Journal of Psycholinguistic Research, Vol. 30, No. 3, 2001

Syntactic Working Memory and the Establishment of Filler-Gap Dependencies: Insights from ERPs and fMRI

Christian J. Fiebach,[1,3] Matthias Schlesewsky,[2] and Angela D. Friederici[1]

In this contribution, we review an ERP experiment and an fMRI experiment which investigated the processing of German wh-questions. On the basis of the ERP results, we will discuss current models of sentence processing and resource distribution during sentence comprehension. We argue that there exists a separate cognitive or neural resource that supports syntactic working memory processes necessary for the temporary maintenance of syntactic information for the parser. In the context of wh-movement, such a memory component is necessary for establishing filler-gap dependencies. The data obtained from the fMRI experiment will be used to discuss the results of previous neuroimaging studies of sentence processing. It is claimed that syntactic working memory, rather than syntactic processing per se, is supported by Broca's Area.

KEY WORDS: syntax; working memory; wh-questions; Broca's area; ERPs; fMRI.

INTRODUCTION

One of the main goals of psycholinguistic research in the last decades has been the search for the locus of processing difficulty in syntactically complex sentences, be it in functional–cognitive terms or with respect to neuroanatomical correlates. Behavioral studies (e.g., Ford, 1983; King & Just,

The authors wish to thank Stefan Frisch, Gabriele Lohmann, Stefan Zysset, and Karsten Müller for their support in methodological issues at various stages of this work. Thanks are also due to Ina Bornkessel for helpful commentaries on an earlier version of this manuscript. The work presented here was supported by Leibniz Science Prize awarded to A. D. F. by the German Research Foundation.

[1] Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany.
[2] Department of Linguistics, University of Potsdam, Leipzig, Germany.
[3] To whom all correspondence should be mailed. email:fiebach@cns.mpg.de

1991), studies using event-related brain potentials (ERPs; e.g., King & Kutas, 1995; Mecklinger *et al.*, 1995), or neuroimaging studies (e.g., Caplan *et al.*, 1998, 2000; Just *et al.*, 1996; Stromswold *et al.*, 1996) have focused on processing differences between relative clause sentences of different syntactic complexity. One classic case, used in many of these studies, is the contrast between sentences with embedded subject relative clauses and embedded object relative clauses (cf. example 1).

(1a) The reporter$_i$ who$_i$ ______ $_i$ attacked the senator admitted the error.
(1b) The reporter$_i$ who$_i$ the senator attacked ______ $_i$ admitted the error.

However, the nature of the processing difficulty differentiating between such sentence types is far from clear. For example, object relative clauses like (1b) differ from subject relative clauses like (1a) in at least two aspects. First, they contain a noncanonical argument order (i.e., object before subject) and, second, the distance between the relative pronoun "who" (which serves as wh-filler) and the position in the sentence structure from which it was extracted (i.e., gap or trace position; indicated in the above examples by "______") is longer in object relative clauses. The same is true for the comparison of right-branching relative clause constructions containing a subject relative clause with center-embedded object relative clauses (cf. Caplan *et al.*, 1998, Exp. I; Stromswold *et al.*, 1996). Both features of object relatives have been assumed to contribute to processing difficulty. However, it is also the case that the word order of subject and object relatives differs on the surface (cf. 1a,b). This feature of English makes it hard to investigate these sentences with online measures, such as ERPs because constituents of different word classes are compared at critical positions. For example, reading time effects reported for the object relative clause verb "attacked" (e.g., King & Just, 1991) were obtained from the comparison with the noun "senator," which is perceived at the same position in the subject relative clause.

It has been recently suggested that two aspects of syntactic processing, syntactic working memory costs, on the one hand, and linguistic integration costs, on the other, contribute to the processing difficulty of sentences that are assumed to be syntactically more complex (Gibson, 1998). In the framework of Syntactic Prediction Locality Theory (SPLT), Gibson (1998) suggested that integration costs arise when new input has to be integrated into the developing mental representation of the sentence that is being processed. Memory costs in the SPLT framework are elicited when syntactic predictions (such as the phrasal categories that are minimally required to form a grammatical sentence) have to be maintained over regions of the sentence. As Gibson (1998) demonstrated, these two aspects of parsing can account for the processing difficulty observed in object relative clause sentences.

We adopt this view by assuming that the cognitive mechanisms responsible for syntactic processing must provide cognitive and/or neural resources (cf. Caplan & Waters, 1999; Just & Carpenter, 1992) for both "local" syntactic operations and long-duration syntactic memory processes. We assume that syntactic memory costs are elicited whenever unintegrated syntactic information has to be kept activated during ongoing sentence processing. This assumption is in line with the Active Filler Strategy (cf. Clifton & Frazier, 1989; Frazier & Flores D'Arcais, 1989) according to which dislocated fillers are actively maintained in memory until they can be associated with their gap positions.

The present paper presents an investigation of the processes that are responsible for maintaining unintegrated syntactic information in working memory by means of ERPs and functional magnetic resonance imaging (fMRI). In the following sections, two experiments studying the processing of German wh-questions are described. Unlike the English relative clause sentences mentioned above, German wh-questions can be employed to manipulate processes eliciting local integration costs and processes subserving sustained working memory mechanisms independently of one another. Importantly, the wh-fillers used in our experiments were not ambiguous (like the English "who"), but were case-marked for nominative or accusative and, therefore, could be easily identified as the subject and object of the sentences. It has been demonstrated using the self-paced reading method that even when there is no ambiguity with respect to the interpretation of the filler as subject or object, processing times in object wh-questions are still increased as compared to subject-first questions (cf. Fanselow *et al.*, 1999). This finding has been interpreted as demonstrating an increased demand on working memory resources in object-first sentences due to the fact that the dislocated object can not be associated with its gap immediately. We took this finding as a starting point for the present experiments.

On the basis of the results of our experiments, we will discuss (1) the role of syntactic working memory in establishing filler-gap relationships, (2) implications of our results for parsing models, especially with respect to the Active Filler Strategy (e.g., Clifton & Frazier, 1989; Frazier & Flores D'Arcais, 1989), to Syntactic Prediction Locality Theory (Gibson, 1998), and to the "Direct Association Hypothesis" put forward by Pickering and Barry (1991), and (3) implications concerning the fractionation and neural substrate of syntactic working-memory resources.

AN ERP STUDY OF GERMAN WH-QUESTIONS

Methods and Results

In a first step, we investigated the processing of German wh-questions with event-related brain potentials (cf. Fiebach *et al.*, submitted, a). Stimulus material were embedded wh-questions with case-marked masculine interrogative pronouns (i.e., "wer"$_{\text{NOM}}$ and "wen"$_{\text{ACC}}$). In this experiment, two factors were varied: First, subject wh-questions were contrasted with object wh-questions, and second, the length of the filler-gap distance within the object wh-questions was varied between one prepositional phrase (PP; two words) and two PPs (six words). The general structure of the stimuli was as follows: [[Matrix clause], [[wer/wen] [1 vs. 2 PPs] [2nd NP] [participle] [aux]]]. Table I gives an overview over the sentence material. The wh-questions were presented visually in a phrase-by-phrase mode (i.e., for NPs and PPs, articles or prepositions and nouns were presented together in one frame; e.g., "*der Doktor*"). After the phrase-wise presentation of each item, an assertion was displayed on the screen that restated certain aspects of the wh-question. In order to control for comprehension, in one half of these assertions, some information delivered by the item was exchanged. The assertions had to be judged for correctness in relation to the preceding item. For further methodological details see Fiebach *et al.* (submitted, a).

In order to investigate syntactic working memory processes that are activated during the processing of wh-questions, we calculated multiword ERPs (e.g., King & Kutas, 1995) that started at the onset of the wh-filler

Table I. Sentence Material of the ERP Experiment[a] with Word-for-Word Translations to English

Short wh-questions
Short subject wh-questions
Thomas fragt sich, wer am Mittwoch den Doktor verständigt hat.
Thomas asks himself, who$_{\text{(NOM)}}$ on wednesday the$_{\text{(ACC)}}$ doctor called has.
Short object wh-questions
Thomas fragt sich, wen$_i$ *am Mittwoch der Doktor*$\_\__i$ *verständigt hat.*
Thomas asks himself, who$_{\text{(ACC)}}$ on wednesday the$_{\text{(NOM)}}$ doctor called has.
Long wh-questions
Long subject wh-questions
Thomas fragt sich, wer am Mittwoch nachmittag nach dem Unfall den Doktor verständigt hat.
Thomas asks himself, who$_{\text{(NOM)}}$ on wednesday afternoon after the accident the$_{\text{(ACC)}}$ doctor called has.
Long object wh-questions
Thomas fragt sich, wen$_i$ *am Mittwoch nachmittag nach dem Unfall der Doktor*$\_\__i$ *verständigt hat.*
Thomas asks himself, who$_{\text{(ACC)}}$ on wednesday afternoon after the accident the$_{\text{(NOM)}}$ doctor called has.

[a]Note that in Table I and in Table II, no gaps are included in subject wh-questions. Although it might be a matter of debate whether or not the gap position is located before or after the PP(s) in subject-first questions, we assume that perceiving a subject in clause-initial position, even if it is a moved constituent, does not elicit processing difficulty as it immediately fulfills the expectations regarding the order of arguments and, thus, can be integrated directly into the phrase structure representation.

and spanned the embedded question. ERPs from 22 participants revealed a sustained negativity for long object wh-questions, as compared to long subject questions (Fig. 1A). In contrast, no sustained ERP differences were found between the two short conditions (Fig. 1B). The negativity for long object wh-questions started at the first element after the wh-filler (i.e., at the first PP) and was distributed over left-anterior electrode sites (cf. Fiebach

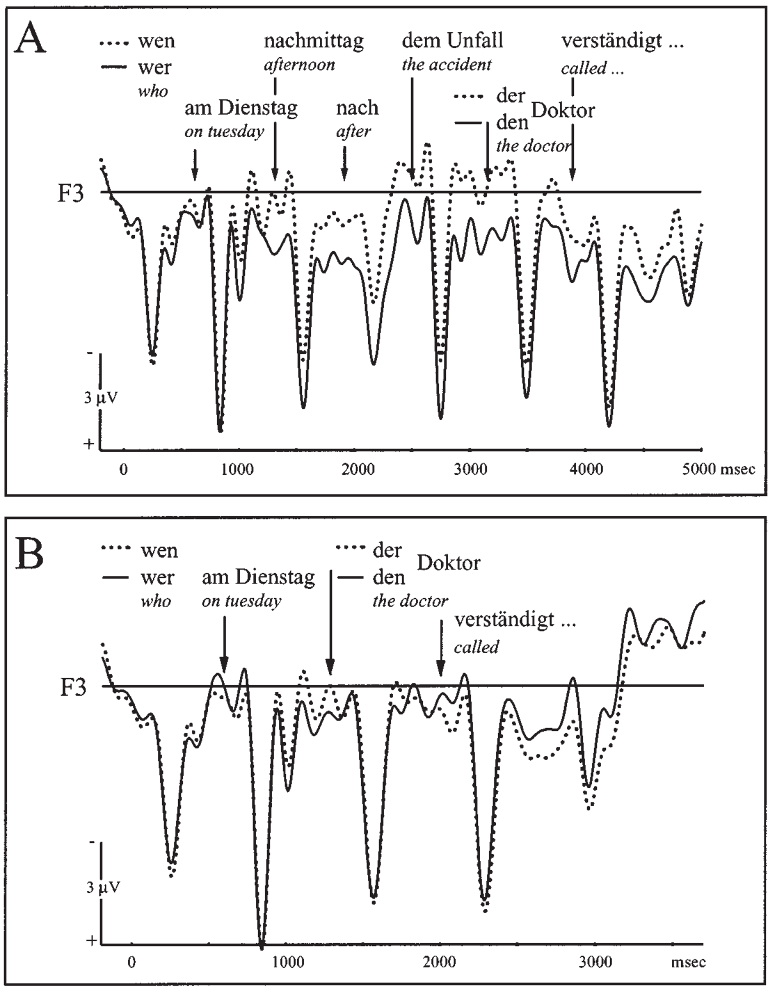

Fig. 1. Multiword ERPs for object wh-questions (dotted line) and subject wh-questions (solid line) over the time course of the embedded questions from one selected electrode position over left anterior brain regions. (A) Long wh-questions; (B) short wh-questions.

et al., submitted, a). Visual inspection of the ERPs to the short questions reveals that there actually was a transient negativity elicited for short object wh-questions at the prepositional phrase (cf. Fig. 1B). A separate analysis using a shorter time window of 200 ms revealed that at the first prepositional phrase, the negativity was present in both long and short wh-questions, distributed mainly over frontal electrode sites. Sentence length (which participants could not know at this point in the sentence) did not influence this negativity. In the long object questions, the negativity became more pronounced as the prolonged PPs had to be processed. In contrast, the transient negativity for short object questions did not develop into a sustained effect. This difference is obviously due to the fact that in short object questions, a noun phrase marked for nominative case (instead of another long PP) is perceived directly after the first prepositional phrase.

Frontal slow-wave negativities have been reported to reflect effort during retention intervals in working memory tasks (e.g., Ruchkin *et al.,* 1990). In line with previous research in sentence processing (e.g., King & Kutas, 1995; Kluender *et al.,* 1998), we take this sustained ERP component to reflect an increased demand on working memory during the processing of object wh-questions. More precisely, we assume that the sustained left-anterior negativity reflects the maintenance of the wh-filler in working memory until the syntactic dependency between the dislocated filler and the gap at its base position can be established. In short object questions, the filler-gap relationship can be established faster and working memory can be freed earlier from the load of the filler than in long questions. Although the distribution of the sustained negativity was very similar to that observed in LAN effects elicited by syntactic violations (e.g., Friederici *et al.,* 1996; Hahne & Friederici, 1999; Neville *et al.,* 1991; Rösler *et al.,* 1993), the temporal and functional characteristics of the LAN component clearly differ from the sustained left-anterior negativity observed here. Nevertheless, the topographical distribution suggests that neural generators of the two components at least partially overlap. The time course of the sustained left-anterior negativity revealed that it increased in strength until the element before the second NP was encountered, thus lending further support to its interpretation as a working-memory effect: the longer the moved wh-filler had to be kept available, the more resources had to be spent. At the second NP, the negativity almost disappeared (cf. Fiebach *et al.,* submitted, a).

Additional support for the assumption that the sustained left-anterior negativity is associated with working memory processes comes from the fact that it was stronger and more broadly distributed in participants with low working memory capacity (as determined with the reading span task; cf. Daneman & Carpenter, 1980) than in participants with a high working memory capacity. Furthermore, in low-span individuals, the negativity appeared earlier than in the high-span group. This interaction of span group with sentence condition in the long wh-questions is strong evidence for an involvement of working memory processes (see also Friederici *et al.*, 1998). Note that the reading span task is not a measure of the capacity of phonological working memory, as specified in the multicomponent working memory model of Baddeley and colleagues (e.g., Baddeley & Logie, 1999). Rather, it was designed to assess working memory for language, encompassing resources for both storage and computation (e.g., Just & Carpenter, 1992).

In addition to the sustained left-anterior negativity, we observed a local positive ERP component for object wh-questions relative to their subject counterparts at the position of the second NP. This positivity was broadly distributed and appeared in the time range of the P600 (i.e., between 400 and 700 ms). We interpreted this positivity as indexing the establishment of the syntactic dependency between the filler and its gap and the integration of this syntactic chain (cf. De Vincenzi, 1991) into the phrase structure representation. This interpretation of the observed positive ERP component follows the assumption put forward by Kaan *et al.* (2000) that the P600 can index the difficulty of syntactic integration processes. In contrast to the sustained negativity, the positivity at the second NP was present for long and short object wh-questions alike. Furthermore, it was not affected by individual differences in working memory capacity.

Discussion

On the basis of our ERP data, we conclude that gaps and the corresponding filler-gap dependencies play an important role during parsing. Furthermore, we suggest that the psychological mechanism that allows filler-gap relations to be established is a syntactic working memory process that maintains the filler actively available until the gap is licensed and the syntactic chain between filler and gap can be established. This conclusion has some implications for issues currently under debate in the psycholinguistic literature. With regard to the Active Filler Strategy (Clifton & Frazier, 1989; Frazier & Flores D'Arcais, 1989), our data support the assumption that the filler, or at least its syntactic fingerprint, is available while new input is continuously being processed. Furthermore, there are two important additional insights to be gained from our results: First, we propose that the mechanism that allows active maintenance of the filler is syntactic working memory. Second, maintenance of the filler in syntactic working memory is not bound to situations where a temporary ambiguity of the filler has to be resolved (like, e.g., in English, Dutch, or Italian). Rather, it is a general feature of the parser that dislocated arguments perceived in noncanonical positions cannot be processed immediately. In such cases, integration into the mental representation of the phrase structure has to be delayed until a sentential element that enables anticipation of the trace position (e.g., a subject NP) is perceived (for similar ideas see, e.g., Crocker, 1994).

At first glance, our results might be taken to contradict the large body of evidence that speaks for reactivation of fillers at the gap position (e.g., Nicol & Swinney, 1989; Nicol *et al.*, 1994). However, this is not the case as it appears to be very likely that it is only the syntactic information that is maintained in working memory. This assumption allows the parser to rapidly assign a gap at a potential gap site rather than analyzing the lexical input at this position (as assumed by the Active Filler Strategy; Clifton & Frazier, 1989). Nevertheless, when the gap is identified, the semantics of the filler have to be reactivated. Antecedent reactivation obviously limps slightly behind the establishment of the syntactic chain between filler and gap. Thus, rather than contradicting the cross-modal priming studies, our data lend further support to the main conclusion from this work, namely, that linguistic gaps have a psychological reality and play an important role in parsing.

Having stated this point, it is necessary to discuss how our data relate to one influential alternative account to trace-based models, namely the Direct Association Hypothesis (cf. Pickering, 2000; Pickering & Barry, 1991). According to this traceless approach, dislocated verbal arguments are directly associated with the corresponding verb. This model was motivated by reinterpreting the above-cited results from cross-modal priming experiments as being caused by reactivation at the verb itself, instead of the gap position directly following the verb. It has been pointed out, however, that the direct association hypothesis does not hold when a broad number of syntactic structures is considered (Gorrell, 1993). As German is an SOV language, for sentences like the wh-questions investigated in our experiment, direct association would predict that processing effects due to direct association occur at the final element of the clause, i.e., after the assumed gap position. Our data, however, clearly demonstrated that integration costs are increased at the second noun phrase position. The conclusion to be drawn is that the linkage between the filler and the gap can already be established before the gap is actually encountered. We have suggested that as soon as the processed input licenses the projection of the phrase structure into the verb phrase, the canonical object argument position (i.e., the gap position) can be predicted and linked to the dislocated wh-filler (cf. Fiebach *et al.*, submitted, a). In German, this is possible when the subject-NP becomes available. For example, Muckel and Pechmann (2000) reported a cross-modal priming study using sentences with topicalized objects. In this study, they demonstrated reactivation effects at the preverbal trace position and concluded, converging with the assumptions stated here, that the parser can anticipate traces. In English, the object trace becomes available when the verb is processed, thus causing increased computational load at the verb (cf. also King & Just, 1991). The evidence that was taken to support the direct association hypothesis, accordingly, appears to be a reflection of an incremental gap prediction mechanism. Consideration of German object-first sentences clearly shows that the direct association hypothesis is not applicable to SOV languages.

The described ERP pattern (i.e., a sustained left-anterior negativity and a local positivity reflecting integration difficulty) is in correspondence with the assumption of two separable causes of processing difficulty as postulated in SPLT (Gibson, 1998). While the sustained negativity reflects syntactic working memory costs that unfold over a longer time domain, the positivity is a marker of locally increased integration costs. However, SPLT in the version published by Gibson (1998) would not predict the pattern of integration costs observed in our data. First, it ascribes the integration costs to the verb. This, however, is due to the SVO structure of English. Our data (as well as other results; e.g., Bader & Lasser, 1994; Muckel & Pechmann, 2000) show that integration processes already take place at the second NP in an SOV language. Second, even at the position of the second NP, SPLT would not predict differences between subject and object wh-questions in German. SPLT's integration cost component is driven by the locality of the integration that takes place, i.e., the integration of two elements like the two argument NPs in the present case, is more costly when a greater distance has to be crossed between the two elements. The distance was suggested to be determined on the basis of intervening discourse referents (Gibson, 1998). In the wh-questions investigated here (cf. Table I), the number of discourse referents (or even the number of intervening words) did not differ between subject and object wh-questions. Thus, according to SPLT, no processing difficulty should be observed at the point where the positivity for object questions occurred in Experiment 1.

As already mentioned in the previous section, we interpret the obtained positivity as reflecting the establishment of a syntactic chain between the filler and its gap and the integration of this chain into the phrase structure representation. It can be assumed that in both subject and object wh-questions, the second NP is integrated into the phrase structure and that this process is more or less equally costly in both conditions. However, in the object questions, in addition to this process, the filler is associated with its gap, causing the greater processing load at this position.

Last, but not least, the question remains of how the two types of syntactic processes that were described relate to the processing resources that are generally postulated in psycholinguistic models. Some approaches claim that there is a general cognitive resource dedicated to the processing of language (cf. Just & Carpenter, 1992). This assumption seems to be contradicted by our data, as the two ERP components observed were differentially influenced by the length of the filler-gap distance and by individual working memory capacity. Others assume a subdivision into (at least) two resource types. For example, Caplan and Waters (1999) recently suggested that there is one component that is responsible for extracting meaning from sentences on the basis of syntactic and prosodic information, but also using sentence-level semantics. This resource component was assumed to be responsible for "interpretive processing." A second "postinterpretive" processing resource is drawn upon when the propositional content of sentences is used for performing a task. In this model, both aspects of syntactic processing that, as suggested within SPLT (Gibson, 1998), contribute to processing difficulty have to be accommodated under the interpretive processing resource. As the different sensitivity to working memory manipulations of the two ERP components described above suggests that they tap different processing resources, we tentatively propose that there has to be more than just one cognitive or neural processing resource subserving syntactic processing. This proposal leads to the assumption of subdivisions within the "interpretive" processing resource postulated by Caplan and Waters (1999), at least into one portion that subserves the actual computational processes of syntactically analyzing and integrating new input and a second component that is responsible for temporary maintenance of unintegrated syntactic information.

AN fMRI STUDY OF GERMAN WH-QUESTIONS

Methods and Results

In order to determine the neural correlates of the syntactic working memory mechanism identified in the first experiment, we conducted a functional magnetic resonance imaging (fMRI) study. In this experiment, exactly the same long wh-questions were used as in the ERP experiment. However, in order to equate the amount of sensory input in long and short sentence conditions, and to allow a direct contrast between short and long object wh-questions, we slightly changed the short sentence conditions. Instead of including just one PP in the short questions, we included the same prolonged prepositional phrases as in the long wh-questions, but placed them after the second noun phrase (cf. Table II).

As already pointed out in the introduction, previous neuroimaging studies of syntactic complexity confounded the two potential sources of processing difficulty in object relative clause sentences, namely the noncanonical argument order (i.e., object before subject) and an increased filler-gap distance. As we have demonstrated in the first experiment, these two manipulations can lead to an increase in processing difficulty. With the present design,

Table II. Sentence Material of the fMRI Experiment

Short wh-questions
Short subject wh-questions *Thomas fragt sich, wer den Doktor am Mittwoch nachmittag nach dem Unfall verständigt hat.*
Short object wh-questions *Thomas fragt sich, wen$_i$ der Doktor ___ $_i$ am Mittwoch nachmittag nach dem Unfall verständigt hat.*
Long wh-questions
Long subject wh-questions *Thomas fragt sich, wer am Mittwoch nachmittag nach dem Unfall den Doktor verständigt hat.*
Long object wh-questions *Thomas fragt sich, wen$_i$ am Mittwoch nachmittag nach dem Unfall der Doktor ___ $_i$ verständigt hat.*

we can examine two important contrasts: (1) canonical vs. noncanonical word order (i.e., subject vs. object questions, both in long and in short sentence conditions), and (2) short vs. long filler-gap distance in object wh-questions. Importantly, in the second contrast, the order of arguments is kept constant. The manipulation of the filler-gap distance was thought to isolate the neural substrates of the above-described syntactic working memory component. The first contrast, on the other hand, was supposed to yield insights into the brain areas subserving the computation of complex, but short-lived, syntactic integration processes. More precisely, the fact that in object conditions the parser has to deal with sentences with noncanonical argument order should make the processing of object questions more difficult. When contrasting long subject and object questions, noncanonical argument order and the long filler-gap distance are confounded in the object questions. However, short object questions have only a very short filler-gap distance but also, naturally, the noncanonical word order. Therefore, by identifying areas similarly activated in both contrasts between object and subject questions, effects of noncanonical word order should be identifiable. Note, however, that the complexity of the phrase structure of German subject and object wh-questions (as determined, for example, on the basis of nonterminal nodes in the phrase structure tree) does not differ.

The results of the calculated statistical contrasts were striking (cf. Fiebach *et al.,* submitted, b). Contrary to our predictions based on the previous neuroimaging evidence discussed above, both contrasts between object and subject wh-questions did not reveal activation differences in the left inferior frontal gyrus. In fact, the only brain area showing differences between subject and object questions at all was a small area within the left-anterior portion of the superior temporal sulcus in which stronger activation was elicited for long object as compared to long subject wh-questions. However, this statement is qualified by the fact that this effect was only observed when a reduced significance threshold (i.e., $p < .05$) was used that did not control sufficiently for the calculation of multiple comparisons, as is generally required in brain imaging research. This result has to be treated as an initial clue regarding a possible neuroanatomical correlate of costs associated with the difficulty of integration processes.

Reliable activation differences in language-relevant cortical areas were obtained exclusively for long object wh-questions when compared to short object wh-questions (cf. Table III). Neural activity was stronger for long object questions in the inferior frontal gyrus (IFG) of the left hemisphere and also in homologous right hemisphere brain areas. The superior portion of the pars opercularis of the IFG, on the border to pars triangularis (i.e., Brodmann areas 44/45; cf. Fig. 2) was activated on both hemispheres. The inferior portion of pars opercularis (BA 44), which has been suggested as the neural correlate of syntactic parsing processes such as initial phrase structure building (e.g., Friederici, 1998), was selectively activated in the left hemisphere. In addition, bilateral areas along the middle portion of the superior temporal sulcus (BA 22/21) were more strongly activated for object questions with a long wh-movement. These activations were more pronounced for the left hemisphere (cf. Fiebach *et al.*, submitted, b).

Discussion

The described fMRI data were surprising in two respects: First, given the consistent findings of increased activation in Broca's area (including pars triangularis/BA 45, pars opercularis/BA 44, and even pars orbitalis/BA 47 of

Table III. Brain Areas that Showed Reliably Increased Activations for Object wh-Questions with Long Filler–Gap Distance as Contrasted with Object wh-Questions with Short Filler–Gap Distance.[a]

Neuroanatomical location		Brodmann area	Hemisphere
Inferior frontal gyrus			
Pars opercularis and pars triangularis (superior portion)		BA 44/45	L/R
Pars opercularis (inferior portion)		BA 44	L
Superior temporal sulcus	Middle portion	BA 22/21	L/R

[a]All reported areas had z values of $z > 3.4$, ($p < .0003$; for more details see Fiebach, *et al.*, submitted, b).

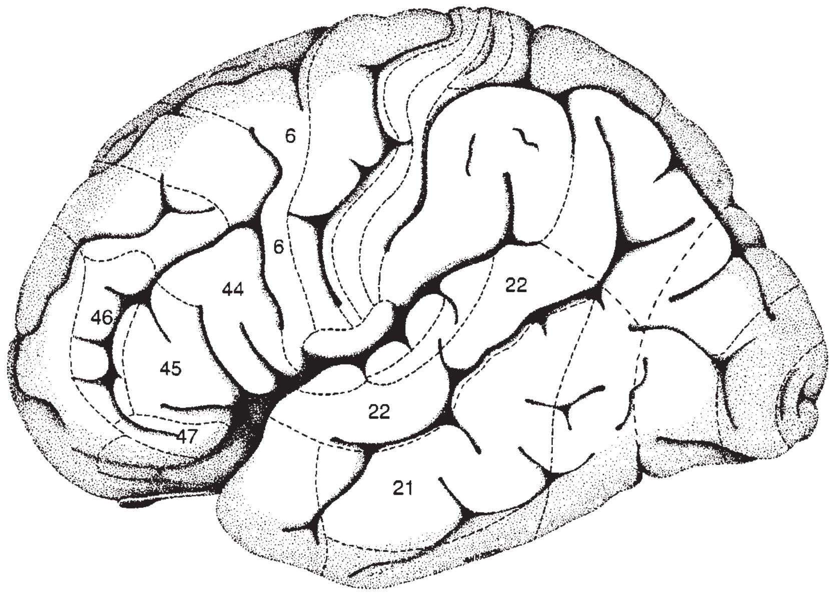

Fig. 2. Brodmann areas of the human brain that were referred to in the text and neighboring cortices (adapted from Duvernoy, 1991).

the left inferior frontal gyrus in different experiments) for object as compared to subject relative-clause sentences (e.g., Caplan *et al.*, 1998; also see Caplan, this volume; Just *et al.*, 1996; Stromswold *et al.*, 1996), our findings were unexpected. Second, given the observation of a sustained left-anterior negativity for long object as compared to subject questions in the ERP experiment, the fMRI finding could not be predicted. However, the difference in the results between the ERP and the fMRI experiment can be due to the fact that these two methodologies strongly differ with respect to the neural mechanisms to which they are sensitive (e.g., Orrison *et al.*, 1995). While fMRI might be insensitive to processes reflected in the negativity observed in the ERP study, the particular contrast that elicited the strong differences in the fMRI experiment could not be examined in the ERP study.[4]

With respect to the previous neuroimaging studies of structural complexity in language comprehension conducted in English, we claim that these studies confounded two aspects contributing to processing difficulty. From the

[4] In the ERP experiment, object wh-questions with short and long filler-gap distance also differed in the absolute length of the sentences (cf. Table I). For this reason, direct contrasts between these two questions could not be calculated.

present fMRI data, we conclude that it is not the syntactic integration cost component that is responsible for brain activation differences between sentences of different complexity, but that the observed activation differences reflect syntactic working memory processes that are required for keeping the dislocated filler active until the filler-gap relationship can be established (cf. Fiebach *et al.,* submitted, b). Such a mechanism is required for processing wh-questions as well as relative clause sentences. Converging evidence comes from two sources: (1) An fMRI study in which subject and object relatives with short or long antecedent-gap linkages were investigated (Grossman, 2000) resulted in a very similar conclusion. It was hypothesized that the "left inferior frontal cortex supports the cognitive resources required to maintain syntactic dependency relations during the comprehension of grammatically complex sentences." (2) Lesion studies also suggest an involvement of the left IFG in syntactic processing. It has been claimed that the linguistic structures that are most impaired in agrammatic aphasics, whose lesions mostly involve Broca's area in the left inferior frontal cortex, are sentences with antecedent-trace relations such as object relatives, object wh-questions, or passives (cf. Grodzinsky, 2000)[5]. As we have argued above, there is strong evidence for an involvement of syntactic working memory in the establishment of filler-gap dependencies in healthy subjects. Accordingly, it is very plausible to assume that Broca's area is critically involved in syntactic working memory processes. While in the context of the trace deletion hypothesis (Grodzinsky, 2000), the deficit of agrammatic aphasics is treated as a representational problem, we suggest on the basis of the present data, that it appears more plausible to ascribe the sentence comprehension deficit of agrammatic aphasics to a problem with the availability or usage of syntactic working memory resources (cf. Friederici & Frazier, 1992; Müller, 2000).

In the last decades, the function of the left inferior frontal gyrus has generally been described as subserving syntactic processing. However, our proposal of the inferior frontal gyrus as contributing to syntactic working memory resources does not necessarily contradict previous results. For example, Stromswold and colleagues (1996) discussed that the effects in Broca's area obtained in their PET experiment might be due to an increased memory demand associated with parsing center-embedded object relative clause sentences as compared to right-branching subject relative clause sentences. In this case, the demand on working memory resources is even greater as dur-

[5] Note, however, that Grodzinsky's trace deletion hypothesis (TDH) does not make the correct predictions for the processing of passives and object-first sentences in agrammatic Broca's aphasics in several languages such as Dutch (Friederici & Graetz, 1987), German (Heeschen, 1980), and Japanese (Hagiwara & Caplan, 1990; cf. also Frisch et al., 2000; Friederici & Gorrell, 1998).

ing the processing of the center-embedded relative clause, the filler has to be maintained and the partially analyzed matrix clause must also be held in working memory. Stowe and colleagues (1998; see also Stowe, 2000) proposed that left IFG supports a verbal working memory capacity that temporarily stores structural information as well as lexical items. The assumption that the activations observed reflect verbal working memory as specified within the multicomponent model of working memory (cf. Baddeley & Logie, 1999) is not unplausible given the fact that neuroimaging studies of verbal working memory (e.g., Paulesu *et al.*, 1993; Smith *et al.*, 1996) consistently associated activation of BA 44 with articulatory rehearsal. However, the assumption that phonological working memory processes are the cause of activation differences during sentence processing is rendered unlikely by the fact that activation increases in Broca's area are also present under conditions of articulatory suppression (Caplan *et al.*, 2000).

We suggest that instead of reflecting working memory processes, the activation of pars opercularis of the IFG (i.e., BA 44) in the context of processing complex sentences is associated with a syntactic working memory resource that is part of the syntactic processing system. This resource is drawn upon whenever phrase structure building has to be delayed and syntactic information cannot be integrated immediately. It is critically involved in establishing syntactic chains between moved constituents and their traces.

The present study also demonstrated that it is not the left inferior frontal gyrus alone that subserves syntactic functions, but that it is a network of inferior frontal and superior temporal regions that is involved. This bilateral network, however, is asymmetric in the sense that left hemispheric activations are more extensive than those of the right hemisphere. A similar involvement of temporal regions has been reported in an fMRI study in which the reading of sentences of increasing structural complexity was investigated (Just *et al.*, 1996) and in an event-related fMRI experiment studying the processing of jabberwocky sentences (Friederici *et al.*, 2000).

CONCLUSION

To conclude, the data reported here from an ERP and an fMRI experiment with German wh-questions provide strong evidence for the existence of a working memory component that is responsible for temporarily maintaining unintegrated syntactic information during ongoing sentence processing. In the context of wh-movement, this syntactic working memory is critically involved in establishing filler-gap dependencies. With respect to the distribution of cognitive resources, our data suggest that syntactic working memory constitutes a separate domain within the resources that subserve interpretive processing of sentences. The neural substrate of the syntactic working memory component appears to be a bilateral network of inferior frontal and superior temporal brain regions, with a left lateralization within the inferior portion of the pars opercularis of the left inferior frontal gyrus (BA 44). Our results provide an interesting starting point for further explorations into the distribution of cognitive and neural resources during the online comprehension of sentences. For this goal, the integration of electrophysiological and neuroimaging methods appears especially fruitful as it provides evidence from normal processing in healthy individuals with a good temporal and spatial (i.e., neuroanatomical) resolution.

REFERENCES

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multi-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge, UK: Cambridge University Press.

Bader, M., & Lasser, I. (1994). German verb-final clauses and sentence processing: Evidence for immediate attachment. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Advances in sentence processing* (pp. 225–242). Hillsdale, NJ: Erlbaum.

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

Caplan, D., Alpert, D., Waters, G., & Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping, 9,* 65–71.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–126.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht: Kluwer Academic.

Crocker, M. W. (1994). On the nature of the principle-based sentence processor. In C. Clifton, L. Frazier & K. Rayner (Eds.), *Advances in sentence processing* (pp. 245–266). Hillsdale, NJ: Erlbaum.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian.* Dordrecht: Kluwer Academic.

Duvernoy, H. (1991). *The human brain.* Wien: Springer.

Fanselow, G., Kliegl, R., & Schlesewsky, M. (1999). Processing difficulty and principles of grammar. In S. Kemper & R. Kliegl (Eds.), *Constraints on language* (pp. 171–201). Dordrecht: Kluwer Academic.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (submitted, a). An ERP investigation of syntactic working memory during the processing of German WH-questions.

Fiebach, C. J., Friederici, A. D., & von Cramon, D. Y. (submitted, b). Syntactic complexity revisited: An event-related fMRI study of syntactic working memory.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 203–218.

Frazier, L., & Flores D'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language, 28,* 331–344.

Friederici, A. D. (1998). The neurobiology of language processing. In A. D. Friederici (Ed.), *Language comprehension: A biological perspective* (pp. 263–301). Berlin: Springer.

Friederici, A. D., & Frazier, L. (1992). Thematic analysis in agrammatic comprehension: Syntactic structures and task demands. *Brain and Language, 42,* 1–29.

Friederici, A. D., & Gorrell, P. (1998). Structural prominence and agrammatic theta-role assignment: A reconsideration of linear strategies. *Brain and Language, 65,* 253–275.

Friederici, A. D., & Graetz, P. (1987). Processing passive sentences in aphasia: Deficits and strategies. *Brain and Language, 30,* 93–105.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1219–1248.

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 and Language, 74,* 289–300.

Friederici, A. D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology, 47,* 193–221.

Frisch, S., Saddy, D., & Friederici, A. D. (2000). Cutting a long story (too) short. *Behavioral and Brain Sciences, 23,* 34–35.

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

Gorrell, P. (1993). Evaluating the direct association hypothesis: A reply to Pickering and Barry. *Language and Cognitive Processes, 8,* 129–146.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences, 23,* 1–21.

Grossman, M. (2000). Neural basis for sentence processing: fMRI studies of healthy adults and patients with frontotemporal dementia. Paper presented at the *13th CUNY Conference on Human Sentence Processing, La Jolla, California.*

Hagiwara, H., & Caplan, D. (1990). Syntactic comprehension in Japanese aphasics: Effects of category and thematic order. *Brain and Language, 38,* 159–170.

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,* 194–205.

Heeschen, C. (1980). Strategies of decoding actor-object-relations by aphasic patients. *Cortex, 16,* 5–19.

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

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science, 274,* 114–116.

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

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

King, J. W., & 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,* 376–395.

Kluender, R., Münte, T., Cowles, H. W., Szentkuti, A., Walenski, M., & Wieringa, B. (1998). Brain potentials to English and German questions. Poster presented at the *Annual Meeting of the Cognitive Neuroscience Society.*

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis of event-related potentials. *Memory & Cognition, 23,* 477–494.

Muckel, S., & Pechmann, T. (2000). Does the parser search for traces? Poster presented at the 13th *CUNY Conference on Human Sentence Processing,* La Jolla, California.

Müller, R.-A. (2000). A big “housing” problem and a trace of neuroimaging: Broca’s area is more than a transformation center. *Behavioral and Brain Sciences, 23,* 42.

Neville, H., 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,* 151–165.

Nicol, J. L., Fodor, J. D., & Swinney, D. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1229–1238.

Nicol, J. L., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–19.

Orrison, W. W., Lewine, J. D., Sanders, J. A., & Hartshorne, M. F. (1995). *Functional brain imaging.* St. Louis: Mosby.

Paulesu, E., Frith, C. D., & Frackowiak, R. (1983). The neural correlates of the verbal component of working memory. *Nature (London), 368,* 342–345.

Pickering, M. J. (2000). No evidence for traces in sentence comprehension (Commentary). *Behavioral and Brain Sciences, 23,* 47–48.

Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6,* 229–259.

Rösler, F., Pütz, P., Friederici, A. D., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5,* 345–362.

Ruchkin, D. S., Johnson Jr., R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and Clinical Neurophysiology, 76,* 419–439.

Smith, E. E., Jonides, J., & Koeppe, R. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex, 6,* 11–20.

Stowe, L. A. (2000). Sentence comprehension and the left inferior frontal gyrus: Storage, not computation (Commentary). *Behavioral and Brain Sciences, 23,* 51.

Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., & Zwarts, F. (1998). Localizing components of complex task: Sentence processing and working memory. *Neuroreport, 9,* 2995–2999.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by Positron Emmission Tomography. *Brain and Language, 52,* 452–473.