COGNITION

ELSEVIER

Cognition 85 (2002) B83–B92

www.elsevier.com/locate/cognit

Brief article

The P600 as an indicator of syntactic ambiguity

Stefan Frisch*, Matthias Schlesewsky, Douglas Saddy, Annegret Alpermann

Institute of Linguistics, University of Potsdam, P.O. Box 60 15 53, 14415 Potsdam, Germany

Received 5 September 2001; received in revised form 1 February 2002; accepted 19 June 2002

Abstract

In a study using event-related brain potentials, we show that the current characterization of the P600 component as an indicator of revision processes (reanalysis and repair) in sentence comprehension must be extended to include the recognition of syntactic ambiguity. By comparing the processing of ambiguous and unambiguous sentence constituents in German, we show that the P600 is elicited when our language processing system has syntactic alternatives at a certain item given in the input string. That the P600 is sensitive to syntactic ambiguity adds crucial evidence to current debates in psycholinguistic modelling, as the results clearly favour parallel models of syntactic processing which assume that ambiguity is recognized and costly. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Event-related brain potentials; Language processing; P600 component; Syntactic ambiguity; Serial processing; Parallel processing

1. Introduction

1.1. Serial versus parallel parsing

One of the key issues in sentence processing is the question of how the human sentence parser deals with syntactic ambiguity, and this is so for at least two reasons. Firstly, the processing of the ambiguous region itself can tell us whether the parser acknowledges – at least temporarily – the presence of more than one syntactic analysis, which would allow us to distinguish between serial and parallel approaches to parsing (cf. Lewis, 2000; Mitchell, 1994). Secondly, the disambiguating region is of interest since it is crucial for deciding whether one alternative is preferred over others or whether all continuations are considered to be equally likely (cf. Mitchell, 1994). On the one hand, serial models (Frazier &

* Corresponding author.
E-mail address: frisch@rz.uni-potsdam.de (S. Frisch).

PII: S0010-0277(02)00126-9

Rayner, 1982) hold that in case of syntactic ambiguity, the human parser chooses only one structural analysis on the basis of structural simplicity. Serial models therefore predict that an ambiguous structure should be equally easy to process as its simplest unambiguous counterpart (Frazier, 1987; Lewis, 2000; Mitchell, 1994). Parallel models, on the other hand, assume that our parser is sensitive to ambiguous material (although this does not imply that all alternative readings are pursued with equal priority) and that this is reflected in enhanced processing cost for an ambiguous element compared to the preferred unambiguous alternative, at least locally (Altmann & Steedman, 1988).

1.2. Grammatical function ambiguity: German as a test case

With respect to syntactic function ambiguities, the processing of an ambiguous argument cannot easily be tested in English since it is difficult to find a suitable unambiguous control condition. We will briefly illustrate this in the following: a nominal phrase (NP) like 'which man', for example, can be the subject of a clause (as in (1a)) or the object (as in (1b)).

(1a) Which man met Ellen?
(1b) Which man did Ellen meet?

There is some consensus in the psycholinguistic literature that the continuation in which 'which man' turns out to be the subject is easier to process than a disambiguation towards the object (cf. Gibson, 1998). Such an assumption, however, must focus on processing differences within or after the disambiguating region following the ambiguous phrase. It is hard to see how one could test whether the processing of the ambiguity of the phrase 'which man' is costly in itself, since it is unclear which phrase could serve as an appropriate unambiguous compare condition.

German, by contrast, offers an excellent testing ground for this question because nominal phrases can be ambiguous as well as unambiguously marked for nominative (subject case) or accusative (direct object case). Since German is a language with a relative freedom of word order, the parser cannot definitely decide whether a constituent is the subject (SUB) or an object (OBJ) just on the basis of linear order. A phrase like 'die Frau' (*the woman*), for example, is case ambiguous and can be disambiguated by subsequent unambiguous information (such as a second argument) either towards the grammatical function of subject (as in (2a)) or object (as in (2b)).

(2a) Die Frau hatte den Mann gesehen.
[the woman]AMB had [the man]OBJ seen
(2b) Die Frau hatte der Mann gesehen.
[the woman]AMB had [the man]SUB seen

Similar to the English examples, disambiguations such as in (2b) are known to be dispreferred and to induce processing difficulties in the disambiguating region (at the second argument 'der Mann'). In contrast to English, however, German allows us to test for whether processing cost is induced by a functional ambiguity itself, because there exist unambiguous (masculine) counterparts to 'die Frau', namely 'der Mann' (*[the man]SUB*) or ‘den Mann’ (*[the man]OBJ*), respectively, and therefore sentences such as (3a) and (3b) are appropriate control conditions for (2a) and (2b).

(3a) Der Mann hatte die Frau gesehen.
[the man]SUB had [the woman]OBJ seen
(3b) Den Mann hatte die Frau gesehen.
[the man]OBJ had [the woman]SUB seen

As introduced above, serial models of parsing would predict that a nominal phrase in German which is ambiguous with respect to its syntactic function as for example ‘die Frau’ (*the woman*) should be equally easy to process as the unambiguous subject alternative ‘der Mann’ (*the man*) since our parser is predicted to immediately assign the structurally simplest (i.e. subject) analysis. Parallel models, on the other hand, assume that our parser acknowledges ambiguity in a sentence and that this is reflected in enhanced processing cost for an ambiguous element compared to the preferred unambiguous alternative. According to parallel models, an ambiguous argument such as ‘die Frau’ should be harder to process compared to its unambiguous subject counterpart (‘der Mann’) because our processing system should be burdened with the presence of alternatives. With respect to the disambiguating region, however, parallel models do not make different predictions compared to serial models as long as they assume that some analyses outrank others for reasons of simplicity (cf. Gorrell, 1987). In this case, both types of models predict that disambiguations towards structurally more complex analyses (such as object-before-subject) are dispreferred and therefore more costly than preferred ones.

1.3. Syntactic ambiguity and event-related potentials

In German, a dispreferred disambiguation of an ambiguous argument towards the object is indeed reflected in longer reading times (Schlesewsky, Fanselow, Kliegl, & Krems, 2000; Schriefers, Friederici, & Kühn, 1995). In addition, studies using event-related potentials (ERPs) have shown that dispreferred disambiguations of ambiguous arguments elicit a positive ERP deflection (Beim Graben, Saddy, Schlesewsky, & Kurths, 2000; Frisch, Beim Graben, & Schlesewsky, in press; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995). This so-called *P600 component* is often seen as an indicator for greater syntactic processing cost due to a necessary revision of a (temporary or persistent) structural mismatch (Friederici, 1995; Osterhout & Holcomb, 1992), which may either consist of an outright syntactic violation (Friederici, 1995; Frisch, 2000; Osterhout & Holcomb, 1992) or of a dispreferred disambiguation of an ambiguous string (Mecklinger et al., 1995; Osterhout & Holcomb, 1992). With respect to the processing of the ambiguity itself, reading time studies have shown that ambiguous arguments take longer to read compared to their unambiguous counterparts (Schlesewsky et al., 2000). What has not been addressed so far, however, is the question of which ERP component – if any – is elicited by a syntactically ambiguous argument. If it is true that it is the P600 component which reflects processing cost associated with the syntactic aspects of language then one would expect to find a positivity for a syntactically ambiguous argument in German compared to an unambiguous one.

2. The present study

In order to test whether syntactically ambiguous arguments are more difficult to process compared to unambiguous ones, we tested German constructions in the following conditions (4a–d) in an experiment using ERPs.

<u>First NP unambiguous/subject before object</u>

(4a) Der Detektiv hatte die Kommissarin gesehen und…
[the detective]MASC.SUB had [the policewoman]FEM.OBJ seen and…

<u>First NP unambiguous/object before subject</u>

(4b) Den Detektiv hatte die Kommissarin gesehen und…
[the detective]MASC.OBJ had [the policewoman]FEM.SUB seen and…

<u>First NP ambiguous/subject before object</u>

(4c) Die Detektivin hatte den Kommissar gesehen und…
[the detective]FEM.AMB had [the policeman]MASC.OBJ seen and…

<u>First NP ambiguous/object before subject</u>

(4d) Die Detektivin hatte der Kommissar gesehen und…
[the detective]FEM.AMB had [the policeman]MASC.SUB seen and…

2.1. Method

2.1.1. Participants

Twenty undergraduate students (mean age 21 years, 11 female) from the University of Potsdam participated. They were paid 13 DM per hour. All were monolingual native speakers of German.

2.1.2. Materials

All sentences consisted of a first clause like the ones in (4a–d) and a conjoined second clause consisting of an auxiliary, an object pronoun, a prepositional phase (PP) and a verb, such as for example 'und wollte sie am Bahnhof treffen' (*and wanted to meet her at the station*). Half of the sentences had a subject-before-object structure (conditions (4a) and (4c)) in the first clause and the other half an object-before-subject structure (conditions (4b) and (4d)). Furthermore, in half of the sentences within each group, the grammatical function of the first argument was unambiguously indicated by the case marking (conditions (4a) and (4b)), whereas in the other half, the grammatical function of the first argument was ambiguous between subject and object and was disambiguated by the second argument (conditions (4c) and (4d)). A total of 320 critical sentences were created out of 40 noun-verb-noun triplets by realizing both noun lexemes in both argument positions in order to exclude lexical influences. From these 320 sentences, two lists were created with 160 sentences (40 in each of the four critical conditions) in each list. Subjects were assigned randomly to one of the lists.

2.1.3. Procedure

Sentences were presented in a randomized order and visually in the centre of a screen, with 500 ms (plus 100 ms *inter-stimulus interval/ISI*) for each of the two arguments and the PP in the second clause and 400 ms (plus 100 ms ISI) for all other words. Subjects were shown a test sentence 1000 ms after the end of each sentence and had to judge whether it was identical to the preceding experimental sentence or not by means of pressing one out of two push buttons (*sentence matching task*).

The EEG was recorded by means of 15 Ag/AgCl electrodes with a sampling rate of 250 Hz (with impedances kept below 5 kOhm) and was referenced to the left mastoid (re-referenced to linked mastoids off-line). The horizontal and the vertical EOG were monitored with two electrodes each.

2.1.4. Data analysis

Single subject averages were computed separately for both critical items, within a 1100 ms time window for the first NP argument and a 1300 ms time window for the second. Both time windows were preceded by a 200 ms pre-stimulus baseline. Only trials with correct answers in the task and without artefacts entered the ERP analysis (68.7% of all trials in the first time window, 68.3% in the second). On the basis of visual inspection and a running *t*-test procedure, the following two time windows were chosen for the statistical analyses of the grand average ERPs: 400–800 ms relative to the onset of the first argument and 600–1000 ms relative to the onset of the second. All ERP effects were statistically computed separately for midline and lateral electrodes. The ANOVA design included two condition factors, namely *ambiguity* of the first argument (*ambiguous* versus *unambiguous*) and *order* of the arguments in the clause (*subject-before-object* versus *object-before-subject*). Furthermore, there was one topographical factor, namely either *electrode* (with the three electrodes *FZ*, *CZ* and *PZ* as levels) for the midline, or *region of interest* (ROI; with four levels) for the lateral electrode analyses, respectively. The four lateral ROIs comprised the following three electrodes each: *left-anterior*: F3, FC5, C3; *right-anterior*: F4, FC6, C4; *left-posterior*: CP5, P3, PO3; *right-posterior*: CP6, P4, PO4.

2.2. Results

Fig. 1 displays the ERPs from the onset of the first argument (at 0 ms) up to 1000 ms thereafter as the main effect between both ambiguous and both unambiguous conditions. As can be seen from the figure, ERPs in the conditions starting with an ambiguous argument are more positive going between 400 and 800 ms compared to the unambiguous conditions. Statistical analyses for the first argument (400–800 ms) revealed main effects of *ambiguity* over the midline ($F(1, 19) = 16.82$, $P < 0.001$) as well as over the lateral electrodes ($F(1, 19) = 14.24$, $P < 0.001$). For the first NP, we found neither main effects of *order* nor interactions between the two condition factors. Over lateral sites, *ambiguity*

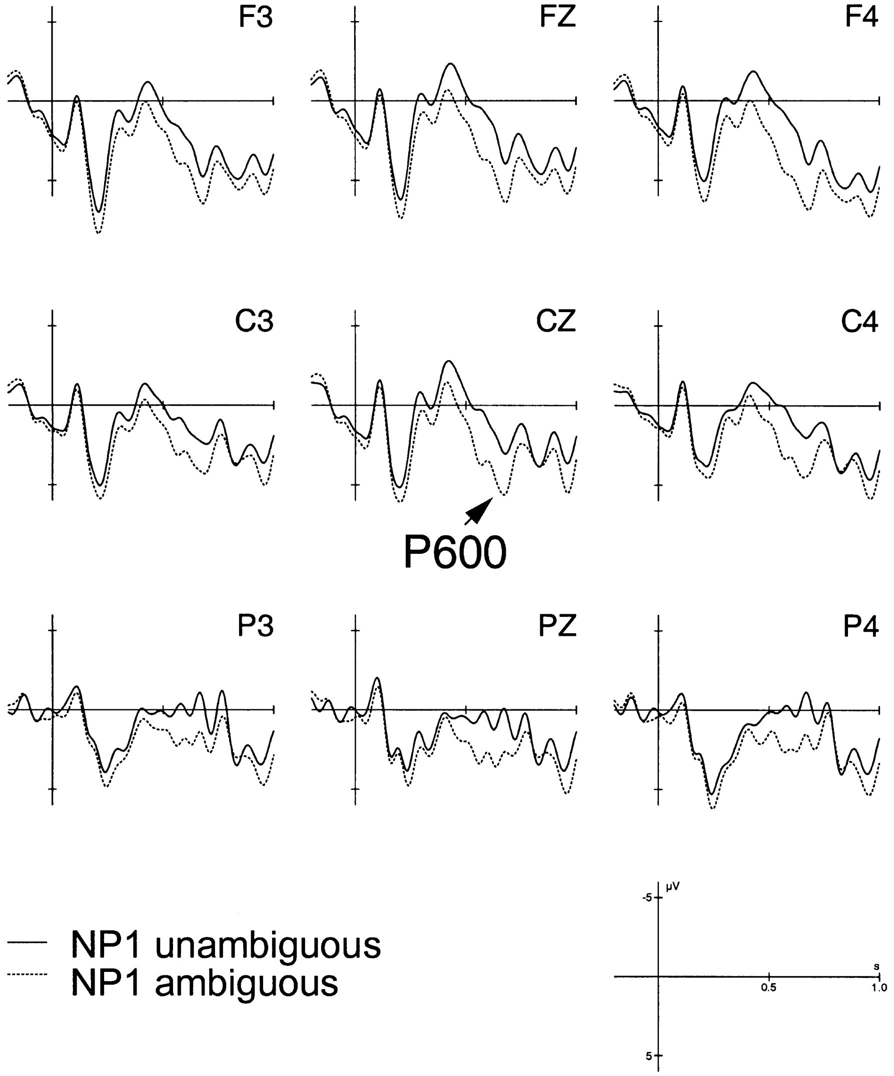

Fig. 1. ERP effects on the first argument (onset at 0 ms) at a subset of nine electrodes. Negativity is plotted upwards. Grand averages have been filtered with 10 Hz low pass (for the plots only). Since there were differences neither between the two ambiguous nor between the unambiguous conditions, only the main effect of *ambiguity* is plotted. Approximately between 400 and 800 ms, the ambiguous conditions are more positive going compared to the unambiguous ones (for details see text).

interacted with *ROI* ($F(3, 57) = 5.88$, $P < 0.01$). Resolving this interaction revealed main effects of *ambiguity* in all four ROIs (all $P < 0.01$).

The ERPs on the second argument (onset at 0 ms) can be seen in Fig. 2. Fig. 2a shows that the two conditions in which the first argument is unambiguous do not differ at the second argument. Fig. 2b contrasts the two initially ambiguous conditions in which the second

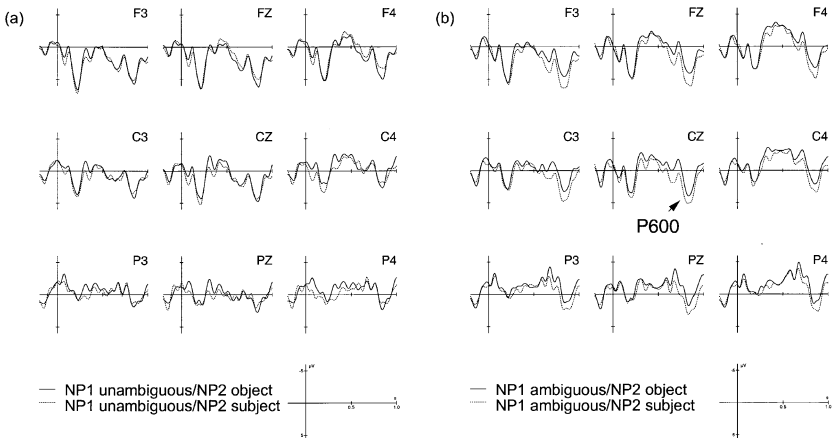

Fig. 2. ERP effects on the second argument (onset at 0 ms) at a subset of nine electrodes. Negativity is plotted upwards. Grand averages have been filtered with 10 Hz low pass (for the plots only). (a) There is no difference on the second argument between the two conditions in which the first argument is already unambiguous. (b) The difference between the two conditions in which the second argument disambiguates. When the disambiguating second argument is the subject, then we see a more positive going waveform between about 600 and 1000 ms compared to a disambiguation towards object (for details see text).

argument disambiguates. As can be seen from the figure, the condition in which the subject preference for the first argument has to be revised (because the second argument is unambiguously marked for subject) is more positive going between about 600 and 1000 ms compared to the condition in which the initial preference is supported. Statistical analyses for the second argument (600–1000 ms) revealed neither main effects nor interactions over the lateral sites. Over the midline, we found an interaction *ambiguity* × *order* ($F(1, 19) = 5.76$, $P < 0.05$) which was due to the fact that there was no difference between the initially unambiguous conditions ($F < 1$), but that the initially ambiguous condition being disambiguated towards the dispreferred reading was significantly more positive going than the ambiguous subject-before-object condition ($F(1, 19) = 10.06$, $P < 0.01$).

3. Discussion

In an experiment using ERPs, we tested the processing of sentence initial nominal constituents which were either ambiguous between subject and object or unambiguous. The former were disambiguated later in the sentence, either in a preferred (subject-before-object) or in a dispreferred way (object-before-subject). Results show that syntactic ambiguity, that is, the (temporary) possibility of more than one syntactic interpretation for an element, as well as its resolution have an impact on online processing and induce clear neurophysiological responses in the form of P600 effects.

When the second argument disambiguated an initially ambiguous sentence towards the dispreferred object-subject order, we found a P600 component compared to a sentence disambiguated in the preferred way (subject-first). This replicates previous findings and can thus be taken as a correlate for the cost associated with the revision of a subject-first preference (Beim Graben et al., 2000; Frisch et al., in press; Mecklinger et al., 1995). By contrast, the P600 observed for the first argument when this argument allowed alternative syntactic analyses is reported here for the first time. This P600 effect for a sentence-initial, ambiguous argument compared to an unambiguous one cannot be taken as a marker of syntactic revision (cf. Osterhout & Holcomb, 1992) seeing that there is no preceding interpretation to be revised. Neither can it reflect the cost of syntactic integration in the sense of Gibson (1998), as proposed in a recent P600 study (cf. Kaan, Harris, Gibson, & Holcomb, 2000). According to Gibson (1998), integration cost occurs at an item as a consequence of a prediction from a preceding item. In order to explain our P600 findings for sentence-*initial* ambiguous arguments in Gibsonian terms one would have to abandon the notion of integration and look to the notion of cost of prediction. One would be led to assert that making predictions itself is costly and that making two structural predictions (i.e. projecting both a subject-first and an object-first structure) is harder in terms of processing than making only one such prediction. Thus, the results strongly call for an extension of the current interpretation of the P600 since this component is obviously sensitive to the complexity of structural predictions possible at an element which is syntactically ambiguous. The finding that both of our P600 effects had a broad topographical distribution over the scalp seems to be in line with earlier findings that the P600 as a reflex of disambiguation or complexity is more evenly distributed than the P600 found for outright violations (Hagoort & Brown, 2000). In this context, the question arises as to how far the P600 component can be conceived of as a homogenous phenomenon or whether it would be more apt to speak of a family of components, reflecting similar, but nevertheless separable aspects of syntactic processing (cf. Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001).

As outlined in Section 1, the psycholinguistic debate between serial and parallel approaches to parsing crucially depends upon the question of whether syntactic ambiguity induces additional processing cost or not (Frazier, 1987; Lewis, 2000; Mitchell, 1994). Thus, our finding of a P600 for an ambiguous argument compared to an unambiguous one is a crucial piece of evidence for deciding between the competing approaches since it clearly speaks in favour of parallel processing models. Parallel parsing is not a homogenous conception, but can be spelled out in different ways, which range from merely acknowledging the presence of alternatives to specifying different structural representations and maintaining them as long as they do not consume too many processing resources (cf. Altmann & Steedman, 1988; Crocker, 2000; Gibson, 1991; Mitchell, 1994). On the basis of the present experiment, we cannot decide between the different possibilities. What we can say, however, is that not all alternatives are equally maintained over the sentence. The finding of a P600 on the second argument when this disambiguates the previous string towards a dispreferred structure shows that, (no later than) at this item, our parser has built up a measurable preference for one continuation (subject first) over the other (object first). This implies that not all alternatives are considered up to the point of disambiguation in an unweighted manner, but that after the ambiguity is acknowledged, we consider some continuations to be more likely than others (in the sense of Gorrell, 1987). An open issue in this context, however, is whether our finding that syntactic ambiguity is costly can be generalized to all instantiations of syntactic ambiguity and to other languages.

3.1. Conclusion

In sum, the finding of a P600 component for a syntactically ambiguous element shows that the view that this component reflects processes of syntactic revision or integration has to be extended. As our data show, the P600 component must be taken as an indicator of syntactic processing cost in general. That such processing cost was enhanced for an ambiguous initial argument as well as for its later disambiguation towards the dispreferred argument order (object-first) adds crucial evidence to the central debate in psycholinguistics between serial and parallel processing. Results clearly favour a processing architecture according to which our parser notices the existence of syntactic alternatives in the case of ambiguity, but according to which not all of these analyses are pursued with the same priority.

Acknowledgements

The present research was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to M.S. and D.S. (FOR 375/1-4). We want to thank Angela D. Friederici and Ina Bornkessel for helpful suggestions and Saskia Kohnen, Kerstin Leuckefeld and Anja Meinke for their support in data acquisition.

References

Altmann, G., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Beim Graben, P., Saddy, D., Schlesewsky, M., & Kurths, J. (2000). Symbolic dynamics of event-related brain potentials. *Physical Review*, *62*, 5518–5541.

Crocker, M. W. (2000). Mechanisms for sentence processing. In S. Garrod & M. Pickering (Eds.), *Mechanisms for sentence processing* (pp. 191–232). Hove: Psychology Press.

Frazier, L. (1987). Sentence processing: a tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). London/Hillsdale, NJ: Erlbaum.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–281.

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: a spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research*, *11*, 305–323.

Frisch, S. (2000). *Argumentstruktur, Kasus und thematische Interpretation beim Sprachverstehen*, *MPI series of cognitive neuroscience*, 12. Leipzig: MPI.

Frisch, S., Beim Graben, P., & Schlesewsky, M. (in press). Parallelizing grammatical functions: P600 and P345 reflect different cost of reanalysis. *The International Journal of Bifurcation and Chaos*.

Gibson, E. (1991). *A computational theory of human linguistic processing: memory limitations and processing breakdown*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA.

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

Gorrell, P. (1987). *Studies in human syntactic processing: ranked-parallel versus serial models*. PhD thesis, University of Connecticut, Storrs.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*, 1531–1549.

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.

Lewis, R. L. (2000). Falsifying serial and parallel parsing models: empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research*, *29*, 241–248.

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

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed), *Handbook of psycholinguistics* (pp. 375–409). San Diego, CA: Academic Press.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–804.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). Preferences for grammatical functions in the processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 65–93). Dordrecht: Kluwer.

Schriefers, H., Friederici, A. D., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, *34*, 499–520.