COGNITION

Cognition 85 (2002) B21–B30

www.elsevier.com/locate/cognit

Brief article

Grammar overrides frequency: evidence from the online processing of flexible word order

Ina Bornkessel[a,*], Matthias Schlesewsky[b], Angela D. Friederici[a]

[a]*Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany*
[b]*Department of Linguistics, University of Potsdam, Potsdam, Germany*

Received 22 August 2001; received in revised form 8 January 2002; accepted 18 April 2002

Abstract

We show that online processing difficulties induced by word order variations in German cannot be attributed to the relative infrequency of the constructions in question, but rather appear to reflect the application of grammatical principles during parsing. Event-related brain potentials revealed that dative-marked objects in the initial position of an embedded sentence do not elicit a neurophysiologically distinct response from subjects, whereas accusative-marked objects do. These differences are predictable on the basis of grammatical distinctions (i.e. underlying linguistic properties), but not on the basis of frequency information (i.e. a superficial linguistic property). We therefore conclude that the former, but not the latter, guides syntactic integration during online parsing.

Keywords: Sentence processing; Event-related brain potentials; Word order; Case marking

1. Introduction

Frequency-based accounts of online sentence processing have become increasingly popular with the advent of probabilistic parsing models (e.g. Crocker & Brants, 2000; Jurafsky, 1996). However, it remains to be demonstrated that frequency-based explanations generalize to cover the whole range of observable processing patterns (i.e. generally patterns of processing difficulty of some sort). The present study aims to provide a counter-example to the assumption of a general applicability of frequency-based approaches. Specifically, we will use event-related brain potentials (ERPs) to demonstrate that online processing difficulties induced by word order variations in German cannot be attributed to

* Corresponding author. Max Planck Institute of Cognitive Neuroscience, P.O. Box 50 03 55, 04303 Leipzig, Germany. Tel.: +49-341-9940-114; fax: +49-341-9940-113.
E-mail address: bornke@cns.mpg.de (I. Bornkessel).

PII: S0010-0277(02)00076-8

the relative infrequency of the constructions in question, but rather appear to reflect the application of grammatical principles during parsing.

1.1. Properties of German: grammar vs. frequency

In contrast to languages such as English, German allows a relatively free word order. In particular, nominal constituents that are objects may occur in front of nominal constituents that are subjects. Here, we will be concerned with structures in which an object precedes the subject *clause-medially*, i.e. immediately following the finite verb in the second position (1a) or the complementizer *dass* ('that', 1b).

(1) a. Gestern hat den Jäger der Gärtner beruhigt.
yesterday has [the hunter]$_{\text{OBJECT}}$ [the gardener]$_{\text{SUBJECT}}$ calmed
'Yesterday, the gardener calmed the hunter.'

b. … dass den Jäger der Gärtner beruhigte.
… that [the hunter]$_{\text{OBJECT}}$ [the gardener]$_{\text{SUBJECT}}$ calmed
'… that the gardener calmed the hunter.'

Despite this apparent word-order freedom, it has consistently been observed that (unambiguous) sentences with an object-before-subject word order are more difficult to process than their subject-initial counterparts (e.g. Fanselow, Schlesewsky, & Kliegl, 2000; Fiebach, Schlesewsky, & Friederici, 2001; Rösler, Pechmann, Streb, Röder, & Hennighausen, 1998). Most importantly for present purposes, previous studies using ERPs have shown that an object immediately following the finite verb (as in 1a) gives rise to processing difficulties which are reflected in a negative deflection approximately 300–450 ms post-onset of the word signalling a non-canonical structure (Rösler et al., 1998; Schlesewsky, Bornkessel, & Frisch, in press). One plausible *grammar-based* interpretation of this negativity is that it signals a mismatch between a predicted structural position and the element encountered in this position (Friederici, Schlesewsky, & Fiebach, in press). This account assumes that the parser makes use of the grammatical principles of German in order to predict that the constituent encountered adjacent to the finite verb in sentences such as (1a) (and also to the complementizer in (1b)) will be the subject of the clause, since this is generally the case in sentences with an unmarked word order.[1] When this prediction is not borne out, for example when an accusative-marked argument is encountered adjacent to the finite verb instead, a mismatch occurs which is reflected in the negative deflection in the ERP.

However, proponents of frequency-based processing models might argue that the negativity elicited by elements unambiguously signalling a non-canonical word order reflects

[1] The word order of a sentence is referred to as *unmarked* if this sentence may be felicitously uttered in the absence of any constraining context (the typical test case being a presumably neutral context such as 'What happened?'). There is general agreement that this is the case for the subject-initial counterparts of (1a) and (1b), but not for sentences involving a clause-medial word order variation (like (1a) and (1b)) (e.g. Lenerz, 2001). Note that this definition of markedness is entirely independent of frequency of occurrence. We will use the terms 'marked' and 'non-canonical' interchangeably in the following.

the rapid application of frequency information during parsing, since structures such as (1a) are much less frequent than their subject-initial counterparts: an analysis of the so called 'W-Pub' corpus archive of written German wordforms (Mannheimer Institut für deutsche Sprache; http://www.ids-mannheim.de) showed that, for transitive sentences, the combination of a finite verb in the second position + *der* ('the$_{\text{NOMINATIVE(SUBJECT)}}$') occurs approximately 96 times more often than the combination of a finite verb and a definite non-nominative determiner.

These two hypotheses with regard to the nature of the processing difficulty arising in structures such as (1a) may be contrasted by means of structures for which the factors markedness and frequency diverge. Specifically, while a nominal constituent marked for an object case (i.e. accusative or dative) unambiguously signals a non-canonical word order when this constituent follows the auxiliary *hat* ('has'), matters are more complex when such a constituent follows the complementizer *dass* ('that'). In this case, a constituent marked for the dative case may either indicate a non-canonical word order (2a) or a passive clause (2b).[2]

(2) a. … dass dem Jäger der Gärtner half.
… that [the hunter]$_{\text{DAT-OBJECT}}$ [the gardener]$_{\text{SUBJECT}}$ helped
'… that the gardener helped the hunter.'

b. … dass dem Jäger geholfen wurde.
… that [the hunter]$_{\text{DAT-OBJECT}}$ helped was
'… that the hunter was helped.'

Example (2b) is a canonically ordered German passive clause, i.e. a dative object adjacent to a complementizer in German need not signal a non-canonical word order (cf., for example, Fanselow, 2000; Lenerz, 1977; Primus, 1999; Wunderlich, 1997). In the theoretical literature on German, this phenomenon has been accounted for by the assumption of a structural position following the complementizer which may be occupied (a) by the (nominative-marked) subject of the clause, or (b) by arguments bearing an oblique case (i.e. dative) (cf. den Besten, 1985). Thus, if the processing difficulty that has been observed for objects encountered before the subject clause-medially is of a grammatical origin (i.e. associated in some way with non-canonicity), a dative-marked argument following a complementizer should allow for a way of circumventing the extra processing cost, since the dative noun phrase (NP) may be interpreted as an argument of a passivized verb. By contrast, accusative-marked constituents do not offer such a way out, since these must always be interpreted as signalling a non-canonical word order irrespectively of whether they follow a finite verb or a complementizer.

Crucially, the grammatical differences between initial datives and initial accusatives following a complementizer are not reflected in the relative frequency of these word orders in German. According to the 'W-Pub' corpus archive (Mannheimer Institut für deutsche Sprache), in transitive sentences, the combination of *dass* ('that') + *der* ('the$_{\text{NOMINATIVE}}$')

[2] Note that the passive reading is ruled out by the choice of auxiliary (*hat*, 'has') in dative-initial main clauses analogous to (1a).

occurs approximately eight times more often than the combination of *dass* ('that') + *den* ('the$_{ACCUSATIVE}$') and the combination of *dass* ('that') + *dem* ('the$_{DATIVE}$'), the latter two occurring with approximately the same frequency.

In this way, it is clear that frequency-based accounts of the processing difficulty for non-canonical word orders in German predict accusative-marked objects and dative-marked objects following a complementizer to behave similarly and to both give rise to measurable processing difficulty in comparison to subjects following a complementizer. A grammatically-based account, by contrast, predicts a divergence between the two types of object case: whereas accusative-marked objects should give rise to clear processing difficulties in contrast to subjects, dative-marked objects should behave like subjects, since they need not be interpreted as signalling a non-canonical structure.

In the following, we present an ERP experiment in which the predictions of these two accounts are tested.

2. The present study

The aim of the present study is to examine whether the processing difficulty observed in previous studies for objects preceding a subject clause-medially in German should be attributed to the fact that these structures are less frequent, or whether this difficulty should rather be seen as resulting from the parser's sensitivity to grammatical distinctions during online processing. The critical experimental conditions are shown in Table 1. If language processing makes use of underlying distinctions in the grammar of German, initial accusative objects (D in Table 1) should give rise to a negativity in contrast to initial subjects (A/C) in accordance with previous findings (Rösler et al., 1998), while no such contrast should be observable between initial dative objects (B) and initial subjects (A/C). On the other hand, if sentence processing is more sensitive to surface properties of linguistic

Table 1
Example sentences for each of the critical conditions in the present study.

Condition	Example
A. Nominative-dative	… dass der Jäger dem Gärtner hilft. … that [the hunter]$_{SUBJECT}$ [the gardener]$_{DAT\text{-}OBJECT}$ helps '… that the hunter helps the gardener.'
B. Dative-nominative	… dass dem Jäger der Gärtner hilft. … that [the hunter]$_{DAT\text{-}OBJECT}$ [the gardener]$_{SUBJECT}$ helps '… that the gardener helps the hunter.'
C. Nominative-accusative	… dass der Jäger den Gärtner besucht. … that [the hunter]$_{SUBJECT}$ [the gardener]$_{ACC\text{-}OBJECT}$ visits '… that the hunter visits the gardener.'
D. Accusative-nominative	… dass den Jäger der Gärtner besucht. … that [the hunter]$_{ACC\text{-}OBJECT}$ [the gardener]$_{SUBJECT}$ visits '… that the gardener visits the hunter.'

utterances (i.e. frequency information), initial dative objects (B) and initial accusative objects (D) should behave similarly and contrast with initial subjects (A/C).

Furthermore, the grammatically-based account would also predict that, in the present experimental sentences, the (canonical) passive analysis initially adopted for dative-initial structures (B) can only be maintained until the second argument is encountered, when it becomes clear that this initial analysis is not correct. Thus, a reanalysis should be observable in the form of a positive deflection in the ERP (cf. Friederici & Mecklinger, 1996, among many others) at the position of the second NP in sentences such as (B) in comparison to sentences such as (A).

2.1. Participants

Sixteen students of the University of Leipzig participated in the experiment (eight female; 20–31 years; mean 24.1 years). All were right-handed, native speakers of German with normal or corrected-to-normal vision. The participants were paid DM 13,- per hour.

2.2. Materials

The experimental sentences comprised a matrix clause of the form *Maria hörte* ('Maria heard') and an embedded clause of the form illustrated in Table 1. As shown in Table 1, the crucial experimental manipulations involved varying the word order of the embedded clause (factor: ORDER; subject-object vs. object-subject) and the case of the object argument (factor: CASE; accusative vs. dative), thus giving rise to four crucial conditions: Nominative-Dative (A), Dative-Nominative (B), Nominative-Accusative (C), and Accusative-Nominative (D). Participants were presented with 80 sentences for each of these conditions in a pseudo-randomized manner. The sentences varied with respect to their continuations following the two NPs in order to distract participants from the critical manipulation at the position of the first NP. Furthermore, participants were required to complete a comprehension task following each experimental sentence. This task involved the presentation of a declarative sentence, for which participants had to decide whether it correctly expressed the content of the preceding sentence or not. The comprehension task required the answer 'yes' equally as often as the answer 'no' and the incorrect sentences either involved a substituted first NP, a substituted second NP or a substituted verb.

2.3. Procedure

Sentences were presented visually in the centre of a computer screen in a phrase-by-phrase manner (segmentation: Matrix Clause–NP–NP–Verb). Single words were presented for 450 ms and phrases for 500 ms with an inter-stimulus interval (ISI) of 100 ms. At the end of a sentence, a blank screen was presented for 1000 ms, after which the comprehension task was presented. Participants were asked to avoid movements and to only blink their eyes between their response to the comprehension task and the presentation of the next sentence. The experimental session began with a short training session followed by eight experimental blocks comprising 40 sentences each, between which the participants took short breaks. The entire experiment (including electrode preparation) lasted approximately 2.5 h.

The EEG was recorded by means of AgAgCl electrodes, which were fixed at the scalp by means of an elastic cap (Electro Cap International). Recording took place from the following electrode sites: F7, F3, FZ, F4, F8, FC5, FCZ, FC6, T7, C3, CZ, C4, T8, CP5, CPZ, CP6, P7, P3, PZ, P4, P8, PO3, POZ, PO4, OZ. The ground electrode was positioned above the sternum. Recordings were referenced to the left mastoid, but re-referenced to linked mastoids offline. In order to control for artefacts resulting from eye movements, the electrooculogram (EOG) was monitored by means of electrodes placed at the outer canthus of each eye for the horizontal EOG and above and below the participant's right eye for the vertical EOG. Electrode impedances were kept below 5 kOhm.

All EEG and EOG channels were amplified using a Twente Medical Systems DC amplifier and recorded continuously with a digitization rate of 250 Hz. The ERPs were filtered offline with 10 Hz low pass for the plots, but all statistical analyses were computed on unfiltered data.

2.4. Data analysis

Average ERPs were calculated per condition per participant from the onset of the critical stimulus items (i.e. the first and the second NP) to 1500 ms post-onset, before grand-averages were computed over all participants. Averaging took place relative to a baseline interval from −200 to 0 ms before the onset of the critical items. Trials for which the comprehension task was not performed correctly were excluded from the averaging procedure, as were trials containing ocular, amplifier-saturation or other artefacts (the EOG rejection criterion was 40 μV).

For the statistical analysis of the ERP data, repeated measures analyses of variance (ANOVAs) using the factors ORDER (subject-object vs. object-subject) and CASE (accusative vs. dative) were calculated for mean amplitude values per time-window per condition. The statistical analysis was carried out in a hierarchical manner, i.e. only significant interactions ($p < 0.05$) were resolved. Additionally, no main effects of or interactions between topographical factors will be reported.

Topographical factors were chosen as follows: for the midline electrodes, the factor Electrode (Elec) with the seven midline electrodes as levels; for the lateral electrodes, the factors Hemisphere (Hemi; left vs. right) and Region (anterior, central, posterior).

For the statistical analysis of the behavioural data, error rates and reaction times for the comprehension task were calculated for each critical condition. Incorrectly answered trials were excluded from the reaction times analysis. We computed a repeated measures ANOVA containing the critical factors ORDER and CASE and the random factors subjects (F_1) and items (F_2).

3. Results

3.1. Behavioural data

The global repeated measures analysis of the error rates for the comprehension task revealed a main effect of ORDER that was marginal in the subjects-analysis ($F_1(1, 15) = 3.94,\ p < 0.07$) and significant in the items-analysis ($F_2(1, 79) = 4.40,$

$p < 0.04$). This effect was due to a higher error rate for object-initial sentences (13.9%) than for their subject-initial counterparts (11.9%). Neither the main effect CASE nor the interaction ORDER × CASE reached significance.

The analysis of the reaction times showed a significant main effect of ORDER ($F_1(1, 15) = 25.55, p < 0.001; F_2(1, 79) = 25.44, p < 0.001$), which resulted from longer reaction times for object-initial (1536 ms) than for subject-initial sentences (1467 ms).

3.2. ERP data

Fig. 1 shows grand-average ERPs at the position of the first NP for nominative-initial, accusative-initial and dative-initial structures. As is apparent from Fig. 1, accusative-initial structures give rise to a broadly distributed negative deflection in comparison to their nominative-initial counterparts. By contrast, no such difference is apparent between dative- and nominative-initial structures.

For the statistical analysis of the effects at the position of the first NP, we chose the time-window 300–450 ms on the basis of previous studies (Rösler et al., 1998). The statistical analysis for the lateral electrodes revealed a main effect of ORDER ($F(1, 15) = 5.33$,

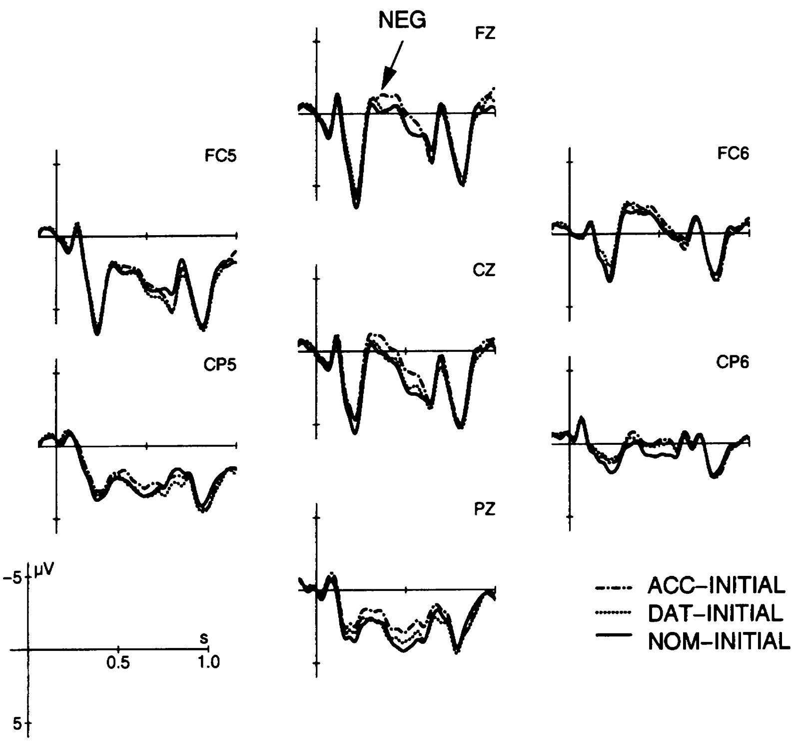

Fig. 1. Grand-average ERPs elicited by the first NP (onset at the vertical line) for nominative-initial, accusative-initial and dative-initial structures. Negativity is plotted upwards.

$p < 0.04$) as well as an interaction ORDER × CASE ($F(1, 15) = 5.62$, $p < 0.04$). Single comparisons for accusative structures showed a main effect of ORDER ($F(1, 15) = 15.45$, $p < 0.002$), while single comparisons for dative structures revealed no such effect ($F < 1$).

With regard to the midline electrodes, there was a significant interaction ORDER × CASE ($F(1, 15) = 9.29, p < 0.001$). Single comparisons again revealed an ORDER effect for accusative structures ($F(1, 15) = 18.17$, $p < 0.001$), but not for their dative counterparts ($F < 1$).

In order to test the hypothesis that dative-initial sentences should elicit a reanalysis effect at the position of the second NP, grand-averages for dative-nominative vs. nominative-dative structures at the position of the second NP are shown in Fig. 2. Descriptively, it appears from Fig. 2 that dative-initial structures elicit an early posterior positivity in comparison to nominative-dative structures. On the basis of previous studies in which such an early positivity was reported (Mecklinger, Schriefers, Steinhauer, & Friederici, 1995), we chose the time-window of 300–400 ms for the statistical analysis of this component. Since no accusative structures entered this comparison, only the condition factor ORDER is of relevance.

At posterior electrode sites only, there was a significant effect of ORDER ($F(1, 15) = 4.79$, $p < 0.05$), which resulted from more positive wave forms for dative-initial structures.

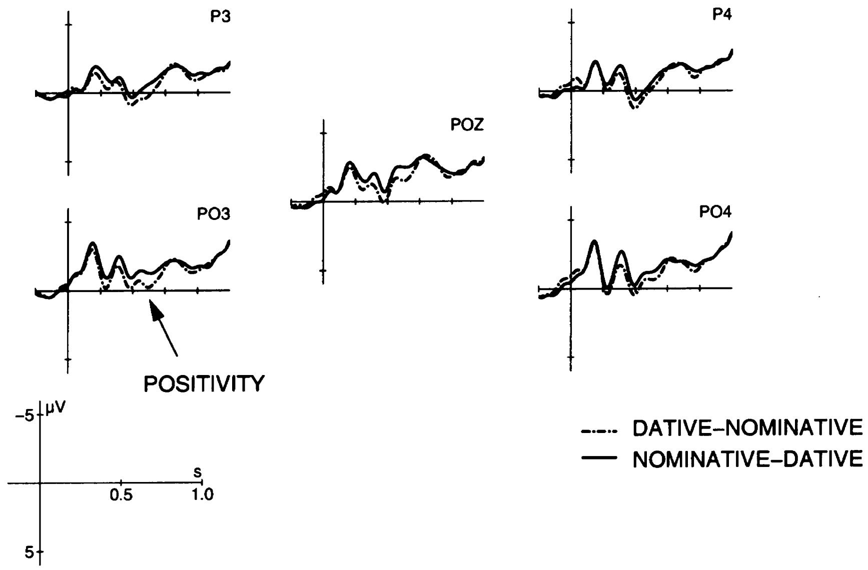

Fig. 2. Grand-average ERPs elicited by the second NP (onset at the vertical line) for dative-nominative vs. nominative-dative structures. Negativity is plotted upwards.

4. Discussion

We have presented an ERP experiment contrasting two possible sources of processing difficulty in German sentences with a non-canonical word order, namely frequency and (grammatical) markedness. Our results show that, in embedded clauses introduced by the complementizer *dass* ('that'), initial accusative-marked objects give rise to a centrally distributed negativity between 300 and 450 ms post-onset in comparison to initial subjects. Dative-initial sentences, by contrast, do not differ from subject-initial sentences at the position of the first NP, but rather show a posterior positivity between 300 and 400 ms post-onset of the second NP.

These data demonstrate the sensitivity of the parser to underlying grammatical distinctions: in the constructions used in the present experiment, initial accusatives unambiguously signal a non-canonical word order, while initial datives may also be interpreted as introducing a (canonically ordered) passive clause. Thus, the finding of a negativity for initial accusatives but not for initial datives shows that the processing difficulty associated with clause-medial word order variations does not result from the relative infrequency of such constructions, but rather appears to reflect the application of fine grained linguistic distinctions during the course of language processing. It therefore appears that the processing difficulties observed for clause-medial word order variations cannot be easily accounted for in frequency-based processing approaches.

Rather, the present data support a grammatically-based account of these processing difficulties as outlined in Section 1 (Friederici et al., in press). When a complementizer is processed in German, the parser makes use of its knowledge with regard to grammatical regularities to predict an upcoming structural position which may be occupied either by a (nominative) subject or by a dative-marked argument. In this way, dative-marked arguments adjacent to the complementizer do not give rise to processing difficulties in comparison to nominatives, whereas accusative-marked arguments, which do not match the properties of the predicted position, do. The assumption made by this approach that an initial dative is interpreted as the argument of a passivized verb is supported by the ERP component observed at the position of the second NP in dative-initial sentences: an early positivity. Though this positivity differs in latency from the P600 component, which is usually associated with reanalysis operations, a similar early positivity has been reported for reanalyses of subject-object ambiguities in German relative clauses (Mecklinger et al., 1995). In light of these latency differences, Friederici and Mecklinger (1996) proposed that the latency of a positivity reflects the degree of difficulty associated with the reanalysis. This also fits well with the present data, since a reanalysis of dative-initial structures at the position of the second NP only requires a right-adjunction of a new structural position for the nominative-marked NP, i.e. a type of operation that is supposedly rather low in extra processing cost (Sturt & Crocker, 1996).

Interestingly, the distinction between dative-initial and accusative-initial structures observable in the ERP data is not reflected in the behavioural data (i.e. participants' responses to the comprehension task): here, object-initial structures generally gave rise to higher error rates and longer reaction times than their subject-initial counterparts. In this way, it appears that factors such as frequency may indeed play a role with regard to *global* properties of sentence comprehension, while they do not appear to affect the online ease or difficulty of the incremental integration of incoming elements into the existing sentence structure. Rather, our ERP data are a strong piece of evidence that such processes of online (syntactic) integration reflect the application of principles of the grammar.

Acknowledgements

The research reported here was supported by the grants FOR 375/1-4 and FI 848/1-1 from the Deutsche Forschungsgemeinschaft (DFG). We would like to thank Stefan Frisch, Gisbert Fanselow and three anonymous reviewers for helpful suggestions, Kristin Wittich for the corpus analysis and Cornelia Schmidt for help with the data acquisition.

References

Crocker, M. W., & Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, *29*, 647–669.

den Besten, H. (1985). The ergative hypothesis and free word order in Dutch and German. In J. Toman (Ed.), *Studies in German grammar* (pp. 23–64). Dordrecht: Foris.

Fanselow, G. (2000). Optimal exceptions. In B. Stiebels & D. Wunderlich (Eds.), *Lexicon in focus* (pp. 173–209). Berlin: Akademie Verlag.

Fanselow, G., Schlesewsky, M., & Kliegl, R. (2000). Processing difficulty and principles of grammar. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: aging, grammar and memory* (pp. 171–201). Dordrecht: Kluwer.

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*, 321–338.

Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: first-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, *25*, 157–176.

Friederici, A. D., Schlesewsky, M., & Fiebach, C. J. (in press). Wh-movement vs. scrambling: the brain makes a difference. In S. Karimi, & T. Langendoen (Eds.), *Word order and scrambling*. Oxford: Blackwell.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194.

Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Lenerz, J. (2001). Word order variation: competition or co-operation? In G. Müller & W. Sternefeld (Eds.), *Competition in syntax* (pp. 249–281). Berlin: De Gruyter.

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

Primus, B. (1999). *Cases and thematic roles*. Tübingen: Niemeyer.

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 and Language*, *38*, 150–176.

Schlesewsky, M., Bornkessel, I., & Frisch, S. (in press). The neurophysiological basis of word order variations in German. *Brain and Language*.

Sturt, P., & Crocker, M. W. (1996). Monotonic syntactic processing: a cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes*, *11*, 449–494.

Wunderlich, D. (1997). Cause and the structure of verbs. *Linguistic Inquiry*, *28*, 27–68.