that each infant oriented to the loudspeaker on each string was recorded. Looking time to ungrammatical strings was longer than for grammatical items. The test session was accomplished in four different ways, leading to four separate experiments. Ungrammatical strings in the first experiment were produced while changing the fist and the last item of a correct string. In experiment 2 the first and the last item were taken from a correct sting, but the intermediate items were incorrect. In both tests, infants were able to discriminate grammatical from violated items. The same ungrammaticalities were applied in the third experiment. Additionally, a second type of artificial grammar rule was utilized using equal violations that produced similar behavioral data. In experiment 4 infants were able to discriminate new strings formed by their training rule from stings generated by a different rule despite of a change in vocabulary between training and test. The results were taken as a support for an extraction of an abstract form of information. The authors stated: "Although abstraction of such patterns does not reflect abstraction of the grammar per se, it does reflect generalization of a complex form (namely, abstraction of identity structure)." p. 131.

Marcus and co-workers (1999), on the other hand, proposed that infants are able to abstract the grammatical structure beyond the specific word input. The authors conducted a study with seven-month-old children. Three minute continuous speech samples were presented using strings with the structure ABA (e.g. wi-di-wi) or ABB (e.g. wi-di-di). The underlying structure was the same for training and for test sessions, only the expressions differed. As a result the children were able to discriminate strings from the training session from those with a different pattern. The findings were interpreted as a rule-based abstraction of the structure underlying the stimuli.

Comparable to the major question in implicit learning the problem of a rule-based or a pattern-based representation of the incidental statistical learning remains an open issue (see Gomez & Gerken, 2000 for an overview). While the research community in implicit learning seems to tend to a rule-based interpretation (e.g. P. J. Reber & Squire, 1999), statistical learning researcher appear to favor a pattern-based interpretation (e.g. Gomez & Gerken, 2000).

1.3.3. Hierarchical rules

Hierarchical dependencies between elements in a structure have been analyzed only by a small number of studies dealing with implicit or statistical learning.

Implicit learning

In a SRT task, Restle (1970) analyzed the type of chunking during the implicit learning of long sequences of digits. For instance the sequence (5 6 5 3 4 3 3 2 3 1 2 1) can be differentiated in unambiguous sub-systems via chunking of three elements (5 6 5) (3 4 3) (3 2 3) (1 2 1). In comparison the sequence (2 1 2 3 4 3 4 5 6 5) can not be separated unambiguously into triplet chunks. When participants chunk these sequences they cause fewer errors during learning of the pattern. The author proposed a structural tree that generates a simple hierarchy for such chunks. Restle (1970) even claimed that the difficulty of learning a transition within such a pattern depends on how high it is situated in the tree. However, the author only provided examples for a binary tree. Thirty years later, Koch and Hoffmann (2000) adopted the idea of hierarchical dependencies in a SRT task. The authors compared implicit learning of hierarchically structured items (e.g. ((1 2 3) (3 2 1)) ((4 5 6) (6 5 4)) ((1 2 3) (2 3 4)) ((3 4 5) (4 5 6))) with sequences that can also be chunked into triplets and comprised the same number of items but not in a hierarchical way (e.g. (1 2 3) (3 2 1) (4 5 6) (1 2 3) (2 3 4) (3 2 1) (4 5 6) (6 5 4). As a result reaction times diminished significantly more for hierarchically structured items in comparison to linearly chunked sequences during learning of the trials. Every chunk represents a minimal type of hierarchical structure, since a chunk represents multiple elements (at least two) from a low level on a higher level. The term hierarchy in implicit learning paradigms could be seen as a second-order chunking (Koch & Hoffmann, 2000).

Statistical learning

A number of statistical learning experiments analyzed whether adult participants can use predictive dependencies in order to facilitate the learning of an artificial grammar (e.g. Morgan & Newport, 1981; Saffran, 2001). The aim of these studies was to investigate what kind of statistical or other information is needed in order to derive dependencies between phrases in a sentence. Many artificial grammar studies applied a linear structure, which is one important aspect of language. But a linear structure does not by itself adequately describe the rule underlying natural language (Chomsky, 1957). As mentioned above, Chomsky (1957) proposed that human language is simultaneously organized into linear (FSG) and hierarchical (PSG) structures. In the experiments by Saffran (2001) and Morgan & Newport (1981) the linguistic structure learned from the statistical information should hence not be managed via local dependencies. It should rather be represented by a hierarchy between phrases in the artificial grammar task. Morgan & Newport (1981) constructed an artificial language with a minimal phrase structure including different word classes consisting of a set of pseudo-words. Word classes were ordered according to the following abstract rules:

S → AP + BP + (CP)
AP → A + (D)
BP → E or CP + F
CP → C + (D)

A sentence (S) consists of an A phrase (AP) and a B phrase (BP) and an optional C phrase (CP). An A phrase includes a pseudo-word of class A and optional a D-class word. The B phrase is coded by an E class word or a C phrase plus a word of class F. The C phrase comprises of a C class word and an optionally D class word. Participants were first exposed to a set of sequences of pseudo-words formed by artificial grammar rules. In addition a set of symbols were presented visually. Four types of relations between symbols and words were designed: no grouping between symbols and pseudo-words, arbitrary grouping, constituent grouping, and syntax incorporated grouping. Each class of pseudo-words referred to a specific class of symbols on the screen (despite of the last condition in which no reference of symbols and words was applied). Consequently, each pseudo-word was associated with a specific type of symbol within the symbol class. Arbitrary grouping refers to a visual adjacency between symbols that was not consistent with the phrase grouping implemented in the artificial grammar. In the constituent grouping condition symbols were grouped according the probability of occurrence of words, as employed in the artificial grammar. In a second step participants were presented with a new set of sequences. Participants were furthermore presented with a set in which some of the sequences contained a transformation of the phrases. As a result, all participants learned the distributional cues within the phrases, i.e. participants induced the finite state organization of the input. In contrast, only participants presented with the additional cuing information (constituent grouping condition) also learned the transformations of phrases, indicating a learning of the structure of the dependencies between the phrases. Morgan & Newport (1981) proposed that the additional cuing information provided in the constituent grouping condition was used by the participants in order to successfully induce a coherent grammatical system.

1.3.4. Non-adjacent dependencies

There is a growing amount of statistical learning literature dealing with infant learning of non-adjacent dependencies in the auditory modality (see e.g. Gomez & Gerken, 2000 for an overview). In this design the relation between elements in remote temporal distances is analyzed. For example, in a structure of the form AXC there would be a predictive relation between the elements A and C, whereas the intervening element X can freely vary. This type of design does not fit satisfactorily into the distinction between linear and hierarchical rules, since depending on the task one could interpret the rule as linear or as hierarchical. A linear rule was applied by Newport and Aslin (2004), since the distance between A and C was held constant (and the similarity between A and C was varied). This rule could be seen as linear, as no hierarchical structure needed to be build due to the constant distance of the elements.

1.4. The role of working memory in syntactic processing

The processing of a sentence involves the sequencing of its elements and the processing of their hierarchical relations. These elements are concatenated via numerous syntactical rules. A common method to investigate the mechanisms of syntactic processing is to analyze the parsing of sentences with different structural complexity. The analysis of the complexity of different sentences is based on several psycholinguistic theories centered around the model of working memory. The influence of working memory on parsing processing could be illustrated in the following example: “The man who said that a cat that a dog that a boy owns chased killed the cat is a liar.” (from Miller & Isard, 1964). A sentence could consist of an infinite number of center-embeddings without generating ungrammaticalities. However, this sentence is not acceptable, which was generally explained as a lack of working memory capacity of the parser required for storage and integration of the numerous embeddings.

The Syntactic Prediction Locality Theory (SPLT) proposed by Gibson (1998) is one of several models that describes the distribution of resources during sentence comprehension. These models are based on the relationship between the individual working memory capacity and sentence processing. For instance the Capacity Theory of Comprehension by (Just & Carpenter, 1992) proposes a single recourse for both the maintenance of linguistic elements and the symbolic computational processes that perform language comprehension. In contrast, Caplan & Waters (1999) questioned the single recourse approach and proposed two components of working memory during sentence comprehension. The authors distinguished interpretive from post interpretive processes. While the former process deals with the actual computations during sentence comprehension, the latter refers to all kinds of secondary processes that employ the meaning of the computed input.

In the present work the SPLT was chosen to be explained in more detail, since this model represents a more quantitative approach to sentence comprehension in comparison to the models by Just & Carpenter (1992) and Caplan & Waters (1999). Additionally, it attempts to describe a broad range of linguistic phenomena (i.e. complexity of non-canonical versus canonical sentences, unacceptability of multiply center-embedded dependencies, ambiguity effects, and cross-serial dependencies versus center-embedded structures). For the present study, the explanation of the processing of center-embedded structures given in this model is of particular interest (see Chapter 4).

The SPLT draws a distinction between integration costs and memory costs. Structural integration (integration cost) refers to the fact that the comprehension of sentences necessitates the integration of new words into the existing syntactic structure. During syntactic integration structural elements are attached in various ways (e.g. matching of syntactic categories, or linking of elements in long distance dependencies). The semantic and discourse integration realize the assignment of thematic roles. Gibson stated "It is assumed that each linguistic integration requires a fixed quantity of computational resources to perform the integration plus additional resources proportional to the distance between the elements being integrated" (p. 11). While the integration of long distance dependencies is assumed to be more expensive, the integration of local dependencies is seen as accomplished easier. Hence, the integration of long distance dependencies requires more resources than integration of adjacent dependencies.

Memory cost was explained as follows: "…there is a memory cost associated with remembering each category that is required to complete the current input string as a grammatical sentence." (p. 13). As for structural integration a sentence-comprehensive (discourse-based) locality function was proposed. The notion of locality function is based on the short-term memory theory (see e.g. Anderson, 1999 for an overview). It describes the case that stored information decays the more inferring items are to process. Consequently, Gibson (1998) proposed an increased computational effort correlated with the building of an intervening new discourse structure. The intervening structure (e.g. an interposed subordinate clause) makes it harder to keep the current syntactic predictions in mind. The more intervening structures are to be processed, the greater the decay of stored information in short-term memory, the higher the effort to parse the sentence.

In line with Just and Carpenter (1992), Gibson (1998) suggests one single unit in working memory for both the storage of linguistic information and the integration of the sentential elements. Hence, the working memory resources can be used for integration (computation) and maintenance of the linguistic input. Moreover, the author segregated the elements of a sentential structure in single memory units. The several memory units of a structure are additive and result in the total memory cost of a sentence. This procedure enables a quantitative examination of the complexity of sentences: the more complex a sentence the more memory units are to process.

The memory cost of a sentence via the quantitative computation of memory units per elements was applied to several phenomena of syntactic complexity. For instance the unacceptability of double-nested relative clause structures was explained with a limited capacity of working memory load in most participants, since to many memory units had to be maintained in memory. Moreover, a comparison of center-embedded and cross-serial dependencies in sentences was conducted. Consider the following sentences:

(a) Cross-serial (Dutch): NP_1 NP_2 NP_3 NP_4 V_1 V_2 V_3

Jeanine heeft de mannen Hans de paarden helpen leren voeren.
Joanna has the men Hans the horses helped teach feed.
Joanna helped the men teach Hans to feed the horses.

(b) Center-embedded (German): NP_1 NP_2 NP_3 NP_4 V_3 V_2 V_1

Johanna hat den Männern Hans die Pferde füttern lehren geholfen.
Joanna has the men Hans the horses feed teach helped.
Joanna helped the men teach Hans to feed the horses.

Cross-serial dependencies are easier to process than structures with center-embedded dependencies (Bach, Brown, & Marslen-Wilson, 1985). In German the verbs are nested, whereas in Dutch the verbs are arranged cross-serially. The complexity differences between center-embedded and cross-serial structures were explained by Gibson (1998) according to the associated memory costs. In the Dutch sentence, three memory units are required until the last (most embedded) noun phrase (NP_4). Since in Dutch the next element is the verb phrase corresponding directly to the first noun phrase (V_1), the memory cost decreases from this point on. In contrast, in the German embedded structure, the next element after the last noun phrase (NP_4) is the most embedded verb (V_3), equal to the least expensive current local prediction. For the next two elements the prediction cost increase, since the elements need to be retrieved from the memory that was stored earlier. Accordingly, the SPLT would predict a high complexity of center-embedded structures due to the distance between the incoming element and the dependent one to which it should be attached (integration cost), and due to the maintenance of the predicted category in working memory until the prediction is satisfied (memory cost).

1.5. Summary

In this chapter the controversial discussion about recursion as a possibly human-specific feature of language processing was briefly described. It is an ongoing debate whether the processing of hierarchical structures and recursion is the crucial trait that makes human language specific, or if it is one of several aspects of language processing. Although the experiments presented in this thesis won’t solve the debate they will add important information to the ongoing discussion by characterizing the brain areas which are involved in the processing of linear and hierarchical structures. Fitch & Hauser (2004) presented behavioral data, suggesting a differentiation between humans and non-human primates according to their ability to process recursive structures. While humans easily processed both, a Finite State Grammar rule (FSG) and a Phrase Structure Grammar rule (PSG), non-human primates were only able to master the FSG, not the PSG. In the second section of this chapter a classification system for syntactic rules was introduced. Syntactic computations were furthermore divided into linear (FSG) and hierarchical (PSG) rules. Both rule types are said to be present in human language. Next, applications of both rule types were inspected. It was shown that numerous behavioral studies employed the concept of the linear rule, hierarchical rules, on the other hand, were only considered in some studies. Implicit learning and statistical learning experiments utilizing a linear rule mostly addressed the question of the type of learning. Some authors propose an abstract rule learning, other suggest exemplar-specific learning, and a further group of authors imply learning of chunk information. Additionally, implicit versus explicit sequence learning is also a topic of an ongoing debate. The application of hierarchical rules in implicit learning tasks aimed to figure out whether participants use a second-level chunking (meta-chunking) during the processing of letter or digit sequences. Statistical learning experiments illustrated that predictive dependencies facilitate the learning of transformations between phrases in an artificial grammar task. Finally, the question of working memory recourses during sentence processing was addressed in the Syntactic Prediction Locality Theory, suggesting higher working memory demands during the processing of embedded hierarchical structures.

Based on the findings by Fitch & Hauser (2004) the present study aims to investigate neural substrates of the processing of two different grammar types, namely the FSG and the PSG.

Chapter 2

Methods of Investigation

This chapter will give an overview of the methods used in the present experiments. In the first part of this chapter a brief summary of principles behind the Electroencephalogram (EEG) will be given. For more detailed descriptions of this method the introductory textbooks by Rugg & Coles (1995) and by Handy (2005) are recommended. The second part of the chapter describes the functional Magnetic Resonance Imaging (fMRI) relatively more comprehensive than the EEG method, since the main focus of this thesis is located on fMRI experiments. The present description of the fMRI method is based on the introductory textbooks of Moonen & Bandettini (2000), Buxton (2002), Toga & Mazziotta (2002), Jezzard, Matthews, & Smith (2002), and Frackowiak et al. (2003). The description of the fMRI method comprises characteristics of magnetic resonance signal, an outline of the blood oxygen level dependency (BOLD) contrast, and a delineation of the analysis of fMRI-data. In the third part of this chapter a short introduction to Diffusion Tensor Imaging (TDI) technique will be given, also based on chapters of the textbooks mentioned above.

2.1. Electroencephalogram

2.1.1. Event-Related Brain Potentials

The electroencephalogram consists of continuous waveforms measuring difference in potentials between scalp electrodes. The EEG reflects electrical activity of the cortex plus surrounding noise. Reference electrodes are usually placed at the nose, the mastoid behind the ears, or can be calculated from the average of scalp electrodes (e.g. Regan, 1989). The EEG signal is amplified and digitized by sampling the continuous signal, which usually consists of a sampling rate of 250 Hz (as used in the present study), yielding to a measurement of the signal at every 4 ms. The structural organization of the human cortex allows the generation of large extracellular field potentials. The pyramidal cells in cortical layers III, IV, and V (mainly the large somas in layer V) that are vertically aligned to the cortex surface largely reflect neural activity picked up in EEG. The dipolar field needs to pass through different types of tissues, including three meninges, cerebrospinal fluid, and scalp, leading to a distance of a few centimeters between the produced dipolar field and electrodes. Hence, activity of a sufficient number of pyramidal cells (at least 10 000) must summate in order to produce a measurable electric field potential on the scalp surface (for more detailed information see for example Nunez, 1981).

Brain activity and artefacts are related to the EEG signal. Potential changes correlated with cognitive events reveal much smaller amplitudes (around ± 5 µV) than background activity (e.g. Rugg & Coles, 1995). Thus, in spontaneous EEG single events can hardly be identified and no implication of a particular brain response to an experimental manipulation can be made (yet). Averaging of the EEG, time-locked to the onset of repeated events can extract the characteristic pattern of electrical activity to an event from the uncorrelated background noise. Hence, event-related brain potentials (ERPs) lead to an increase of the signal-to-noise ratio (SNR) of electric activity measured in the EEG.

2.1.2. Components of the ERP

ERPs are characterized by their peak latency, polarity, and distribution. Latency refers to the time between stimulus onset and peak of the averaged signal. Since the temporal resolution of ERPs are on the order of milliseconds, this method is useful for the investigation of temporal dynamics of cognitive processes. Polarity is related to the direction of deflection (positive or negative). A negative deflection was argued to be correlated with excitatory postsynaptic potentials and positive deflection might be caused by inhibitory postsynaptic activity (Birbaumer, Elbert, Canavan, & Rockstroh, 1990). Exogenous and endogenous ERP components have been distinguished. While exogenous components refer to early sensory stimulus properties that arise before about 80 ms, endogenous components were elicited later (about 200 ms after stimulus onset). Endogenous components refer to 'higher' cognitive functions such as allocation of attention, task relevance of the stimuli, or processing strategy of the participant. In addition to the latency and the polarity of the signal, the distribution represents another attribute of the ERP components. However, it is important to keep in mind that the implications from the location of the observed deflection stand for a rather week tool for the understanding of the underlying cognitive processes. The poor spatial resolution of the ERPs results (among others) from the distance between the recording electrodes and the neural generators of the signal. By convention, ERPs are named according to the three components, for instance 'P600' refers to a positive deflection of the signal (P), with its peak at about 600 ms after stimulus onset, and having a broad parietal distribution.

2.2. Functional Magnetic Resonance Imaging

The foundation of the Magnetic Resonance Imaging (MRI) was disclosed by Bloch (1946) and Purcell (1946). The first MR-Image was produced by Lauterbur (1973) and since 1984 the MRI technique is used in clinical practice. The coupling of the blood oxygen level dependency (BOLD, Ogawa, Lee, Nayak, & Glynn, 1990) with the MRI method made it possible to use functional Magnetic Resonance Imaging (fMRI) in the brain. The analysis of the fMRI data is conducted in two steps, namely the preprosessing and the parametric estimation using the general linear model (GLM, Friston, Holmes, Worsley et al., 1995).

2.2.1. The Characteristics of the Magnetic Resonance Signal

Basic physical principals

The rotation of the atomic nucleus is a law of nature. It is part of every elementary particle. This spin is an intrinsic feature of the protons and neutrons of the atom. The magnitude can not be increased or decreased; only the axis of the spin can be changed. The rotation induces a magnetic dipole moment, which is specific to every type of nucleus. The hydrogen nucleus is far more abundant in the body than any other nucleus and features positive nucleus-specific characteristics (net spin). Hence, the hydrogen nucleus is used in the measurement of magnetic resonance imaging (MRI). Atom-nuclei are not only moving free in space (dependent on the state of aggregation) they also rotate around their axis (i.e. precession). The frequency of the precession is defined by the Larmor relation: $\omega_0 = \gamma B_0$ (ω_0 = Larmor frequency, γ = geometric relation, B_0 = magnetic field strength). The components of the magnetic field are sub-divided by the three axes in space (x, y, and z). The magnetization component of the z-axis is the longitudinal magnetization; on the xy-plane it is called transversal magnetization. In normal case, the net spin in the spin ensemble is zero, which means that all magnetic components were compensated (thermodynamic balance). Putting the spin ensemble into a strong magnetic field (B_0) induces a thermodynamic imbalance, the net spin becomes greater than zero, and the nuclei rotate in parallel to the B_0 vector. The radiation of an electromagnetic pulse (B_1) orthogonally into the magnetic field (B_0) causes a tilting of the rotation vector of the nuclei, an energy consumption of the spin ensemble, and a temporary phase coherence emerges. Phase coherence is given when all nuclei rotate not only in parallel but also with equal phase. Thereby, in the xy – plane (transversal magnetization) a voltage is induced, which is measurable by a receiver. The receiver in the MRI is called detector coil. The relaxation time reflects the transition from the imbalance state to the balance state.

Relaxation of the signal

Three types of relaxation times are to be differentiated. The T_1-relaxation reflects the decomposition of the longitudinal magnetization and is called spin-lattice relaxation. The energetic interaction consists of the activated spin ensemble and the spin lattice (tissue). The characteristic function illustrates the reestablishing of the balance state. The T_1-relaxation is about eight to ten times larger than the T_2-relaxation. The energetic interaction of the T_2-relaxation exists within the activated spin ensemble (spin-spin interaction). The characteristic function describes the gradient of the transversal relaxation and hence, the process of dephasing. The T_2^*-relaxation describes the decay of the longitudinal magnetization via the spin-spin interaction, the additional inhomogeneity of the magnetic field, and the different susceptibility of diverse tissues. The decay of the T_2^*-relaxation is much faster than the decay of the T_2-relaxation. The brain tissue differs according to the number of hydrogen nuclei. The MR-signal of the three relaxation processes differs depending on the brain tissue and the time point of measurement.

Free Induction Decay

The basic principle of the MR measurement is the Free Induction Decay (FID). A 90° radio frequency (RF) pulse (B_1) tips the dephased, but parallel spinning atoms from the longitudinal magnetization (along the B_0 magnetic field) into the transverse plane (perpendicularly to the B_0). The signal decays away, because the precessing components of the magnetization itself decay away. The decay follows an exponentially function and the coil measures the oscillating signal. A Gradient Echo (GE) Puls Sequence is a FID, it depends on the repetition time (TR) and the flip angle (α). A Spin Echo (SE) Puls Sequence is based on the phenomena that when a second RF pulse is applied after the first one, an echo (spin echo) of the original FID is created. By convention, the time to the spine echo is called TE. The radiofrequency coil in a MRI scanner is used for both, transmission and receiving of the signal. The RF coil used to detect the MR signal is sensitive to a large volume of tissue. In order to measure the signal of the several different small parts of the 3D tissue, the concept of magnetic field gradients and gradient echoes are applied.

Field encoding of the signal

The mapping of the MR-signal is realized by applying different magnetic field gradients. In an MR scanner there are three gradient coils in addition to the RF coils and the B_0 coil.

The gradient echo produces a magnetic field (B_0) that varies linearly along each axis. The combination of the three gradient coils produces field gradients along the three dimensions. Hence, for each voxel a particular field gradient can be applied. The localization of the MR signal is done with slice selection (z-axis), frequency encoding (x-axis), and phase encoding (y-axis). Therefore, the resolution of the fMRI image in the present experiment depends on the number and the thickness of slices, and the repetition time.

2.2.2. The BOLD contrast

All organs of the body need to metabolize energy in order to keep the basic cellular processes running. Glucose and oxygen are taken from the blood in order to produce adenosine triphosphate (ATP), which is the common free energy of the biological system. The waste products of the energy metabolism carbon dioxide and water return into the blood. The production of ATP in the cells of the body also happens in the neurons during functional activity. The basic dissociation between electrical and chemical information transfer segregates the axonal action potential from the synaptic processes. Energy metabolism is higher during synaptical processing than during the saltatoric information transfer along the myelinated axons of the neuron. It was hypothesized that the sending of an action potential is less costly due to the efficient propagation along the myelinated fibers (Buxton, 2002). Axonal connections between the brain areas are mostly situated in the white matter. In the gray matter, information transfer mostly occurs via chemical transmitters, moving from the pre-synapse to the post-synapse. Hence, the metabolism of ATP energy is three to four times higher in gray matter in comparison to white matter. Since most synaptic connections are situated in layer IV of the cortex, glucose metabolic rate is highest in this layer. The neuro-transmitters are produced in Astrocytes surrounding the synapses. The pre-synapse releases the transmitter and the post-synapse detects the transmitter. Thereafter, the neuro-transmitter is removed from the synaptic gap and the adjacent Astrocytes uptake them again. Glucose is taken from the blood in order to produce ATP-energy that produces the transmitter (Gusnard & Raichle, 2001). Animal studies demonstrated an association between glucose metabolism and neural functional activity (e.g. Schwartz et al., 1979).

Cerebral blood flow (CBF), cerebral metabolic rate of oxygen ($CMRO_2$) and cerebral metabolic rate for glucose (CMRGlc) increase in activated brain areas (Ginsberg, Dietrich, & Busto, 1987). Oxygen metabolism increases much less than CBF and glucose metabolism. The ratio of oxygen consumption to glucose consumption at rest is usually at a rate of about 5:1 (Frackowiak, Lenzi, Jones, & Heather, 1980). The underlying reasons for this imbalance are still unclear. The BOLD-effect results from this imbalance. If CBF increases more than oxygen metabolism, less oxygen is removed from the blood and the venous blood oxygenation must increase (Ogawa et al., 1990). Deoxyhemoglobin (dHb) is paramagnetic and oxyhemoglobin (Hb) is diamagnetic. Hence, the MR-signal is sensitive to changes of dHb. About 40% of the oxygen is extracted from the blood in resting state in the normal brain (Marchal et al., 1992).

The oxygen limitation model

As mentioned above, the mechanisms underlying the imbalance of CBF and CMO_2 are still under debate. Buxton (1997) assumed a limited oxygen delivery underlying this phenomena (oxygen limitation model, balloon model). The author hypothesized that, due to a limited oxygen delivery during rest, a tight coupling of CBF and oxygen metabolism occurs at this stage. It was proposed that not all of the oxygen delivered to the capillary bed become available for metabolism. Hence, oxygen delivery to the tissue is “barrier-limited” and extraction depends on the time in which the oxygen remains in the capillary bed. Moreover, velocity of blood in the capillary bed is responsible for CBF increase. A CBF increase leads to a decreased capillary transit time and a reduced extraction of oxygen. The increase of CBF decreases the fraction of oxygen delivered to the capillary bed. Hence, blood flow must be increased substantially to produce a small increase of the fraction. Taking together, the oxygen limitation model suggests a coupling of CBF and oxygen metabolism: a large blood flow change is necessary to support a small increase in oxygen metabolism. Hoge (1999) found a linear relationship between $CMRO_2$ and CBF during activation. The ratio of this dependency is at a level of 2:1 (CBF: $CMRO_2$). Moreover, Uludag et al. (2004) also showed this relationship for deactivation.

Correlation between oxygen and MR-signal

Despite of the association between CBF and oxygen metabolism, the change in glucose metabolic rate is relatively independent of CBF changes. Hence, regulation of blood flow during activation is not manipulated by the need for glucose delivery. This is supported by the fact that oxidative metabolism is more efficient at generating ATP energy. Furthermore, most of the additional ATP production during activation occurs due to metabolism of oxygen (75%).

The BOLD effect is based on synaptic activity leading to a higher consumption of ATP energy. This local neural activity induces changes in metabolism of glucose and oxygen and vascular changes (including increase of CBF, cerebral blood volume, CBV, and blood velocity). Logothetis (2001) found a direct link between the BOLD-contrast and neural activity. This linear relationship is the basement on the correlation between the MR-signal in fMRI and experimental manipulations. Changes in metabolism and vascular changes produce a decrease of oxygen extraction from the capilar bed. $CMRO_2$ increases much less than CBF. The result of this imbalance in changes of CBF and $CMRO_2$ is a substantial drop in oxygen extraction fraction and a corresponding increase of oxygen in the capilar bed. The enhancement of oxygen in the capilar bed causes a drop in dHb content of venous blood. The decrease of paramagnetic dHb causes a drop in inhomogeneity of the magnetic field and therefore, produces BOLD change in T_2^*-weighted MR signal. The fact that oxygenation level of blood has measurable effect on the MR signal from the surrounding tissue was discovered by Ogawa (1990) imaging rats at 7 Tesla. The human brain consumes about 20% of the oxygen delivered via breathing. The portion of hemoglobin on total volume of the brain is very low. Moreover, change of the hemoglobin value is also very low. Though the GE weighted T_2^*-effect of the MR-signal during activation in a particular brain region increases at a level of about 1 to 10%.

The BOLD response is characterized by a hemodynamic delay. I.e. the hemodynamic response function increases about 2 seconds after onset of the stimulus. Time to peak can vary from 1 second for very short stimuli, up to 8 seconds for a continuous robust stimulus (Jezzard et al., 2002, p. 153). The duration of BOLD varies as a function of stimulus length. Moreover, a saturation effect occurs for stimuli larger than 4 seconds, indicated by a stable increased hrf-function that can last over as far as 20 minutes (Bandettini et al., 1997). After stimulation the signal increases significantly more than at the beginning (post stimulus undershoot). It was speculated that the undershoot arises from either an enhanced $CMRO_2$ even if CBF returned to baseline, requiring an increased oxygen extraction fraction (Frahm, Kruger, Merboldt, & Kleinschmidt, 1996), or CBV remains elevated while CBF increases again (Buxton, Wong, & Frank, 1998). Some optical imaging studies reported an initial dip (e.g. Malonek & Grinvald, 1996), while other fMRI studies could not replicate these findings (e.g. Marota et al., 1999).

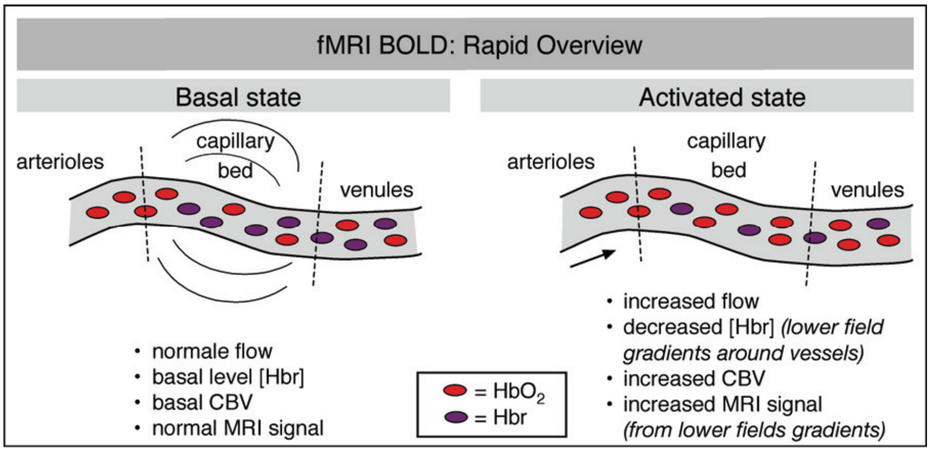

Figure 2.1: Basic principles of the BOLD-response (adapted from www.fmrib.ox.ac.uk/)

Factors manipulating the MR-signal of the BOLD contrast

The MR-signal of GE-weighted images is primarily sensitive to the veins. The GE-BOLD technique shows predominantly signal contrast from capillary bed and in and around venous vessels, independent of velocity. Whereby SE-BOLD images are sensitive to intravascular vein vessels and also to the capilar bed independent of velocity (hence, GE images cover a bigger spectrum of the blood volume).

The BOLD response depends on several factors, including magnetic field strength, TE, TR, α, voxel size, and, inter stimulus interval (ISI) and the measurement technique. The strength of the magnetic field is directly connected with BOLD. The higher B_0 the larger magnetization within the brain and the field gradients, due to the increase of magnetic susceptibility differences. The BOLD signal furthermore correlates with TE, i.e. the shorter TE the higher the signal. The optimal TE is about equal to T_2^*-relaxation time. The typical T_2^*-value in a 3T magnet is at about 40 – 60 ms; hence an echo time of this range should be used. The voxel resolution also affects the SNR of the BOLD. Nevertheless, the optimal voxel size is difficult to predict, since sizes of functional regions vary over space and across experimental conditions. A matching of voxel resolution with cortical thickness (about 3mm) was suggested (Bandettini, 2001). The SNR of the BOLD is indirect proportional to TR and to α. A short TR and a relative small α produces a better SNR (see also Buxton, 2002, p. 436).

Temporal and spatial resolution of the BOLD-contrast

Temporal resolution of the BOLD-response is clearly weaker than temporal resolution of EEG measurement. Yet, it was demonstrated that the BOLD-signal in human visual cortex could be detected using isolated stimuli as short as 34 ms, even thought the duration of the signal lasts over several seconds (Rosen, Buckner, & Dale, 1998). Pollmann, Wiggins, Norris, von Cramon, & Schubert (1998) reported a linear relationship between different inter stimulus intervas (ISI's, between 15 and 4 seconds) and the amplitude of the hemodynamic response. Also Dale & Buckner (1997) found support for a linear relationship between the amplitude of BOLD and ISI. For ISI's < 1 second a non-linear relationship was observed (Friston, Josephs, Rees, & Turner, 1998). In cognitive experiments a realistic ISI should be greater than 2 seconds (see Mandeville & Rosen, 2002 p. 340), since signal detection to multiple stimuli can not be resolved with a lower ISI and vascular heterogeneity leads to different regional distributions in the brain (e.g. larger vessels lead to larger signal change). Randomization of ISI (jittering) significantly improves the power of signal detection (Burock, Buckner, Woldorff, Rosen, & Dale, 1998). Jittering increases temporal resolution and statistical power, since the BOLD of each trial is measured at a number of different time points (Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000).

2.2.3. Analysis of fMRI-data

The fMRI data set consists of a huge number of voxels with different gray values measured at different time points during scanning. These time series of voxels are part of a 3D-matrix of signal intensity that is linked to particular experimental manipulations. The aim of the analysis of these data sets is the detection of brain regions that show a specific signal distribution. The signal comprises not only of BOLD functions correlated with experimental variables but also of artefacts originated from different sources. Signal changes associated with experimental manipulations represent only a few percent of the mean value of the signal, most of the signal consists of noise. Noise contains of systematical and stochastical errors. Artefacts arise from baseline drift, image deformation, ghost, regional loss of sensitivity due to susceptibility differences, body movements of participants and physiological movements (pulse, breathing, swallowing). In order to reduce the influence of artefacts on parametric estimation of the signal a preprocessing of functional data is applied.

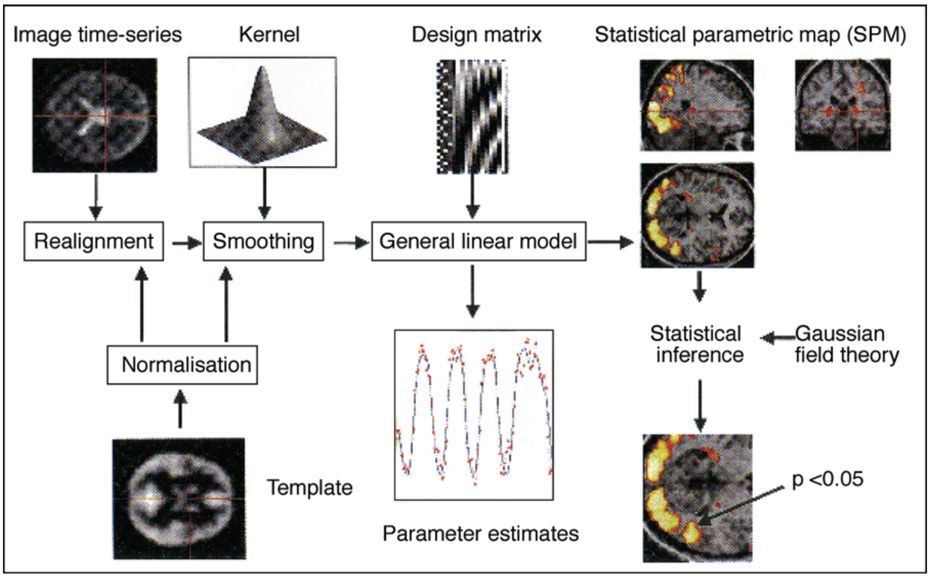

Figure 2.2: Schematic description of the transformation steps from preprocessing to statistical parametric map (SPM), from Frackowiak et al. (2003)

Preprocessing

In the two fMRI experiments of this work the following data preprocessing methods are used: movement correction, slice time correction, temporal filtering, and spatial filtering.

Movement correction

While analyzing time series of signal intensity, it is assumed that an identical region of the brain is measured at every time point of the experiment. Movements of the head render this assumption incorrect, since a displacement of the voxel from the original location takes place. Hence, if the head movement correlates with the experimental manipulation, signal associated with movement can not be distinguished from signal associated with the experiment. On the other hand, movement of the head incoherent with the experimental manipulation represents additional noise, decreasing the SNR.

In the present two fMRI experiments offline motion correction is applied using Siemens motion correction protocol (Siemens, Erlangen, Germany). The mean signal intensity difference between two successive images is calculated using the first image as a template. If participants move signal intensity varies to a grater extend. Translational (linear to the x, y, and z axes) and rotational (rotation among the x, y, and z axes) movements of the head are taken into account during calculation. The resulting values are used to compute estimated displacements of each voxel related to the template image. This procedure allows removing the intensity changes related to movement from time series.

Slice time correction

One image of an fMRI time series contains several 2D slices measured within a particular time. In the present two experiments 16 slices are measured within a TR of 2 seconds. Time points of scans of these 16 slices slightly differ within one image (or volume). Though, the later statistical analysis assumes that all slices of a volume are measured at the same time. Statistical analysis will be improved if each voxel's time series is adjusted in order to pretend that all voxels of a volume are scanned at the same time. The slice time correction procedure interpolates time series of the voxels so that it appears as if all voxels are scanned at the same time.

Temporal filtering

The expression baseline correction comprises several operations for reduction of signal drift, including temporal filtering or global scaling. Temporal filtering is applied in order to remove components of time series that are not correlated with signal of interest. Two types of filtering can be applied, namely a high-pass filter and a low-pass filter. A high-pass filter efforts to remove slow drifts from the signal. This slowly varying signal could arise from physiological effects like heart beat or breathing, or scanner-related drifts. If these slow drifts are not filtered out from the signal, SNR decreases and the residual of the model fit in the statistical analysis increases. Hence, in the present two fMRI experiments high-pass filters are applied. A low-pass filter reduces high frequency noise of time series. The low-pass filter works like a Gaussian kernel that smoothes the signal. The danger of the use of a low-pass filter is that the underlying stimulus related to signal can be corrupted. Hence, in the present two studies no high-pass filter is applied.

Spatial filtering

Signal intensity in a voxel differs, even if no stimulation is applied. Commonly, signal change induced by stimulation ranges between 0.5 and 5 percent and noise level is between 0.5 and 1 percent. The aim of spatial filtering is the reduction of noise, whilst the underlying signal remains equal. Moreover, it is required to smooth the data in order to reduce the effect of false positive activation in latter statistical analysis. Error distribution is rendered more normal after smoothing and validity of the statistical analysis will increase. Smoothing is also necessary to facilitate the averaging of fMRI data of several participants. It is important that the smoothing kernal is not bigger than the expected activated area. In fMRI experiments the common filter ranges between 3 and 10 mm full-width-half-maximum (FWHM). The expression FWHM refers to the 68 percent under the Gaussian curve of the smoothed area, i.e. the range between lower inflection point and upper inflection point of the Gaussian function of the smoothing kernel. In the present two fMRI experiments two one dimensional smoothing kernels are applied, resulting in an in-plain blurring of the data.

Registration and Normalization

Registration

The concept of registration comprises transformation of an image in order to fit it maximally to another volume. A spatial transformation changes position or orientation or shape of coordinates of the image. In the SPM software functional images are registered to a high resolution anatomical image. In LIPSIA, which is used in the present studies, registration is applied at anatomical 2D slices, measured immediately before functional volumes. This could be critical, since the potential movement of the participants head between anatomical and functional scans is not taken into account.

Transformation of the MR-image can be calculated using linear or non-linear algorithms. Non-linear transformation is described as a warping algorithm. This type of transformation is useful for high resolution images, since a huge degree of freedom (DOF) can be applied. However, only few implementations of these geometric models are available and computational demands are very high. Linear transformation is differentiated in rigid body (6 DOF, linear translation and linear rotation), similarity (7 DOF, translation, rotation, and single global scaling), and affine (12 DOF, translation, rotation, scaling, and skewing).

In LIPSIA a rigid body transformation is implemented. During the first step of registration a transformation matrix is computed between the 2D anatomical slices and the 3D high resolution data set. Factors of translation and rotation for 2D coordinates are computed and converted into the transformation matrix. The second step contains application of this matrix to preprocessed functional data.

Normalization

In order to specify locations of particular brain areas across participants a reference framework is necessary. At present (at least) two common brain templates are used, namely Talairach and Tournoux Atlas (1988) and Montreal Neurological Institute (MNI) Atlas. Templates comprises of a standard coordinate system. The Talairach and Tournoux

Atlas uses eight different landmarks assigned to the coordinate system. Among others, points on the brain surface along anterior commissure and posterior commissure (AC-PC) serve as a landmark. The Talairach template is based on a post mortem brain of a 60 years old female subject having an alcoholic problem in her lifetime. This shall be deemed as a disadvantage of the Talairach & Tournoux Atlas. In contrast, the MNI atlas delivers a brain that is more representative to the population (assuming that a brain deformed by alcohol abuse is not characteristic). For example SPM99 uses an average of 152 brain scans of the MNI template. An average of 305 brain scans is also available. MNI brains are slightly larger than the Talairach brain.

Normalization is calculated in the present studies using linear and non-linear scaling, implemented in LIPSIA.

Statistical analysis of activation

Strategies of the analysis

The most common design types in fMRI experiments are block design and event-related design. The block design was the first method that was applied in fMRI studies. Several trials of one condition are repeated in one block. SNR increases with every new recording that enters computation of the average. It is assumed that all trials activate the same brain region. In order to detect the brain region that contributes to specific cognitive process, a comparison of different experimental conditions is necessary. This can be implemented while subtracting a baseline condition from an experimental condition or by subtracting two different conditions. According to the choice of baseline task, Binder et al. (1999) suggested that a simple resting baseline should be avoided. During rest, one can not control the cognitive process in which a participant is engaged, since it is simply impossible not to think. Unspecific noise related to resting baseline increases error variance artificially and makes a resting baseline less efficiency. Nonetheless, Gusnard & Raichle (2001) proposed that brain activity during rest does not vary unpredictably. BOLD during rest is rather intrinsically constrained by default functionality during baseline. Default functionality varies between brain regions.

The block design delivers relatively robust results for global effects. Many experiments, however, need to detect stimuli in shorter time. For instance in oddball paradigms novel stimuli need to be distinguished from a sequence of standards or in memory tasks one may want to calculate contrasts between remembered and forgotten items. A block design does not serve this purpose well. Due to the development of scanners with higher field strengths and following electrophysiological studies the event-related design was introduced to neuroimaging research (Josephs, Turner, & Friston, 1997). In event-related fMRI the BOLD response is time locked at every trial (event). This enables the presentation of different trial types in a randomized order, preventing effects of expectation and habituation. Moreover, behavioral date (reaction times, error rates) can be correlated with brain activity. The introduction of null-events can improve SNR in event-related designs (Friston, Zarahn, Josephs, Henson, & Dale, 1999). Null-events could be seen as an equivalent to resting baseline in block designs (without disadvantages of the latter). Null-events increase SNR while preventing BOLD from a potential ceiling effect induced by frequent trial presentation.

Statistical Parametric Mapping

The analysis of fMRI data is usually calculated in a mass-univarate framework (Friston, Holmes, Worsley et al., 1995). At first glance this approach seems to be rather contra-intuitive. Since the dependent variable comprises of several thousands of voxels, one would implicate a multivariate analysis. Conversely, the multivariate approach requires more observations than cells of factors (i.e. number of voxels), which is not realistic in common psychological experiments using fMRI, as the number of repetitions would be too large. Moreover, in order to estimate the error covariance during the multivariate analysis, much more parameters need to be calculated (e.g. error covariance between all pairs of voxels) than during the univarate analysis. Hence, the multivariate analysis lacks of power in comparison to the univariate approach.

Statistical Parametric Maps (SPM's) are matrices of voxel-values that are distributed according to Students t or F probability functions. Values of the SPM rose under the null hypothesis that no difference exist between experimental conditions. The time series of each voxel is analyzed separately (univariate) and are taken together into one SPM. The analysis of time series of every voxel is accomplished using the General Linear Model (GLM):

$$Y = X\beta + \varepsilon$$

This equation expresses all observed response variables included in matrix Y equals the explanatory variable comprised in matrix X combined with a specific parametric weight β plus an error term ε. Equations of all voxels of a data set are comprised in the design matrix (X). In the present studies the design matrix is modeled using a box-car function convolved with a hemodynamic response function. This function consist of two gamma-functions which models the post stimulus undershoot more accurate than a single gamma-function (Friston, Fletcher et al., 1998). The GLM estimates the optimal fit between the hemodynamic response function of the design matrix with the real (preprocessed) time series of the fMRI experiment. In the present studies the least-square estimation approach is used. This procedure calculated β-values and random error terms for each participant to a specific experimental manipulation. The random error explains how the observed BOLD response to a particular treatment varies even if the experiment is repeated on the same participant under the same conditions. But this is not the only error produced by the MR-signal. In one experiment several trials related to the explanatory variable are usually repeated. For this reason, the time series of a voxel correlates not only with the experimental manipulation, but also with itself. At least two methods exist that deal with this temporal auto-correlation. The pre-coloring method uses a temporal smoothing algorithm, usually with a Gaussian kernel of some seconds FWHM. Pre-whitening removes any temporal smoothing in the data (Worsley et al., 2002). In the first experiment of the present work the "traditional" approach of pre-coloring is employed. In the second experiment both types of dealing with auto-correlation were tested. Since the pre-whitening method is rather new it was implemented in LIPSIA after conducting the first experiment. The comparison of the two methods exposed higher z-values (ca. 10%) and more activated voxels per cluster (ca. twice as much) utilizing the pre-whitening method.

After estimating β-values and its error-variance statistical tests between β's of particular experimental manipulations are applied. T-Tests between conditions of interest are implemented in contrast vectors. The resulting contrast images of each participant enter the second-level analysis. In this stage the null hypothesis that contrasts are zero is tested (implemented by a single-sample t-test). This procedure is called random-effects analysis, since randomness of different activations between participants is taken into account while comparing mean activation with variability of activations of different participants. In contrast, the fixed-effects analysis treats β-values of different participants as if they would stem from one participant, as inter-individual variance is not considered. Therefore, inference about the population is not appropriated with this method. The random-effects analysis is more conservative and allows inference to be generalized to the population.

2.3. Diffusion Tensor Imaging

Diffusion-weighted imaging

The distribution of diffusion of a water molecule in free space can be expressed by a Gaussian function. The probability to detect the molecule at the centre of the distribution is highest and gradually decreases in symmetrical manner. Diffusion-weighted imaging uses diffusion of water in the brain for measurement. The mapping of the MR-signal is realized by employing different magnetic field gradients. Gradients along the three axes produce local, measurable inhomogeneity of the B_0 linear to the axes. Duration, strength, and polarity of the gradient can be controlled. This advantage is used in order to assign different dephasing functions of numerous water protons to specific voxels that form a MR-image. In contrast to the fMRI method, at which usually single shot GE-weighted images are used, diffusion-weighted imaging applies usually SE-sequences simply to avoid the T_2^*-decay in the signal (Mori, 2002). Detection of movement of the water molecules occurs after the second gradient of the SE-signal. This gradient produces a rephasing of protons. But the phase of a moved proton differs from the phase of a unmoved proton after the second HF-impulse. Hence, the diffusion-weighted MR-technique detects the imperfect rephasing by signal loss in one voxel with lots of moving water molecules in comparison to another voxel including less movement of protons. The resulting apparent diffusion constant (ADC) is coded in brightness of each voxel. An area with high ADC is for instance situated within ventricles. Diffusion within ventricles tends to be high and movement of protons is high, inducing a reduced signal after the rephasing gradient, since moved protons do not rotate in phase anymore. The resulting picture represents the diffusion-weighted image.

White Matter Tractography

Two types of water diffusion, namely isotrophic and anisotrophic, are present in the brain. Isotrophic diffusion describes a random spatial distribution. If a barrier comes into play, water tends to diffuse along this structure, leading to anisotrophic diffusion. Hence, the probability of finding this water molecule after a certain time depicts an elliptic function. In particular, diffusion within fibers of the white matter follows a diffusion ellipsoid function (anisotrophic diffusion). On the other hand, diffusion between fibers follows a spherical diffusion functions (isotropic diffusion).

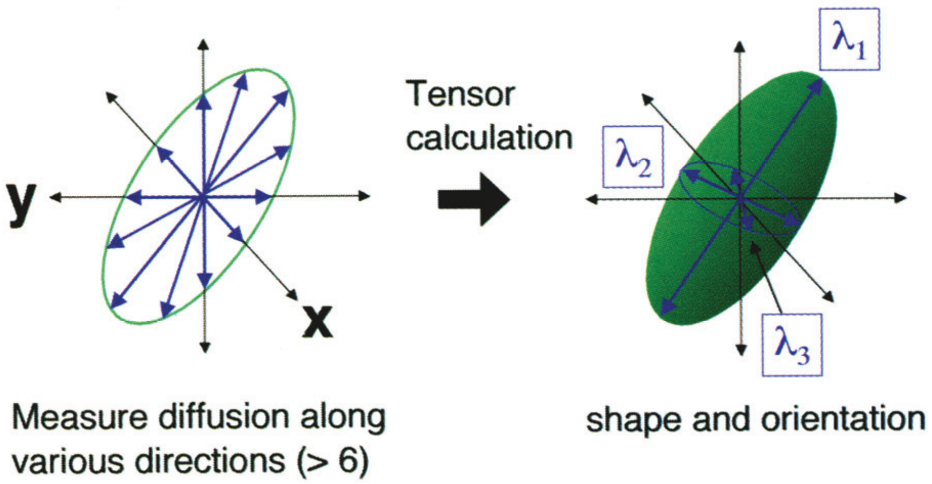

Figure 2.3: Diffusion ellipsoid characterized from measurements along six independent axes from Toga & Mazziotta (2002), p. 385.

In order to detect the orientation of fibers in the white matter, the anisotrophic diffusion of at least six parameters have to be measured. Parameters consist of three numbers that define the shape of the ellipsoid and three vectors that define the orientation of the ellipsoid. Note that the ellipsoid represents a simplified model taken due to its useful mathematical implementation. These values result in a 3 x 3 tensor, called diffusion tensor. To determine the six elements of the tensor, at least six diffusion contrasts along the six independent axes need to be measured. The direction of each voxel can be calculated from this matrix. The probability of a specific direction of fibers underlying one voxel can be calculated using a maximum likelihood model. Nevertheless, it is important to note that the direction of fibers is always a potential mix of many fiber directions within a certain voxel. The maximum likelihood method produces an eigenvalue of each voxel, representing its direction. The eigenvalue of direction enters a random walk algorithm proposed by Koch and colleagues (2002). In this model transition probability of a voxel to a neighboring voxel depends on a local probability density function. This function is based on the local eigenvalue that is modeled by the diffusion tensor imaging measurement. The presumed fiber direction is calculated while 'walking' from voxel to voxel with high probability along directions with high anisotrophic diffusivity. While repeating this random walk several times, a relative measure of anatomical connectivity between a start and a target voxel is generated. The probability of connectivity between two neighboring voxels is calculated from the product of the diffusion coefficient of two voxels. The resulting 3D connectivity distribution starting at a given seed point is called tractogram.

Chapter 3

Natural and artificial Grammar and the Brain

3.1. Artificial Grammar and ERPs

Only few ERP experiments exist that have examined the computation of different types of rule-based structures (see Table 3.1 for an overview). Recent studies have investigated speech-related artificial grammar rules (Friederici, Steinhauer, & Pfeifer, 2002), syllable sequences with transformational rules (Hoen & Dominey, 2000), and arithmetical sequences (Nunez-Pena & Honrubia-Serrano, 2004). Hoen & Dominey (2000) compared transformations in letter sequences with a memory control condition. The transformations followed the rule 123-312 (e.g. ABC-CAB), whereas the memory condition involved recognizing the repetition of a target element at the end of the sequence (e.g. ABC**Z**DEF**Z**). In the ERPs a LAN and a later positivity were found for transformations. The authors suggested that the LAN component represents an intersection of general structure building processes.

Nunez-Pena & Honrubia-Serrano (2004) conducted an experiment with numerical series and varied the violation of the series in terms of error salience. A sequence (e.g. 7, 10, 13, 16, 19, 22) could end with a correct digit (i.e. 25) or a clear incorrect (50), or a confusing incorrect digit (24). An early negativity and a late positivity were reported for the processing of violations. Moreover, the amplitude of the positivity varied as a function of salience: clear incorrect endings of a sequence showed a higher amplitude than confusing endings; the more salient the error in the sequence the larger the positivity. This saliency effect was interpreted to reflect different degrees of difficulty in integrating the violated arithmetical structure.

Furthermore, the computation of visually presented sentences and rule-based sequences of letter strings was compared in a study by Lelekov-Boissard & Dominey (2002). In the language condition (French sentences), the word order was changed in a phrase to create a word category violation in a sentence. A translated example of a correct sentence is the following: “Philippe was presented to Yves by Frank.”; an example of an incorrect sentence: “Philippe was presented to by Yves Frank.” In the sequencing condition, the letter strings followed a simple transformation rule (e. g. A-B-C B-C-B), and violations of this rule were produced by inserting a new element (e.g. A-B-C B-C-**D**). Both, sentences and letter string sequences elicited a late positivity for violations, but the effects differed with regard to scalp distribution between the two sequence types. The authors proposed related and overlapping neurophysiological processes for structure building in speech and non-speech processes. In contrast to the simple transformational rules in the studies by Lelekov-Boissard & Dominey (2002) and Hoen & Dominey (2000), Friederici et al. (2002) used a language-like artificial grammar task to investigate more complex rule processing. This artificial grammar featured pseudowords in different word classes (e.g. nouns, verbs, and determiner) combined via different syntactical rules. The rules were derived from natural languages, comprising of “sentences” with a noun and a verb phrase. The rules were first learned by participants until a certain performance level was reached. Then participants were tested with correct and violated sentences. Violations of these structures (“word category violations”) revealed an early negativity and a late positivity. Overall, the studies discussed above suggest that the processing of both transformational rules and language-like artificial grammar rules show similarities to native language processing. In artificial grammar tasks, ERP components were reported comparable to components found in language processing studies.

The serial reaction time task (SRT, Nissen & Bullemer, 1987) represents a different paradigm in the investigation of structure building and sequencing. In this task, the assessment of sequence learning is analyzed (see Chapter 1 for a description). Some ERP studies have employed variations of an SRT, resulting in somewhat conflicting results. While Eimer et al. (1996) and Rüsseler & Rösler (2000) found an enhancement of the N200 and P300 component for violations of the regular sequence, Baldwin & Kutas (1997) only found a higher amplitude of the P300 for grammatically structured orders of the letters. Also, the influence of explicit versus implicit learning of the task on the P300 is still questionable. Both Eimer et al. (1996) and Rüsseler et al. (2003) found the N200 and P300 effects only for explicit learners, whereas Baldwin et al. (1997) reported no learning type effect whatsoever. Therefore, the degree to which the processing of the SRT and the processing of natural language rely on overlapping cognitive mechanisms remains an open question. In this respect, it is interesting to note that patient studies have demonstrated that aphasic and agrammatic patients master the SRT. In contrast, they fail to learn transformations of the rule (Dominey, Hoen, Blanc, & Lelekov-Boissard, 2003), and they are impaired in learning auditorily presented rule-based phoneme sequences (Goschke, Friederici, Kotz, & van Kampen, 2001). For example, agrammatic patients showed impairments in a transformation task in which the order of the letter sequence had to be shifted (Dominey et al., 2003). Thus, if the N200 and P300 components are assumed by Hoen et al. (2000) to be language independent and if aphasic and agrammatic patients, which are impaired in certain aspects of language processing, successfully perform the SRT, then this suggests different neural generators for the type of sequencing required in the SRT and language-related sequencing.

Table 3.1: Overview of ERP studies investigating artificial grammar rules.

AG rules and ERPs			
Authors	Design	Comparison	Result
Baldwin and Kutas (1997)	SRT	Rule-based sequences vs. violation	P300
Eimer et al. (1996)	SRT	Violation vs. rule-based sequences	N200 and P300
Friederici et al. (2002)	Language-related AG rules	Structure violation vs. correct rule	Early negativity and late posterior positivity
Hoen and Dominey (2000)	Syllable sequences with transformation	Transformation vs. Recognition	LAN and late posterior positivity
Lelekov-Boissard and Dominey (2002)	Syllable sequences with transformation	Violation vs. rule-based sequences	Late posterior positivity
Nunes-Pena et al. (2004)	Arithmetic sequences	Sequence violation with different error salience	Late posterior positivity
Rüsseler and Rösler (2000)	SRT	Violation vs. rule-based sequences	N200 and P300

3.2. Artificial Grammar and fMRI

Some imaging studies exist that have examined the computation of artificial rule-based structures (see Table 3.2 for an overview). Recent PET and fMRI studies that investigated the learning and processing of artificial grammar rules addressed the question of explicit versus implicit rule learning (Fletcher, Buchel, Josephs, Friston, & Dolan, 1999; Forkstam & Petersson, in press; Peigneux et al., 2000; Petersson, Forkstam, & Ingvar, 2004; Seger, Prabhakaran, Poldrack, & Gabrieli, 2000; Skosnik et al., 2002; Strange, Henson, Friston, & Dolan, 2001), chunk versus pattern learning (Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004), or identification of language related brain areas (Musso et al., 2003; Opitz & Friederici, 2003, 2004).

Most imaging studies using an artificial grammar task, applied Reber’s artificial grammar (Reber, 1967, see Chapter 1). Brain activations reported for the processing of Reber's artificial grammar (Fletcher et al., 1999; Lieberman et al., 2004; Peigneux et al., 2000; Petersson et al., 2004; Seger et al., 2000; Skosnik et al., 2002) varied as a function of the used design and of the type of the comparisons. In most of the studies, rule-based letter strings were presented visually. Participants were instructed to judge the strings for the existence of a rule in the sequence (grammatical judgment task). For instance Fletcher et al. (1999) used an implicit learning paradigm in their fMRI experiment. Single letters were presented on a screen. Grammatical sequences of the letters were repeated within one block, at which recognition of the strings should be rated. The authors reported that areas in the right prefrontal cortex were engaged in recognition and left prefrontal regions were active while judging of the grammaticality of the letter strings. Peigneux et al. (2000) conducted a PET study and presented single letters on different positions on the screen. The sequence of the letters either followed a grammatical rule or occurred in random order (50%). Participants were instructed to press a button as fast as possible when the next letter appeared on the screen. Reaction time decreases over time significantly more in the grammatical condition than in the non-grammatical condition. A random effects analysis of non-grammatical versus grammatical trials revealed activation in the Striatum (caudate nucleus and putamen) and in the inferior frontal gyrus (among others).

It is an ongoing controversy whether the advances in reaction times during the artificial grammar task were due to the learning of an abstract rule or because of superficial similarities between the letter strings. Vokey & Brooks (1992) postulated an exemplar-specific categorization: these superficial similarities are based on the whole pattern. In contrast, Perruchet & Pacteau (1990) suggested a fragment-based categorization (i.e. exemplar parts are taken as a chunk). These considerations where taken into account by the fMRI study of Lieberman and colleagues (2004). The authors computed an artificial grammar that implemented different chunk strengths. Letter strings with a high transitional probability within bigrams and trigrams were categorized as high chunk strength and strings with low transitional probability were classified as low chunk strength. Items with high chunk strength where associated with an increased degree of superficial similarity. The authors reported that rule learning was correlated with activation in right caudate and chunk strength was associated with the medial temporal gyrus.

In sum, several different regions were reported to be involved in the learning of a Reber grammar. Most common activations were found in different parts of the left inferior frontal gyrus (Forkstam & Petersson, in press; Lieberman et al., 2004; Peigneux et al., 2000; Petersson et al., 2004; Seger et al., 2000). Moreover, the basal ganglia were also engaged in artificial grammar learning. Activation in the caudate nucleus was reported by Peigneux et al. (2000), Liebermann et al. (2004), and Forkstam et al. (in press). Lieberman et al. (2004) additionally found activation in the putamen and the claustrum. Furthermore left medial frontal activations were reported by Fletcher et al. (1999), Seger et al. (2000), and Skosnik et al. (2002). Additional activation was reported in the medial (Fletcher et al., 1999) and superior (Forkstam, in press) temporal gyrus, the frontal operculum/anterior insula (Forkstam, in press), the cerebellum (Fletcher et al., 1999), and the left occipital gyrus (Seger et al., 2000). The functional interpretations of these findings seem to appear difficult, because the type of analysis differed between the studies. Skosnik et al. (2002) and Lieberman et al. (2004) found their brain activation pattern computing the contrast grammatical versus non-grammatical items. Petersson et al. (2004) applied the opposite contrast (non-grammatical versus grammatical). Forkstam et al. (in press) calculated both contrasts. Seger et al. (2000) contrasted grammatical items with a baseline condition. Peigneux et al. (2000) and Fletcher et al. (1999) used the individual performance as a regressor in the general linear model. Taken together, artificial grammar experiments using finite state grammar rules delivered heterogeneous brain activation patterns.

Another group of studies working with artificial grammars used rules containing a more language-like structure than in the Reber grammar (Musso et al., 2003; Opitz & Friederici, 2003, 2004). Opitz et al. (2003) conducted an fMRI study with a complex artificial grammar. This grammar comprised of a phrase structure with different pseudo-word categories (determiner, nouns, verbs, adjectives, and adverbs) which were combined with different types of grammatical rules. Learning of the artificial language came along with activation in the hippocampal formation in the early learning stage; the left inferior frontal gyrus was engaged during the later part of the task. Musso et al. (2003) conducted two fMRI experiments in which grammatical rules that exist in natural languages (real grammatical rules) where compared with rules which are not present in natural languages (unreal grammatical rules). Native German speaking participants learned a small set of Italian words (first experiment) or Japanese words (second experiment) previous to the fMRI session. During scanning syntactical rules were learned by trial and error. The contrast between real grammatical rules and unreal grammatical rules was correlated with activation in Broca’s area in both experiments. These results suggest that language-like artificial grammar processing relies on common activation patterns as observed in studies exploring syntactical processing in natural language.

Table 3.2: Overview of fMRI studies investigating artifical grammar rules. L = left, R = right, MFG = middle frontal gyrus, MTG = middle temporal gyrus, IFG = inferior frontal gyrus, OG = occipital gyrus, FOP = frontal operculum.

AG rules and fMRI

Authors	Design	Comparison	Result
Fletcher et al. (1999)	Reber grammar	Rule acquisition vs. baseline	L MFG, L MTG, cerebellum
Forkstam and Petersson (in press)	Reber grammar	Violation vs. rule-based sequences	L IFG, FOP, STG, Basal ganglia,
Lieberman et al. (2004)	Reber grammar, different probabiltiy	Rule-based vs. violation	R caudate and R MTG
Musso et al. (2003)	Minimal natural grammar rules	Natural rules vs. unreal (not language) rules	L IFG
Opitz and Friederici (2003)	Language-related AG rules	Rule acquistion vs. baseline	Hippocampal formation and L IFG
Opitz and Friederici (2004)	Language-related AG rules	Rule change vs. word change	L ventral premotor cortex
Peigneux et al. (2000)	SRT	Violation vs. rule-based sequences	L IFG and L caudate nucleus
Petersson et al. (2004)	Reber grammar	Violation vs. rule-based sequences	Left IFG
Seger et al. (2000)	Reber grammar	Rule-based vs. baseline	L IFG, L MFG, and L OG
Skosnik et al. (2002)	Reber grammar	Rule-based vs. violation	L MFG

3.3. Natural Grammar and ERPs

Electrophysiological investigations of structural language processes have utilized different types of violations in sentences (Table 3.3). Language studies that examined phrase structure violations have reported an early left anterior negativity (ELAN), which was interpreted as a reflection of an initial phrase structure building process, and a late positivity called the P600, representing syntactic reanalysis or repair (Friederici, Pfeifer, & Hahne, 1993; Hahne & Friederici, 1999, 2002; Neville, Nicol, Barss, Forster, & Garrett, 1991). Morphosyntactic agreement violations have been associated with a left anterior negativity (LAN) and a late positivity (Coulson, King, & Kutas, 1998; Friederici et al., 1993; Gunter, Stowe, & Mulder, 1997; Münte, Heinze, Matzke, Wieringa, & Johannes, 1998; Osterhout & Nicol, 1999). For instance, Münte et al. (1998) tested semantic, morphosyntactic, and orthographic violations in German sentences. While semantic violations correlated with a biphasic N400, P600 pattern, morphosyntactic violations caused an early negativity and a late positivity. The early negativity was interpreted as a detection mechanism in response to the morphosyntactic error, and the late positivity was hypothesized to reflect a syntactic reanalysis process. Friederici et al. (1993) compared normal sentences ("Der Finder wurde belohnt.") with semantic ("Die Wolke wurde begraben."), morphological ("Das Parkett wurde bohnert."), and word category violations ("Der Freund wurde im besucht."). An N400 component was obtained by semantic violations, while morphological violations elicited a negativity and a late positivity, and word category violations elicited an early left anterior negativity (ELAN).

In contrast to the studies with violations, other experiments used correct structures and varied the dependencies between phrases in a sentence (Fiebach, Schlesewsky, & Friederici, 2001, 2002; Kaan, Harris, Gibson, & Holcomb, 2000; King & Just, 1991; King & Kutas, 1995; Kluender & Kutas, 1993; Rösler, Pechmann, Streb, Röder, & Hennighausen, 1998). For example, Kaan et al. (2000) found a late positivity with a higher amplitude for a lexical element in a complex syntactic structure in comparison to a less complex syntactic structure. The complexity of the syntactic structure was also manipulated in a study by Fiebach and colleagues (2001). In this study, subject versus object wh-questions were used. A persistent left frontal negativity and a late positivity were reported for the more complex object wh-questions only. The late positivity related to the variation of the long distance dependencies was assumed to reflect difficulties in integrating the more complex syntactic structure. A LAN was also found in English sentences with wh-questions in comparison to yes/no questions (Kluender & Kutas, 1993). Rösler et al. (1998) discovered a LAN on German non-canonical sentences when compared to canonical sentences. In both cases, the authors interpreted the LAN to reflect the additional load imposed on working memory during the processing of long distance dependencies in sentences.

Table 3.3: Examples of ERP studies investigating natural language.

Natural language and ERP			
Authors	Design	Comparison	Result
Coulson et al. (1998)	Morphosyntactic agreement	English sentences violation vs. correct	LAN and P600
Fiebach et al. (2001)	Syntactic complexity/working memory	German sentences object-related vs. subject related	Sustained LAN and P600
Friederici et al. (1993)	Phrase structure violation	German sentences Violation vs. correct	ELAN and P600
Kaan et al. (2000)	Syntactic complexity	English sentences complex vs. less complex	P600
Kluender & Kutas (1993)	Syntactic complexity/working memory	English sentences wh-questions vs. yes/no question	LAN
Münte et al. (1998)	Morphosyntactic agreement	German sentences violation vs. correct	LAN and P600
Neville et al. (1991)	Phrase structure violation	English sentences violation vs. correct	ELAN and P600

3.4. Natural Grammar and fMRI

In recent years a growing number of studies were published, dealing with different types of syntactic manipulations in natural sentences in fMRI and PET settings. Most of these studies included not only syntactic manipulations, but also semantic variations. In some cases pragmatic or phonological features of the sentences were modified. In the following section these experiments are first classified by the experimental manipulation (i.e. complexity versus violation), second by the modality (i.e. auditory versus visual presentation). Experiments dealing with local syntactic violations are summarized in Table 3.3 and studies manipulating the complexity of a sentence are reviewed in Table 3.4.

3.4.1. Syntactic violation studies

Violations in auditory presented sentences

Ni et al. (2000) conducted an fMRI study in which correct sentences were compared with local syntactical anomalies (e.g. “Trees can grew”) and semantic errors (e.g. “Trees can eat”). Correct sentences correlated with activity in several regions, mainly in the perisylvian brain areas. The activation pattern of syntactic and semantic violations showed largely overlapping brain regions. These included the bilateral IFG (BA 44, 45, and 47), the bilateral middle frontal gyrus (BA 46/9), and bilateral superior and middle temporal gyri. A direct comparison between the syntactic and semantic violation conditions revealed several small clusters of activated brain regions only for the semantic violations. No brain regions specifically for the syntactic condition could be identified.

Pragmatic (e.g. “The woman painted the insect.”), semantic (e.g. “My mother ironed a kiss.”), and subcategory violations (e.g. “His father chattered the umbrella.”) were included in auditory presented English sentences by Kuperberg et al. (2000). The three conditions exposed activation in the left inferior temporal and fusiform gyrus. Again, no brain region could be specified for the local syntactic violation.

Friederici, Rüschemeyer, Hahne, & Fiebach (2003) compared normal sentences with syntactic incorrect, and semantic incorrect sentences (examples similar to the stimuli used in Friederici et al., 1993). Both incorrect conditions showed activity in the superior temporal regions in comparison to correct sentences. A ROI analysis revealed that semantic violations were correlated with a higher percentage of signal change in the bilateral middle portion of the superior temporal cortex. In contrast, the authors reported more signal change for syntactic anomalies in the anterior portion of the left superior temporal gyrus, the left posterior frontal operculum (FOP) and the putamen.

Violations in visually presented sentences

Moro and co-workers (2001) conducted a PET study with visually presented Italian sentences. Different types of violations were included, i.e. local syntactic violations (change of the word order), morphosyntactic anomalies (suffix disagreement), and phonological errors (including strings of consonants). All three conditions exposed activation in left prefrontal regions (including the IFG), bilateral cerebellar areas, temporal, and parietal regions. Syntactic and morphosyntactic anomalies engaged Broca’s area (BA 44/45) and only syntactic violations were reported to activate the bilateral IFG, the left caudate nucleus and the left insula.

English sentences containing grammatical errors were compared with sentences containing spelling errors in the experiment by Embick, Marantz, Miyashita, O'Neil, & Sakai (2000). Example of a grammar error: “Bill wrote paper a about the discussion of the treaty.” Example of a spelling error: “Bill wrote a paper about the discussion of the treaty.” Left perisylvian regions were detected during the processing of both types of violations. Moreover, grammatical anomalies showed a higher BOLD response in Broca’s area, in the left superior temporal gyrus (Wernicke, BA 22), and the left angular and supramarginal gyrus.

3.4.2. Syntactic complexity

Complexity variation in auditory presented sentences

Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky (2003) manipulated syntactic features in Hebraic sentences in a 2 x 2 design. Grammatical versus ungrammatical sentences and sentences containing transformations versus no transformations were applied. The transformation was realized by an embedded object-relative clause (e.g. “I helped the girl [that Mary saw in the park].”). In sentences without transformation a sentential complement was used (e.g. “I told Mary [that the girl ran in the park].”). The baseline contrast revealed several different brain areas to be activated during listening to the sentences. The ROI analysis in these regions exposed sensitivity to transformation only in the left inferior frontal gyrus (Broca’s area) and in the bilateral posterior superior temporal sulcus. The authors concluded that these regions represent the brain basis of phrase movements.

Caplan, Alpert, & Waters (1999) presented English sentences containing object relative clauses versus subject relative clauses. The object relative clauses were considered to be more complex (e.g. “It was the juice that the child enjoyed.”) in comparison to subject relative clauses (e.g. “It was the child that enjoyed the juice.”). The direct comparison between object related versus subject related sentences correlated with activity in the left pars triangularis of the left IFG (Broca’s area) and was interpreted as an engagement of this particular brain region in the processing of syntactically more complex sentences.

Syntactic complexity and comprehension difficulty was examined in the study by Peelle, McMillan, Moore, Grossman, & Wingfield (2004). Syntactic complexity was manipulated while presenting subject relative clauses (e.g. “Men that assist women are helpful.”) or object relative clause structures (e.g. “Women that men assist are helpful.”). Comprehension difficulty was assessed with three different speech rates of the sentences. Syntactic complexity was associated with activity in Broca’s area and speech rate was correlated with activity in frontal brain regions such as anterior cingulate and premotor cortex.

Röder, Stock, Neville, Bien, & Rösler (2002) used German sentences and compared syntactic complexity with meaningfulness. Syntactic complexity was parameterized while applying four different sentence types with varying object – subject relations between phrases. An example of a less complex sentence is the following: “Jetzt wird der Astronaut dem Forscher den Mond beschreiben.”. A sentence with a more difficult/complex grammatical structure: “Jetzt wird den Mond dem Forscher der Astronaut beschreiben.”. Meaningfulness comprised of the factor-levels normal words versus pseudo-word sentences. Pseudo words were computed via replacing letters in each normal word. As a result syntactic difficulty activated Broca’s area. Additionally this effect was stronger for normal word sentences in comparison to pseudo word sentences.

Complexity variation in visually presented sentences

Stromswold, Caplan, Alpert, & Rauch (1996) directly compared English center-embedded sentences with right-branching sentences in a word and in a non-word condition. The centre-embedded sentences (e.g. “The juice [that the child spilled] stained the rug.”) were considered to be more complex than right-branching sentences (e.g. “The child spilled the juice that stained the rug.”). In the non-word condition the verb was displaced by a pseudo-word (e.g. “spilled” = “chorried”). The increased activity in Broca’s area (pars opercularis) in the more complex sentences (centre-embedded) in comparison to less complex sentences (right-branching) was interpreted as greater memory load associated with higher syntactic complexity. In the PET studies by Caplan and collaborators (1998) the same sentence material (excluding the non-word condition) was applied, testing eight female rather than eight male participants as in the former study. Again, Broca’s area was engaged during the processing of syntactic complexity, replicating the findings by Stromswold and colleagues (1996).

Syntactic complexity and working memory demands during sentence reading was investigated in studies using German (Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005) and English (Cooke et al., 2001) material. Syntactic complexity was manipulated by applying object-first versus subject-first sentences. Working memory demands were altered via the variation of the filler-gap dependency in the sentence. Both authors reported greater activity in Broca’s area during the processing of sentences with a longer filler-gap dependency. The results were interpreted as a crucial role of the left IFG for syntactic working memory during sentence comprehension.

Fiebach, Vos, & Friederici (2004) and Bahlmann, Rodriguez-Fornells, Rotte, & Münte (in press) compared syntactic complexity and syntactic ambiguity in German sentences. In both studies, the factor syntactic complexity comprised the levels subject relative (canonical) and object relative (non-canonical) clauses. The sentences of the experiment by Fiebach et al. (2004) were all ambiguous and contained different disambiguation points.

Whereas Bahlmann's et al. (in press) design contained ambiguous as well as unambiguous sentences. Hence, in half of the items a reanalysis of the sentences was necessary.

Table 3.4: Overview of fMRI studies investigating natural language. L = left, R = right, ant STG = anterior portion of the superior temporal gyrus, MTG = medial temporal gyrus, IFG = inferior frontal gyrus, FOP = frontal operculum, post FOP = posterior frontal operculum, IFJ = inferior frontal junction, IS = intraparietal sulcus, preSMA = presupplementary motor area, SG = supramarginal gyrus

Natural language and fMRI

Authors	Design	Comparison	Result
Syntactic anomaly studies			
Embick et al. (2000)	Visully English sentences	Syn violation vs. spelling errors	L IFG, L STG, L angular G, L SG
Friederici et al. (2003)	Auditory German sentences	Syn violation vs. correct	L post FOP and L ant STG
Kuperberg et al. (2000)	Auditory English sentenses	Syn violation vs. correct	n.s.
Moro et al. (2001)	Visually Italien sentences	Syn violation vs. phonotactic task	Bilateral IFG, L caudate, L insula
Ni et al. (2000)	Auditory English sentences	Syn violation vs. sem violation	n.s.
Syntactic complexity studies			
Bahlmann et al. (in press)	Visually German sentences	O-first vs. S-first	L IFG and SG
Ben-Shachar et al. (2003)	Auditory Hebraic sentences	Transformation vs. no transformation	L IFG and bilateral post STS
Caplan et al. (1998)	Visually English sentences	Center-embedded vs. right-branching	L IFG,
Caplan et al. (1999)	Auditory English sentences	O-first vs. S-first	L IFG
Cooke et al. (2002)	Visually English sentences	Long filler-gap and O-first vs. baseline	L IFG, bilateral post STG, MTG, and fusiform G
Fiebach et al. (2004)	Visually German sentences	O-first vs. S-first	n.s.
		Long disambiguation	bilateral IFG, IS
Fiebach et al. (2005)	Visually German sentences	O-first vs. S-first	n.s.
		Long filler-gap	bilateral IFG, STG
Friederici et al. (2006)	Visually German sentences	O-first vs. S-first (parametric)	L IFG, preSMA
Grewe et al. (2005)	Visually German sentences	O-first vs. S-first	bilateral IFG and FOP, L IFJ
Pelle et al. (2004)	Auditory English sentences	O-first vs. S-first	L IFG, R caudate
Röder et al. (2002)	Auditory German sentences	O-first vs. S-first	L IFG, L STG, R insula
Stromswold et al. (1996)	Visually English sentences	Center-embedded vs. right-branching	L IFG,

The reanalysis process was suggested to be reflected by an increased activity in the left supramarginal (BA 40) and left medial temporal gyrus (BA 21/37) and syntactic complexity was correlated with a higher BOLD response in Broca's area (BA 44). The length of the disambiguation in Fiebach's et al. (2004) study was associated with enhanced activity in BA 44.

Grewe and co-workers (2005) compared the factors permutation and NP-type with German sentences. Permutation comprised the factor levels object before subject versus subject before object sentences. NP-type represents a pronominal (e.g. "er") versus a non-pronominal noun phrase (e.g. "der Lehrer"). Object before subject sentences with non-pronominal noun phrase (e.g. "Dann hat der Gärtner dem Lehrer den Spaten gegeben.") activated the left IFG, pars opercularis.

Linguistic complexity was analyzed in a study by Friederici, Fiebach, Schlesewsky, Bornkessel, & von Cramon (2006). Complexity was manipulated while using "scrambled" sentences. In these sentences one or two object noun-phrases moved in front of the subject noun-phrase in order to generate increasing complexity of the sentences. The left pars opercularis (BA 44) was engaged during the processing of complex sentences. Moreover, the BOLD in the left IFG varied as a function of the number of permutations, i.e. the more complex a sentence the higher the hemodynamic response function in Broca's area.

3.5. Summary

The learning and processing of rule-based artificial structures was associated with different ERP components. Violations of numerical sequences elicited a late positivity, whereby the amplitude differed as a function of the error salience. Most studies using transformational rules reported a late positivity. A LAN, or an early negativity was additionally found. The SRT task caused a N200 and P300 component for violations, but also for correctly structured letter sequences.

Several different brain regions were reported to be activated during an artificial grammar task. The different designs used in the artificial grammar task could be categorized as language-like and non language-like grammar rules. While the former was associated with several different brain regions (including prefrontal, temporal, subcortical, and occipital brain areas) the latter seems to engage left perisylvian regions, comparable to studies exploring natural language.

ERP experiments employing syntactic violations could be segregated into studies applying local violations and studies dealing with different syntactic complexity of sentences. Local anomalies comprise phrase structure violations, which usually generated an ELAN and a P600, while morphosyntactic violations were linked to a LAN and a P600. Syntactic complexity was correlated with a LAN, a P600, or a sustained left-anterior negativity and a late positivity.

The literature review of fMRI and PET studies managing syntactic processing in natural sentences suggests (at least) two conclusions. First, syntactic complexity seems to provide a clear cut picture of involved brain areas. That is, all described studies reported activity in Broca's area for syntactic complex sentences, independent of the modality (auditory versus visual) or the task (grammaticality judgment versus implicit tasks). These results were interpreted in terms of transformations during phrase movements, working memory, linearization, or hierarchical structure building. Second, syntactic anomaly studies delivered an inhomogeneous picture of correlated brain regions. The bilateral IFG, basal ganglia, insula, Broca's area, superior temporal gyrus, and angular gyrus were correlated with local syntactic errors. Two studies identified no brain area for the processing of syntactic anomalies.

Chapter 4

Aims of the Study

The present work comprises comparisons between linear and hierarchical structures in artificial grammar tasks. Derived from the interesting study by Fitch & Hauser (2004) a FSG rule (linear) and a PSG rule (hierarchical) was applied. The aim of this thesis is to characterize the underlying neuronal mechanism during the processing of FSG and PSG rules.

Experiment 1: The first study was designed in order to investigate ERP correlates during the processing of the FSG and the PSG. As illustrated in the literature review (Chapter 3) a P600 was correlated with the processing of natural language as well as with the computation of numerical sequences and artificial grammars using transformational rules. Hence, it was hypothesized that both rule types (FSG and PSG) elicited a P600-like late positivity, similar to the P600 found during the processing of natural sentences and artificial grammar processing. Additionally, due to the different complexity of the two rules, the amplitude of the late positivity should vary as a function of rule type.

Experiment 2: This study was designed in order to evaluate brain regions associated with the two rule types. In a number of fMRI experiments Broca's area was associated with the processing of long distance dependencies in natural sentences. Based on the above reviewed literature, it was hypothesized that violations of the long distance dependencies produced in the artificial PSG rule should also cause activity in Broca's area. On the other hand, the processing of violations of adjacent elements in the artificial FSG rule should engage the frontal operculum, a region that was in some cases correlated with local syntactic anomalies.

Experiment 3: The aim of the third experiment was to investigate the processing of hierarchical structures in a more elaborated PSG rule in comparison to the PSG rule used in the first two experiments. This structure should eliminate the possible confound of a simple counting mechanism instead of using a hierarchical way of processing of the PSG sequences. Moreover, a direct comparison of FSG and PSG sequences was accomplished in a within-subject design. These experimental manipulations should replicate the findings of the second experiments, namely that the left IFG is activated during the processing of the PSG rule. It was additionally hypothesized that not only violations of the PSG engage Broca’s area (as demonstrated in Experiment 2), but also that PSG sequences in direct comparison to FSG sequences activate this region.

Chapter 5

Experiment 1: An ERP Study[3]

5.1. Abstract

The present study investigated the processing of two types of artificial grammars by means of event-related brain potentials (ERPs). Two categories of meaningless consonant-vowel syllables were applied in each grammar type. The two grammars differed with regard to the type of the underlying rule. The FSG followed the rule $(AB)^n$, thereby generating local transitions between As and Bs (e.g. n = 2: ABAB). The PSG followed the rule A^nB^n, thereby generating center-embedded structures in which the first A and the last B embed the middle elements (e.g. n = 2 [A[AB]B]). Two sequence lengths (n = 2, n = 4) were used. Violations of the structures where introduced at different positions of the syllable sequences. Early violations were situated at the beginning of a sequence, and late violations where placed at the end of a sequence. A posteriorly distributed early negativity elicited by violations was only present in FSG. This effect was interpreted as the possible reflection of a violated local expectation. Moreover, both grammar type violations elicited a late positivity. This positivity varied as a function of the violation position in PSG, but not in FSG. These findings suggest that the late positivity could reflect difficulty with integration in PSG sequences.

5.2. The present study

This ERP experiment was designed to compare violations in locally organized rule-based sequences with violations in hierarchically organized rule-based sequences. The rules underlying these structures are based on Chomsky's theory of generative grammar (Chomsky, 1957, see Chapter 1). The generative grammar applies the idea that in a natural language an infinite number of sentences can be generated by a finite set of grammatical rules. One of the simplest formal models of the generative grammar is a regular grammar

---

[3] This work was published in Bahlmann, Gunter, & Friederici (in press)

(type 3 in the Chomsky hierarchy). A regular grammar generates a mechanism for the specification of a certain type of structural regularities in a linear way. For instance, the formula $(AB)^n$ generates linear transitions between two types of categories (A and B). A more complex grammar model represents the context free grammar (type 2 in the Chomsky hierarchy). As an example, the rule A^nB^n generates center-embeddings of the two categories. Moreover, a hierarchical dependency between single members can be produced.

In a recent behavioral study by Fitch & Hauser (2004) a comparison of these two types of artificial grammar rules was conducted. The authors raised the question whether humans and non-human primates (i.e. cotton-top tamarins) differ in their ability to process linear and hierarchical syntactic structures. It was reported that humans easily learned both types of rules. In contrast, monkeys learned only the FSG, but not the PSG.

A simple example of a PSG is the rule A^nB^n in which the first A and the last B embed the middle elements (e.g. n = 2 [A[AB]B]). Such artificially generated embeddings could be compared to natural sentences like [The football player [that scored the goal] was celebrated]. It is worthwhile to note that in natural speech, the generation of long distance dependencies is not only provided by the relation between word categories. Other features like animacy, case, case in combination with gender, verb argument structure, etc. also provide mechanisms for the interpretation of long distance dependencies. Like in most studies with natural sentences, in this study only the aspect of category was explored, while other features were held constant.

In the present experiment, syllable sequences following the FSG were compared with syllable sequences following the PSG (see Figure 5.1). In response to violations of both grammatical rule types, the prediction was a late positivity similar to the P600 component for violations in natural language tasks. Two types of sequence lengths were applied: short sequences (n = 2) and long sequences (n = 4). The violations of the syllable sequence structures were situated at different positions in the sequences. The violation positions were counter-balanced over the structure types (FSG and PSG) and sequence lengths (short and long). The manipulation of the violation positions ensured that participants needed to focus on the whole syllable sequence structure and not just part of it to complete the error detection task required in the experiment. In addition, this variation made it possible to investigate the ERP correlates at different points in time while processing the different structures. This manipulation allows to compare violations in the first half of the sequence (early violations) to violations in the second half (late violations). With regard to the respective rule types, it was hypothesized that FSG and PSG would elicit differences in the late positivity for early and late violations. The processing of violations at different positions in a local structure (FSG) should cause no differences in the ERP components. In contrast, violations at an early position in a hierarchical structure (PSG) should elicit a different late positivity, than violations at late positions. A violation at the first position of a sequence should not cause differences in the ERPs for FSG and PSG.

5.3. Methods

5.3.1. Participants

Twenty-four healthy, right-handed subjects participated in this study (12 female, mean age 23, SD = 2.4). All participants were native German speakers and had normal or corrected-to-normal vision. All participants gave their written informed consent prior to testing. The data were handled confidentially.

5.3.2. Stimuli

According to the experiment by Fitch & Hauser (2004), the FSG was coded as $(AB)^n$, while the PSG followed the rule A^nB^n. Sequences of CV syllables were visually presented to the participants. Syllables were assigned to two categories. In the study by Fitch & Hauser (2004), the two categories where distinguished by pitch (A = high pitch, B = low pitch). Due to the visual presentation in the present study, the two categories were coded phonologically. Category A syllables contained the vowels i or e (e.g. de, gi, le, ri, se, ne, ti, mi) while category B syllables were comprised of o or u vowels (e.g. bo, fo, ku, mo, pu, wo, tu, gu). Examples for syllable sequences are given in Table 5.1 and Figure 5.1. For both types of artificial grammar, the same syllables were used. To prevent pattern learning, the probability of occurrence of the frequency of the syllables was balanced in the material. If one syllable occurred by chance more frequently in a certain position of one sequence, then the participant could assume a rule behind this chunk. Hence, all syllables appeared with equal frequency in the experiment.

Table 5.1: Stimuli of the testing period. Length of the sequences, number of items, and positions of included violations (bold and underlined). In each sequence one of the possible error was included.

	Sequence length	Correct items	Incorrect items	Positions of violation
FSG	4 syllables	175	175	**A**B**AB**
	8 syllables	175	175	**A**B**ABABA**B
PSG	4 syllables	175	175	**A**A**BB**
	8 syllables	175	175	**A**A**AABBB**B

5.3.3. Procedure

The experiment was divided into two sessions. One session contained the learning and testing of the FSG, and in the other session, the PSG was trained and inquired. Each experimental session contained a learning period followed immediately by a testing period. The sessions were separated by 7 (± 1) days. Due to technical reasons, two participants had a delay between the experimental sessions was shorter (3 and 4 days) and two participants had a longer delay (13 and 14 days). However no differences in behavioral data were found for these four participants compared to the others.

Learning: The learning period consisted of several blocks. At the beginning of each block, ten correct sequences were presented. Afterwards, five correct and five incorrect sequences were shown for which participants were required to respond with a button press whether the sequences were grammatical or ungrammatical. For each sequence, feedback was given. Participants were instructed to extract the rule underlying the syllable sequences. Training ended when participants answered 90% of the items correctly in two successive blocks. For the learning period, 504 syllable sequences were computed. Sequences of four, six, and eight syllables were employed (168 items each sequence length).

Testing: During the testing period, the actual ERP measurement was conducted. Participants were seated in a dimly lit, sound-attenuated chamber facing a computer screen. Participants were instructed to judge whether the sequences were rule-based or not. As in the training, feedback was given. For the testing period, 700 new sequences were presented, i.e. 350 short sequences (four syllables) and 350 long sequences (eight syllables) were computed, and half of them (175 for each length) were violated. The start of a sequence was indicated by a fixation cross (1000 ms). Each syllable was presented for 300 ms with an inter-stimulus interval of 200 ms between the syllables. At this point, participants could make their judgment for 1000 ms, which was followed by feedback for

500 ms. The trial length was set to 4500 ms (four syllables per sequence), 5500 ms (six syllables), or 6500 ms (eight syllables).

Key assignment, sex of the participants, and order of the grammar type (FSG first versus PSG first) was counterbalanced across participants. In order to prevent participants from focusing on one particular position within a sequence during the detection of a violation, the position of the violation was systematically changed.

5.3.4. ERP recording:

The EEG activity was recorded with 56 Ag/AgCL electrodes mounted in an elastic cap. Electrode impedance was kept below 5 kΩ. Bipolar horizontal and vertical EOGs were recorded for artifact rejection purposes. All scalp electrodes were referenced to the left mastoid. The signals were recorded continuously with a band pass filter from DC to 30 Hz at a sampling rate of 250 Hz and stored on hard disc for off-line analysis.

5.3.5. Data analysis

Average ERPs of correct and incorrect syllables were computed for positions one, three, four, six, and seven in both the FSG and PSG. Each average started at 200 ms pre-stimulus and lasted 1000 ms after the presentation of a syllable and was computed for each electrode position. Only correctly judged sequences entered the analysis. Approximately 3% of the trials had an ocular artifact. Thus, approximately 6% of all trials were excluded from further analyses. In order to avoid a loss of statistical power when repeated measure ANOVAs are used to quantify large numbers of electrodes (Oken & Chiappa, 1986), ten Regions of Interest were computed out of the 56 electrodes, each containing the mean of three electrodes. Statistical analyses were then performed using five anterior and five posterior ROIs in two latency windows (negativity: 300-400 ms; late positivity: 400-750 ms). The analyses differed in the positions of violations within a sequence (see below). Main effects were interpreted at a significance level of $p < .05$. The Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was always applied when evaluating effects with more than one degree of freedom in the numerator. In such cases, the uncorrected degrees of freedom (df), the corrected p-values, and the correction factor ε are reported.

For all significant effects, the partial effect size measure ω^2 was calculated[4]. The partial effect size reflects the amount of variance of the dependent variable that is declared by the independent variable (Hays, 1973) and ranges between 0 and 1.

5.4. Results

5.4.1. Behavioral Data

The Phrase Structure Grammar learning took longer than the Finite State Grammar learning did (PSG: 15.45 minutes, FSG: 8.25 minutes, $t(23) = 3.073$, $p < .01$). No transformation effects of the order of the learning of the two grammar types ($t(22) = .79$, n.s.), gender ($t(22) = 1.19$, n.s.), or key assignment ($t(22) = 1.62$, n.s.) were found. In order to verify differences in the processing of short and long sequences of FSG and PSG, an ANOVA with the factors grammar type and sequence length (both two-level, within-subject) was conducted for error rates. The processing of the PSG resulted in slightly more errors (4%) than processing of the FSG (2%), indicated by a main effect of grammar type: $F(1, 23) = 4.95$, $p < .05$ $\omega^2 = .06$. On the basis of the significant interaction between grammar type and sequence length ($F(1, 23) = 6.93$, $p < .05$, $\omega^2 = .11$), separate t-tests for paired samples were conducted for both grammar types. Error rates for long sequences were significantly higher than for short sequences ($t(23) = 2.58$, $p < .05$ in PSG (short: 3%, long: 5%, but not in FSG (short: 3%, long: 2%) ($t(23) = 1.16$, n.s.).

5.4.2. ERP Data

Long sequences in FSG and PSG

In order to explore the differences between both rule types, the violation positions were compared in detail. In both rule types the violations elicited a posterior negativity around 350 ms post stimulus followed by a large positivity with its maximum around 700 ms. The pattern of the positivity clearly differs between the two types of grammatical rules (see Figure 5.1B and Figure 5.2). The amplitude of the positivity in the FSG seems to remain equal, independently from the position of the violation within a sequence. Violation at

---

[4] The calculation of the effect size was based on the formula

$$\omega^2 = \frac{df_{num}(F-1)}{df_{num}(F-1) + n_{obs}}$$

in which df_{num} stands for the degrees of freedom of the numerator, F for the empirical F-value, and n_{obs} for the numbers of observations.

position four seems to have the same positivity effect as violations at the late positions six and seven. In contrast, the position of the violation in a PSG sequence appears to have an influence on the amplitude of the positivity. An error of the rule on an early position (four) elicits a smaller positivity than on a late position (six and seven). Due to the temporal ambiguity of violations at position three in the PSG, a direct comparison of FSG and PSG at this position would be insufficient. Consequently, each rule type was separately analyzed. First, a distinct analysis in the FSG condition was calculated to show that the position of violation within a sequence had no influence on the ERP pattern. An ANOVA with the factors position (four), violation (two), anterior/posterior (two) and ROI (five) was conducted. For the factor position, the positions included were three, four, six, and seven. In the PSG condition, positions four, six, and seven were analyzed to demonstrate that the ERP pattern systematically changed over the violation positions in a sequence. In this analysis, positions four, six, and seven formed the factor position and an ANOVA with the factors position (three), violation (two), anterior/posterior (two) and ROI (five) was carried out. The third position did not enter the analysis in the PSG condition, because this is an ambiguous position. Thus, the third position was analyzed separately (see below). In both analyses (FSG and PSG) the first position was excluded from the factor position, because no differences of the ERP pattern of the first position of the two rule types were expected. Hence the first position was analyzed separately (see below).

FSG: Positions three, four, six, and seven

a) 300 – 400ms: posterior negativity

The main analysis of the 300-400 ms interval showed no two-way, three-way, or four-way interactions with the factors position and violation (all $F < 2.65$, n.s.). A two-way interaction of anterior/posterior and violation ($F(1, 23) = 24.69$, $p < .0001$, $\omega^2 = .32$) was observed. Inspections of the ERPs showed that this interaction was caused by the posterior scalp distribution of the negativity (see Figure 5.2A and 5.2C). These results indicate that the negativity in FSG is independent of the violation position in a sequence. The negativity does not differ between violations in the first half and the second half of a FSG sequence.

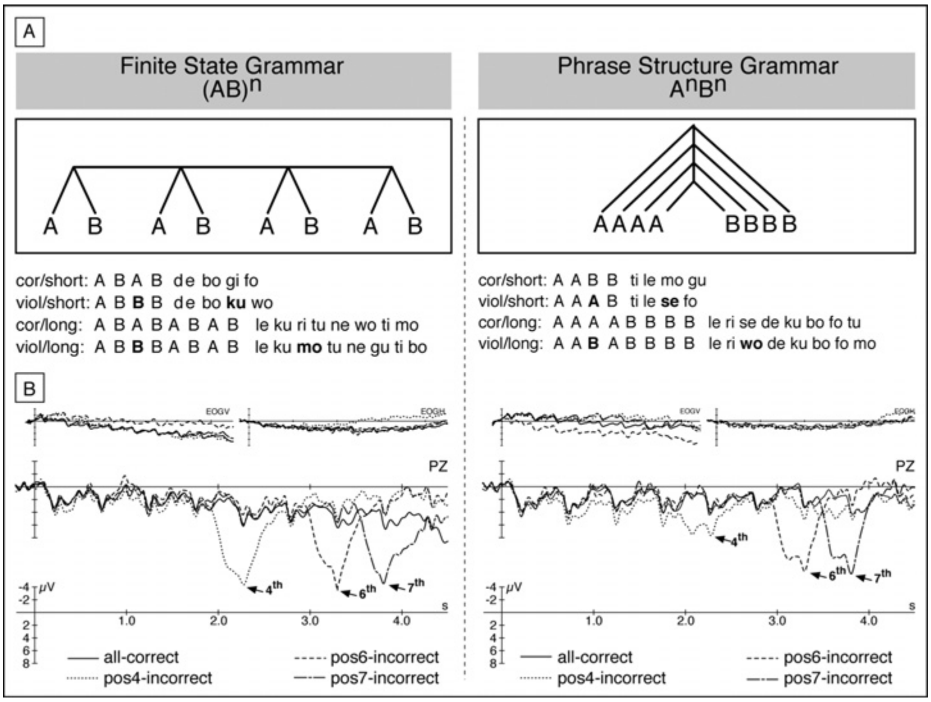

Figure 5.1:

A: General structure and examples of the stimuli in FSG and PSG. Short and long sequences were applied. Violations of the structure were situated at different positions in a sequence for both rule types (FSG and PSG). In the example, the violations are placed at the third position for FSG and PSG and for short and long sequences (bold letters).

B: Example of the grand average ERPs for long sequences from 24 participants. Negative voltage is always plotted upwards. One selected electrode is displayed for FSG (left panel) and for PSG (right panel). ERPs of the whole sequence for a correct sequence (all-correct, solid line), a sequence with a violation at the fourth position (pos4-incorrect, dotted line), a violation at the sixth position (pos6-incorrect, dashed line), and a violation at the seventh position (pos7-incorrect, dotted and dashed line).

b) 400 – 750ms: late positivity

The main analysis for the 400-750 ms interval showed no two-way, three-way, or four-way interactions with the factors position and violation (all $F < 2.75$, n.s.). Significant interactions between anterior/posterior and violation ($F(1, 23) = 14.86$, $p < .001$, $\omega^2 = .21$) and between ROI and violation ($F(4, 92) = 25.01$, $\varepsilon = .48$, $p < .0001$, $\omega^2 = .49$.) were found. Visual inspections of the ERPs indicate that these interactions were caused by a centro-parietal scalp distribution of the late positivity (see Figure 5.2A and 5.2C). This analysis shows the late positivity in the FSG condition is found regardless of violation position. Violations in the first and second half of a FSG sequence show a similar late positivity in the ERPs.

PSG: Positions four, six, and seven

a) 300 – 400ms: posterior negativity

The main analysis of the 300-400 ms interval showed a significant three-way interaction of position, anterior/posterior, and violation ($F(2,46) = 15.82$, $\varepsilon = .82$, $p < .0001$, $\omega^2 = .32$). On the basis of this interaction, separate analyses with the factors anterior/posterior and violation were performed for each position in the PSG condition:

Position four: No main effect or interaction reached significance at the fourth position (all $F < .92$).

Position six: An interaction of anterior/posterior and violation was found ($F(1, 23) = 19.06$, $p < .0005$, $\omega^2 = .27$). On the basis of this interaction, separate analyses were conducted for anterior and posterior regions. At posterior ROIs, the main effect of violation ($F(1, 23) = 1.82$, n.s.) was not significant, confirming that no posterior distributed negativity was found for violations of position six of the PSG. At anterior ROIs, the main effect of violation reached significance $F(1, 23) = 12.8$, $p < .005$, $\omega^2 = .19$). Visual inspection of the ERPs revealed that the main effect of violation at this late position was caused by the earlier onset of the late positivity, previous to the positivity at an earlier position (i.e. position four, see Figure 2D).

Position seven: A significant three-way interaction between anterior/posterior, ROI, and violation ($F(4, 92) = 3.44$, $\varepsilon = .60$, $p < .05$, $\omega^2 = .08$) was observed. On the basis of this interaction, separate analyses were carried out for anterior and posterior ROIs at position seven. In anterior ROIs, an interaction between ROI and violation was observed ($F(4,92) = 3.18$, $\varepsilon = .51$, $p < .05$, $\omega^2 = .08$). Based on this interaction, separate analyses were conducted for each of the five ROIs. The two left ROIs and the very right ROI showed a significant main effect of violation ($5.5 \leq F \leq 14.44$, $.001 \leq p \leq .05$, $.08 \leq \omega^2 \leq .21$). The main effect of violation did not reach significance in the middle and the middle-right ROIs (all $F < 2.11$, n.s.). Inspections of the scalp distributions showed a similar pattern as in the anterior ROIs of position six: violations at the late positions (six and seven) show an earlier onset of the late positivity than violations at position four.

b) 400 – 750ms: late positivity

The main analysis for the 400-750 ms interval showed a significant three-way interaction between position, ROI, and violation ($F(8,184) = 7.29$, $\varepsilon = .52$, $p < .0001$, $\omega^2 = .33$). On the basis of this interaction, separate analyses with the factors ROI and violation were performed for each position.

Position Four: A significant main effect of violation was found (F(1, 23) = 11.04, $p < .005$, $\omega^2 = .17$). Furthermore an interaction between ROI and position was observed (F(4, 94) = 4.35, $\varepsilon = .51$, $p < .05$, $\omega^2 = .11$). Inspections of the ERPs indicate that this interaction was produced by a centro-parietal scalp distribution of the late positivity (see Figure 5.2B).

Position Six: A significant main effect of violation was obtained (F(1, 23) = 97.2, $p < .0001$, $\omega^2 = .67$). A three-way interaction between anterior/posterior, ROI, and violation was also significant (F(4, 92) = 6.05, $\varepsilon = .59$, $p < .005$, $\omega^2 = .17$). On the basis of this interaction, separate analyses in the five anterior and five posterior ROIs were performed. All five anterior ROIs showed significant main effects of violation ($26.0 \leq F \leq 49.5$, all $p < .0001$, $.34 \leq \omega^2 \leq .50$) as did all five posterior ROIs ($100.79 \leq F \leq 114.4$, all $p < .0001$, $.67 \leq \omega^2 \leq .70$). The larger effect at posterior ROIs ($.67 \leq \omega^2 \leq .70$) in comparison to anterior ROIs ($.34 \leq \omega^2 \leq .50$) assumed a posteriorly distributed late positivity on violations at the sixth position. Inspections of the ERPs showed a broadly distributed late positivity with its peak on the posterior electrodes for violations on the sixth position (see Figure 5.2D).

Position seven: A significant main effect of violation was elicited F(1, 23) = 125.7, $p < .0001$, $\omega^2 = .72$). On the basis of a significant three-way interaction between anterior/posterior, ROI, and violation (F(4, 92) = 5.42, $\varepsilon = .56$, $p < .001$, $\omega^2 = .14$), separate analyses in the five anterior and five posterior ROIs were conducted. All anterior and posterior ROIs showed significant main effects of violation (anterior: $31.77 \leq F \leq 55.4$, all $p < .0001$, $.38 \leq \omega^2 \leq .53$; posterior: $83.9 \leq F \leq 153.8$. all $p < .0001$, $.63 \leq \omega^2 \leq .76$). Posterior ROIs showed a larger violation effect ($.63 \leq \omega^2 \leq .76$) than anterior ROIs ($.38 \leq \omega^2 \leq .53$), hence it was assumed that the late positivity showed a posteriorly distributed late positivity. Inspections of the ERPs showed a broadly distributed late positivity with its peak on the posterior electrodes for violations at the seventh position.

Taken together, violations on the sixth and seventh position show similar scalp distributions and effect sizes in the PSG condition. In comparison to the two positions later in a sequence, violations on the forth position revealed a different scalp distribution and effect size. Positions six and seven showed a broadly distributed late positivity, while position four displayed a posterior-parietal distributed late positivity. The size of the violation main effect on position four ($\omega^2 = .17$) is smaller than that for positions six and seven ($\omega^2 = .67$ and $\omega^2 = .72$).

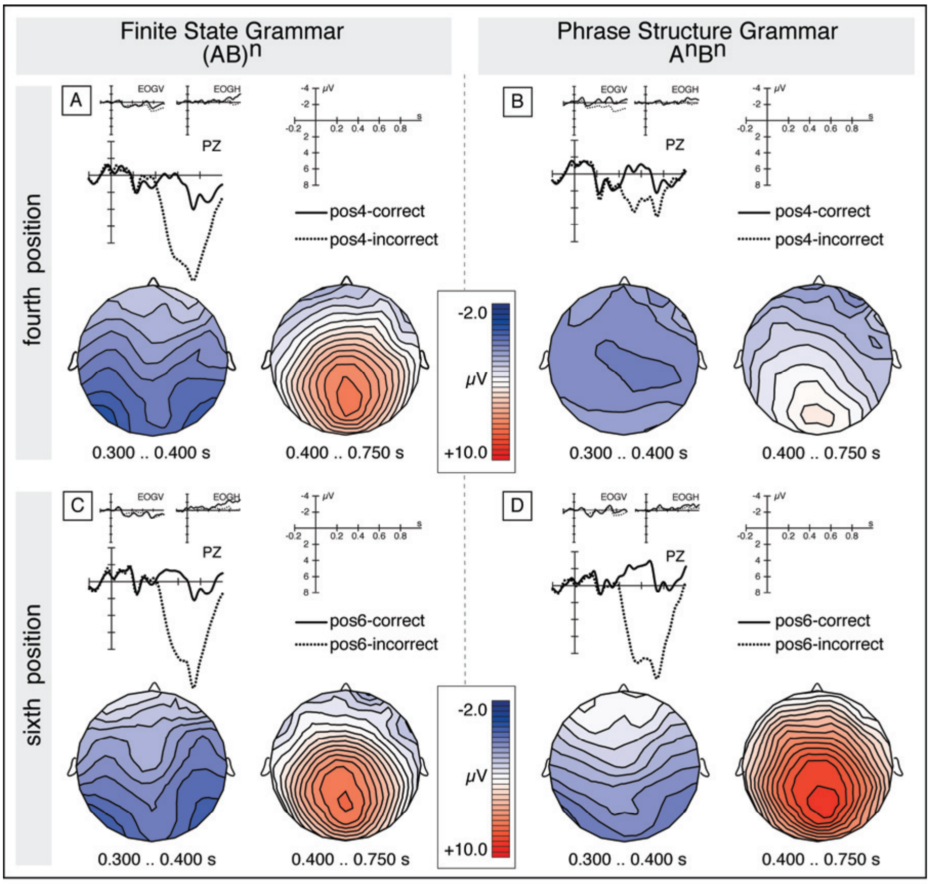

Figure 5.2:

Upper plot: ERPs elicited by FSG (A) and PSG (B) at the fourth position for correct (pos4-correct, solid line) and for incorrect (pos4-incorrect, dotted line) sequences as shown by a selected electrode. Topographic maps of the differences between effects for correct and incorrect sequences at the fourth position are shown for the two time windows that entered analyses. Red areas indicate positive differences between the conditions, and blue areas indicate negative differences.

Lower plot: ERPs for the sixth position in FSG (C) and PSG (D) for one selected electrode. ERPs of the sixth position for correct (pos6-correct, solid line) and for incorrect (pos6-incorrect, dotted line) sequences are shown. Topographic maps of the ERPs of correct minus incorrect sequences at the sixth position are shown for the two time windows that entered analyses.

To summarize, the effect of violations that were introduced at an early position in the sequences were compared with violations occurring at later positions. This comparison was done in FSG on positions three and four (early violations) and on positions six and seven (late violations). Due to the temporal ambiguity at the third position in PSG, this position did not enter the analyses of early and late violations in the PSG condition. Hence, position four (early violation) was compared with positions six and seven (late violations). In the FSG condition, violations at positions three, four, six, and seven show a similar ERP pattern. Independent of the position of the violation within a sequence, a negativity followed by a late positivity was elicited. In contrast, in the PSG condition the ERP pattern changes as a function of the violation position in a sequence.

Short vs. long sequences in FSG and PSG

FSG: Positions three and four in short and long sequences

In order to verify whether short and long sequences differ in the ERPs in FSG, the third and forth position of short and long sequences were compared. An ANOVA with the factors anterior/posterior (two), ROI (five), short/long (two) and violation (two) was conducted. The main analysis for the 300-400 ms interval showed no main effect of short/long (F(1, 23) = 0.11, n.s.) and no interaction with the factor short/long (all F < .32, n.s.). Also, the ANOVA on the 400-750 ms interval showed no main effect of short/long (F(1, 23) = 0.57, n.s.) and no short/long interactions (all F < 1.3, n.s.). These analyses demonstrate that there is no difference in the ERPs for short and long sequences in FSG.

PSG: Position three in short and long sequences as a case of ambiguity

a) Position of the violation in PSG: Third position:

In the PSG, a violation at the third position of the sequence should cause ambiguity because a category A syllable at this position could be a correct long sequence or an incorrect short sequence. Upon visual inspection of the ERPs, there seems to be no difference between incorrect, short sequences (AA**A**B) and correct, long sequences (AA**A**ABBBB) at the third position. Apparently, also correct, short (AA**B**B) and incorrect, long sequences (AA**B**ABBBB) show no differences in the ERPs (see Figure 5.3). Therefore it was assumed that participants treat this ambiguous position as if it were a long sequence. If this post-hoc hypothesis is true, then the ERPs found in the incorrect short vs. correct long (AA**A**B vs. AA**A**ABBBB) and correct short vs. incorrect long (AA**B**B vs. AA**B**ABBBB) should show no difference. To test this, two separate analyses were carried out, one with the factors anterior/posterior (two), ROI (five), and incorrect short/correct long (two), and the second ANOVA with correct short/incorrect long (two) as the third factor.

Incorrect short sequences: The analysis for the 300-400 ms interval for the first analysis showed no main effect of incorrect-short/correct-long (F(1, 23) = 0.01, n.s.) and no interactions with this factor. Also, the analysis for the 400-750 ms interval revealed no main effect of incorrect short/correct long (F(1, 23) = .02, n.s.) and no interaction effect with this factor, meaning that the ERPs for correct short and incorrect long sequences at the third position in the PSG are not different.

Incorrect long sequences: In the second analysis at the 300-400 ms interval, there was no main effect of correct short/incorrect long (F(1, 23) = .67, n.s.), but an interaction of correct short/incorrect long with anterior/posterior (F(1, 23) = 4.44, $p < .05$, $\omega^2 = .06$). The step-down analysis revealed no main effect of correct short/incorrect long in the anterior ROIs (F(1, 23) = 3.24, n.s.) and in the posterior ROIs (F(1, 23) = 0.2, n.s.). The analysis for the 400-750 ms interval revealed no main effect of correct short/incorrect long (F(1, 23) = 0.03, n.s.) and no interactions with this factor. Together, the ERPs for incorrect short and correct long sequences and for correct short and incorrect long sequences at the third position in the PSG do not differ. The above analyses therefore support the assumption that participants dealt with the ambiguous third position by processing it as a correct long sequence.

b) Position of detection of the violation: Fourth position

The previous analyses showed that participants process information at the ambiguous third position as belonging to the long sequence. Although the system managed to deal with this ambiguous position, when does the detection of a violation for a short PSG takes place? In order to analyze the position of violation detection, the ERPs of the third position were lengthened by 2000 ms, thereby providing information about when the violation was actually detected. Inspections of the ERPs revealed two different positive components elicited from violations for short and long sequences (see Figure 3). The first positivity occurred at 950 to 1250 ms, and the second positivity appeared from about 1450 to 1850 ms after the onset of the third item. In order to explore the significance of these components, the same type of analyses as in the previous section (2.2.1.) were conducted in the time windows 950 -1250 ms and 1450 – 1850 ms.

Incorrect short sequences: An ANOVA with the factors anterior/posterior (two), ROI (five), and incorrect short/correct long (two) was conducted. The main analysis for the 950-1250 ms interval showed an interaction between anterior/posterior and incorrect short/correct long (F(1, 23) = 7.2, $p < .05$, $\omega^2 = .11$). On the basis of this interaction, separate analyses in anterior and posterior ROIs were carried out. The main effect of incorrect short/correct long was only significant in posterior ROIs (F(1, 23) = 13.97, p <

.001, $\omega^2 = .20$), but not in anterior ROIs ($F(1, 23) = .39$, n.s.). This effect revealed that in short sequences, the positivity showed a posterior scalp distribution.

The main analysis for the 1450-1850 ms interval also displayed a significant interaction of anterior/posterior with incorrect short/correct long ($F(1, 23) = 8.59$, $p < .01$, $\omega^2 = .13$), leading to the break-down analysis in anterior and posterior ROIs. In anterior ROIs, the main effect of incorrect short/correct long was not significant ($F(1, 23) = 4.01$, n.s.). In posterior ROIs, this effect was significant ($F(1, 23) = 54.55$, $p < .0001$, $\omega^2 = .52$). Thus, the second positivity showed a posterior distribution and a higher effect size than the first positivity on short sequences.

Incorrect long sequences: An ANOVA with the factors anterior/posterior (two), ROI (five), and correct short/incorrect long (two) was conducted.

The main analysis for the 950-1250 ms interval showed a main effect of correct short/incorrect long ($F(1, 23) = 39.14$, $p < .0001$, $\omega^2 = .44$). No other effect was significant. Inspections of the ERPs indicate a broadly distributed first negativity for long sequences.

The main analysis for the 1450-1850 ms interval showed a three-way interaction between anterior/posterior, ROI, and correct-short/incorrect-long ($F(4, 92) = 6.89$, $\varepsilon = .58$, $p < .001$, $\omega^2 = .17$). On the basis of this interaction, separate analyses were carried out in each of the five anterior and each of the five posterior ROIs. No main effect of correct short/incorrect long reached significance ($.5 \leq F \leq 1.62$, n.s.). This analysis indicated that the second positivity was not significant for long sequences. Taken together, it was shown that the detection of an error that occurred at the third position took place at the forth position, independent of the length of the sequence. This effect was substantiated by a positivity at 950 to 1250 ms after onset of the third item. Moreover, a second positivity at about 1450 to 1850 ms was elicited by short sequences, but not by long sequences.

Table 5.2: Summary of the effect sizes of the ambiguous third position of violation.

		950 –1250 ms	1450 –1850 ms
violation	Violation of A AA**A**B	$\omega^2 = .20$ posterior distribution	$\omega^2 = .52$ posterior distribution
	Violation of B AA**B**ABBBB	$\omega^2 = .44$ posterior distribution	$.5 \leq F \leq 1.62$, n.s.

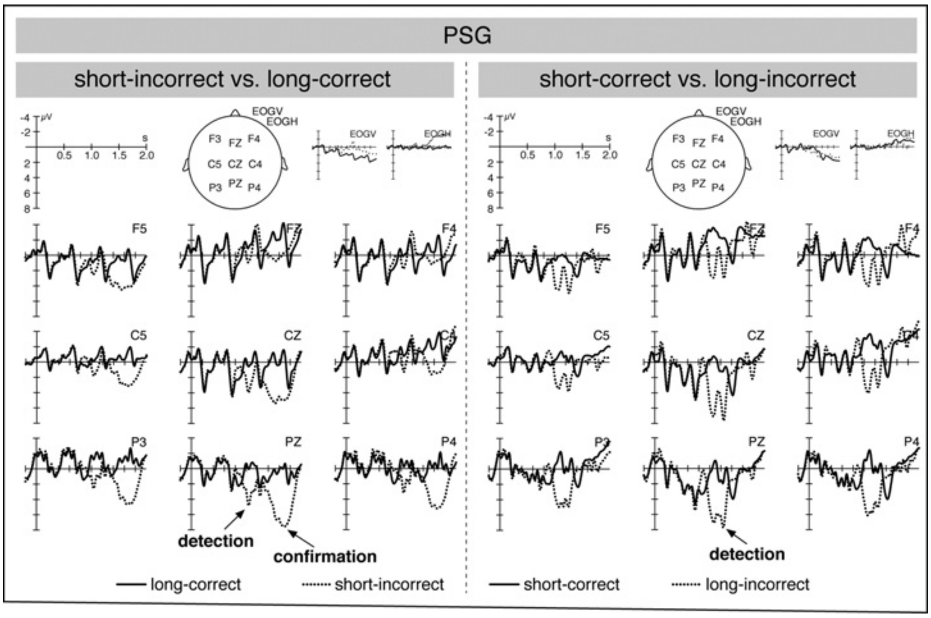

Figure 5.3: ERPs for the ambiguous third position in the PSG are illustrated.

Left panel: ERPs for short incorrect (AA**A**B, solid line) and long correct (AA**A**ABBBB, dotted line) sequences are shown at the third position.

Right panel: ERPs for short correct (AA**B**B, solid line) and long incorrect (AA**B**ABBB dotted line) sequences are illustrated at the third position.

First position in FSG and PSG

In order to explore whether differences between the rule types already occur at the first position of the sequences, an ANOVA with the factors rule type (two), violation (two), anterior/posterior (two) and ROI (five) was carried out for the first position of the sequences.

The main analysis for the 300-400 ms interval showed a main effect of violation ($F(1, 23) = 20.6$, $p < .0001$, $\omega^2 = .28$). The interaction between rule type and violation was not significant ($F(1, 23) = 0.00$, n.s.). The main analysis for the 400-750 ms interval showed a main effect of violation ($F(1, 23) = 7.41$, $p < .001$, $\omega^2 = .11$). No other interactions were significant. Taken together, no differences of the ERPs at the first position were found for FSG and PSG (see Figure 5.4).

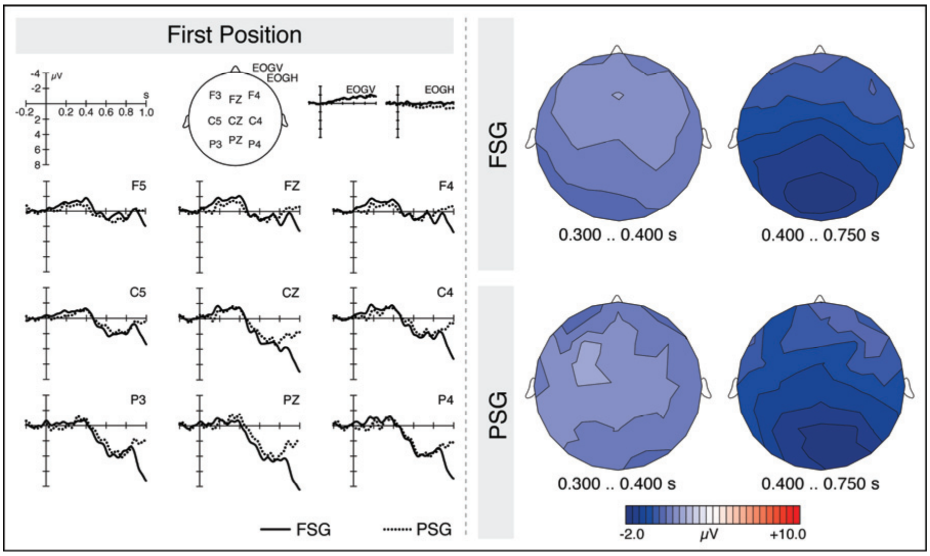

Figure 5.4:

Left panel: Difference waves of grand average ERPs from 24 participants for nine selected electrodes. Correct and violated sequences of the first position in a sequence in both rule types are shown. Solid lines represent the difference waves of the ERPs from correct minus violated sequences of the FSG, while dotted lines represent the difference waves of the ERPs from correct minus violated sequences of the PSG.

Right panel: Topographic maps of the differences between the ERPs of correct and incorrect sequences of the first position in the sequences. Two time windows are shown for both rule types (FSG in the upper right and PSG in the lower right panel). Red areas indicate positive differences between the conditions, and blue areas indicate negative differences.

5.5. Discussion

In the present study, two types of artificial grammar rules were compared in which two types of categories (A and B) of meaningless syllables were used. The sequences of syllables followed either a locally organized Finite State Grammar rule (FSG) or a hierarchically organized Phrase Structure Grammar rule (PSG). The number of syllables per category (eight syllables in each category) and the length of the syllable sequences (four syllables in short and eight syllables in long sequences) were held constant between the two types of grammar rules. Furthermore, due to the within-subject design, the inter-individual variance was held constant. The behavioral data suggest that no transfer effect was found for the learning of the grammar rules. Violations of the rules were introduced at different positions of the sequences. A negativity at about 300 to 400 ms was elicited only in FSG and a positivity between 400 to 750 ms was found with different amplitudes for FSG and PSG.

5.5.1. The late positivity

The late positivity for violations of structured sequences are in line with several studies on the processing of language (Coulson et al., 1998; Friederici et al., 1993; Gunter, Friederici, & Schriefers, 2000; Gunter et al., 1997; Münte et al., 1998; Osterhout & Nicol, 1999), language-like artificial grammar (Friederici et al., 2002), syllable sequences with transformational rules (Hoen et al., 2000, Lelekov-Boissard et al., 2002), arithmetical sequences (Nunes-Pena, et al., 2004), and harmonic violations in music (Patel, Gibson, Ratner, Besson, & Holcomb, 1998). In the experiment by Friederici et al. (2002), a late positivity was observed for category violations in a complex artificial grammar with phrase structure. Hoen and Dominey (2000) reported similar results for non-linguistic abstract structures, i.e. a late positivity was elicited for transformations in letter sequences. Lelekov-Boissard et al. (2002) compared syntactic errors in sentences with errors in a transformation structure while Patel et al. (1998) compared syntactic anomalies in sentences with harmonic violations in music. Although both experiments by Lelekov-Boissard et al. (2002) and Patel et al. (1998) used completely different stimuli, they both report a late positivity evoked by anomalies of the structure. Studies using natural language that have reported a late positivity manipulated the phrase structure by introducing a word category violation (Friederici, 1993; Hahne, 1999, 2002) or the morphosyntactic structure (Coulson, 1998; Münte, 1998; Friederici, 1993; Gunter, 1997, 2000; Osterhout, 1999). The similarity of the late positivity in the present study to those found in natural language studies suggests at least partly overlapping neuronal generators of structure building in natural language and in an artificial grammar.

5.5.2. The negativity

Another interesting finding obtained in the present experiment is that the processing of the FSG elicited a posteriorly distributed negativity at about 300 – 400 ms. A similar negativity was also reported when an additional item violated the structure of a language-like artificial grammar (Friederici et al., 2002). Eimer et al. (1996) reported a larger negativity for deviant stimuli than for the standard stimuli in a serial reaction task (SRT) for explicit learners, but not for implicit learners. Rüsseler et al. (2003) found a broadly distributed negativity with a centro-parietal maximum for deviants in a SRT paradigm. Again, this was shown only for a participant group that was explicitly aware of the rule. The authors suggested that the conscious recognition of the unfulfilled expectation of a specific item could be reflected by the negativity. This effect could also account for the early posterior negativity in the FSG in the present study. The processing of the local probabilities between the syllable categories lead to an expectation of a specific category, i.e. an A is always followed by a B. If this category did not appear, the expectation was not fulfilled and the violation was detected. The posteriorly distributed early negativity in the FSG could reflect this unfulfilled expectation of the specific B category during the processing of the syllable sequences. In addition, violations at the first syllable position (B instead of A) elicited a broadly distributed early negativity and a centro-parietally distributed late positivity for both FSG and PSG. This activation pattern can be taken so support the hypothesis that an unfulfilled expectation of a category caused the early negativity, as both rule types correctly start with an A category. The finding that violations of PSG elicited no negativity (except for violations at the first position of a sequence) suggest that the processing of the dependencies in the PSG might require more complex mechanisms than the simple matching of transitional probabilities of syllable categories.

5.5.3. Effects of violation position: early versus late

Violations of the syllable sequence structures were introduced at different positions in the sequences. Early violations were located in the first half of the sequence at position three and four, while late violations occurred in the second half at position six and seven. The late positivity varied as a function of the structure type. The FSG revealed no differences of the amplitude of the positivity at varying violation positions. In contrast, the PSG elicited different amplitudes of the positivity for early versus late violations within a sequence, with the later positions demonstrating a larger amplitude than the early one.

The FSG is specified as probabilistic transitions between a finite set of elements (Chomsky & Miller, 1958). In the present case, the local transition remains the same independent of the position. This is reflected in the ERPs. The early negativity and the late positivity in the FSG do not differ in the amplitude or scalp distribution for early and late violations of the structure. In the PSG, however, the position of the violation matters functionally. This is also reflected in the ERPs: There were differences in the positivity size when comparing early and late violation positions, i.e. late positions elicited a larger positivity, but I assume that this was not merely an expectancy effect.

In terms of the traditional approach of context updating, variation of the P300 amplitude points toward variations of subjective probability or of modifications in task relevance (Donchin & Coles, 1988). Subjective probability (Johnson & Donchin, 1982) or expectancy ((Squires, Wickens, Squires, & Donchin, 1976) was shown to be reciprocally proportional to the P300 amplitude, i.e. the higher the expectation of an item, the smaller the amplitude. The amplitude variation of the positivity in the PSG condition, however, was most probably not caused by differences in expectancies. In the present study, the number of violations was counterbalanced over the positions in a sequence for both rule types. Due to this manipulation, the expectancy was held constant over the positions in a sequence. Moreover, if an increased expectation would be the basis of the reduced amplitude of the positivity at the forth position in PSG, the same decrease of the amplitude would be expected in the FSG condition. This was not the case. Hence, differences in subjective probability can be excluded as a generator of the variation of the amplitude. Task relevance as a predictor of the varying amplitude can also be excluded. The attention to a task is directly proportional to the amplitude (where ignoring of a task elicits no P300). In the present study, all stimuli were relevant to the task, and participants reached a very high performance level (correct answers in FSG = 98%, PSG = 96%). Thus, task relevance could not have generated the amplitude differences of the late positivity in the PSG condition. However, another attribute of the P300 is the variability of the amplitude as a function of the conditional probability of an event, i.e. the less expected the stimulus the larger the amplitude (Squires et al., 1976). Applying this to the present study, one could assume that the occurrence probability of a category B item at the fourth position (incorrect item for long sequences and correct item for short sequences) is higher than a category A item at the seventh position (only incorrect long sequence). An A at the end of the sequence is less likely than a B at the fourth position and the amplitude of the positivity is higher at the end of the sequence than at the fourth position of the sequence. This could also have had an influence on the modulation of the positivity's amplitude.

The data at hand are in agreement with those reported by Nunez-Pena et al. (2004), Patel et al. (1998), and Coulson et al. (1998), who showed that the late positivity varied as a function of the saliency of the violation. Nunez-Pena et al. (2004) reported that the positivity elicited by the processing of numerical series varied as a function of the violation type, i.e. the more salient the error, the higher the amplitude. In a study with anomalies in music, Patel et al. (1998) showed that the more distant a violating tone in a harmonic sequence from the expected tone, the higher the amplitude of the late positivity. Last, Coulson et al. (1998) reported that pronoun-case violations in a sentence caused a higher amplitude of the late positivity than verb-argument violations, i.e. the more salient the morphosyntactic violation, the higher the amplitude of the positivity. In line with these findings, the results of the present study support the hypothesis that the late positivity reflects difficulty of integration, and as will be discussed below, the degree of integration difficulty affects the size of the effect.

5.5.4. Effect of sequence length: long versus short

In the present study, short and long sequences where used in both rule types FSG and PSG. The manipulation of the sequence length caused a temporary ambiguity at the beginning of the sequence in PSG, but not in FSG, i.e. a category A syllable at the third position could be processed as a correct long sequence (AAAABBBB) or an incorrect short sequence (AA**A**B), and vice versa for category B. Post- hoc analyses showed that participants always treated those positions as members of long sequences. Moreover, the detection of an error took place at the item that immediately followed the violation at the third position, as revealed by a positivity at 950 to 1250 ms after the onset of the third position. This was the case for both short and long sequences.

Two differences were observed in the ERP pattern between short and long sequences for the detection of a violation at the ambiguous third position. First, a second positivity was displayed at about 1450 to 1850 ms after onset of the third position for short sequences only. As this component showed up at the end of the short sequence, this positivity could represent a confirmation effect of the error detection. Second, the effect of violation detection for long sequences ($\omega^2 = .44$) is higher than for short sequences ($\omega^2 = .11$). This result might suggest a difficulty of integration effect. Since it was shown that the third position was processed as if it were a long sequence, a B at the third position (incorrect/long) is more difficult to integrate than an A (incorrect/short) at this position. Hence, the greater amplitude of the error detection positivity for long sequences could have been caused by increased difficulty with integration in long sequences at the ambiguous third position.

5.5.5. Summary

To summarize, different ERP components were found for the processing of sequences of different structural types, namely sequences characterized by transitional probabilities (FSG) on the one hand, and sequences characterized by hierarchical dependencies (PSG) on the other hand. The processing of the FSG elicited a posteriorly distributed early negativity and a late positivity in response to a local structural violation, independent of its position in the sequence. In contrast, the PSG revealed a positivity that varied as a function of the violation position within a hierarchically structured sequence. Violations introduced late in a PSG sequence elicited a higher amplitude of the positivity than violations at the beginning of a sequence. The positivity was therefore interpreted to reflect difficulty of integration. In contrast, the early negativity observed in the FSG seems to reflect an unfulfilled expectation of an upcoming category in a local transition.

Chapter 6

Experiment 2: an fMRI Study[5]

6.1. Abstract

A series of fMRI experiments found Broca's area to be involved in language processing. The nature of this involvement is still not fully characterized. The present study addressed the question whether syntactical processing and hierarchical structure building in general is processed in the left inferior frontal gyrus. For this purpose the same artificial grammar task was applied as in the previous study. Sequences of senseless syllables were used that followed either a PSG rule or a FSG rule. Short sequences (n = 2) and long sequences (n = 4) were utilized and correct and violated sequences were used. These manipulations resulted in a 2 x 2 x 2 design with the between subject factors rule type (PSG versus FSG), and the within subject factors length (short versus long) and violation (correct versus incorrect). As a result, the processing of PSG sequences as compared to transitional probabilities of adjacent elements (FSG sequences) are supported by different areas in the human brain. Processing of violations of local transitions is subserved by the frontal operculum (FOP). In contrast, Broca's area was involved during the processing of violations of the PSG. Moreover, tractography data are presented suggesting different structural connectivity signatures for these two brain areas. These findings suggest that Broca's area is involved in the processing of hierarchical structures and the frontal operculum holds responsible for the computations of local transitions, not only in natural language, but also in general.

---

[5] This work was published in Friederici, Bahlmann, Heim, Schubotz, & Anwander (2006).

6.2. The present study

In the previous chapter the processing of two different grammar types was explored in an ERP setting. Violations elicited a posteriorly distributed late positivity for both grammar types. Additionally, the amplitude varied as a function of integration difficulty in the PSG condition. In order to determine possible neural generators of these components an fMRI experiment was conducted. In the present study, the same artificial grammar rules were used as in the previous one. Two groups of participants learned and processed either of the rule types.

As described in the third chapter, an involvement of the left IFG (BA 44/45) was found for the processing of natural sentences requiring a hierarchical reordering of the argument due to a non-canonical surface structure. This was reported for English (Caplan, et al., 1998, 1999; Cooke et al., 2002; Peelle, et al., 2004; Stromswold, et al., 1996), German (Friederici, et al., 2006; Grewe et al., 2005; Röder, et al., 2002), and Hebrew (Ben-Shachar, et al., 2003). In contrast, local phrase structure violations revealed activity in the left frontal operculum (Friederici et al., 2003). In line with the findings of syntactical complexity studies, it was hypothized that violations of the PSG engage Broca's area. In contrast, the FSG should be associated with activity in the frontal operculum.

6.3. Methods

6.3.1. Participants

Forty healthy, right-handed subjects participated in this study (19 male, mean age 26 years, SD = 2.7 years). No participant of the present study participated in the previews ERP experiment. They were native German speakers and had a normal or corrected to normal vision. No participant had a known history of neurological, major medical, or psychiatric disorder. Prior to scanning, participants were informed about the potential risks and gave a declaration of consent.

6.3.2. Stimuli

The stimuli of the present experiment comprised of the same structure as discribed in the first experiment (see Chapter 5). Sequences of consonant-vowel syllables were visually presented to the participants. The syllables were assigned to two classes (A and B), which were coded by different consonants and vowels (see Figure 6.1). For both types of artificial grammar the same syllables were used. As in Experiment 1, the probability of occurrence of the frequency of the several syllables was balanced in order to prevent pattern learning. Violations occurred at different positions in the sequence thus forcing the participants to parse the entire sequences according to the rule as learned. In contrast to the ERP experiment, violations occurred at positions 3 and 4 in the four-element short sequences and at the crucial positions 1, 2 and 7, 8 in the eight-element long sequences for the FSG and at the positions 1, 8 and 4, 5 in the long sequences for the PSG. In order to process these sequences, participants had to judge their grammaticality and to identify the class-membership of a given consonant-vowel-syllable and to match the class-membership of the incoming element with the rule-based-predicted class.

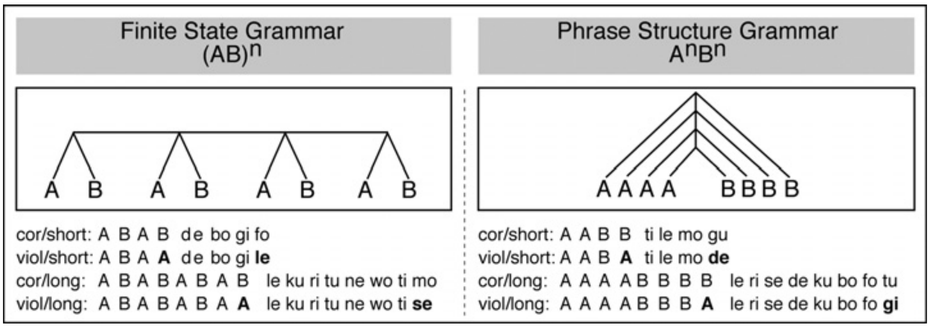

Figure 6.1: Structure and examples of stimuli in FSG and PSG. The same syllables were used for both types of grammar. The positions of the violations in the sequences were systematically changed. Examples of correct (corr) and violation (viol) sequences are given for the short and the long condition in each grammar.

6.3.3. Procedure

In contrast to the ERP experiment, participants were randomly assigned to one of two groups either learning FSG or PSG in the present study. A between subject design was chosen to prevent interference of the two grammars during learning and testing.

Learning: The learning procedure of the ERP experiment was also employed in the present study. Learning took place 2 days prior the fMRI experiment. The learning period was segmented into 12 blocks. In each block 10 correct sequences were presented. Thereafter 5 correct and 5 incorrect sequences were shown. Participants were instructed to extract the rule underlying the syllable sequences. In total 240 stimuli were presented. Sequences of 4, 6 and 8 syllables were employed in both the Finite State Grammar (FSG) and the Phrase Structure Grammar (PSG); i.e. FSG: $(AB)^2$, $(AB)^3$, $(AB)^4$; PSG: A^2B^2, A^3B^3, A^4B^4 (80 items each). Participants were required to respond to sequences by indicating with button press whether the sequences were grammatical or ungrammatical. Feedback was given. The learning lasted 23 minutes.

Testing: The testing differed from the previous experiment in the amount of stimuli and in the positions of included violations in the sequences. In the fMRI session 160 new items were presented (80 correct, 80 incorrect). Half of the sequences contained 4 syllables and half of them 8 syllables. Participants judged whether the sequences were rule-based or not. Again, feedback was given. 40 null-events were included and presented in a pseudo-randomized order with the other trials.

Each sequence started with a fixation cross (500 ms). Each syllable was presented for 300 ms with an inter stimulus interval of 200 ms between the syllables. Then, participants could respond for 1.000 ms, followed by feedback for 500 ms. Trials started with a jitter of 0, 500, 1000, or 1.500 ms.

6.3.4. fMRI Data Acquisition

Imaging was performed at 3T Trio scanner (Siemens, Erlangen, Germany) equipped with the standard birdcage head coil. Stabilization cushions were used in order to reduce head motion. For registration purpose, two sets of two-dimensional anatomical images were acquired for each participant immediately prior to the functional imaging. An MDEFT and an EPI-T1 sequence were used. T1-weighted MDEFT (Ugurbil et al., 1993) images (data matrix 256 x 256, TR = 1.3 s, TE = 7.4 ms) were obtained with a non slice-selective inversion pulse followed by a single excitation of each slice (Norris, 2000). Anatomical images were positioned parallel to AC-PC. Functional data were acquired from 16 axial slices (thickness = 3 mm; gap = 0.6 mm) using a gradient-echo EPI sequence with a TE = 30 ms, flip angle = 90°, TR = 2.000 ms, acquisition bandwidth = 100 kHz. The matrix acquired was 128 x 128 with a FOV of 25.6 cm, resulting in an in-plane resolution of 2 mm x 2 mm. One functional run with 666 time points was measured.

6.3.5. Functional Imaging Data Analysis

The fMRI data processing was performed using the software package LIPSIA (Lohmann et al., 2001). Functional data were motion-corrected offline with the Siemens motion correction protocol (Siemens, Erlangen, Germany). To correct for the temporal offset between the slices acquired in one scan, a cubic-spline-interpolation was applied. A temporal highpass filter with a cutoff frequency of 1/60 Hz was used for baseline correction of the signal and a spatial Gaussian filter with 3.768 mm FWHM was applied. The increased autocorrelation caused by the filtering was taken into account during statistical calculation by an adjustment of the degrees of freedom. To align the functional slices with a 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was performed. The rotational and translational parameters were acquired on the basis of the MDEFT and EPI-T1 slices to achieve an optimal match between these slices and the individual 3D reference data set. This 3D reference data set was acquired for each participant during a previous scanning session. The MDEFT volume data set with 160 slices and 1 mm slice thickness was standardized to the Talairach stereotactic space (Talairach & Tournoux, 1988). The rotational and translational parameters were subsequently transformed by linear scaling to a standard size. The resulting parameters were then used to transform the functional slices using trilinear interpolation, so that the resulting functional slices were aligned with the stereotactic coordinate system. This linear normalization process was improved by a subsequent processing step that performed an additional nonlinear normalization.

The statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations (Friston, Holmes, Poline et al., 1995; Friston, Holmes, Worsley et al., 1995). The design matrix was generated with a synthetic hemodynamic response function and its first and second derivative (Worsley & Friston, 1995). The model equation, including the observation data, the design matrix and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM to deal with the temporal autocorrelation. Contrast-images of the differences between the specified conditions were calculated for each participant. Each individual functional dataset was aligned with the standard stereotactic reference space. Due to poor behavioral data in the fMRI session (less than 57% correct answers), three participants were excluded from the analysis. Group analyses (random-effects model) based on the contrast-images were subsequently performed. The individual contrast-images were then entered into a second-level random effects analysis (one-sample t-test). Subsequently, t-values were transformed into Z-scores. To protect against false positive activations, only regions with Z-score greater than 3.09 ($P < .05$; corrected for multiple comparisons) and with a volume greater than 120 mm^3 (10 voxels) were considered.

The time course analysis for the region-of-interest in the FOP was performed in the voxel (X, Y, Z of the Talairach system) yielding the highest activation (contrast incorrect vs. correct) in both the PSG and the FSG group (-34, 20, -2). For the region-of-interest in Broca's area, the highest activated voxel in the PSG (-46, 16, 8) was chosen, and time courses were calculated for both groups at this voxel coordinate.

6.3.6. Structural Data Acquisition

Diffusion-weighted data and high-resolution 3-D T1-weighted as well as 2-D T2-weighted images were acquired in four participants (two from the FSG and two from the PSG group) on a Siemens 3T Trio Scanner with an 8-channel array head coil.

The diffusion-weighted data were acquired using a spin-echo EPI sequence (TR=8.100 ms, TE=120 ms, 44 axial slices, resolution 1.7 x 1.7 x 3.0 mm, gap = 0.3 mm, two acquisitions, maximum gradient strength 40 mT/m). The diffusion weighting was isotropically distributed along 24 directions (b-value=1000 s/mm2). Additionally a data set with no diffusion weighting was acquired. The total scan time was approximately 7 minutes.

The T2-weighted images (RARE; TR=7.800 ms; TE=105 ms, 44 axial slices, resolution 0.7 x 0.7 x 3.0 mm, gap = 0.3 mm, flip angle 150°) were co-registered to the 3D T1-weighted (MPRAGE; TR=100 ms, TI=500 ms, TE=2.96 ms, resolution 1 x 1 x 1 mm, flip angle 10°, two acquisitions) images. Subsequently, the diffusion-weighted images were non-linearly registered onto the T2-weighted images to reduce distortion artefacts. Based on the registered diffusion weighted images a diffusion tensor image (DTI) was calculated.

6.3.7. White Matter Tractography

For DTI analysis, a 3 dimensional extension of the random walk method proposed by Koch et al. (2002) was developed. The algorithm was applied to all white matter voxels in the functional activated area. The target space for the tractography was the whole white matter volume with a resolution of 1 x 1 x 1 mm. The fiber tracts for all start voxels in the region were averaged.

The algorithm can be described by a model of particles, moving randomly from voxel to voxel. The transition probability to a neighboring voxel depends on a local probability density function based on the local diffusivity profile that is modeled from the DTI measurement. The particle will move with a high probability along directions with high diffusivity, i.e. the presumed fiber direction. While repeating this random walk many times a relative measure is calculated for the anatomical connectivity between the start and the target voxel. For each elementary transition, the probability for a movement is computed from the product of the diffusion coefficient of the two neighboring voxels in the direction of the connecting line. The product is raised to the power 7 to focus the probability distribution to the main fiber direction and suppress the influence of transverse diffusion. This value was empirically chosen in such a way that the trajectories of most particles follow the main fiber directions. A total of 500 000 particles were tested for the start region. To compensate for the distance dependent bias of connectivity, each value is normalized to the distance to the start region. After reducing the dynamic range of the connectivity values by logarithmic transformation, the values were scaled to the range between zero and one. To remove random artefacts, only voxels with connectivity values bigger than 0.6 were displayed. This 3-D distribution of connectivity values in the brain is called tractogram.

6.4. Results

6.4.1. Behavioral results

Learning of the PSG was more difficult than of the FSG (error rate of the last block of FSG: 12.1% and of PSG: 27.0%; main effect grammar type F(1, 34) = 12.05, $p < .001$). Two days later, each group was submitted to a test session during which brain activation was measured by means of MRI. Behavioral data from the present fMRI experiment revealed that the FSG was slightly easier to process (error rates for the judgment task were FSG: 6.5% and PSG: 15.5%; main effect of grammar type: F(1, 34) = 5.01, $p < .05$).

6.4.2. Functional imaging results

Analyses of the BOLD contrast were performed for correct and incorrect sequences during the test session at the individual participant level and averaged over individuals for each group. Different brain activation patterns were observed for the two grammar types. Violations of either grammar caused activation in the FOP (Figure 6.2, lower panel, Talairach coordinates FSG: -36, 16, 0 and PSG: -36, 20, -2). In contrast, only violations of PSG induced additional activation in Broca's area (Figure 6.2, upper right panel; -46, 16, 8). Additional activation not included in the figure was found for both grammar types in the right temporal region (BA 21/22) and left BA 21 for the PSG (see Table 1).

6.4.3. Time course analysis

An ANOVA with the factors Grammar Type (two) x Violation (two) x Length (two) x Time Step (four: 5, 7, 9 and 11 seconds after sequence onset) was conducted for the two critical ROIs in the frontal cortex, FOP and Broca's area.

Table 6.1: Talairach coordinates and *z* value for the peak location in different ROIs for the two grammar types; *p* value was corrected at cluster level for the statistically significant differences of the corresponding activated regions; BA = approximate Brodmann's Area; L = Left hemisphere, R = Right hemisphere, G = Gyrus, S = Sulcus.

		Stereotactic coordinates				
Brain region	BA	x	y	z	t_{max}	p_{corr}
FSG						
L frontal operculum		-36	16	0	5.47	.001
R superior temporal S	21/22	52	-42	2	6.04	.001
PSG						
L frontal operculum		-36	20	-2	5.96	.001
L inferior frontal G	44	-46	16	8	5.06	.01
R superior temporal S	21/22	52	-40	8	5.90	.001
L middle insula		-38	-8	4	5.17	.008
R middle insula		34	0	-2	4.56	.001
R middle temporal G	21	-46	-30	-4	5.26	.003

For the FOP an ANOVA with the same factors revealed a significant main effect of Violation (F(1, 34) = 24.76, p < .001), of Length (F(1, 34) = 5.22, p < .05), of Time Step (F(3, 102) = 10.96. p < .001) and a Time Step by Length interaction (F(3, 102) = 23.82, p < .001). The interaction was further evaluated by t-tests indicating positive values for Time Step 5 (t(18) = 6.29, p < .001) and 7 (t(18) = 2.20, p < .05) and negative values for the Time Steps 9 (t(18) = -4.04, p < .001) and 11 (t(18) = -5.26, p < .001). The absence of a significant Grammar Type by Violation interaction (F(1, 34)=1.92, n.s.) indicates that the FOP subserves the processing of both the FSG and the PSG. For Broca's area the ANOVA revealed a significant main effect of Violation (F(1, 34) = 10.36, p < .01) and of Length (F(1, 34) = 10.51, p < .01) and a significant Grammar Type by Violation interaction (F(1, 34) = 4.96, p < .05).

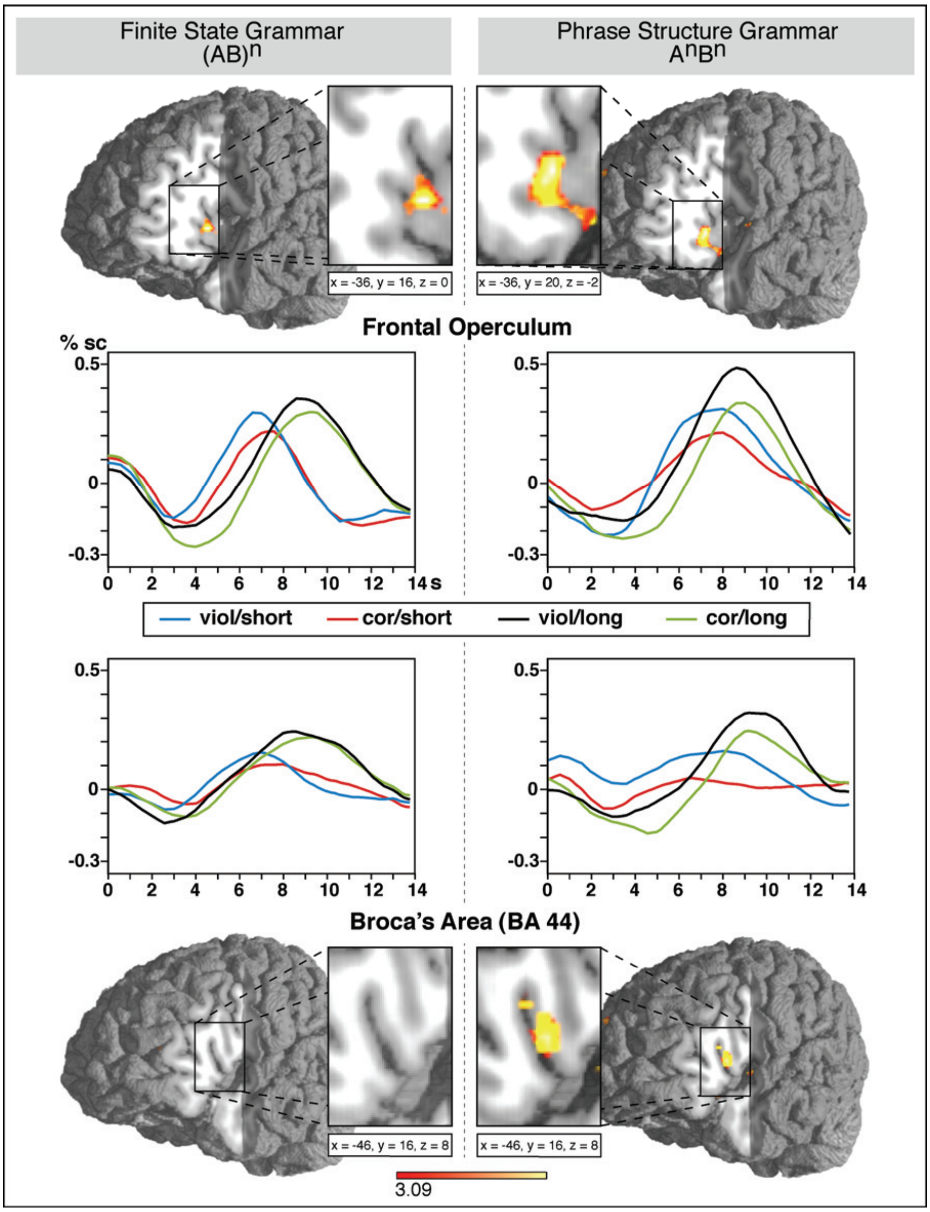

Figure 6.2: Brain activation pattern for the two grammar types. Statistical Parametric Maps of the group-averaged activation during processing of violations of two different grammar types ($P < .001$, corrected at cluster-level). Left panel shows the contrast of incorrect vs. correct sequences in the FSG, right panel shows the same contrast in the PSG for Broca's Area (upper part) and FOP (lower part). Middle panel displays time courses (% signal change) in corresponding voxels of maximal activation.

This interaction reflects an increased activation of Broca's area for the PSG compared to the FSG indicating that the processing of the PSG (but not of the FSG) additionally recruits Broca's area.

To test whether the activity in Broca's region is due to the difficulty of the task in the PSG an additional analysis in BA 44/45 with the two-level factors Performance x Violation was conducted. Individual performance of the PSG group was taken into account: High performance (n = 9, mean value = 95% correct answers, SD = 2.8) and low performance (n = 8, mean value = 71% correct answers, SD = 7.4) formed the between subject factor Performance. No effect was found (Performance x Violation: F(1, 15) = 0.47, n.s.). These results suggest that difficulty is not correlated with the activity in Broca's area.

6.4.3. Structural imaging results

Analysis of the DTI data of representative participants from the two different grammar groups reveal differential tractograms for the FOP and BA 44 (Figure 6.3). The figure illustrates the 3D-rendering of the distribution of the connectivity values of two start regions with all voxels in the brain volume. The two participants from the FSG group showed activation in the FOP in the critical contrast correct versus incorrect sequences ($p < .005$). The individual FOP activations were taken as the seed points for the individual tractograms (Talairach coordinates of Participant 1: -39, 25, 0 and Participant 2: -38, 20, -4). Both participants demonstrated a structural connectivity of the FOP with the anterior temporal lobe via the fasciculus uncinatus (Figure 3, foreground). These connectivity profiles were also present in the two participants from the PSG group, providing additional support for the generality of this connection (Figure 3, background). The two participants from the PSG group had shown individual activation maxima in Broca's area (green) in the critical contrast incorrect vs. correct sequences ($P < .005$). Talairach coordinates of Participant 3: -46, 18, 21 and Participant 4: -49, 18, 4. These activations were taken as the seed points for individual tractograms. For both participants from the PSG group, tractograms with seed points in Broca's area for these participants indicated connectivity with the posterior and middle portion of the superior temporal region via the fasciculus longitudinalis superior (Figure 3, foreground). A look at the two participants from the FSG group (showing no activation in BA 44) indicates that these demonstrate very similar tractograms as those from the PSG group (Figure 3, background).

Thus for all four participants, two distinct connectivity profiles were found, one connecting Broca's area via the fasciculus longitudinalis superior to the posterior and middle temporal lobe and one connecting the FOP via the fasciculus uncinatus to the anterior temporal lobe. Given this structural similarity, the activation of Broca's area and the FOP thus is clearly a function of the input, namely the two different grammar types.

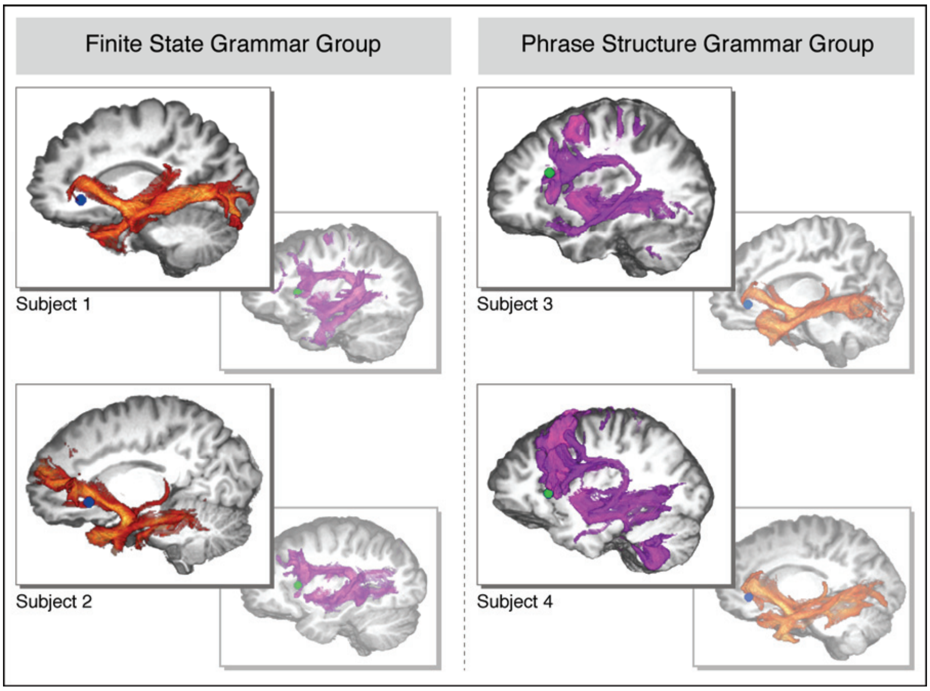

Figure 6.3: Tractograms for two brain regions: Broca's area and FOP. Orange: tractograms from FOP, purple: tractograms from Broca's area.

Left panel: For the FSG group, for which the FOP tracts (orange) is most relevant, Broca's area tracts (purple) are displayed.

Right panel: For the PSG group, for which the Broca's area tract (purple) is most relevant, FOP tracts (orange) are displayed. For additional information, tractograms of each participant are given for the respective other brain area. For these tractograms, the group mean activation voxels (see Table 6.1) were used as starting points.

6.5. Discussion

The data at hand suggest a functional differentiation between two brain areas in the left frontal cortex. The processing of violations of transitional dependencies in sequences generated by an FSG rule caused activity in the FOP. In contrast, the evaluation of violations in hierarchical dependencies in sequences generated according to a PSG rule, additionally correlated with activity in Broca’s area (BA 44/45). This result is in line with findings across different imaging studies showing that BA 44/45 is crucial for the processing of syntactical complex sentences in natural languages (Ben-Shachar et al., 2003; Caplan et al., 1998, 1999; Cooke et al., 2002; Friederici, Fiebach et al., 2006; Peelle et al., 2004; Stromswold et al., 1996), but not for the processing of local syntactic anomalies (e.g. Friederici & Kotz, 2003; Ni et al., 2000). In the present study a functional differentiation between Broca's area and the FOP was shown. While the involvement of the FOP was observed for both grammar types, the computation of the hierarchically structured sequences only activated Broca's area. It was hypothized that the FOP supports the check of the incoming element against the predicted element. Hence, this area might be involved during the processing of ungrammaticalities independent of the structure of the sequence. Broca's area is additionally recruited when hierarchical structures are to be processed. Moreover, the data suggest that it is not processing difficulty per se that activated Broca's area (no interaction effect of Performance x Violation). It rather represents a correlation with the presence of working memory demands in order to process structural hierarchies. However, it remains an open question to what extend the computations assigned to Broca's area and FOP are specific to the language domain or not.

In addition to functional data, DTI measurements were performed in a subgroup of participants and connectivity outlines for the two regions of interest (BA 44 and FOP) were calculated. A distinctive connectivity pattern was observed for both areas. The tractogram that captured the FOP as a seed point revealed connections to the anterior temporal lobe via the fasciculus uncinatus. In contrast, Broca's area is connected to the posterior and middle portion of the superior temporal region via the fasciculus longitudinalis superior.

For a detailed discussion of the role of Broca's area and FOP see Chapter 9.

Chapter 7

Experiment 3: A second fMRI Study

7.1. Abstract

Perruchet & Rey (2005) recently claimed that the rules used in the study by Fitch & Hauser (2004) can be processed by a simple counting mechanism as it does not produce center-embedded structures via corresponding categories in a syllable sequence. Regarding the proposal by Perruchet & Rey (2005) a different type of categorization was designed in the present study. Concatenations between categories of syllables were produced via corresponding sub-categories (e.g. FSG: $A_1B_1A_2B_2$, PSG: $A_2A_1B_1B_2$). Moreover, due to the within subject design in the study at hand a direct comparison between PSG and FSG sequences was conducted. Short sequences and long sequences were constructed, both in rule adequate and rule violating sequences. These manipulations resulted in a within subject design with the factors rule type (PSG versus FSG), length (short versus long) and violation (correct versus incorrect). As a result the crucial contrast between hierarchical (PSG) and linear (FSG) sequences caused activity in bilateral inferior frontal gyrus and the basal ganglia. These findings indicate that Broca's area is involved in the processing of hierarchical structures independent of whether these are part of a natural language or not.

7.2. The present study

The third study of this thesis was designed in order to directly compare hierarchical with linear structures. In the two previous studies the artificial grammar rule used by Fitch & Hauser (2004) were adapted and the processing of the respective rules were investigated in an ERP setting (see Chapter 5) and an fMRI setting (see Chapter 6). These studies differed from the behavioral experiment by Fitch & Hauser (2004) in several respects. First, the processing and not the learning of the rules was investigated in the ERP and the fMRI study. Learning of the rules took place several days prior scanning and participants entered the ERP/fMRI session in a proficiency state. Second, a visual and not auditory presented action mode was applied in the present work. The categories A and B where coded phonologically, and not by pitch of the voice. This coding made the A^nB^n rule of the present work more complex than in the behavioral study by Fitch & Hauser (2004). Third, the embeddings in the PSG where generated and parameterized by using different lengths of the syllable sequences. Short sequences (i.e. n = 2: AABB) showed less embedding of syllables of one category than long sequences (i.e. n = 4: AAAABBBB). In the previous fMRI experiment each grammar type (FSG and PSG) was learned by one participant group resulting in a between subject design. Two different regions of interest were correlated with the processing of the two grammar rules: The processing of FSG activated the FOP only, whereas the processing of the PSG engaged the FOP and in addition Broca's area. Therefore, it was suggested that Broca's area might be crucial for the processing of hierarchical structures in language.

Nevertheless it remains an open issue to what extend the processing of the PSG, as applied in the experiment by Fitch & Hauser (2004), necessarily relies on the generation of hierarchical representational structures. Perruchet & Rey (2005) claimed that the rule A^nB^n used in the study by Fitch & Hauser (2004) does not imply a hierarchical processing of the structure. Instead, the task could be solved via a simple counting mechanism.

In reference to Perruchet & Rey's (2005) claim the present study used a novel paradigm discouraging the application of a simple counting mechanism. In this paradigm two categories of syllables were produced (A and B). These categories A and B were further subdivided via plosives into A_1 - B_1, A_2 - B_2, and A_3 - B_3 categories (e.g. $A_3A_2A_1B_1B_2B_3$ = gi de be po tu ko). This categorization makes the built up dependency of corresponding syllable categories necessary (see Figure 7.1). Moreover, a generalization of the dependency should be induced, since not single items, but members of categories of syllable sequences produced the conjunction of syllables. Furthermore, the A^nB^n rule in the present study necessarily generates long distance dependencies between categories and center-embedding. The processing of this rule should induce the built up of long distance dependencies via hierarchical structure building. The processing of this hierarchical rule was directly compared with a linear rule (AB^n) in a within-subject design. The main question of the present study was, whether the processing of a hierarchical structure with dependencies between corresponding categories in comparison to a linear artificial grammar rule engages Broca's area. If so, evidence was herewith provided that the activation in Broca's area observed in the previous study is indeed due to the processing of hierarchical dependencies, and not to a simple counting mechanism. Moreover, such a result would indicate that Broca's area is responsible for the processing of hierarchical dependencies independent of whether these are part of a natural language or part of an artificial language.

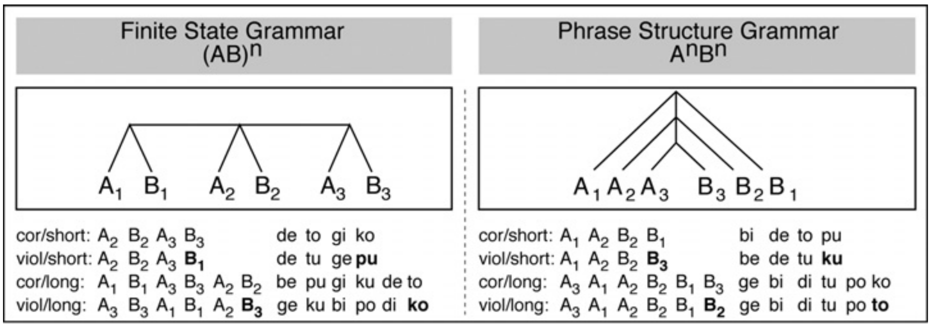

Figure 7.1: General structure and examples of the two rule types. FSG was generated by simple transitions between categories of consonant-vowel syllables. PSG was produced by embeddings between the two syllable categories. Short and long sequences were applied. Violations of the structure were situated at the last 3 or 4 positions (short sequences) and at the last 4, 5, or 6 positions (long sequences). In the example, the violations are placed at the third position for short sequences and at the sixth position for long sequences (bold letters).

7.3. Methods

7.3.1. Participants

Sixteen native German speaking participants participated in this study (9 male, mean age 25.5 years, SD = 3.7 years). They were right-handed and had a normal or corrected to normal vision. No participant has a known history of neurological, major medical, or psychiatric disorder. Prior scanning, participants were informed about the potential risks and gave a declaration of consent.

7.3.2. Stimuli

Sequences of consonant-vowel syllables were visually presented to the participants. Two types of categorizations of the syllables were applied. First, the syllables were assigned to two classes (A and B), which were coded by different vowels. Class "A" syllables ended with "e" or "i" (be, bi, de, di, ge, gi) and class "B" syllables ended with "o" or "u" (po, pu, to, tu, ko, ku). Second, conjunctions between pairs of A's and B's were coded by plosives:

$A_1 - B_1$ = b – p; $A_2 - B_2$ = d – t; $A_3 - B_3$ = g – k (see Figure 7.1). For both types of rules the same syllables were used. As in the previous studies, the probability of occurrence of the syllable frequency was balanced in order to prevent pattern learning.

A 2 x 2 x 2 design was applied with the within subject factors rule type (linear vs. hierarchical), length (short vs. long) and violation (correct vs. incorrect). In the hierarchical rule, the structure of the syllable sequence followed the rule A^nB^n. The local rule was generated via the formula $(AB)^n$. For both rule types, short sequences (n = 2) and long sequences (n = 3) were applied and correct and violated sequences were used (50% each).

Table 7.1: Structure of the design. Short versus long and correct versus incorrect sequences were applied in both rule types. Violations are shown in bold and underlined symbols.

	violation	FSG	PSG
short	correct	$A_1B_1A_2B_2$	$A_1A_2B_2B_1$
	incorrect	$A_1B_1A_2\underline{\mathbf{A_2}}$	$A_1A_2B_2\underline{\mathbf{B_3}}$
long	correct	$A_1B_1A_2B_2A_3B_3$	$A_1A_2A_3B_3B_2B_1$
	incorrect	$A_1B_1A_2B_2A_3\underline{\mathbf{A_3}}$	$A_1A_2A_3B_3\underline{\mathbf{B_1}}B_2$

PSG

An example for an embedded structure (PSG) is the following: $A_3A_2A_1B_1B_2B_3$ = gi de be po tu ko. Though, this structure could also be learned by simply memorizing the syllables. In order to prevent this pattern learning, the positions of the categories in a sequence were changed: $A_3A_2A_1$ $B_1B_2B_3$; $A_2A_1A_3$ $B_3B_1B_2$; etc. Six combinations of the variation of the positions are possible. Furthermore $2^6 = 64$ combinations of the permutations of the categories A(e, i) $\wedge$ B(o, u) were possible. Hence, 64 x 6 = 384 items for the different syllable sequences were generated in order to force the learning of the rule instead of learning of the syllable pattern. The testing of the correct processing of the embedded structure was implemented by introducing errors into the structure. In order to prevent the focusing of the participants at one particular position of the sequence, the errors were placed at different positions of the syllable sequence (as in the previous experiments). In contrast to the two former studies of the present work, violations were placed at one of the last 2 positions in short sequences (AA**BB**) or at one of the last 3 positions in long sequences (AAA**BBB**).

FSG

An example for a FSG structure is the following: $A_3B_3A_2B_2A_1B_1$ = ge ko de to be po). As in the hierarchical rule, the order of the sub-categories was changed systematically: $A_2B_2A_1B_1A_3B_3$; $A_1B_1A_3B_3A_2B_2$; etc. The variation of the positions (6 variations) and the combination of permutations of the categories (64 items) resulted in a total amount of 384 possible items for the syllable sequences. Thus, in the local and in the hierarchical rule the same amount of items were used. In order to realize a direct comparison of the two rule types, errors of the local rule were also included at one of the last 2 positions in short sequences ($\mathrm{AB}\underline{\mathbf{AB}}$) or at one of the last 3 items of the long sequences ($\mathrm{ABA}\underline{\mathbf{BAB}}$).

Violations

For both rule types, two different classes of violations were placed at the last positions of a sequence. The consonant-vowel violation represents the first type of error. In this case, class A and B syllables were violated (e.g. FSG, short sequence: $A_1B_1A_2\underline{\mathbf{A_2}}$; PSG, long sequence: $A_1A_2A_3B_3\underline{\mathbf{A_2}}B_1$). The second type of error violated the plosive-concatenations (e.g. FSG, long sequence: $A_1B_1A_2B_2A_3\underline{\mathbf{B_2}}$; PSG, short sequence: $A_1A_2B_2\underline{\mathbf{B_3}}$). This manipulation was done in order to force the participants to process both types of categorization. The type of violation (A => B or A_x => B_y), the positions of errors within a sequence (position 4, 5, or 6 in long sequences and position 3 or 4 in short sequences), the length of the sequences (short sequences = 4 syllables, long sequences = 6 syllables), and the amount of violations (50% correct) were counter balanced over the rule types.

Table 7.2: Overview of types of errors used in the present study. A violation was included in each type of categorization. Violations are indicated in bold and underlined letters.

	Error type	FSG	PSG
short	B => A	$A_1B_1A_2\underline{\mathbf{A_2}}$	$A_1A_2B_2\underline{\mathbf{A_3}}$
		$A_1B_1\underline{\mathbf{B_3}}B_2$	$A_1A_2\underline{\mathbf{A_2}}B_1$
	B_x => B_y	$A_1B_1A_2\underline{\mathbf{B_1}}$	$A_1A_2B_2\underline{\mathbf{B_3}}$
		$A_1B_1\underline{\mathbf{A_1}}B_2$	$A_1A_2\underline{\mathbf{B_3}}B_1$
long	B => A	$A_1B_1A_2B_2A_3\underline{\mathbf{A_3}}$	$A_1A_2A_3B_3B_2\underline{\mathbf{A_1}}$
		$A_1B_1A_2B_2\underline{\mathbf{B_1}}B_3$	$A_1A_2A_3B_3\underline{\mathbf{A_2}}B_1$
		$A_1B_1A_2\underline{\mathbf{A_2}}A_3B_3$	$A_1A_2A_3\underline{\mathbf{A_3}}B_2B_1$
	B_x => B_y	$A_1B_1A_2B_2A_3\underline{\mathbf{B_2}}$	$A_1A_2A_3B_3B_2\underline{\mathbf{B_3}}$
		$A_1B_1A_2B_2\underline{\mathbf{A_2}}B_3$	$A_1A_2A_3B_3\underline{\mathbf{B_1}}B_1$
		$A_1B_1A_2\underline{\mathbf{B_3}}A_3B_3$	$A_1A_2A_3\underline{\mathbf{B_1}}B_2B_1$

7.3.3. Procedure

Two experimental sets were performed. One set comprised the learning and testing of the FSG, in the other set the PSG was trained and inquired. Hence, each participant learned and processed both of the rule types. The sets were separated by 7 days. Each experimental set was subdivided into learning and a testing session.

Learning

The learning of the rules took place 2 days prior scanning. The learning period was subdivided into 4 parts which consisted of several blocks with different sequence lengths. First, sequences with 2 items were learned. At the beginning of each block, 10 correct sequences were presented. Afterwards, 5 correct and 5 incorrect sequences were shown for which participants were required to respond with a button press whether the sequences were grammatical or ungrammatical. For each sequence, feedback was given. Participants were instructed to extract the rule underlying the syllable sequences. The first part of the training ended when participants answered 90% of the items correctly in 2 successive blocks for the sequences with 2 items. Thereafter sequences with 4 items were learned with the same procedure as with 2 items. Again, when participants reached a level of 90% correct answers in 2 successive blocks, the third level of learning was entered. In the third part of learning, participants were inquired to learn sequences with 6 items until 90% of the items were answered correctly in 2 successive blocks. In the last part of the learning, 60 sequences with 4 and 6 items were presented randomly. Half of them were violated and participants were instructed to judge these sequences.

Testing

During testing of the rules fMRI measurements were conducted. Participants were instructed to judge whether the sequences were rule-based or not. As in the training, feedback was given. For the testing period, 144 new sequences were presented, i.e. 72 short sequences (four syllables) and 72 long sequences (six syllables) were computed, half of them (36 for each length) were violated. The 36 violated sequences in each length consisted of 18 plosive-concatenation errors and 18 consonant-vowel errors. In the short sequences half of the errors (9 sequences each error type) were situated at the third position and half of them were included at the fourth position. In the 36 long violated sequences one third (6 sequences each error type) were included at the fourth position, on third at the fifths, and on third at the sixth position. The start of a sequence was indicated by a fixation cross (500ms). Each syllable was presented for 800 ms with an inter-stimulus interval of 200 ms between the syllables. Participants could make their judgment for 1000 ms, which was followed by feedback for 500 ms. Afterwards a fixation cross was shown for 1000ms. Trials started with a jitter of 0, 500, 1000, or 1.500 ms.

Key assignment, gender of the participants, and order of the grammar type (linear rule first versus hierarchical rule first) was counterbalanced across participants.

7.3.4. fMRI Data Acquisition

Imaging was performed at 3T Trio scanner (Siemens, Erlangen, Germany) equipped with the standard birdcage head coil. Stabilization cushions were used in order to reduce head motion. For registration purpose, two sets of two-dimensional anatomical images were acquired for each participant immediately prior to the functional imaging. An MDEFT and an EPI-T1 sequence were used. T1-weighted MDEFT (Ugurbil et al., 1993) images (data matrix 256 x 256, TR = 1.3 s, TE = 7.4 ms) were obtained with a non slice-selective inversion pulse followed by a single excitation of each slice (Norris, 2000). In addition, a set of T1-weighted spin-echo EPI images (TE 14 ms, TR 3000 ms) were taken with the same geometrical parameters (slices, resolution) and the same bandwidth as used for the fMRI data. A slice-selective inversion pulse was applied with an inversion time of 1200 ms. Anatomical images were positioned parallel to AC-PC. Functional data were acquired from 16 axial slices (thickness = 3 mm; gap = 0.6 mm) using a gradient-echo EPI sequence with a TE = 30 ms, flip angle = 90°, TR = 2000 ms, acquisition bandwidth = 100 kHz. The matrix acquired was 128 x 128 with a FOV of 25.6 cm, resulting in an in-plane resolution of 2 mm x 2 mm. One functional run with 780 time points was measured. Due to technical problems, two of the 16 participants had to be excluded from the analysis.

7.3.5. Functional Imaging Data Analysis

The fMRI data processing was performed using the software package LIPSIA (Lohmann et al., 2001). Functional data were motion-corrected offline with the Siemens motion correction protocol (Siemens, Erlangen, Germany). To correct for the temporal offset between the slices acquired in one scan, a cubic-spline-interpolation was applied. A temporal high-pass filter with a cutoff frequency of 1/30 Hz was used for baseline correction of the signal and a spatial Gaussian filter with 5.65 mm FWHM was applied. To align the functional data slices with a 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was performed. The rotational and translational parameters were acquired on the basis of the MDEFT and EPI-T1 slices to achieve an optimal match between these slices and the individual 3D reference data set. The MDEFT volume data set was standardized to the Talairach stereotactic space (Talairach & Tournoux, 1988). The rotational and translational parameters were subsequently transformed by linear scaling to a standard size. The resulting parameters were then used to transform the functional slices using trilinear interpolation, so that the resulting functional slices were aligned with the stereotactic coordinate system. This linear normalization process was improved by a subsequent processing step that performed an additional nonlinear normalization.

Statistical evaluation

The statistical evaluation was based on a least-squares estimation using the general linear model dealing with serially autocorrelated observations using pre-whitening of each voxel's time series (Worsley et al., 2002). The design matrix was generated with a box-car function, convolved with the hemodynamic response function. The model equation, including the observation data, the design matrix and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM to deal with the temporal autocorrelation. In the following, contrast-images, i.e. estimates of the raw-score differences between the six conditions, were generated for each participant. The single-participant contrast-images were then entered into a second-level random effects analysis for each of the contrasts. The group analysis consisted of a one-sample t-test across the contrast images of all participants that indicated whether observed differences between conditions were significantly distinct from zero. Subsequently, t-values were transformed into Z-scores. To protect against false positive activations, only regions with z-score greater than 3.09 ($p < .001$), with a corrected spatial extent threshold of $p < .05$, and with a volume greater than 96mm^3 (12 measured voxels) were considered.

This non-arbitrary voxel cluster size was determined by using the program AlphaSim (http://afni.nimh.nih.gov/afni/doc/manual/AlphaSim), which provides a mean of estimating the probability of false detection ($p < .05$). Figures show the resulting zmap overlaid onto an anatomical reference image.

Time course analysis

The time course analysis of the region-of-interest (ROI) was performed in the voxels that yielded the highest activation in the contrast PSG versusFSG (main effect rule type). In order to detect a possible functional influence of the hemisphere, right-hemisphere homologues were also included in the analysis. Hence, ROI's were situated in the left inferior frontal gyrus (-46, 4, 16; BA44/6), the right ventral premotor cortex (46, -2, 16; BA6), the left anterior insula (-28, 18, 4), and the right anterior insula (32, 14, 4). Since the length of the stimuli was unequal (long sequences: 6 seconds; short sequences: 4 seconds) and to facilitate a possible latency effect of the BOLD response, time courses were analyzed at four different time points. Time points 5, 7, 9, and 11 seconds were chosen, since the hemodynamic response is known for a temporal delay (approximately 2-3 seconds) and an acquisition time of 2 seconds was used in the present study. Hence, starting at 5 seconds after stimulus onset, 4 time points every 2 seconds were selected (5, 7, 9, and 11). An ANOVA on the hemodynamic response at the specified ROI's was conducted with the within subject factors hemisphere (right versus left hemisphere), rule type (PSG versus FSG), violation (correct versus incorrect sequences), length (short versus long sequences), and time step (5, 7, 9, and 11 seconds after stimulus onset). The Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was always applied when evaluating effects with more than one degree of freedom in the numerator. In such cases, the uncorrected degrees of freedom (df) and the corrected p-values, and the correction factor ε are reported.

The Bayesian statistics

In a recent paper Poldrack (2006) suggested a new type of analysis of non-expected activity in fMRI experiments. In most cases of brain imaging studies the non-hypothesized brain activity in a study was compared with experiments reporting also these areas and the same cognitive process correlated with this activity was implicated. The author argued that these “reverse inference” statements are somewhat week from the position of propositional logic, since it is not deductively valid. A way out of this “reverse inference” problem could be a probabilistic approach to non-expected activity. The selectivity of a particular brain region to a specific cognitive process was estimated via the Bayes’ statistics calculated using the BrainMap database (Laird, Lancaster, & Fox, 2005, http://brainmap.org). This procedure was adapted in the present study. For each activated brain region, a ROI with 20 mm width was defined in the BrainMap database. Next, a search was performed for experiments in the language domain that reported activity in these ROI’s. Subsequently, the same locations were linked with all experiments not in the language domain and a search was conducted in the same ROI. The resulting posterior probability represents an estimate of the language-relatedness of the particular brain region (see Table 7.3). A posterior greater than 0.75 reflects evidence for confidence over the prior. The conditional probability of activity in language studies $P(Z|X)$ was calculated from the number of language experiments reporting activity in the ROI (Z), divided by the total number of language experiments (X) contained in the database (899 in February 2006). The conditional probability of non-language studies was computed by dividing the number of studies with activation by the total number of studies (2254 in February 2006). The prior probability P(X) was set to 0.5, since no prior hypothesis exists about the particular brain region (one is equally confident that language processes are engaged or not). The posterior probability to language-relatedness was calculated as follows:

$$\mathrm{P}\,(\mathrm{X}\,|\,\mathrm{Z}) = \frac{P(Z\,|\,X)P(X)}{P(Z\,|\,X)P(X) + P(Z\,|\!\sim X)P(\sim X)}$$

7.4. Results

7.4.1. Behavioral results

The learning of the rules took about 32 minutes. In order to investigate, whether the learning of one of the two rule types takes longer than the other a pared t-test was conducted at which the amount of learning-blocks that were needed to reach the criteria was analyzed. No significant differences of the learning time was found (t(13) = 1.77, n.s.). Also no session effect was found: the amount of learning-blocks in the first session did not significantly differ from the second session (t(13) = 0.31, n.s.). Moreover, the rule type is not a predictor of the learning time in the first or the second session, i.e. no transformation effect was found (F(1, 12) = 1.48, n.s.). To investigate the performance in the test session an ANOVA on the error rates was conducted, including the within subject factors rule type (PSG versus FSG), length (short versus long sequences), and violation (correct versus incorrect sequences). A significant main effect of rule type was found (F(1, 13) = 12.97, n.s.), indicating that the PSG induced more errors (5.14%, SD = 3.4) than the FSG (2.16%, SD = 2.5). No other effect reached significance.

7.4.2. fMRI Results: Activation Patterns

In the present study a 2 x 2 x 2 design was applied with the factors rule type (PSG versus FSG), violation (correct versus incorrect sequences), and length (short versus long sequences). In order to investigate the three main effects, contrast images for rule type, violation, and length were computed. For details see Table 7.3 and Figure 7.2.

The main effect of rule type (PSG versus FSG) yielded activity in Broca’s area. Additional activation was found in the bilateral anterior insula, the left inferior frontal sulcus, left fusiform gyrus, and bilateral subcortical activations (including caudate nucleus). The opposite contrast (FSG versus PSG) revealed activity in the bilateral superior temporal sulcus and the left hippocampal gyrus. The BOLD contrast of violation (incorrect versus correct) elicited several bilateral activations, i.e. in the anterior insula, superior temporal sulcus, inferior occipital gyrus, cerebellum, thalamus, and in the right cingulate gyrus. The contrast of correct versus incorrect sequences displayed activation in the left cingulate gyrus. The length contrast (long versus short) correlated with activity in the bilateral anterior insula. In the opposite contrast (short versus long) activity in the right frontomedial gyrus, the bilateral posterior insula, bilateral hippocampus, and the left retrosplenial cortex was observed.

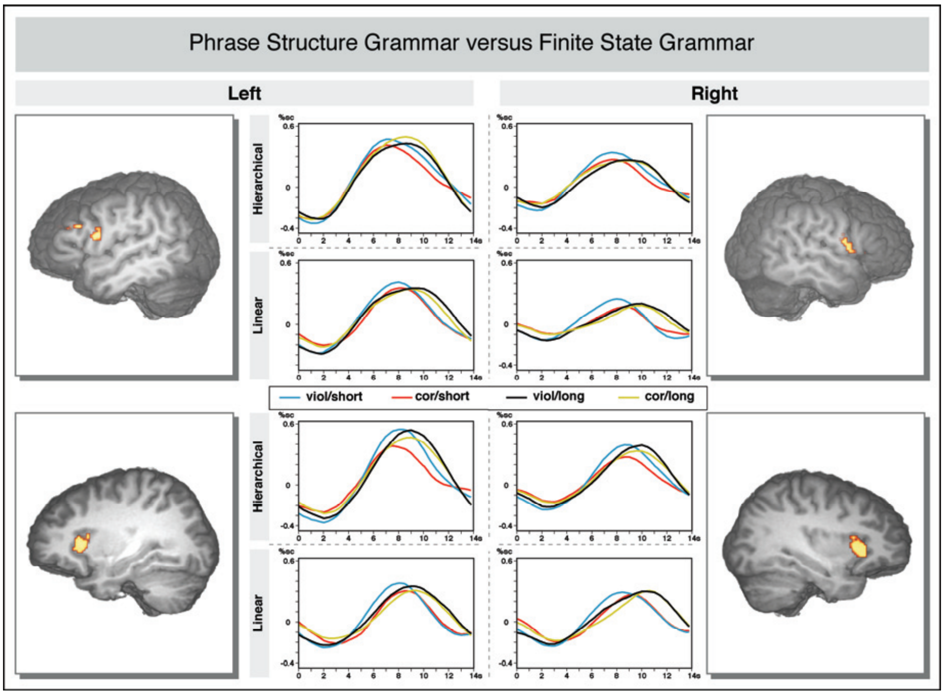

Figure 7.2.: BOLD response on the main effect of PSG versus FSG in BA 44/6 (upper panel) and anterior insula (lower panel). Left-hemispheric activity on the left side of the plot; right-hemispheric activity on the right side.

Middle panel: Time courses of the different variables obtained by the ROI's in BA44/6 (upper panel) and anterior insula (lower panel). Hierarchical sequences show a higher BOLD response than linear sequences in bilateral BA44/6 and in bilateral anterior insula, independent of violation (correct or incorrect) and of the length of the sequences (short or long). Left-hemispheric ROI's revealed a higher hemodynamic response in comparison to right-hemispheric ROI's in BA44/6 and in anterior insula, independent of rule type (hierarchical or linear), violation (correct or incorrect), or length (short or long).

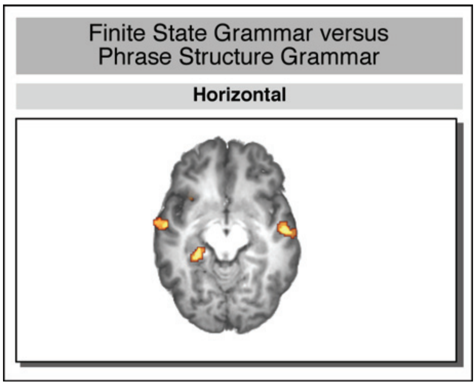

Figure 7.3.: Brain activation pattern revealed by the contrast FSG versus PSG sequences.

Table 7.3a: Anatomical areas, approximate Brodmann's Area (BA), mean Talairach coordinates (x, y, z), maximal z values, frequency of experimental comparisons in language-related studies (L) and not language-related studies (NL) as searched in the BrainMap database, and posterior probability of the activated area (P(X|Z)) of the significant activations in the main effect rule type.

		Stereotactic coordinates						
Brain region	BA	x	y	z	z_{max}	L	NL	P(X|Z)
PSG versus FSG								
L inferior frontal G	44/6	-46	4	16	$3.72^{\#}$	239	190	$.76^{*}$
R ventral premotor C	6	46	-2	16	3.70	129	132	.71
L anterior Insula		-28	18	4	4.23	182	220	.67
R anterior Insula		32	14	4	4.54	126	232	.57
L inferior frontal S	45/9	-34	28	22	3.80	154	166	.69
L fusiform G	37	-42	-56	-18	4.20	199	161	$.75^{*}$
L fusiform G	18	-22	-98	-18	4.07	40	19	$.84^{\$}$
R caudate head		14	12	6	4.08	61	97	.61
L globus pallidus		-16	-6	4	4.38	99	187	.57
FSG versus PSG								
L superior temporal S	21	-58	-7	-8	3.87	101	67	$.79^{*}$
R superior temporal S	21	52	-11	-8	4.14	98	74	$.77^{*}$
L Hippocampus G	37	-24	-33	-8	4.51	65	73	$.69^{\$}$

Notes.

* … posterior probability provides evidence for moderate confidence over the prior, i.e. the cognitive domain of language processing is specifically engaged in this particular brain region.

$... less than 5% of the overall experimental comparisons reported activity in the ROI of this brain region, i.e. L + NL < 5% (threshold). L = 899, NL = 2254: 899 + 2254 = 3153, 5% of 3153 = 157. Hence no prediction can be implicated over the posterior ROIs with less than 157 experimental comparisons.

#...pcorr = .08

L, left hemisphere; R, right hemisphere; G, gyrus; S, sulcus; Ncl, nucleus

Table 7.3b: Main effect violation.

		Stereotactic coordinates						
Brain region	BA	x	y	z	z_{max}	L	NL	P(X|Z)
incorrect versus correct								
R anterior Insula		32	20	4	5.78	138	236	.59
L anterior Insula		-32	18	0	5.83	185	229	.67
L superior temporal S	22/42	-42	-46	14	5.31	110	70	.79[*]
R inferior occipital G	18	38	-84	-8	5.06	103	112	.69
L fusiform G	37	-38	-50	-14	4.80	191	149	.76[*]
L Thalamus (lateral posterior Ncl)		-8	-22	16	5.00	84	133	.61
R Thalamus (ventral lateral Ncl)		12	-10	12	4.33	50	162	.43
R Cerebellum		12	-68	-20	4.88	99	118	.67
correct versus incorrect								
L cingulate G	32	-6	19	-8	4.39	55	54	.71[$]

Notes. See Table 7.3a

Table 7.3c: Main effect length.

		Stereotactic coordinates						
Brain region	BA	x	y	z	z_{max}	L	NL	P(X|Z)
long versus short								
L anterior Insula		-26	18	10	4.23	147	176	.67
R anterior Insula		28	18	4	5.09	150	213	.63
short versus long								
R frontomedial G[§]	32	6	41	2	6.21	42	95	.52[$]
L posterior insula		-36	-14	16	5.86	36	108	.45[$]
R posterior insula		34	-18	12	5.44	34	81	.52[$]
L Thalamus (posterior Ncl)		-16	-36	6	5.08	27	64	.51[$]
L hippocampus		-26	-28	-6	5.42	40	85	.54[$]
R hippocampus		26	-26	-8	5.24	24	74	.44[$]
L retrosplenial C	30	-2	-54	12	5.07	36	75	.55[$]
L Thalamus (dorsal Ncl)		-2	-15	10	3.94	51	195	.39

Notes. See Table 7.3a

§… since the cluster of activated voxels (calculated with the threshold of pcorr = .05 and 12 voxel per cluster) was too big, a more conservative threshold was applied and the resulting brain regions were reported below.

7.4.3. fMRI Results: Time course

In order to specify the possible influence of the different factors (violation, length, and hemisphere) on the activity detected for the rule type in the inferior frontal gyrus and in the anterior insula, a time course analysis was conducted in these brain regions. Since for both regions (inferior frontal gyrus and in the anterior insula) the right-hemisphere homologous were also identified, these regions were included in the ANOVA.

Inferior frontal gyrus: The main analysis in the inferior frontal and ventral premotor cortex (BA44/6) revealed a marginally significant 5-way interaction of hemisphere, rule type, violation, length, and time step (F(3, 39) = 2.78, ε = 0.58, p = .08). On the basis of this interaction separate analyses were performed for the factors hemisphere, rule type, violation, and length. The pared t-Test on hemisphere (t(13) = 2.41, p < .05) illustrated a general higher BOLD response for the left sided activation in comparison to the right-hemisphere activity. The higher hemodynamic response to PSG sequences versus FSG sequences in Broca's area was indicated by the significant rule type effect (t(13) = 5.20, p < .001). Also the paired t-Test of violation showed a significant effect (t(13) = 3.27, p < .01), pointing toward a higher activity for incorrect versus correct sequences. The length effect was only marginally significant (t(13) = 1.86, p = .08), showing a slightly increased BOLD response on long versus short sequences. An additional analysis for short versus long sequences in each of the rule type, however revealed no length effect in either of the rules (FSG: t(13) = 0.98, n.s.; PSG: t(13) = 1.56, n.s.).

Anterior insula: The main analysis in the anterior insula showed a marginally significant 4-way interaction of hemisphere, rule type, violation, and length (F(1, 13) = 3.95, p = .06). On the basis of this interaction separate analyses (paired t-Tests) on each of the factors were carried out. The factor hemisphere showed a significant effect (t(13) = 2.58, p < .05), displaying a higher activity in the left anterior insula in comparison to the right anterior insula. Also the paired t-Test of rule type was significant (t(13) = 3.68, p < .001), indicating a higher BOLD response in the left and right anterior insula for PSG versus FSG structures. Incorrect sequences also showed a higher activity in comparison to correct sequences, independent of rule type, length, or hemisphere, revealed by an significant violation effect (t(13) = 8.09, p < .001). As in the analysis on BA44/6, the effect of length was only marginally significant (t(13) = 1.94, p = .07).

7.5. Discussion

The present study investigated the brain basis of processing hierarchical structures, as one of the core features of the human language faculty by directly comparing hierarchically structured syllable sequences (PSG) with linearly constructed sequences (FSG). As hypothesized, the processing of the hierarchical rule engaged Broca's area more than the processing of the linear rule. Moreover, several prefrontal, temporal, and subcortical brain areas were additionally correlated with the three main effects, tested by a 2 x 2 x 2 design with the within-subject factors rule type (PSG versus FSG), length (short versus long sequences), and violation (correct versus incorrect sequences).

Additionally, in the present study a new approach to the analysis and interpretation of non-expected/not hypothesized brain areas was applied. Recently, Poldrack (2006) suggested a Bayesian method to deal with non-expected brain areas in fMRI experiments. The author proposed a quantification of the handling of the activation patterns; instead of simply infer the cognitive process underlying the activation from other studies that also found this specific brain area active. In the present study the *posterior probability* of language-relatedness of each activated brain region was calculated using the BrainMap database (see Table 1). A high *posterior probability* of a specific brain region represents an indicator that language processes are highly correlated with this area. This additional analysis represents an attempt to quantify the non-hypothesized brain regions of the present study while deriving related functions from a database of experiments, not only from single published studies. Note, that this post-hoc analysis was not employed as a filter in order to emphasize the (hypothesized) role of Broca's area in the experiment at hand. Instead, this approach was used to deal with the non-expected brain regions found to be activated, while clustering these areas according to the experiments published in the BrainMap database. Alternatively, one could argue that the term language-relatedness represents such a broad range of experiments that is hardly to compare with the specific domain of rule-based computations in the present study. However, as every psychophysiological experiment, the present study does not only involve brain regions associated with rule-based processing. The processing of the artificial grammar task rather engages a number of cognitive processes that may or may not be part of the language system. Hence, in order to classify the activated brain regions (at least) according to the cognitive domain of language processing the Bayesian approach denotes an adequate method.

The following discussion covers only the non-expected activations in fusiform gyrus and the subcortical areas. For detailed discussion of Broca's area, anterior insula, and superior temporal regions see Chapter 8.

7.5.1. The fusiform gyrus

The left fusiform gyrus was activated in the main effect of rule type (PSG versus FSG) and in the main effect of violation (incorrect versus correct sequences). The rule type effect revealed two different areas to be activated in the left fusiform gyrus, namely in BA 18 and in BA 37, whereas the violation effect activated only BA 37. Interestingly, the posterior probability of both ROI's in BA 37 for the rule type effect and for the violation effect were considerably high (ROI rule type: 0.75; ROI violation: 0.76), suggesting a specificity of BA 37 to language-linked experimental conditions. The posterior probability in BA 18 could not be interpreted, because the amount of studies in the database reporting activity in this region was below the threshold of 5% of all studies (see Table 1). Moreover, a further analysis of the database in the ROI of BA 37 exposed that 180 of the 199 experimental conditions in the language domain used visual presentation as the stimulus modality requiring a letter-to-phoneme conversion. A recent meta-analysis using the activation likelihood estimation (ALE) method suggested that the left fusiform gyrus was uniquely associated with varying phonological processing demands on the syllable or phoneme level (Tan, Laird, Li, & Fox, 2005). The expression *phonological demand* was often associated with the difference in reading pseudowords versus words. Also, the computation of the pronunciation should be more demanding during the processing of pseudowords in comparison to words (Dietz, Jones, Gareau, Zeffiro, & Eden, 2005).

In the present study in which senseless syllable sequences were presented visually BA 37 was correlated with the factors rule type and violation. That is, this area was more activated for hierarchically than for linearly formed senseless syllable sequences (e.g. 'de gi ku to' versus 'de to gu ku'), and likewise, it showed more activation for syllable sequences that do not follow either of the rules than rule-based sequences (e.g. 'bi de to ku' versus 'gi de to ko'). In these syllable sequences all syllables represent pseudowords, since no known word can be decoded. Hence, the present data suggest a widening of the interpretation of the cognitive processes underlying the activity in BA 37. The activation in BA 37 may not only reflect the phonological demands during letter-to-phoneme integration when processing pseudowords compared to words. BA 37 also seems to reflect increased phonological demands during the processing of more complex (hierarchically) structured syllable sequences and in syllable sequences in which the structure violates the participants' expectation. Hence, it is suggested that activation in the left fusiform gyrus may be revealed by the higher phonological demands of hierarchical and incorrect sequences.

7.5.2. Subcortical areas

The processing of PSG sequences in comparison to FSG sequences engaged the left globus pallidus and the right caudate head and the main effect of violation revealed thalamic activations. The caudate nucleus was also activated in other artificial grammar learning experiments (Forkstam & Petersson, in press; Lieberman et al., 2004; Peigneux et al., 2000). Peigneux et al. (2000) suggested that the striatum is involved in the implicit automatization of serial information and in the selection of the most appropriate response during motor-linked cognitive learning. Lieberman and colleagues (2004) pointed that the caudate nucleus may be recruited during the application of implicitly learned rules in artificial grammar tasks. The increased activity in the caudate nucleus in the experiment conducted by Forkstam et al. (in press) was suggested to reflect automatic integration processes of the artificial grammar strings. Taking together, the caudate nucleus activity was interpreted as automatization, application, or integration of the processing of rule-based sequences. In contrast to these findings, neuropsychological studies demonstrated that patients with striatal impairment such as Huntington's disease or Parkinson's disease showed normal performance in artificial grammar tasks (Knowlton & Squire, 1996; Reber & Squire, 1999; Smith, Siegert, McDowall, & Abernethy, 2001; Witt, Nuhsman, & Deuschl, 2002). These findings suggest that the striatum is not necessarily engaged in artificial grammar learning. Nevertheless, caudate nucleus may be sufficiency correlated with the processing of rule-based structures. Ullman and colleagues (1997) indeed reported impairments in rule application during verb inflection for Huntington's and Parkinson's disease patients, but not for patients with Alzheimer's disease. These results are in line with the Ullman model (2004) proposing a distinction between a frontal-striatal circuit (procedural memory) and a temporal lobe circuit (declarative memory), with the former underlying the use of grammatical rules, the latter being implicated in retrieving stored information. In contrast, Teichmann et al. (2005) proposed a domain-general rather than a domain-specific involvement of the striatum during the application of rules. The Bayesian statistic on the BrainMap database revealed a low *posterior probability* of language relatedness for the ROI in the right caudate head, the left globus pallidus, and the bilateral thalamus. These results are in line with the heterogeneous findings of artificial grammar experiments with patient and with neurologically healthy subjects. The latter ones reported activity in the caudate nucleus. Patients suffering from Parkinson's or Huntington's desease would be expected to show impairments in functions that rely on the striatum and thus should have difficulties with an artificial grammar task, which was - in contrast - not the case.

7.5.3. Summary

In the present study, a direct comparison of hierarchical versus linear structures was conducted using visually presented syllable sequences in an artificial grammar task. As a result, Broca’s area and its right hemispheric homologue were significantly more engaged during the processing of the hierarchical structure in comparison to the linear structure. Moreover, several other language-related brain areas were activated, suggesting a close link between the neural correlates of natural language processing and the structures used in the artificial grammar task.

Chapter 8

General Discussion and Summary

The present work aimed to identify neural correlates of the processing of two different types of artificial grammars. In a recent study, Hauser & Fitch (2004) found that humans and non-human primates differ in their ability to learn and generalize artificial grammar rules. Humans easily master both, a complex hierarchical and a simple linear artificial grammar rule. Non-human primates also master a linear rule. In contrast, they fail to generalize a hierarchical rule. In the present thesis it was questioned, which brain regions contribute to the processing of these two different artificial grammar rules. Moreover, it was examined whether the processing of the artificial grammars share common neural networks as during the processing of natural language. In order to answer these questions, three different experiments were conducted. Different methods (ERPs and fMRI) were used and the structure of the rules was systematically varying over the experiments. ERP measurement was applied in order to investigate the electrophysiological correlates during the computations of the rules. In addition, due to the high temporal resolution of this method, ERPs were assessed at different stages (i.e. early and late positions) during the processing of the artificial grammar sequences. Metabolic changes in the brain were estimated with the aim of identifying brain regions correlated with the computations of the two different rule types.

In the present work a comparison between the processing of linear (FSG) and hierarchical (PSG) rules was conducted. These rules were derived from the behavioral study by Hauser & Fitch (2004). Meaningless consonant-vowel syllables were used and separated into different categories that followed either of the rules. As a result, a late positivity for violations of both structures (FSG and PSG), with an amplitude that varied as a function of rule type was found. Violations of the PSG rule engaged Broca's area more than violations of the FSG rule. Moreover, the direct comparison between PSG and FSG sequences also correlated with activity in Broca's area.

Despite of the behavioral study mentioned above, the processing of these types of artificial grammar rules has not been investigated so far. Hence, a comparison between artificial grammar rules used in previous studies and the rules employed in this work will now be given. Behavioral studies applying an artificial grammar task were commonly motivated by the theoretical constructs of implicit learning or statistical learning (see e.g. Perruchet & Pacton, 2006 for an overview). Many ERP and fMRI studies were in line with the theoretical motivations of these behavioral studies. Eimer et al. (1996) or Rüsseler et al. (2003) applied for instance a SRT task in an ERP setting and reported a broadly distributed early negativity for deviants. Also many fMRI studies used derivations of a Reber grammar (e.g. Fletcher et al., 1999; Lieberman et al., 2004; Seger et al., 2000) and found several brain regions to be activated (e.g. prefrontal, subcortical, parietal areas). Some experiments on the other hand, used artificial grammar rules that were more similar to natural languages. The artificial grammar by Friederici et al. (2002) and Opitz et al. (2004) featured pseudowords in different word classes and syntactical rules. Friederici et al. (2002) reported an early negativity and a late positivity for violations of the rules. Opitz and colleagues (2004) found an involvement of the left IFG during rule acquisition. Musso and co-workers (2003) used minimal natural grammar rules. The authors reported activity in the left IFG during processing of "natural rules" versus "un-real rules" (not present in natural language). Hence, a differentiation between language-like and non language-like artificial grammar rules could be implicated (as described in Chapter 3).

In the present work, learning of the rules was managed with feedback. Participants learned the rules by trial and error. Hence, the learning process was not implicitly as in studies applying a Reber grammar. The learning could therefore be described as an explicit problem solving. Hence, the present work could not be embedded in the "traditional" theoretical frameworks of artificial grammar learning, namely implicit or statistical learning. It was assumed that the inter-individual differences of learning performance could be a major confound of the ERP and fMRI results in the present work. I.e. if participants learn the same rule differently then the given feedback would vary across participants. This would lead to varying motivational stages that cause activity variations in particular brain regions across participants. In other words, motivation as an additional cognitive component should be avoided. Moreover, recent work (e.g. Opitz et al., 2003, 2004) has already shown that the hippocampus is engaged during early stages of learning of artificial grammar rules. Thus, neural correlates of learning of the rules were not from particular interest in the present work. For this reason, ERP and fMRI measurements were accomplished during the processing of the rules in a proficiency state, not during learning of the rules. Due to the types of structure used and due to the ERP and fMRI results, the artificial grammars of the present work could regard as a member of language-like artificial grammar rules.

8.1. The role of Broca's area

Broca's area was associated with various language related phenomena, including syntax, semantic, and phonology (see Bookheimer, 2002 for an overview). Semantic and phonological features were held constant in all conditions of the present experiments. Hence, these aspects of language processing could be excluded to be responsible for the activity in Broca's area in the work at hand. During natural language processing the increase of activation in the left IFG after syntactic manipulations was interpreted as correlate of higher syntactic complexity (Caplan, 2001), higher syntactic working memory load (Fiebach, et al., 2005), or costs due to movement in hierarchical structures (Ben-Shachar et al., 2003; Grodzinsky, 2000). Furthermore, procedural memory (Ullman, 2004), unification (Hagoort, 2005), or hierarchical structure building (Friederici, 2004) have been discussed as functional underpinnings. Caplan et al. (2001) reported increased activity in Broca's area for sentences with object relative clauses (SO) in comparison to subject relative clauses (SS) in English. The authors interpreted their findings as a result of different demands on syntax relevant memory due to differences in the structure of the two sentence types. More complex SO sentences needed additional operations and processes to determine the meaning of the sentences in comparison to SS sentences. In German, Fiebach and colleagues (2005) investigated the relationship between syntactic integration and syntactic working memory in sentences with wh-questions in a 2 x 2 design. Working memory demands were manipulated while varying the filler-gap distance in sentences (short versus long distance). Syntactic integration was controlled with subject-initial versus object-initial sentences (subject-first versus object-first). As a result, object-initial sentences showed no significant higher activity in the brain in comparison to subject-initial sentences. In contrast, greater filler-gap distances correlated with a higher BOLD response in the left IFG and the bilateral superior temporal gyrus (STG). These findings were interpreted as higher syntactic working memory demands for longer filler-gap dependencies. It was suggested that not complexity per se accounts for activity in Broca's area, but syntactic working memory which is required for the construction of the filler-gap dependency. This interpretation was supported by Bahlmann et al. (in press). The authors proposed that greater working memory demands during the processing of non-canonical sentences was associated with activity in Broca’s area (BA 44).

In linguistic terms Ben-Shachar et al. (2003) defined Broca's area as the brain basis for phrase-movement. It was demonstrated that Broca's area was activated, in particular, for the processing of non-canonical sentences. In line with this definition Friederici et al. (2006) were able to show that Broca's area increased parametrically with the degree to which the non-canonical sentence deviated from normal word order (operationally defined as the number of permutations/movements necessary to achieve the non-canonical structure).

Ullman (2004) proposed a model segregating syntactic from semantic structures by the procedural (syntax) and the declarative (semantic) memory systems. Based on behavioral data on patients suffering from Huntington’s and Parkinson’s diseases, the author suggested that Broca’s area is part of a neural network including the basal ganglia, the parietal lobe and the cerebellum. The procedural memory was described as a rule building combination of lexical items (retrieved from the declarative memory, situated in the temporal lobe) into complex representations.

The unification model by Hagoort (2005) suggest three components of language processing, namely memory, unification and control. The different types of language information are stored in and retrieved from the long term memory. The unification process comprises the integration of retrieved lexical information in greater representations. Unification should come along with semantic, syntactic, and phonological processes and is claimed to be associated with Broca’s area activity.

Based on a literature review of language studies Friederici (2004) proposed that Broca’s area is engaged in hierarchical structure building. This process is involved during managing of long-distance dependencies and transformational structures. The author suggested that Broca’s area (BA44/45) is recruited when structural hierarchies are to be computed whereas the ventral premotor cortex supports the processing of local probabilities. The ventral premotor cortex was repeatedly shown to be involved in the serial prediction of stimuli (e.g. Schubotz & von Cramon, 2002; or Schubotz & von Cramon, 2003 for an overview). The authors suggested that the premotor cortex is engaged in anticipatory processing of both motor as well as perceptual events. Several sub-regions of the premotor cortex could be distinguished as a function of the sequence modality that participants were attending to (e.g. position, object, rhythm, or pitch of the stimuli, Schubotz & von Cramon, 2003). The ventral premotor cortex seems thus to be engaged in the linear, domain-independent sequencing, while Broca's area might be crucial for the integration of a higher order sequencing particularly related to the language domain.

In the present work Broca's area was engaged during the processing of violations of hierarchical structured artificial grammar rules. In contrast, the computation of violations of transitional dependencies in sequences generated by an FSG rule caused no activity in Broca's area. Moreover, the ROI analysis in this region suggests that it is not processing difficulty per se that activated Broca's area, it rather represents a correlation with the presence of structural hierarchies. In the Syntactic Prediction Locality Theory (SPLT) Gibson (1998) proposed a description of the distribution of resources during sentence comprehension. In this model the author suggested that the integration of long distance dependencies requires more working memory and integration recourses than adjacent dependencies. A detailed depiction of an example of center-embedded sentences was given, implicating a higher complexity of this structure. This complexity results from the distance between the incoming element and the dependent to which it should be attached (integration cost) and the maintenance of the predicted category in working memory until the prediction is satisfied (memory cost). According to this framework, I would propose that the chunking of the elements of the artificial structure and the maintenance of the particular element in working memory until the matching category occurred describes the long distance dependencies of the PSG structure at hand.

The present EPR data would also speaks in favor to the use of working memory capacity during the processing of the PSG, but not FSG. A posteriorly distributed late positivity was shown for rule violating sequences in comparison to rule adequate stimuli in both rule types. Moreover, in order to use the advantage of the high temporal resolution of ERPs, violations were situated at different positions of the sequences. I. e., early violations were located in the first half of the syllable sequences (positions three and four) and late violations were placed in the second half of the sequences (positions six and seven) for both rule types. The size of the violation effect of the late positivity systematically changed over error positions in the PSG, but not in the FSG. In the PSG, the violation effect at the fourth position ($\omega^2 = .17$) was smaller than the effect of violation at later positions (sixth position: $\omega^2 = .67$; seventh position: $\omega^2 = .72$). In contrast, in the FSG the violation effect was fund to be independent of the error position in a sequence. At the early positions of the sequences in the PSG, a structural violation is not as salient as it is at later positions. Due to the initial ambiguity, inserting a B at the beginning of the structure is not as salient an error as placing an A at the end of the structure. Hence, the integration of a syllable at the end of the structure is more difficult than at an earlier position. The difference observed for the two grammar types indicates that FSG and PSG are processed differently. The FSG can be processed on the basis of local rules (local probability), but the PSG cannot. Thus, the change in the amplitudes on the different violation positions in the PSG could reflect the hierarchical processing of the sequences. The longer the distance between the elements that needs to be matched, the higher the amplitude of the late positivity, and thus, the greater the use of working memory in order to process the long distance dependencies.

Still, this is a necessary, but not sufficient argument for the hierarchical processing of the A^nB^n structure in the first two experiments. Interestingly, in a recent study Perruchet & Rey (2005) questioned the involvement of hierarchical structure building during the processing of the PSG rule used by Fitch & Hauser (2004). More precisely, Perruchet & Rey (2005) claimed that the rule A^nB^n can be processed by a simple counting mechanism as it does not produce center-embedded structures by the use of corresponding categories. However, it was shown that the design used in the first two experiments of the present work differed from the experimental manipulations of Fitch & Hauser's study (see Chapter 7). But still, the use of a hierarchical way of sequence processing in the first two studies might be questionable. Hence, a novel paradigm discouraging the application of a simple counting mechanism was used in the third study. Additionally, a direct comparison of the FSG versus the PSG rule was possible in a within subject design, avoiding a violation design. As a result the direct comparison of hierarchically structured versus linearly structured syllable sequences revealed activity comprising Broca's area and the anterior prefrontal cortex (BA44/6) and its right hemisphere homologue. A ROI analysis in these regions corroborated the findings of the whole brain analysis, such as a main effect of rule type elicited a higher BOLD response for the hierarchical rule in comparison to the linear rule. Moreover, the hemisphere effect exposed a higher hemodynamic response in the area comprising the left pars opercularis and adjacent precentral sulcus (BA44/6), suggesting a left-dominant activation pattern in the brain. In order to interpret these findings functionally, the exact structure of the sequences has to be considered. In the hierarchical artificial rule at hand the dependent elements of the syllable categories were concatenated over a long distance. Two types of categories were produced. One category segregated the syllables by different vowels (i.e. A and B); the other category was built up via different types of plosives (e. g. A_1 and B_1). Hierarchy was produced by the combination of several concatenations of categories that were not built up by adjacency (e. g. $A_3A_2A_1B_1B_2B_3$ = ge di be pu to ko). The processing of the long distance dependencies of the PSG structure at hand can be described by chunking of the two classes of the artificial structure and maintenance of the particular element in working memory until the matching category occurs. The computation of this structural hierarchy engaged BA44/6. Hence, the work at hand would suggest that Broca's area, including the directly adjacent rim of the premotor cortex, is associated with the working memory required for the processing of embedded, hierarchical structures and the resulting hierarchical dependencies. Additionally, this experiment confirms the findings of the previous study, while using a more elaborated structure to test Broca's role in grammar processing more rigorously. In both studies Broca's area was more engaged during the processing of the PSG than of the processing of the FSG.

These findings suggest that Broca's area is mainly engaged in processing hierarchical as compared to non-hierarchical grammars, an idea similarly outlined in a literature review by Friederici (2004). From a broader perspective, the findings at hand are in line with the literature on language-like artificial grammar learning (Musso et al., 2003; Opitz & Friederici, 2003, 2004). Moreover also a number of fMRI studies investigating syntax processing in natural language reported activity in Broca's area (Bahlmann, Rodriguez-Fornells et al., in press; Ben-Shachar et al., 2003; Caplan, 2001; Fiebach et al., 2005; Friederici, Fiebach et al., 2006; Stromswold et al., 1996). Thus, it is not surprising that the Bayesian analysis revealed a high posterior probability ($P(X|Z)$) for the activity in Broca's area (and, to a lesser extent, in its right hemisphere homologue) in language-related paradigms (see Chapter 7).

8.2. FOP/anterior insula

The anterior insula receives afferent connections from several brain areas, such as the primary somatosensory cortex, the retroinsular area, the STG, the amygdaloid body, and the entorhinal cortex. Additionally, this brain region is efferently connected to the supplementary motor area, the ventral granular frontal cortex, the temporopolar cortex, and the retroinsular area (Preuss & Goldman-Rakic, 1989). Several cognitive and autonomous functions were associated with the anterior insula, including heart beat, breathing, temporal processing, cognitive control, subjective feeling, or language processing (see Augustine, 1996 for an overview).

With respect to language processing an involvement of the anterior insula has been reported in particular for speech. For example, lesions in the anterior insular are correlated with impairments in the processing of speech (Kreisler et al., 2000; Saygin, Wilson, Dronkers, & Bates, 2004). Saygin and co-worker (2004) reported significant deficits in linguistic tasks, but not in non-linguistic tasks in patients suffering from lesions in the inferior anterior insula, the anterior superior temporal gyrus, and the supramarginal gyrus. Lesions in the insula and the external capsula were described by Kreiler et al. (2000) to be correlated with a repetition disorder of speech stimuli.

In the second experiment of the present work, the comparison of sequences built according to a rule versus sequences violating that rule revealed activity in the left FOP. This was shown independent of the rule type. The ROI analysis in the FOP confirmed these findings. Moreover, no interaction of rule type and violation was found (see Table 8.1). This result suggests an involvement of the FOP during the processing of ungrammaticalities independent of the structure of the sequence. It was hypothesized that the FOP supports the check of the incoming element against the predicted element.

The third experiment exposed activity in the bilateral dorsal anterior insula in all three main effects, namely rule type, violation, and length. The ROI analysis in these brain regions confirmed these findings. PSG syllable sequences revealed a higher BOLD response than the FSG sequences. Incorrect stimuli also engaged the anterior insula more than correct stimuli. Furthermore, the ROI in the left hemisphere elicited a higher hemodynamic response than the right hemispheric anterior dorsal insula (see Table 8.1). Note that this experiment was designed in order to further elaborate the type of hierarchical structure of the PSG. No additional predictions about the nature of the local processing of the FSG were formulated. Hence, conclusions about adjacent dependencies (FSG) from the results of the third experiment should be considered as post-hoc explanations. The significantly higher engagement of this region during processing of violations as during the processing of rule-based sequences speaks in favor to the hypothesis that the FOP/anterior insula is involved in the computation of ungrammaticalities. On the other hand, the main effect of rule type would not confirm the interpretation of a structure type independent computation of ungrammaticalities. A grammar type independent local checking of the incoming element against the predicted one would implicate a similar increase of the hemodynamic response for both rule types. This is not the case, since the PSG exposed a higher hrf-function than the FSG in the anterior insula. However, it also does not falsify the prediction, since the ROI analysis revealed that this region is employed in both rule types, with a characteristic BOLD response for the FSG condition and the PSG condition (no null effect for FSG). The FOP is significantly more engaged during the computation of PSG structures in comparison to FSG structures. In other words, the FOP is not not engaged during the detection of ungrammaticalities of the FSG rule; it is less engaged in the processing of the FSG as compared to the PSG ungrammaticalities.

Table 8.1: Overview of contrasts calculated that revealed activations in FOP in the two fMRI experiments. ROI analysis conducted in this area confirmed the findings of the contrast images.

	Activity in FOP	ROI analysis
Experiment 2	FSG: inco versus corr	F(1, 34) = 24.76, P < .001
	PSG: inco versus corr	
		Interaction FSG x PSG: n.s.
Experiment 3	PSG versus FSG (bilat)	t(13) = 3.68, p < .001
	Inco versus corr (bilat)	t(13) = 8.09, p < .001
	Long vesus short (bilat)	t(13) = 1.94, p = .07
		Interaction rule type x violation x length x hemisphere: F(1, 13) = 3.95, p = .06

From the neuroanatomical perspective, the comparison of the brain region found in the first fMRI experiment with the second experiment’s counterpart makes the depiction of (at least) two physiological considerations necessary. First, the individual variability within brain structures renders the comparison of a brain area between groups of small samples as rather imprecise. Second, the adjacent areas FOP and anterior insula are situated in a small in plain area of about 2 x 2 cm in the brain. On the z-axis this brain region has a length (from dorsal to ventral) of approximately 4 cm. In the second study (first fMRI experiment) the local maxima of activity was at -36, 16, 0 (FSG) and -36, 20, -2 (PSG). In the third study (second fMRI experiment) the local maxima ranged between -28, 18, 4 (PSG versus FSG) and -32, 18, 0 (incorrect versus correct). All maxima were clustered in the dorsal part ($-2 \leq$ z-axis ≤ 4). However, the in plain loci varied up to 8 mm. Due to the individual variability and because of the narrow adjacency it is difficult to disentangle the FOP and the anterior insula neuroanatomically.

Two issues would speak in favor of an alternative interpretation of the function of the anterior insula. First, the main effect of rule type (PSG over FSG) found in the third experiment could be discussed from a different viewpoint. This result would not fit into the interpretation of a rule independent processing of ungrammaticalities supported by this region since the rule type hade still an influence on activity in this area. Second, the diversity of afferent and efferent connections from/to numerous brain regions renders a syntax specific function of this region unlikely. Related to this, a number of different autonomous and cognitive functions were associated with this region. The low posterior probability (P(X|Z)) of language relatedness (see Chapter 7) is in line with this notion, since a low degree of language relatedness would implicate many other cognitive functions to be associated with activity in this brain area. In order to further evaluate the nature of this region the microstructure of the anterior insula should be considered. Mesulam & Mufson (1982) suggested a segregation between three distinct regions in the macaque insula cortex. This parcelation was based on cytoarchitectonical and fiber tracing studies. The anterior ventral insula was labeled agranular sector due to its indistinct laminar structure and apparent lack of granule cells. The anterior dorsal insula was labeled dysgranular sector. This area shows the first emergence of a granular layer IV. Moreover, the anterior dorsal insula blends into the frontal operculum. The posterior portion of the insula cortex contains granular layers II and IV and was hence classified as granular sector. Based on the cytoarchitectonical distinction of the ventral and dorsal part of the anterior insula, Wager and colleagues (in press) proposed a functional separation of these regions in a meta-analysis. While the anterior ventral insula should be associated with core affects (e.g. motivational stages) the anterior dorsal insula was postulated to particularly reflect the mental effort of a cognitive task which is accompanied by autonomic reactions. The authors argued: “When a task is difficult, error signals are generated that the organism is doing the wrong task or doing the right task in a substandard way. Attention must be reallocated or the strategy changed.” p. 15. In the present experiments the processing of the PSG was more difficult than the processing of the FSG. More attention has to be redistributed during the detection of errors of the PSG in comparison to error computing in the FSG. In the studies at hand all local maxima clustered in the *dorsal anterior insula/FOP*. According to this notion, an alternative interpretation of the anterior insula activation in the tasks of the present work could be an autonomous reaction triggered by a signal for optimizing an ongoing action. Behavioral data confirm this interpretation, as far as error rates were increased for hierarchical sequences compared to linear sequences; however, no behavioral differences were recorded for violated versus correct and long versus short sequences.

8.3 The superior temporal sulcus

DTI data of the second experiment suggests a neuroanatomical differentiation between Broca’s area and FOP. While the former was found to be connected to the posterior and middle portion of the superior temporal region via the fasciculus longitudinalis superior, the latter is connected via the fasciculus uncinatus to the anterior temporal lobe.

In the third experiment the main effect of violation exposed left posterior superior temporal (posterior STS) activation for incorrect as compared to correct sequences in either of the rule types. Language studies which reported activation in this area investigated either syntactic complexity (Ben-Shachar et al., 2003; Röder et al., 2002) or local syntactic violations (Rüschemeyer, Fiebach, Kempe, & Friederici, 2005). Activation in this area was functionally connected to integration processes. It was proposed that during the processing of ungrammatical sentences (local syntactic violations) the integration of the violating element into the prior structure is impossible. During the processing of (syntactically complex) object-first sentences the integration of the next phrase into the prior sentential structure is difficult. Fiebach and colleagues (2005) illustrated that in object-first sentences not syntactic complexity (per se) is responsible for activity in the STG, but that this area reflects a manipulation of the long distance dependencies within a sentence. In the study at hand violations were processed either within the (local) linear structure, or within the hierarchical structure. The latter one was produced via violations of the long distance dependencies of the structure; violations of the linear sequences represent local mismatches. Hence, our findings are in line with the literature on natural languages. Violations of either rule type (PSG or FSG) activated the posterior STS. The varying syntactic violation paradigms using natural speech (manipulation of the long distance dependencies or local syntactic violations) also reported activity in the posterior STS/STG. It is therefore suggested that in the present study the posterior STS was engaged during the integration of the incoming syllable into the established grammatical structure, be it a complex hierarchical or a simple linear one.

The contrast of FSG versus PSG sequences exposed activations in the bilateral anterior STS. Activation in the anterior portion of the STG in addition to its posterior portion was observed for local structural mismatches (Friederici et al., 2003). Using MEG to disentangle different syntactic processing stages it was found that in the STG only the anterior portion was seen to be activated during an early phase of structure building (Friederici, Wang, Herrmann, Maess, & Oertel, 2000). This would mean that the posterior

STG may rather be involved in a late integration processing phase. The notion that the anterior STG is involved in local structure building is also supported by the finding that this region showed increased activation for sentences containing syntactic structure as compared to word lists (Friederici, Meyer, & von Cramon, 2000; Stowe et al., 1998; Vandenberghe, Nobre, & Price, 2002).

8.4. A brain basis of the evolution of the ability to process hierarchical structures?

From an evolutionary perspective a nice link could be drawn between the findings by Fitch & Hauser (2004) and the date at hand. Tamarin monkeys were shown not to be able to process hierarchical structures in an artificial grammar task. The present work demonstrated that Broca's area was more engaged during the processing of hierarchical structures than during the computation of linear structures.

In order to bridge the gap between the human ability to process hierarchical structures and it's ancestors that do not show this capability, homologies between different cortical areas in human and non-human primates should be considered. The comparability of brain regions between species is a highly discussed topic in the research community. The present contribution can not cover the immense number of diverse opinions about this research field. Nonetheless, I herewith try to implicate some speculations about the few facts that exist regarding homologies between humans and macaques. Some authors suggest that a homologue of the human prefrontal cortex appeared during the evolution of primates, but not prior to the evolution of primates (e.g. Aboitiz & Garcia, 1997). While some authors emphasis on the enlargement of frontal cortex in general (e.g. Deacon, 1997), other authors highlight on the differential development of specific subsystems which support the evolution of language (e.g. Arbib & Bota, 2003). More precisely, some authors suggest that area F5 in the monkey brain represents a pre-stage of BA 44 in the human brain. Petrides, Cadoret, & Mackey (2005) showed that F5 in macaques is responsible for orofacial musculature innervations. The authors suggested that BA 44 is already present in macaques (with extend of 4-9 mm) and evolved into the larger BA 44. Hence, F5 and BA 44 could be characterized as cytoarchitectonical homologues between the human and the monkey brain. In a review article Rizzolatti & Luppino (2001) presented a number of studies using intracranial electrode measurements that suggested that F5 in monkeys contains a movement representation on the hand and the mouth, as part of the mirror neuron system. In line with Petrides and co-workers (2005) it was proposed that the human BA 44 represents a homologue to monkey area F5 and the ventral premotor cortex corresponds to F4 in the monkey brain. Taking together, one could speculate that the enlargement of F5 into BA 44 (in terms of cell density, thickness of the layers, folding, and size) during evolution plays a crucial role in the evolution of certain language properties. Note that the tamarin monkeys examined by Fitch & Hauser (2004) represent a species that is about 40 million years older than the macaques investigated by Petrides et al. (2005) and Rizzolatti & Luppino (2001). No analysis about brain area homologous between tamarins and humans was accomplished, yet. In a recent review, Aboitiz and colleagues (2006) claimed that: “linguistic recursion demands significant working memory resources, and that at least partly, neural networks that participate in recursion were gradually elaborated from simpler networks involved in active memory in the primate brain.” p. 41. In line with this notion, one could conclude that the evolution of Broca’s area played a key role in the ability to use working memory capacities in order to process hierarchical/recursive structures and therefore in the evolution of language.

8.5. Summary and open questions

In the present work, comparisons of hierarchical versus linear structures were conducted using visually presented syllable sequences in artificial grammar tasks. Three different experiments were conducted varying the methods of investigation (ERP and fMRI) and the design of the structure of the artificial grammars. In the first experiment (ERP) violations of the structure were included at different positions of the syllable sequences, constructed according to either a PSG or a FSG. As a result, a late positivity was found for violations for both rule types. Additionally, only in the PSG the amplitude of the late positivity varied as a function of violation position. These results could indicate higher integration difficulty during the processing of the PSG in comparison to the FSG. In the second experiment (fMRI) the frontal operculum was engaged during the processing of violations of both rule types. In contrast, only violations of the PSG additionally activated Broca’s area. In the third experiment (also fMRI), the structure of the rules was changed in order to further elaborate the production of long distance dependencies and hence, a hierarchical processing of the PSG. As a result, Broca’s area and its right hemispheric homologue were significantly more engaged during the processing of the hierarchical structure (PSG) in comparison to the linear structure (FSG). Moreover, several other language-related brain areas were activated, suggesting a close link between the neural correlates of natural language processing and the structures used in the artificial grammar task at hand. These results suggest that Broca's area is associated with working memory demands in order to build hierarchical structures, independent of whether these hierarchies are part of a natural language or not. A number of open questions remain, that need to be discussed in detail.

Foremost, the extent to which the A^nB^n structure can be taken to represent human language remains open. Natural grammars contain more complex syntactic rules than the rules used in the present artificial grammar design. A^nB^n sequences feature center-embedded syllables, comparable with human languages. The processing principles for this type of center-embedded structures should be the same in the A^nB^n structure as in embedded sentences and should therefore recruit similar brain structures. Moreover, in the ERP the processing of syntactic violations is usually correlated with the E/LAN effect. As described in Chapter 3, a LAN was sometimes associated with memory processes (Kluender & Kutas, 1993, Rösler et al., 1998). In the present study, no LAN was found. The ELAN component has been interpreted to reflect automatic aspects of local structure building during language processing (Neville et al., 1993; Hahne et al., 1999). Hoen et al. (2000) even assumes that the LAN reflects structure-building processes in general. Further research is needed to explore whether the processing of the A^nB^n structure shares the early, automatic, violation detection mechanisms reflected by the ELAN in natural language processing.

Next, the present findings confirm that non-natural (artificial) grammar processing suffices to engage Broca's area. An open question remains whether this is due to the fact that language-specific regions are exploited for the processing of language-like stimuli or, conversely, natural grammar can be conceived as one of many possible sequential and/or hierarchical structures to be processed (inter alia) in Broca's area. With respect to this issue a further experiment employing the PSG structure used in the present study could be carried out using a non-language related modality. Visually presented symbol sequences instead of syllable sequences could be applied in order to test the hypothesis whether Broca's area is generally engaged in the processing of hierarchies despite of the language domain.

A further open question remains according to the highly discussed field of homologies between brain areas in different species. The comparability of the microstructure of Broas's area and the frontal operculum could be questioned. To date no direct comparison of the cytoarchitecture between Broca's area and the anterior dorsal insula was accomplished, jet. Amunts and colleagues (1999) categorized the human BA 44 as dysgranular and BA 45 as granular. The areas differed mainly in layer IV according to the cell density and the thickness of the layer. Mesulam & Mufson (1982) categorized different sectors in the anterior insula cortex according to cytoarchitectonic properties. The authors classified the anterior ventral insula as agranular and the anterior dorsal insula also as dysgranular. However, the categorization between agranular, dysgranular, and granular represents a relative quantity which only holds for the brain region of interest in one particular analysis. Different types of methods and different kinds of samples lead to insufficient comparability (see for example Zilles, Schleicher, Palomero-Gallagher, & Amunts, 2002 for an overview of the methods). While Amunts et al. (1999) used an observer-independent analysis of postmortem human brains, Mesulam & Mufson (1982) used visual inspection to categorize the brain areas of macaque monkeys. Hence, the cytoarchitectonically distinction between Broca's area and the frontal operculum remains an open issue.

Another open aspect of the implications of this work concerns the claim of recursion as a human-specific trait. Gentner and co-workers (2006) recently showed that songbirds are able to process hierarchical structures after intensive training. These finding are in sharp contrast to the hypothesis that the computation of recursion is the only trait that makes the faculty of language human specific (Hauser et al., 2002). However, it is still questionable whether the stimuli presented by Gentner et al. (2006) represent real recursive structures. Perruchet & Rey (2005) expressed concerns about the type of processing of the structure used by Fitch & Hauser (2004). Accordingly, a simple counting mechanism, instead of the processing of hierarchical structures, could also hold for the stimuli presented to songbirds by Gentner et al. (2006). Further research is needed in order to test whether non-human primates or other species are able to process hierarchical structures.

To conclude, in line with Pinker & Jackendoff (2005) one could argue that recursion is one necessary aspect of human language but not the only feature that enables humans to speak. Maybe recursion is not the only feature of human cognitive facilities that initiated the evolution of human language; it could rather be part of a set of different cognitive abilities which together enables humans to process language. Further comparative research (e.g. humans versus non-human primates) is needed to evaluate the contribution of characteristics like phonology, morphology, case, or working memory on the evolution of language.

References

Aboitiz, F., & Garcia, R. (1997). The evolutionary origin of the language areas in the human brain - a neuroanatomical perspective. *Brain Research Reviews, 25*(3), 381-396.

Aboitiz, F., Garcia, R. R., Bosman, C., & Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain & Language, 98*(1), 40-56.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B. M., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology, 412*(2), 319-341.

Anderson, J. R. (1999). *Learning and memory : An integrated Approach* (2nd ed.). New York: Wiley.

Arbib, M., & Bota, M. (2003). Language evolution: neural homologies and neuroinformatics. *Neural Networks, 16*(9), 1237-1260.

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews, 22*(3), 229-244.

Bach, E., Brown, C., & Marslen-Wilson, W. (1985). Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Language and Cognitive Processes, 1*(4), 1985-1986.

Bahlmann, J., Gunter, T. C., & Friederici, A. D. (in press). Hierarchical and linear sequence processing: An electrophysiological exploration of two different grammar types. *Journal of Cognitive Neuroscience*.

Bahlmann, J., Rodriguez-Fornells, A., Rotte, M., & Münte, T. F. (in press). An fMRI study of canonical and non-canonical word order in German. *Human Brain Mapping*.

Baldwin, K. B., & Kutas, M. (1997). An ERP analysis of implicit structured sequence learning. *Psychophysiology, 34*(1), 74-86.

Bandettini, P. A. (2001). Selection of the optimal pulse sequence for functional MRI. In P. Jezzard, P. M. Matthews & S. M. Smith (Eds.), *Functional MRI: an introduction to methods* (pp. 123 - 143). New York: Oxford University Press.

Bandettini, P. A., Kwong, K. K., Davis, T. L., Tootell, R. B., Wong, E. C., Fox, P. T., et al. (1997). Characterization of cerebral blood oxygenation and flow changes during prolonged brain activation. *Human Brain Mapping, 5*(2), 93-109.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. *Psychological Science, 14*(5), 433-440.

Berry, D. C. (1994). Implicit learning: Twenty-five years on. A tutorial. In C. Umilita & M. Moscovitch (Eds.), *Attention and performance XIV: Conscious and nonconscious information processing* (Vol. 30, pp. 755-782). Cambridge: MIT Press.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *Journal of Cognitive Neuroscience, 11*(1), 80-95.

Birbaumer, N., Elbert, T., Canavan, A. G., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Review, 70*(1), 1-41.

Bloch, F. (1946). Nuclear Induction. *Physical Review, 70*(7 - 8), 460 - 474.

Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2006). How to hit Scylla without avoiding Charybdis: comment on Perruchet, Tyler, Galland, and Peereman (2004). *Journal of Experimental Psychology: General, 135*(2), 314-321

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151-188.

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport, 9*(16), 3735-3739.

Buxton, R. B. (2002). *Introduction to functional magnetic resonance imaging: principles and technques*. Cambridge: Cambridge University Press.

Buxton, R. B., & Frank, L. R. (1997). A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *Journal of Cerebral Blood Flow & Metabolism, 17*(1), 64-72.

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine, 39*(6), 855-864.

Caplan, D. (2001). Functional neuroimaging studies of syntactic processing. *Journal of Psycholinguistic Research, 30*(3), 297-320.

Caplan, D., Alpert, N., & Waters, G. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience, 10*(4), 541-552.

Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *Neuroimage, 9*(3), 343-351.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77-94.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1959). On certain formal properties of grammars. *Information & Control, 2*, 137-167.

Chomsky, N. (1965). Aspects of the theory of syntax. *Oxford, England: M.I.T. Press.*

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, Massachusetts: MIT Press.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*(3), 235-253.

Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., Detre, J., et al. (2002). Neural basis for sentence comprehension: grammatical and short-term memory components. *Human Brain Mapping, 15*(2), 80-94.

Coulson, S., King, J. W., & Kutas, M. (1998). ERPs and domain specificity: Beating a straw horse. *Language & Cognitive Processes, 13*(6), 653-672.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping, 5*(5), 329-240.

Deacon, T. W. (1997). What makes the human brain different. *Annual Review of Anthropology, 26*, 337-357.

Dietz, N. A., Jones, K. M., Gareau, L., Zeffiro, T. A., & Eden, G. F. (2005). Phonological decoding involves left posterior fusiform gyrus. *Human Brain Mapping, 26*(2), 81-93.

Dominey, P. F., Hoen, M., Blanc, J. M., & Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: evidence from simulation, aphasia, and ERP studies. *Brain & Language, 86*(2), 207-225.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(3), 357-427.

Eimer, M., Goschke, T., Schlaghecken, F., & Sturmer, B. (1996). Explicit and implicit learning of event sequences: evidence from event-related brain potentials. *Journal of Experimental Psychology : Learning, Memory and Cognition, 22*(4), 970-987.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences of the United States of America, 97*(11), 6150-6154.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. *Journal of Psycholinguistic Research, 30*(3), 321-338.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory & Language, 47*(2), 250-272.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping, 24*(2), 79-91.

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*(5656), 377-380.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: clarifications and implications. *Cognition, 97*(2), 179-210.

Fletcher, P., Buchel, C., Josephs, O., Friston, K., & Dolan, R. (1999). Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cerebral Cortex, 9*(2), 168-178.

Forkstam, C., & Petersson, K. M. (in press). Syntactic Classification of Aquired Structural Regularities.

Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., et al. (2003). *Human brain function* (2nd ed.). Amsterdam: Elsevier.

Frackowiak, R. S. J., Lenzi, G. L., Jones, T., & Heather, J. D. (1980). Quantitative measurement of regional cerebral blood flow and oxygen metabolism in man using 15O and positron emission tomography: theory, procedure, and normal values. *Journal of Computer Assisted Tomography, 4*(6), 727-736.

Frahm, J., Kruger, G., Merboldt, K. D., & Kleinschmidt, A. (1996). Dynamic uncoupling and recoupling of perfusion and oxidative metabolism during focal brain activation in man. *Magnetic Resonance in Medicine, 35*(2), 143-148.

Friederici, A. D. (2004). Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Science, 8*(6), 245-247.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America, 103*(7), 2458-2463.

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing Linguistic Complexity and Grammaticality in the Left Frontal Cortex. *Cerebral Cortex.*

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage, 20 Suppl 1*, S8-17.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain & Language, 75*(3), 289-300.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*(3), 183-192.

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex, 13*(2), 170-177.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 99*(1), 529-534.

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. *Human Brain Mapping, 11*(1), 1-11.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: characterizing differential responses. *Neuroimage, 7*(1), 30-40.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage, 2*(1), 45-53.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D., & Frackowiak, R. S. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping, 2*, 189-210.

Friston, K. J., Josephs, O., Rees, G., & Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magnetic Resonance in Medicine, 39*(1), 41-52.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. *Neuroimage, 10*(5), 607-619.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204-1207.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition, 68*(1), 1-76.

Ginsberg, M. D., Dietrich, W. D., & Busto, R. (1987). Coupled forebrain increases of local cerebral glucose utilization and blood flow during physiologic stimulation of a somatosensory pathway in the rat: demonstration by double-label autoradiography. *Neurology, 37*(1), 11-19.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109-135.

Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science, 4*(5), 178-186.

Goschke, T., Friederici, A. D., Kotz, S. A., & van Kampen, A. (2001). Procedural learning in Broca's aphasia: dissociation between the implicit acquisition of spatio-motor and phoneme sequences. *Journal of Cognitive Neuroscience, 13*(3), 370-388.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrica, 24*, 95-112.

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlesewsky, M. (2005). The emergence of the unmarked: a new perspective on the language-specific function of Broca's area. *Human Brain Mapping, 26*(3), 178-190.

Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *Behavioral and Brain Sciences, 23*(1), 1-21; discussion 21-71.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*(4), 556-568.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34*(6), 660-676.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Review Neuroscience, 2*(10), 685-694.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Science, 9*(9), 416-423.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*(2), 194-205.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*(3), 339-356.

Handy, T. C. (2005). *Event-related potentials: A method handbook.* Cambridge: MIT press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science, 298*(5598), 1569-1579.

Hays, W., L. (1973). *Statistics for the social sciences.* (2nd ed.). New York: Holt, Rinehart and Winston.

Hoen, M., & Dominey, P. F. (2000). ERP analysis of cognitive sequencing: a left anterior negativity related to structural transformation processing. *Neuroreport, 11*(14), 3187-3191.

Hoge, R. D., Atkinson, J., Gill, B., Crelier, G. R., Marrett, S., & Pike, G. B. (1999). Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. *Proceedings of the National Academy of Sciences of the United States of America, 96*(16), 9403-9408.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition, 97*(2), 211-225.

Jezzard, P., Matthews, P. M., & Smith, S. M. (2002). *Functional MRI: an introduction to methods*. Oxford: Oxford university press.

Johnson, R., Jr., & Donchin, E. (1982). Sequential expectancies and decision making in a changing environment: an electrophysiological approach. *Psychophysiology, 19*(2), 183-200.

Josephs, O., Turner, R., & Friston, K. (1997). Event-related fMRI. *Human Brain Mapping, 5*(4), 243-248.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language & Cognitive Processes, 15*(2), 159-201.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory & Language, 30*(5), 580-602.

King, J., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*(3), 376-395.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*(2), 196-214.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 79-91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology : Learning, Memory and Cognition, 22*(1), 169-181.

Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychol Res, 63*(1), 22-35.

Koch, M. A., Norris, D. G., & Hund-Georgiadis, M. (2002). An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage, 16*(1), 241-250.

Kreisler, A., Godefroy, O., Delmaire, C., Debachy, B., Leclercq, M., Pruvo, J. P., et al. (2000). The anatomy of aphasia revisited. *Neurology, 54*(5), 1117-1123.

Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., et al. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. *Journal of Cognitive Neuroscience, 12*(2), 321-341.

Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics, 3*(1), 65-78.

Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature, 242*, 190-191.

Lelekov-Boissard, T., & Dominey, P. F. (2002). Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Clinical Neurophysiology, 32*(1), 72-84.

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *Journal of Cognitive Neuroscience, 16*(3), 427-438.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Lohmann, G., Muller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., et al. (2001). LIPSIA--a new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics, 25*(6), 449-457.

Malonek, D., & Grinvald, A. (1996). Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science, 272*(5261), 551-554.

Mandeville, J. B., & Rosen, B. R. (2002). Functional MRI. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Methods* (pp. 315 -349). Amsterdam: Academic Press.

Marchal, G., Rioux, P., Petit-Taboue, M. C., Sette, G., Travere, J. M., Le Poec, C., et al. (1992). Regional cerebral oxygen consumption, blood flow, and blood volume in healthy human aging. *Archives of Neurology, 49*(10), 1013-1020.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77-80.

Marota, J. J., Ayata, C., Moskowitz, M. A., Weisskoff, R. M., Rosen, B. R., & Mandeville, J. B. (1999). Investigation of the early response to rat forepaw stimulation. *Magnetic Resonance in Medicine, 41*(2), 247-252.

Mesulam, M. M., & Mufson, E. J. (1982). Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *Journal of Comparative Neurology, 212*(1), 1-22.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage, 11*(6 Pt 1), 735-759.

Miller, G., & Isard, S. (1964). Free recall of self-embedded English sentences. *Information & Control, 7*, 293-303.

Miller, G. A. (1967). The psychology of communication: seven essays. In G. A. Miller (Ed.), *The psychology of communication.* (pp. 125-187). New York: Basic Books.

Moonen, C. T. W., & Bandettini, P. A. (2000). *Functional MRI.* Berlin: Springer.

Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial grammar language. *Journal of Verbal Learning & Verbal Behavior, 20*, 67-85.

Mori, S. (2002). Principles, Methodes, and Applications of Diffusion Tensor Imaging. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Methods*. San Diego: Academic Press.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage, 13*(1), 110-118.

Münte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36*(3), 217-226.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., et al. (2003). Broca's area and the language instinct. *Nature Neuroscience, 6*(7), 774-781.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*(2), 1-32.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127-162.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience, 12*(1), 120-133.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1-32.

Norris, D. G. (2000). Reduced power multislice MDEFT imaging. *Journal of Magnetic Resonance Imaging, 11*(4), 445-451.

Nunez, P. L. (1981). *Electrical fields of the brain*. New York: Oxford University Press.

Nunez-Pena, M. I., & Honrubia-Serrano, M. L. (2004). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research, 18*(2), 130-141.

Ogawa, S., Lee, T. M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine, 14*(1), 68-78.

Oken, B. S., & Chiappa, K. H. (1986). Statistical issues concerning computerized analysis of brainwave topography. *Annals of neurology, 19*(5), 493-497.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *Neuroimage, 19*(4), 1730-1737.

Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. *Journal of Neuroscience, 24*(39), 8436-8440.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language & Cognitive Processes, 14*(3), 283-317.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717-733.

Peelle, J. E., McMillan, C., Moore, P., Grossman, M., & Wingfield, A. (2004). Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: evidence from fMRI. *Brain & Language, 91*(3), 315-325.

Peigneux, P., Maquet, P., Meulemans, T., Destrebecqz, A., Laureys, S., Degueldre, C., et al. (2000). Striatum forever, despite sequence learning variability: a random effect analysis of PET data. *Human Brain Mapping, 10*(4), 179-194.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119*(3), 264-275.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Science, 10*(5), 233-238.

Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychological Bulletin Review, 12*(2), 307-313.

Perruchet, P., Vinter, A., Pacteau, C., & Gallego, J. (2002). The formation of structurally relevant units in artificial grammar learning. *The Quarterly Journal of Experimental Psychology: Section A, 55*(2), 485-503.

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science, 28*(3), 383-407.

Petrides, M., Cadoret, G. V., & Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature, 435*(7046), 1235-1238.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition, 95*(2), 201-236.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Science, 10*(2), 59-63.

Pollmann, S., Wiggins, C. J., Norris, D. G., von Cramon, D. Y., & Schubert, T. (1998). Use of short intertrial intervals in single-trial experiments: a 3T fMRI-study. *Neuroimage, 8*(4), 327-339.

Preuss, T. M., & Goldman-Rakic, P. S. (1989). Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. *Journal of Comparative Neurology, 282*(2), 293-316.

Purcell, E. M., Torrey, H. C., & Pound, R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review, 69*(1-2), 37 - 38.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior, 6*(6), 855-863.

Reber, P. J., & Squire, L. R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience, 113*(2), 235-242.

Regan, D. (1989). *Human Brain Electrophysiology. Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York: Elsevier.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77*(6), 481-495.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31*(6), 889-901.

Röder, B., Stock, O., Neville, H., Bien, S., & Rösler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance Imaging study. *Neuroimage, 15*(4), 1003-1014.

Rosen, B. R., Buckner, R. L., & Dale, A. M. (1998). Event-related functional MRI: past, present, and future. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 773-780.

Rösler, F., Pechmann, T., Streb, J., Röder, B., & Hennighausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory & Language, 38*(2), 150-176.

Rugg, M., & Coles, M. (1995). *Electrophysiology of Mind: Event-related Brain Potentials and Cognition*. New York: Oxford University Press.

Rüschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping, 25*(2), 266-286.

Rüsseler, J., & Rösler, F. (2000). Implicit and explicit learning of event sequences: evidence for distinct coding of perceptual and motor representations. *Acta Psychologica, 104*(1), 45-67.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language, 44*(4), 493-515.

Saygin, A. P., Wilson, S. M., Dronkers, N. F., & Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia, 42*(13), 1788-1804.

Schubotz, R. I., & von Cramon, D. Y. (2002). A blueprint for target motion: fMRI reveals perceived sequential complexity to modulate premotor cortex. *Neuroimage, 16*(4), 920-935.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. *Neuroimage, 20 Suppl 1*, S120-131.

Schwartz, W. J., Smith, C. B., Davidsen, L., Savaki, H., Sokoloff, L., Mata, M., et al. (1979). Metabolic mapping of functional activity in the hypothalamo-neurohypophysial system of the rat. *Science, 205*(4407), 723-725.

Seger, C. A., Prabhakaran, V., Poldrack, R. A., & Gabrieli, J. D. E. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychobiology, 28*(3), 283-292.

Shieber, S. M. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy, 8*(3), 333-343.

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Reber, P. J. (2002). Neural correlates of artificial grammar learning. *Neuroimage, 17*(3), 1306-1314.

Smith, J., Siegert, R. J., McDowall, J., & Abernethy, D. (2001). Preserved implicit learning on both the serial reaction time task and artificial grammar in patients with Parkinson's disease. *Brain & Cognition, 45*(3), 378-391.

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science, 193*(4258), 1142-1146.

Stowe, L. A., Broere, C. A., Paans, A. M., Wijers, A. A., Mulder, G., Vaalburg, W., et al. (1998). Localizing components of a complex task: sentence processing and working memory. *Neuroreport, 9*(13), 2995-2999.

Strange, B. A., Henson, R. N., Friston, K. J., & Dolan, R. J. (2001). Anterior prefrontal cortex mediates rule learning in humans. *Cerebral Cortex, 11*(11), 1040-1046.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain & Language, 52*(3), 452-473.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotactic atlas of the human brain*. New York: Thieme.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Human Brain Mapping, 25*(1), 83-91.

Teichmann, M., Dupoux, E., Kouider, S., Brugieres, P., Boisse, M. F., Baudic, S., et al. (2005). The role of the striatum in rule application: the model of Huntington's disease at early stage. *Brain, 128*(Pt 5), 1155-1167.

Toga, A. W., & Mazziotta, J. C. (2002). *Brain Mapping: The Methods* (Second edition ed.). San Diego: Academic press.

Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology*. New York: Wiley.

Ugurbil, K., Garwood, M., Ellermann, J., Hendrich, K., Hinke, R., Hu, X., et al. (1993). Imaging at High Magnetic Fields: Initial Experiences at 4T. *Magnetic Resonance Quarterly, 9*, 259.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*(1-2), 231-270.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., & et al. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*(2), 266-276.

Uludag, K., Dubowitz, D. J., Yoder, E. J., Restom, K., Liu, T. T., & Buxton, R. B. (2004). Coupling of cerebral blood flow and oxygen consumption during physiological activation and deactivation measured with fMRI. *Neuroimage, 23*(1), 148-155.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550-560.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 328-344.

Wager, T. D., & Barrett, L. F. (in press). From affect to control: Functional specialization of the insula in motivation and regulation.

Witt, K., Nuhsman, A., & Deuschl, G. (2002). Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia, 40*(9), 1534-1540.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited--again. *Neuroimage, 2*(3), 173-181.

Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., et al. (2002). A general statistical analysis for fMRI data. *Neuroimage, 15*(1), 1-15.

Zilles, K., Schleicher, A., Palomero-Gallagher, N., & Amunts, K. (2002). Quantitative analysis of cyto- and receptor architecture of the human brain. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping: The Methods* (second edition ed., pp. 573-602). San Diego: Academic press.

List of Figures

List of Tables

Curriculum Vitae

Jörg Bahlmann

18.07.1975	Geboren in Frankfurt/Oder
1994	Abitur am Gerhart-Hauptmann-Gymnasium, Wernigerode
1994-1995	Grundwehrdienst
1995-1997	Studium der Informatik an der TU Dresden
1997-2002	Studium der Psychologie an der Otto-von-Guericke-Universität, Magdeburg
1999	Vordiplom in Psychologie
2000	Studium der Psychologie an der Westminster University, London, UK
2002	Diplom in Psychologie
2002-2003	Wissenschaftlicher Mitarbeiter an der Universität Magdeburg (Bereich Neuropsychologie)
seit 10/03	Doktorand am MPI für Kognitions- und Neurowissenschaften, Leipzig (Bereich Neuropsychologie)

Bibliographic Details

Bahlmann, Jörg
Neural correlates of the processing of linear and hierarchical artificial grammar rules: Electrophysiological and neuroimaging studies.
Universität Leipzig
138 pages, 187 references, 14 figures, 13 tables

Paper The aim of the dissertation presented here was to investigate neural correlates of the processing of two different types of artificial grammar rules. The rules differed in complexity. The Finite State Grammar (FSG) rule was specified by local adjacent probabilities. In contrast, the Phrase Structure Grammar (PSG) rule generated hierarchical dependencies via center-embeddings. Electrophysiological changes were recorded using event-related brain potentials (ERPs) and hemodynamic changes were measured applying functional Magnetic Resonance Imaging (fMRI). Within three experiments the structures of the rules were systematically changed.

The processing of the two artificial grammar types revealed differences in the late positivity on violations of the structures. The amplitude of the late positivity varied as a function of the position of a violation of the sequence in the PSG rule. In contrast, no changes of the amplitude of the varying violation positions of the FSG were found. FMRI results revealed an increased activity in Broca's area on violations of the PSG, but not of the FSG. Additionally, the frontal operculum was associated with anomalies independent of the rule type. Moreover, the direct contrast of PSG sequences versus FSG sequences also correlated with an increased hemodynamic response function in Broca's area and it's right hemispheric homologue.

The chunking of the elements of the artificial structure and the maintenance of the particular element in working memory until the matching category occurred describes the long distance dependencies of the PSG structure at hand. This process was correlated with activity in Broca's area in the present work. Hence, Broca's area could be associated with working memory demands during the processing of artificial hierarchical structures. These finding could be interpreted as a potential key role of Broca's area in a network of the processing of hierarchical structures.

MPI Series in Human Cognitive and Brain Sciences:

1 Anja Hahne

Charakteristika syntaktischer und semantischer Prozesse bei der auditiven Sprachverarbeitung: Evidenz aus ereigniskorrelierten Potentialstudien

2 Ricarda Schubotz

Erinnern kurzer Zeitdauern: Behaviorale und neurophysiologische Korrelate einer Arbeitsgedächtnisfunktion

3 Volker Bosch

Das Halten von Information im Arbeitsgedächtnis: Dissoziationen langsamer corticaler Potentiale

4 Jorge Jovicich

An investigation of the use of Gradient- and Spin-Echo (GRASE) imaging for functional MRI of the human brain

5 Rosemary C. Dymond

Spatial Specificity and Temporal Accuracy in Functional Magnetic Resonance Investigations

6 Stefan Zysset

Eine experimentalpsychologische Studie zu Gedächtnisabrufprozessen unter Verwendung der funktionellen Magnetresonanztomographie

7 Ulrich Hartmann

Ein mechanisches Finite-Elemente-Modell des menschlichen Kopfes

8 Bertram Opitz

Funktionelle Neuroanatomie der Verarbeitung einfacher und komplexer akustischer Reize: Integration haemodynamischer und elektrophysiologischer Maße

9 Gisela Müller-Plath

Formale Modellierung visueller Suchstrategien mit Anwendungen bei der Lokalisation von Hirnfunktionen und in der Diagnostik von Aufmerksamkeitsstörungen

10 Thomas Jacobsen

Characteristics of processing morphological structural and inherent case in language comprehension

11 Stefan Kölsch

Brain and Music
A contribution to the investigation of central auditory processing with a new electrophysiological approach

12 Stefan Frisch

Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen

13 Markus Ullsperger

The role of retrieval inhibition in directed forgetting – an event-related brain potential analysis

14 Martin Koch

Measurement of the Self-Diffusion Tensor of Water in the Human Brain

15 Axel Hutt

Methoden zur Untersuchung der Dynamik raumzeitlicher Signale

16 Frithjof Kruggel

Detektion und Quantifizierung von Hirnaktivität mit der funktionellen Magnetresonanztomographie

17 Anja Dove

Lokalisierung an internen Kontrollprozessen beteiligter Hirngebiete mithilfe des Aufgabenwechselparadigmas und der ereigniskorrelierten funktionellen Magnetresonanztomographie

18 Karsten Steinhauer

Hirnphysiologische Korrelate prosodischer Satzverarbeitung bei gesprochener und geschriebener Sprache

19 Silke Urban

Verbinformationen im Satzverstehen

20 Katja Werheid

Implizites Sequenzlernen bei Morbus Parkinson

21 Doreen Nessler

Is it Memory or Illusion? Electrophysiological Characteristics of True and False Recognition

22 Christoph Herrmann

Die Bedeutung von 40-Hz-Oszillationen für kognitive Prozesse

23 Christian Fiebach

Working Memory and Syntax during Sentence Processing.
A neurocognitive investigation with event-related brain potentials and functional magnetic resonance imaging

24 Grit Hein

Lokalisation von Doppelaufgabendefiziten bei gesunden älteren Personen und neurologischen Patienten

25 Monica de Filippis

Die visuelle Verarbeitung unbeachteter Wörter.
Ein elektrophysiologischer Ansatz

26 Ulrich Müller

Die katecholaminerge Modulation präfrontaler kognitiver Funktionen beim Menschen

27 Kristina Uhl

Kontrollfunktion des Arbeitsgedächtnisses über interferierende Information

28 Ina Bornkessel

The Argument Dependency Model: A Neurocognitive Approach to Incremental Interpretation

29 Sonja Lattner

Neurophysiologische Untersuchungen zur auditorischen Verarbeitung von Stimminformationen

30 Christin Grünewald

Die Rolle motorischer Schemata bei der Objektrepräsentation:
Untersuchungen mit funktioneller Magnetresonanztomographie

31 Annett Schirmer

Emotional Speech Perception: Electrophysiological Insights into the Processing of Emotional Prosody and Word Valence in Men and Women

32 André J. Szameitat

Die Funktionalität des lateral-präfrontalen Cortex für die Verarbeitung von Doppelaufgaben

33 Susanne Wagner

Verbales Arbeitsgedächtnis und die Verarbeitung ambiger Wörter in Wort- und Satzkontexten

34 Sophie Manthey

Hirn und Handlung: Untersuchung der Handlungsrepräsentation im ventralen prämotorischen Cortex mit Hilfe der funktionellen Magnet-Resonanz-Tomographie

35 Stefan Heim

Towards a Common Neural Network Model of Language Production and Comprehension: fMRI Evidence for the Processing of Phonological and Syntactic Information in Single Words

36 Claudia Friedrich

Prosody and spoken word recognition: Behavioral and ERP correlates

37 Ulrike Lex

Sprachlateralisierung bei Rechts- und Linkshändern mit funktioneller Magnetresonanztomographie

38 Thomas Arnold

Computergestützte Befundung klinischer Elektroenzephalogramme

39 Carsten H. Wolters

Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain

40 Ansgar Hantsch

Fisch oder Karpfen? Lexikale Aktivierung von Benennungsalternativen bei der Objektbenennung

41 Peggy Bungert

Zentralnervöse Verarbeitung akustischer Informationen Signalidentifikation, Signallateralisation und zeitgebundene Informationsverarbeitung bei Patienten mit erworbenen Hirnschädigungen

42 Daniel Senkowski

Neuronal correlates of selective attention: An investigation of electro-physiological brain responses in the EEG and MEG

43 Gert Wollny

Analysis of Changes in Temporal Series of Medical Images

44 Angelika Wolf

Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern

45 Kirsten G. Volz

Brain correlates of uncertain decisions: Types and degrees of uncertainty

46 Hagen Huttner

Magnetresonanztomographische Untersuchungen über die anatomische Variabilität des Frontallappens des menschlichen Großhirns

47 Dirk Köster

Morphology and Spoken Word Comprehension: Electrophysiological Investigations of Internal Compound Structure

48 Claudia A. Hruska

Einflüsse kontextueller und prosodischer Informationen in der auditorischen Satzverarbeitung: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

49 Hannes Ruge

Eine Analyse des raum-zeitlichen Musters neuronaler Aktivierung im Aufgabenwechselparadigma zur Untersuchung handlungssteuernder Prozesse

50 Ricarda I. Schubotz

Human premotor cortex: Beyond motor performance

51 Clemens von Zerssen

Bewusstes Erinnern und falsches Wiedererkennen:
Eine funktionelle MRT Studie neuroanatomischer Gedächtniskorrelate

52 Christiane Weber

Rhythm is gonna het you.
Electrophysiological markers of rhythmic processing in infants with and without risk for Specific Language Impairment (SLI)

53 Marc Schönwiesner

Functional Mapping of Basic Acoustic Parameters in the Human Central Auditory System

54 Katja Fiehler

Temporospatial characteristics of error correction

55 Britta Stolterfoht

Processing Word Order Variations and Ellipses: The Interplay of Syntax and Information Structure during Sentence Comprehension

56 Claudia Danielmeier

Neuronale Grundlagen der Interferenz zwischen Handlung und visueller Wahrnehmung

57 Margret Hund-Georgiadis

Die Organisation von Sprache und ihre Reorganisation bei ausgewählten, neurologischen Erkrankungen gemessen mit funktioneller Magnetresonanztomographie – Einflüsse von Händigkeit, Läsion, Performanz und Perfusion

58 Jutta L. Mueller

Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study

59 Franziska Biedermann

Auditorische Diskriminationsleistungen nach unilateralen Läsionen im Di- und Telenzephalon

60 Shirley-Ann Rüschemeyer

The Processing of Lexical Semantic and Syntactic Information in Spoken Sentences: Neuroimaging and Behavioral Studies of Native and Non-Native Speakers

61 Kerstin Leuckefeld

The Development of Argument Processing Mechanisms in German. An Electrophysiological Investigation with School-Aged Children and Adults

62 Axel Christian Kühn

Bestimmung der Lateralisierung von Sprachprozessen unter besonderer Berücksichtigung des temporalen Cortex, gemessen mit fMRT

63 Ann Pannekamp

Prosodische Informationsverarbeitung bei normalsprachlichem und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

64 Jan Derrfuß

Functional specialization in the lateral frontal cortex: The role of the inferior frontal junction in cognitive control

65 Andrea Mona Philipp

The cognitive representation of tasks
Exploring the role of response modalities using the task-switching paradigm

66 Ulrike Toepel

Contrastive Topic and Focus Information in Discourse – Prosodic Realisation and Electrophysiological Brain Correlates

67 Karsten Müller

Die Anwendung von Spektral- und Waveletanalyse zur Untersuchung der Dynamik von BOLD-Zeitreihen verschiedener Hirnareale

68 Sonja A.Kotz

The role of the basal ganglia in auditory language processing: Evidence from ERP lesion studies and functional neuroimaging

69 Sonja Rossi

The role of proficiency in syntactic second language processing: Evidence from event-related brain potentials in German and Italian

70 Birte U. Forstmann

Behavioral and neural correlates of endogenous control processes in task switching

71 Silke Paulmann

Electrophysiological Evidence on the Processing of Emotional Prosody: Insights from Healthy and Patient Populations

72 Matthias L. Schroeter

Enlightening the Brain – Optical Imaging in Cognitive Neuroscience

73 Julia Reinholz

Interhemispheric interaction in object- and word-related visual areas

74 Evelyn C. Ferstl

The Functional Neuroanatomy of Text Comprehension

75 Miriam Gade

Aufgabeninhibition als Mechanismus der Konfliktreduktion zwischen Aufgabenrepräsentationen

76 Juliane Hofmann

Phonological, Morphological, and Semantic Aspects of Grammatical Gender Processing in German

77 Petra Augurzky

Attaching Relative Clauses in German – The Role of Implicit and Explicit Prosody in Sentence Processing

78 Uta Wolfensteller

Habituelle und arbiträre sensomotorische Verknüpfungen im lateralen prämotorischen Kortex des Menschen

79 Päivi Sivonen

Event-related brain activation in speech perception: From sensory to cognitive processes

80 Yun Nan

Music phrase structure perception: the neural basis, the effects of acculturation and musical training

81 Katrin Schulze

Neural Correlates of Working Memory for Verbal and Tonal Stimuli in Nonmusicians and Musicians With and Without Absolute Pitch

82 Korinna Eckstein

Interaktion von Syntax und Prosodie beim Sprachverstehen: Untersuchungen anhand ereigniskorrelierter Hirnpotentiale

83 Florian Th. Siebörger

Funktionelle Neuroanatomie des Textverstehens: Kohärenzbildung bei Witzen und anderen ungewöhnlichen Texten

84 Diana Böttger

Aktivität im Gamma-Frequenzbereich des EEG: Einfluss demographischer Faktoren und kognitiver Korrelate