Repair, Revision, and Complexity in Syntactic Analysis: An Electrophysiological Differentiation

Edith Kaan[1] and Tamara Y. Swaab[2]

Abstract

■ One of the core aspects of human sentence processing is the ability to detect errors and to recover from erroneous analysis through revision of ambiguous sentences and repair of ungrammatical sentences. In the present study, we used event-related potentials (ERPs) to help identify the nature of these processes by directly comparing ERPs to complex ambiguous sentence structures with and without grammatical violations, and to simpler unambiguous sentence structures with and without grammatical violations. In ambiguous sentences, preference of syntactic analysis was manipulated such that in one condition, the structures agreed with the preferred analysis, and in another condition, a nonpreferred but syntactically correct analysis (garden path) was imposed. Nonpreferred ambiguous structures require revision, whereas ungrammatical structures require repair. We found that distinct ERPs reflected different characteristics of syntactic processing. Specifically, our results are consistent with the idea that a positivity with a posterior distribution across the scalp (posterior P600) is an index of syntactic processing difficulty, including repair and revision, and that a frontally distributed positivity (frontal P600) is related to ambiguity resolution and/ or to an increase in discourse level complexity. ■

INTRODUCTION

Language comprehension is a seemingly straightforward and easy process, and, yet, the linguistic input is often extremely complex, containing many ambiguities, speech errors, repeats, and hesitations. One of the core aspects of human sentence processing is the ability to rapidly process complex sentences and to detect errors and recover from erroneous analysis through revision and repair. We use the term ''revision'' to refer to the processes engaged to yield a correct analysis when the linguistic input contains a syntactically correct but non-preferred continuation of an ambiguous sentence fragment. In contrast, ''repair'' refers to processes that attempt to construct a representation of the sentence when a sentence continuation is ungrammatical under all possible analyses of the preceding fragment and therefore cannot yield a syntactically correct analysis. The exact nature of the mechanisms that underlie the processes of revision and repair is a matter of debate (Fodor & Ferreira, 1998). In the present study, we used event-related potentials (ERPs) to investigate if revision and repair processes are distinct, as reflected by specific ERPs. Differences in the scalp distributions of ERPs during revision and repair, for example, would be con- sistent with the idea that separable neuronal sources are active, which is also suggestive of functional differences

in the processes reflected by the ERPs (e.g., Rugg & Coles, 1995).

Since the early 1990s, several studies have shown that ERPs are sensitive to syntactic aspects of the linguistic input (e.g., Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992). In particular, the P600 or ''syntactic positive shift'' has been related to the pro- cesses of revision and repair in sentence processing and will therefore be the focus of the present research.[1] The P600 is a positive deflection starting at around 400 msec. It is reportedly elicited by words that are either ungram- matical continuations of the preceding sentence frag- ment and trigger repair processes (e.g., ''The spoiled child *throw the toys on the floor'', Hagoort et al., 1993) or are nonpreferred continuations of the preceding sentence fragment, which trigger revision processes (Osterhout & Holcomb, 1992). An example of the latter is ''The broker persuaded to sell the stock was sent to jail,'' where ''persuaded'' is preferably taken as the main verb of the sentence. At the word ''to,'' this analysis can no longer be maintained, and the analysis is revised, such that ''persuaded'' is the verb of the relative clause ''(who was) persuaded to sell the stock.''

Recent evidence suggests, however, that the scalp distribution of the positivity is different for nonpre- ferred and ungrammatical sentence continuations. Hagoort, Brown, and Osterhout (1999) observed that nonpreferred continuations generally elicit a frontally distributed positive voltage ERP (Hagoort et al., 1999; Van Berkum, Brown, & Hagoort, 1999; Osterhout &

[1]Duke University, [2]University of California–Davis

Holcomb, 1992), whereas the positive voltage ERP observed for ungrammatical continuations generally has a more posterior focus (e.g., Coulson, King, & Kutas, 1998b). On the basis of these observations Hagoort et al. (1999) propose that the frontally distributed positivity reflects costs associated with overwriting the preferred, most active structural representation of the sentence, as is the case with nonpreferred continuations of the preceding sentence fragment (revision), whereas the posteriorly distributed positivity reflects costs associated with a collapse of the structural representation, as is the case in ungrammatical continuations (repair).

However, more general interpretations have been proposed both for the frontal and the posterior scalp positivities on the basis of studies that manipulated sentence complexity (Friederici, Hahne, & Saddy, 2002; Kaan et al., 2000). On the one hand, Friederici et al. (2002) proposed that the *frontal* positivity is not restricted to revision processes for ambiguous sentence structures, but that it is also found with increasing complexity. In a study that compared simple and more complex unambiguous sentence structures, the complex structures elicited a frontally distributed positivity at the word position after the onset of the complexity. Friederici et al. hypothesized that this frontal positivity is associated with processing syntactic complexity. However, because the sentences used in the Friederici et al. study are disambiguated before the point of comparison, the processes indexed by the frontal positivity do not necessarily involve revision.

On the other hand, other evidence indicates that the *posterior* positivity may not be restricted to repair processes. Kaan et al. (2000) found a posterior positivity to words that were complex, although syntactically correct and preferred continuations of the preceding sentence fragments, such as the italicized verb in "Emily wondered who the performers in the concert *imitate*..." versus "Emily wondered whether the performers in the concert *imitate* a pop star...." The "who"-clause is more complex at the position of the verb than the "whether"-clause because the former involves an additional operation: the "who"-phrase

(filler) at the beginning of the clause needs to be related to the verb or to an "empty" position at the verb (gap), in order to be interpreted as its direct object. These results suggest that the posterior P600 reflects more than just repair, but that it is also sensitive to integration difficulty as a function of syntactic complexity of sentence structures. In addition, the Kaan et al. results show that an increase in complexity in general is not sufficient to generate the frontal positivity reported by Friederici et al. (2002).

In short, at present, it is unclear how the posterior and frontal scalp phases of the positivity differentiate between processes related to repair, revision, and complexity in syntactic analysis. Therefore, the current study will further investigate if posterior and frontal phases of a positive scalp potential indeed reflect different aspects of sentence processing. First, in order to test the proposal by Hagoort et al. (1999), we explored if the P600 to ungrammatical continuations, which require repair, is indeed posteriorly distributed. Second, to identify if a more frontally distributed positive ERP is related to processes of revision, we manipulated syntactic preference and compared preferred continuations that do not require revision with nonpreferred ambiguous structures that do. Finally, we tested the effects of complexity on the distribution of the positivity by comparing ambiguous, complex sentence structures with unambiguous simple structures.

We addressed the foregoing issues using the experimental paradigm illustrated in Table 1. The critical word, to which the ERPs were measured, was the finite verb in the relative clause, italicized for clarity in Table 1. This critical verb in the relative clause is immediately preceded by either one or two noun phrases (NPs). To explore the effect of *repair*, grammaticality is manipulated by number (dis)agreement. In the One NP condition, the verb in the relative clause either agrees (Grammatical) or does not agree (Ungrammatical) in number with the modified NP ("hamburger(s)"). In the Two NP condition, the verb in the Ungrammatical condition does not agree with either of the preceding NPs (e.g., "the cake" and "the pizza" in Table 1), whereas the verb in the grammatical Preferred and Nonpreferred

Table 1. Overview of the Experimental Conditions

One NP conditions	
Grammatical	The man in the restaurant doesn't like the hamburgers that *are* on his plate.
Ungrammatical	The man in the restaurant doesn't like the hamburger that *are* on his plate.
Two NP conditions	
Preferred (Grammatical)	I cut the cake beside the pizzas that *were* brought by Jill.
Nonpreferred (Grammatical)	I cut the cakes beside the pizza that *were* brought by Jill.
Ungrammatical	I cut the cake beside the pizza that *were* brought by Jill.

conditions agrees in number with one of the two NPs. In both the One NP and Two NP cases, Ungrammatical continuations trigger repair processes.

To explore *revision* processes, syntactic preference is manipulated in the Two NP condition. Sentences in the Two NP condition are ambiguous up to the critical verb since the relative clause can modify either NP. Behavioral studies have shown that native speakers of English prefer to have a relative clause modify the last NP (Gilboy, Sopena, Clifton, & Frazier, 1995; Cuetos & Mitchell, 1988). The verb in the Preferred condition agrees in number with the last NP and hence confirms this interpretation. The verb in the Nonpreferred condition agrees only with the first NP and is therefore not compatible with the preferred analysis and triggers revision.

Finally, we explore the effects of *complexity* by comparing the Two NP conditions to the One NP conditions. In the One NP conditions, only one NP can be modified by the relative clause. In the Two NP conditions, the relative clause can potentially modify either of the two NPs. Up to the critical verb, the Two NP conditions are therefore ambiguous, whereas the One NP condition is not. In addition to the presence of an ambiguity, the Two NP conditions are more complex than the One NP conditions in terms of the syntactic structure of the NP before the relative clause (one simple NP in the One NP conditions vs. a NP modified by a prepositional phrase containing another NP in the Two NP conditions). This syntactic complexity corresponds to a complexity at the discourse level. In the One NP conditions, only one new discourse referent needs to be set up before the relative clause (e.g., "hamburger(s)" in Table 1); in the Two NP conditions, two new referents need to be set up (one for "cake(s)" and one for "pizza(s)"), which may be more effortful (Garrod & Sanford, 1994). Moreover, the referent of the second NP is used to further restrict the reference of the first (e.g., the cake(s) that are cut are the cake(s) beside the pizza(s)). The relative clause then introduces a second restriction in the Two NP conditions. In the One NP condition, no such additional restriction takes place.

ERPs were recorded from 26 subjects while they performed an acceptability judgment task. The results from the judgment task allow us to see to what extent participants judge the Nonpreferred condition as acceptable, indicating that they have successfully revised the structure.[2]

RESULTS

Behavioral Data

The percentage of accurate responses and the mean latency of accurate responses are given in Table 2. Performance was less accurate in the Nonpreferred compared with the Preferred, one NP Grammatical, and both Ungrammatical conditions [$F(4,22) = 18.78$, $p < .001$, pairwise t tests: p's < .001]. The Preferred, one NP Grammatical, and both and Ungrammatical conditions did not differ from each other [p's = .075]. The mean judgment latency also differed across conditions [$F(4,22) = 7.55$, $p < .001$]. Reaction times (RTs) in the Nonpreferred condition were longer than in each of the other conditions [p's < .01]. The Preferred condition was associated with longer RTs compared with the Two NP Ungrammatical [$p < .01$] and both One NP conditions [vs. Grammatical: $p < .05$; vs. Ungrammatical, $p < .01$]. The Two NP Ungrammatical, One NP Ungrammatical, and One NP Grammatical conditions did not differ from each other [p's > .071].

Event-Related Potentials

#### ERP Effects of Revision and Repair: Verbs in the Ambiguous Two NP Condition

The grand average ERPs for the Two NP conditions are given in the left and middle columns of Figure 1; isovoltage maps indicating the distribution of the effects are given in Figure 2A and B. Significant results of the ANOVA are displayed in Table 3. Between 100 and 200 msec, the Nonpreferred and Ungrammatical conditions showed a negativity at frontal sites and a positivity at posterior sites, relative to the Preferred condition. However, analyses restricted to frontal and posterior sites yielded no significant difference among the conditions. No significant differences were obtained between 200 and 300 msec. Between 300 and 400 msec, the average ERPs for both Nonpreferred and Ungrammatical conditions showed an anterior negativity relative to the Preferred condition. This is reminiscent of a LAN effect (Coulson et al., 1998b; Friederici, 1995). However, results were not significant (p's > .2 for frontal electrodes). We will now focus on 200-msec time windows from 500 to 1100 msec after onset of the critical verb, covering the P600 component (see Methods). Starting at 500 msec, both the Nonpreferred and the

Table 2. Behavioral Data

	One NP		Two NP		
	Grammatical	*Ungrammatical*	*Preferred*	*Nonpreferred*	*Ungrammatical*
Accuracy	94% (1.2)	95% (0.1)	91% (2.1)	70% (4.4)	96% (0.1)
Latency	970 (.10)	883 (.01)	1154 (.14)	1409 (.16)	924 (.11)

Mean percentage of accurate responses, and latency to correct responses in msec (*SE* in parentheses).
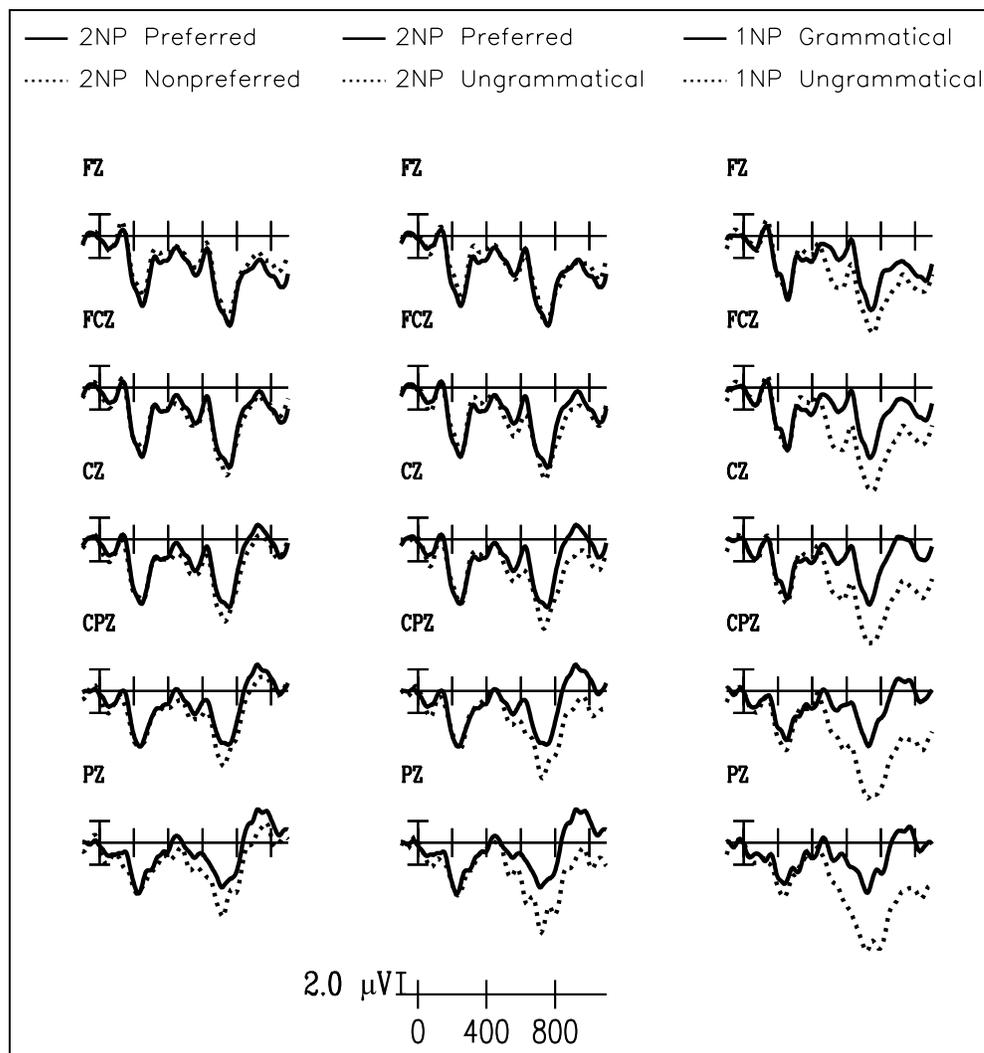
Ungrammatical conditions showed a positive deflection relative to the Preferred. The positivity was maximal for posterior electrode sites in the 700–900-msec time window and was largest for the Ungrammatical condition in the 700–1100-msec intervals.

We conducted pairwise comparisons, because we were particularly interested in the differences among the three conditions. Results are given in Table 4. Both the Nonpreferred and Ungrammatical conditions elicited a larger positivity at posterior sites compared with the Preferred condition, leading to significant Sentence Type × Electrode interactions. Compared with the Nonpreferred condition, the positivity elicited in the Ungrammatical condition had a larger amplitude especially at midline posterior sites in the 700–900- and 900–1100-msec intervals, leading to significant effects of Type and Type × Electrode interaction for these comparisons. To test whether there were distributional differences between the Nonpreferred and Ungrammatical conditions, we computed an ANOVA on the z score normalized difference waves obtained by subtracting the ERPs to the Preferred condition from those to the

Ungrammatical and Nonpreferred conditions (Rösler, Heil, & Glowalla, 1993). The conditions did not show any difference in distribution within the three time windows [typically, p's > .2]. We therefore have no evidence that the positivity elicited in the Nonpreferred condition has a different distribution from the positivity in the Ungrammatical condition.

In sum, relative to the Preferred condition, both Nonpreferred and the Ungrammatical conditions elicited a P600 that was larger and lasted longer for the Ungrammatical condition. As expected, the Ungrammatical continuation, which requires repair, elicited a posteriorly distributed P600 relative to the grammatical Preferred continuation. However, the Nonpreferred continuation, which triggers revision, also elicited a posteriorly distributed P600 relative to the Preferred continuation.

ERP Effects of Repair and Complexity: Effect of Grammaticality for the One and Two NP Conditions

The grand average ERPs at the verb in the One NP conditions are given in the rightmost column in Figure 1.

Figure 2. Isovoltage maps for the critical verb in the Two NP conditions (A: Nonpreferred minus Preferred; B: Ungrammatical minus Preferred) and One NP conditions (C: Ungrammatical minus Grammatical).
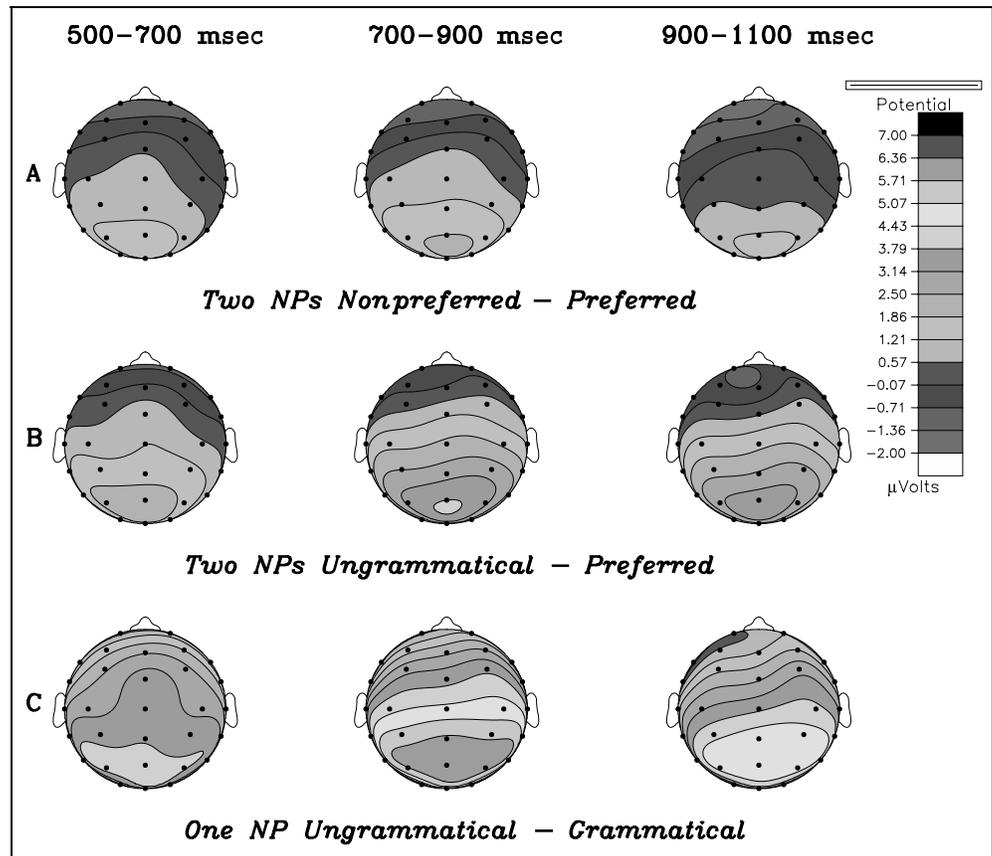

Figure 2C displays the isovoltage maps for these conditions. A comparison of the middle and right columns in Figure 1 shows that the positivity related to ungrammaticality was larger for the One NP compared with the Two NP conditions. This observation was confirmed by an ANOVA (Table 5) with the factors Number of NPs (One NP vs. Two NPs) and Grammaticality (Grammatical/Preferred vs. Ungrammatical). The positivity was largest at posterior sites, which was confirmed by a significant interaction of Grammaticality and Electrode. Although the positive effect of Ungrammaticality in the Two NP conditions appeared to be more restricted to posterior sites (cf. Figure 3), no significant three-way interactions of Grammaticality by Number of NPs with the factors Electrode and/or Hemisphere were found in any of the time windows [typically p's > .2]. An analysis on z score difference waves (Two NP Ungrammatical minus Preferred, vs. One NP Ungrammatical minus

Table 3. F Values for the ANOVA in the Two NP Conditions (Preferred, Nonpreferred, Ungrammatical)

Analysis	Effect (df)	Time Windows			
		100–200 msec	500–700 msec	700–900 msec	900–1100 msec
Midline	T (2,50)	–	–	8.84***	5.92**
	T × E (8,200)	3.77*	5.24**	13.62***	13.65***
Parasagittal	T (2,50)	–	–	5.48**	6.52**
	T × E (10,250)	4.89**	8.42***	13.59***	10.21***
Temporal	T (2,50)	–	–	3.30*	4.89*
	T × E (8,200)	–	5.58**	13.09***	11.19***

T = Sentence Type; E = Electrode; H = Hemisphere.

*p < .05.

**p < .01.

***p < .001.

Grammatical) also did not yield any significant differences in scalp distribution of the effect of Grammaticality for the One and Two NP conditions [p's > .2].

In sum, both the ambiguous Two NP and unambiguous One NP structures yielded a clear effect of grammaticality, that is, the P600 was larger in amplitude to ungrammatical than to grammatical sentence continuations. This effect was larger in the unambiguous One NP than in the ambiguous Two NP condition. For both ambiguous and unambiguous sentence structures the effect of grammaticality (repair) was maximal over posterior electrode sites.

ERP Effects of Complexity: One NP Grammatical versus Two NP Conditions

ERPs to the more complex and ambiguous Two NP conditions were more positive at frontal sites compared with the simpler unambiguous One NP Grammatical condition between 500 and 900 msec, especially at right electrode sites (see Figure 3). To test this, we calculated difference waves by subtracting the ERPs to the One NP Grammatical condition from the ERPs to each of the three Two NP conditions and conducted an ANOVA on these difference waves using eight frontal sites (F7, F8, FT7, FT8, F3, F4, FC3, FC4). This test confirmed that the positivity to the critical verb in the Two NP conditions was different from the One NP Grammatical condition in the 700–900-msec interval

[Mean difference: $F(1,25) = 5.46, p < .05$], and that the positivity was largest in the right hemisphere [Hemisphere: $F(1,25) = 6.65, p < .025$]. No significant differences were obtained as a function of Sentence Type (Preferred, Nonpreferred, or Ungrammatical) [Sentence Type: $F(2,50) = .3497$; Sentence Type × Electrode: $F(6,150) = 2.50, p = .09$; Sentence Type × Hemisphere: $F < 1$; Sentence Type × Electrode × Hemisphere: $F < 1$], indicating that the size and distribution of the frontal effect was not different among the conditions. The frontal positivity is therefore rather insensitive to type of continuation (Preferred, Nonpreferred, or Ungrammatical). This is not expected under the hypothesis that the frontal positivity reflects revision. Note furthermore that the difference between the Two NP Preferred and One NP Grammatical conditions was smallest at posterior sites where the effect of revision in the Two NP Nonpreferred condition, and the effect of Ungrammaticality in both One and Two NP conditions were largest (see Figure 3, bottom left graph). The frontal positivity found for the Two NP sentences can therefore be distinguished from the posterior positivity between 700 and 900 msec found for Ungrammatical and Nonpreferred continuations.

In sum, a frontally distributed positivity, which we will refer to as frontal P600 (FP600), was found to the critical verbs that were preceded by the more complex Two NP contexts compared with grammatical continuations in the simple One NP context. These results are consistent

Table 4. Pairwise Comparisons, Two NP Conditions

Effect (df)	500–700 msec			700–900 msec			900–1100 msec		
	P versus N	P versus U	N versus U	P versus N	P versus U	N versus U	P versus N	P versus U	N versus U
Midline									
T (1,25)	–	4.77*	–	4.56*	12.74**	6.29*	–	7.91**	9.83**
T × E (4,100)	6.12*	9.63***	–	9.06**	29.97***	4.48*	8.76***	29.22***	5.20*
Parasagittal									
T (1,25)	–	–	–	–	6.94*	6.78*	–	6.00*	14.38***
T × E (5,125)	8.91**	13.51***	–	12.97***	22.50***	–	11.57***	16.76***	–
Temporal									
T (1,25)	–	–	–	–	4.69*	4.32*	–	–	9.64**
T × E (4,100)	5.64*	7.78**	–	11.70**	21.71***	–	13.78***	18.46***	–

P = Preferred; N = Nonpreferred; U = Ungrammatical; T = Sentence Type; E = Electrode; H = Hemisphere.

*p < .05.

**p < .01.

***p < .001.

Table 5. *F* Values for the ANOVA with the Factors Number of NPs (Two NPs vs. One NP) by Grammaticality (Grammatical/ Preferred vs. Ungrammatical)

Analysis	Effect (df)	500–700 msec	700–900 msec	900–1100 msec
Midline	G (1,25)	24.46***	54.43***	24.69***
	G × E (4,100)	7.68**	29.11***	29.77***
	N (1,25)	–	13.97**	29.02***
	N × E (4,100)	–	7.51*	9.04***
	N × G (1,25)	8.06**	12.44**	7.95**
Parasagittal	G (1,25)	11.73**	42.37***	23.30**
	G × E (5,125)	23.47***	44.36***	28.26***
	G × H (1,25)	–	–	7.58*
	G × E × H (5,125)	–	–	3.18*
	N (1,25)	–	6.56*	16.71***
	N × E (5,125)	–	14.28***	11.02***
	N × G (1,25)	12.41**	16.04***	9.71**
Temporal	G (1,25)	9.41**	42.03***	22.09***
	G × E (4,100)	14.02***	57.39***	42.66***
	G × H (1,25)	–	–	9.85**
	N × E (4,100)	–	12.71***	12.84***
	N × E × H (4,100)	3.62*	–	–
	N × G (1,25)	12.66**	17.54***	8.15**

G = Grammaticality, E = Electrode Site, N = Number of NPs, H = Hemisphere.

*p < .05.

**p < .01.

***p < .001.

with the idea that FP600 is not restricted to syntactic revision processes per se, but that it is related to the processing of the more complex ambiguous structure.

DISCUSSION

The experiments reported in the present study investigated the neural correlates of syntactic analysis, focusing on the mechanisms of revision in ambiguous sentences, and of repair in Ungrammatical ambiguous and unambiguous sentences. Positive voltage ERPs with distributional differences were observed as a function of different syntactic operations that were required in the different sentence structures. A *posteriorly* distributed P600 was found to Ungrammatical continuations for both unambiguous and ambiguous sentence structures, and to sentence continuations that were not consistent with the preferred structural interpretation. A *frontally* distributed positivity (FP600) was observed when complex ambiguous sentence structures were compared

with simple grammatically correct unambiguous sentence structures (Two NP conditions vs. One NP Grammatical). Next we will further discuss the pattern of results, the implications for the existing ideas about the functional significance of the ERPs elicited by the different conditions, and the implications for psycholinguistic models.

ERPs to Ungrammatical and Nonpreferred Continuations of Ambiguous Sentence Structures

A comparison of the ERPs to Ungrammatical and Nonpreferred continuations of ambiguous sentence fragments permits an evaluation of the ERPs elicited by repair (Ungrammatical) compared to revision (Nonpreferred). The ERPs to the Ungrammatical condition elicited a larger positivity than to the Nonpreferred condition replicating the results of Osterhout, Holcomb, and Swinney (1994). However, we did not find any distributional differences between the two types of

continuation: The positive voltage ERPs to both Nonpreferred and Ungrammatical continuations were maximal over posterior electrode sites. This finding does not support the proposals by Hagoort and Brown (1994) and Hagoort et al. (1999). They have suggested that a frontal positivity may be a reflection of revision processes in case of Nonpreferred continuations, whereas the posterior positivity reflects repair in cases of Ungrammatical continuations. This predicts that in our study, positivity should be more frontally distributed for the Nonpreferred continuation, and posteriorly for the Ungrammatical continuation, a pattern that we did not find.

A factor that should be considered regarding the Nonpreferred condition in our experiment is that this condition may in fact be perceived as ungrammatical, and that, if true, both the Ungrammatical and Nonpreferred continuations in our experiment might actually involve repair. However, this seems unlikely to be the case, for the following reasons. First, the Nonpreferred conditions were judged to be grammatically correct in about 70% of the cases. Revision is not very difficult in these constructions, especially not when compared to notorious garden path sentences such as "The boat floated down the river sank," which received an acceptability rating of only 4–12% in a recent study (Osterhout, 1997). Second, also when the analysis was restricted to trials that were correctly judged acceptable or unacceptable, the same pattern of results was obtained: ERPs to both Ungrammatical and Nonpreferred conditions elicited a posterior positivity; no increased positivity relative to the Preferred condition was seen at frontal sites (see Figure 4, upper panel). The posterior positivity found in the Nonpreferred conditions can therefore not be caused by repair processes elicited by trials judged to be ungrammatical.

Finally, an analysis restricted to those eight subjects who had an accuracy rate of more than 85% for the Nonpreferred condition also showed a positivity with a posterior focus for both Ungrammatical and Nonpreferred conditions. As can be seen in Figure 4 (lower panel), no increased positivity was seen for the Nonprefer condition at frontal sites. So even in those cases where revision in the Nonpreferred condition is successful, the distribution of the positivity is not significantly more frontal for the Nonpreferred compared with the Ungrammatical conditions.

The present results therefore suggest that the posterior P600 reflects processes involved both in revision and repair.
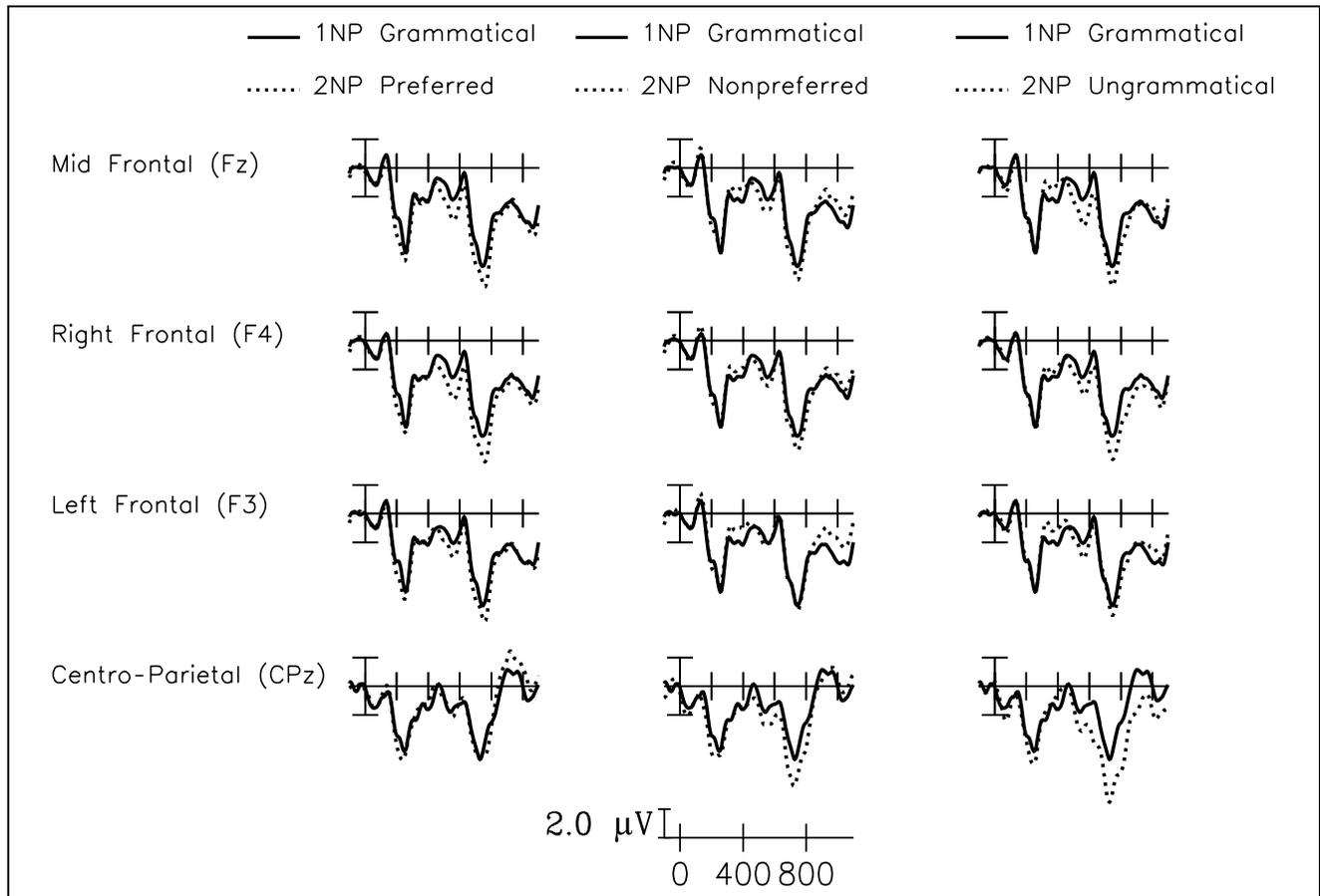

Figure 3. ERPs to the critical verb for midline frontal (Fz), right frontal (F4), left frontal (F3) and midline centro-parietal sites (CPz). Left column: One NP Grammatical condition (solid line) versus Two NP Preferred (dashed); middle column: One NP Grammatical condition (solid line) versus Two NP Nonpreferred (dashed); right column: One NP Grammatical (solid line) versus Two NP Ungrammatical (dashed).

Comparison of Complex Ambiguous and Simple Unambiguous Sentence Structures

The comparison between the Two NP and One NP conditions allowed us to investigate (a) the effects of complexity and ambiguity in grammatical sentences and (b) the effects of ambiguity and complexity on the processing of an Ungrammatical continuation.

Effects of Complexity and Ambiguity in Grammatical Sentences: Frontal Positivity

A comparison of the ambiguous, complex Two NP conditions versus the unambiguous, simple One NP grammatical conditions showed a frontal positivity (FP600). This frontally distributed positivity was found to be insensitive to the type of continuation of the Two NP fragment (Preferred, Nonpreferred, Ungrammatical). It is therefore unlikely that this frontal positivity is restricted to revision processes, as has been suggested (Hagoort & Brown, 1994; Hagoort et al., 1999). Instead, the frontal component could be related to the difference in complexity preceding the critical verb.

As pointed out in the introduction to this study, the Two NP conditions may be more taxing than the One NP conditions because the syntactic and discourse representations of the NP immediately preceding the relative clause are more complex. Although these differences occur before the critical verb, they may have affected processes at this verb. Our results correspond to those reported by Friederici et al. (2002) although our interpretation is slightly different.
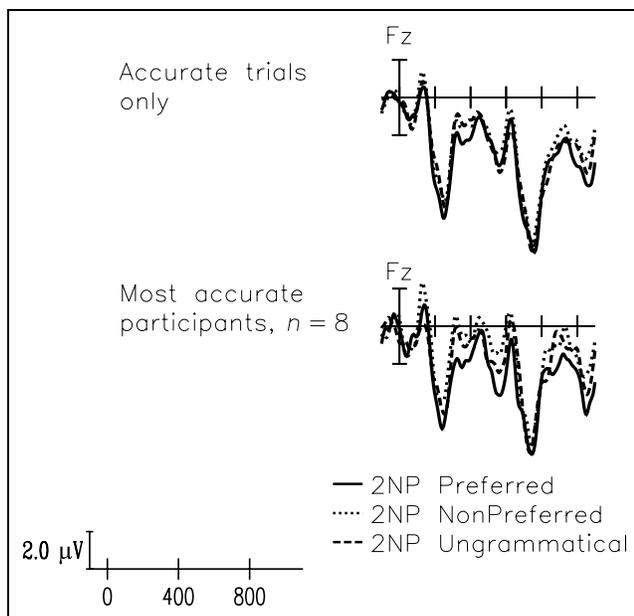

Figure 4. ERPs to the critical verb at Fz. Top: ERPs for trials that were accurately responded to; bottom: ERPs for the eight best-performing participants. Solid line: Two NP Preferred condition; dotted line: Two NP Nonpreferred; dashed line: Two NP Ungrammatical condition.

Note that also in the Friederici et al. study, positivity was found at the word following the word that introduced the complexity. The frontal positivity may then be interpreted as an aftermath of increases in discourse complexity and/or of increases in syntactic complexity due to factors other than integration of distal elements (which elicit a posterior positivity, as shown by Kaan et al., 2000).

Another possibility, which we cannot exclude, is that the frontal positivity is related to the presence of an ambiguity in the Two NP conditions. In the Two NP Preferred conditions, the parser may have retained two analyses in parallel before the verb is encountered (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Gibson, 1991): The preferred interpretation corresponding to the modification of the most recent NP, and the alternative interpretation corresponding to the modification of the first NP. When the clause is disambiguated at the verb, the alternative reading can be dropped from syntactic working memory. The frontal positivity could therefore be related to the removal of alternative analyses at the point of disambiguation. The One NP condition, on the other hand, is unambiguous, so no parallel processing and subsequent removal of alternatives occurs.

If the frontal positivity is indeed related to the presence of an ambiguity in the preceding context, this has some methodological consequences: The positivity found for Nonpreferred continuations may show a different distribution depending on whether an ambiguous or unambiguous structure is used as a baseline condition. This may in part explain the difference in distribution for Nonpreferred continuations reported in the literature: A more frontal/broad distribution of the positivity for Nonpreferred continuations has been reported when these items are compared to an unambiguous baseline (Hagoort et al., 1999; Friederici, Hahne, & Mecklinger, 1996, Experiment 2; Osterhout & Holcomb, 1992; Osterhout et al., 1994, Experiment 1); a positivity with a posterior maximum has been reported when an ambiguous baseline is used (Osterhout et al., 1994; Experiment 2: intransitive preferred vs. transitive preferred, but cf. Van Berkum et al., 1999).

Effects of Complexity and Ambiguity on Ungrammaticality: Posterior Positivity

The second effect of the One NP versus Two NP manipulation was that the posterior P600 for the Ungrammatical relative to the Grammatical conditions was larger for the simple, unambiguous One NP condition compared with the complex, ambiguous Two NP condition, although no differences in distribution of the positivity were found. The decrease in P600 amplitude may have been due to the presence of an ambiguity in the Two NP conditions. In the Two NP conditions, the relative clause can modify either

the first or the second NP. Initially, the relative clause is taken to modify the second NP. However, the verb in the Ungrammatical condition is incompatible in number with this NP. Since there is another NP, the parser may have first tried to combine the relative clause with the first NP before deciding that the verb is incorrect and initiating mechanisms dealing with this ungrammaticality. In the One NP condition, on the other hand, there is only One NP that the relative clause can modify. The ungrammaticality of the verb can be decided upon immediately. This may have led to a more diffuse and, on average, smaller positivity in the Two NP compared with the One NP cases.

Functional Significance of the Positive Components

Combined with the results of previous experiments, the present data support the view that several positive voltage ERPs components can be distinguished, which are related to certain aspects of syntactic processing. First, a positivity between 500 and 1100 msec can be distinguished with a posterior focus (posterior P600). This component can be associated with syntactic processing difficulty related to both repair (Ungrammatical continuations) and revision (Nonpreferred continuations), as well as certain kinds of structural complexity (e.g., filler-gap dependencies). Second, a frontal positivity between 500 and 900 msec (FP600) can be distinguished. The present findings suggest that this component cannot be associated with syntactic revision processes per se, as opposed to what has been proposed in previous studies. Instead, it may be associated with ambiguity resolution and/or increases in discourse complexity. More research is needed to further identify the cognitive processes these components are indices of. This line of investigation will render the P600 and FP600 components, and ERPs in general, an even better tool to investigate language processing.

Psycholinguistic Models

On the basis of the current data, we have no evidence that the processes involved in syntactic repair and revision are distinct. This supports models in which Ungrammatical continuations and Nonpreferred continuations are dealt with in the same way. One such model is the Diagnosis and Repair model proposed by Fodor and Inoue (1998). According to this model, the sentence processing mechanisms only pursues one reading in the case of syntactic ambiguity. Input that is incompatible with the analysis pursued is attached to the representation anyway. Starting from the errors resulting from this illicit attachment, the structure is changed, eliciting a posterior P600 component. This revision can either be successful (in the case of Nonpreferred input) or not (Ungrammatical continuations).

In the latter case, the resulting structure may be changed again. This may account for the fact that the posterior P600 is larger and longer lasting for the Ungrammatical compared with the Nonpreferred continuations. Additional support for this model is that there are no early differences between the Nonpreferred and Ungrammatical conditions, suggesting that also at early stages of processing both are dealt with in the same way (Hopf, Bader, Meng, & Bayer, in press). These results are somewhat problematic for parallel models, according to which multiple analyses are pursued in parallel and ranked according to preference. In such models, Nonpreferred continuations trigger the reactivation of a lower-ranked parse. Ungrammatical continuations, on the other hand, involve completely different mechanisms, because no analysis is compatible with the newly input continuation. Although the present data are consistent with the Fodor and Inoue model, clearly, more research is needed to exclude other models of revision and repair.

METHODS

Subjects

Twenty-six subjects (12 men, age 17–22 years, mean 19.5) participated, either paid or for credit. All were healthy, right-handed, monolingual native speakers of English, had normal or corrected-to-normal vision, and were undergraduate or graduate students at Duke University. Subjects gave informed consent before the experiment.

Materials

Two NP Conditions

Triplets were constructed according to the format illustrated in Table 1. The critical verb was followed by at least two other words. The critical verb was always plural. This was to avoid cases such as "the pizza next to the cakes that was..." where the (stronger) plural features of the second NP may overwrite the singular features of the first NP, which may lead to a sense of ungrammaticality at the verb independently of modification preferences (Deevy, 2000; Bock & Miller, 1991). Furthermore, we used prepositions other than "of" or "with" to separate the two NPs to ensure a strong preference to modify the second NP (Gilboy et al., 1995). A total of 120 triplets were selected on the basis of a paper-and-pencil completion questionnaire. Sixty-eight native speakers of English were asked to complete sentences presented up to and including "that." Items were selected on the basis of the following criteria. First, the "that" clause was completed as modifying the second NP in more than 60% of the cases in which the modified NP could be unambiguously identified; and, second, there

was no significant difference in the percentage of second NP modifications between the versions with a singular first NP and a plural second NP on the one hand (83%), and versions with a plural first NP and a singular second NP on the other (85%). For the ERP study, three main lists were created according to a Latin square design such that each list contained 40 items per condition. Each list contained only one member of each triplet, and each member of a triplet occurred only once in a list.

One NP Conditions

Eighty pairs were created of the format illustrated in Table 1. In order to have these items resemble the Two NP conditions, the critical verb was always plural. In order to counterbalance these control sentences over the experimental lists, two sublists were created for each of the three main lists. The One NP items were Latin-squared over the two sublists, such that each list contained 40 grammatical and 40 ungrammatical One NP sentences.

In addition, 120 filler items were constructed. Forty contained a syntactic ambiguity resolved towards the nonpreferred reading (e.g., "Mary put the strawberries in the fridge in the ice cream"). In this way, the experiment contained and equal number of Nonpreferred and Ungrammatical continuations. Forty additional fillers contained semantic violations, and another 40 were semantically and grammatically correct. Experimental items and fillers were pseudorandomized. Each list was divided into 10 blocks of 32 sentences each. The order of the 10 blocks was different for each participant. A list of experimental and filler items can be obtained from the first author.

Procedure

Participants were comfortably seated in a dimly lit, electrically shielded booth, with a video screen 1.10 m in front of them. The screen was covered with black cardboard, with a 9 × 4 cm slit in the center through which the stimuli could be viewed. Sentences were presented word by word at a rate of 500 msec per word (300 msec word, 200 msec blank screen), Tahoma 14 pts, white letters on a black background. The visual angle was less than 3°. Each sentence was preceded by a fixation cross (1500 msec). The last word of each sentence was followed by a blank screen (1500 msec), followed by a prompt (OK ? BAD). The prompt remained on the screen until the participant responded by pushing a button labeled "ok" or "bad" on the response pad. Response hand was balanced across lists. After the response, the message "Press for next" was displayed, which remained on the screen until the participant responded.

Participants were instructed to read the sentences carefully and not to blink from the first word of the sentence until they saw the prompt. They were asked to judge each sentence for semantic and syntactic acceptability, and to respond accurately and quickly at the prompt. Before the actual experiment, participants read a practice block with seven items and feedback was given when the subject made any incorrect judgments. No feedback was given during the actual experiment. On average, each experimental session lasted 2 hr and 45 min, including preparation.

EEG Recording

EEG was recorded from 31 Ag/AgCl scalp electrodes, mounted in an elastic cap (Neuroscan Quickcap): midline: AFz, Fz, FCz, Cz, CPz, Pz, Oz; Lateral: Fp1/2, F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, O1/2, referenced to the left mastoid. Additional electrodes were placed on the left and right outer canthus, and above and below the left eye to monitor eye movements. EEG was amplified and digitized at a rate of 250 Hz. The signal was filtered on-line between 0.01 and 30 Hz.

Analysis

Behavioral Data

Percentage of correct responses and RTs for accurate responses were analyzed with a repeated-measure General Linear model, with Sentence Type (5 levels) as a within-subjects factor. Outliers longer than 10 sec were treated as missing data. If the main effect of Type was significant, pairwise comparisons were conducted using t tests.

EEG Analysis

Epochs were comprised of the 200 msec preceding and 1600 msec following the critical verb. Trials with excessive eye movements or drift were rejected from analysis. This was 6.7–8.6% of the data in each condition. Data were filtered off-line using a gaussian low-pass filter with a 25-Hz half amplitude cutoff. ERPs were quantified as the mean amplitude relative to a 100-msec prestimulus baseline, using the following latency windows: 100–200 msec (N1), 200–300 msec (P2), 300–400 msec (LAN effect), 500–700 msec (Early P600); 700–900 msec (mid P600), and 900–1100 msec (late P600), based on the literature and visual inspection. All trials were taken into consideration regardless of the subject's response on the acceptability judgment task. This was because some of the participants had only very few correct responses in some of the conditions, and a response-contingent averaging would have decrease the stimulus to noise ratio. However,

an analysis on correct responses only did not qualitatively differ from the analysis reported in the main text.

Separate repeated-measure ANOVAs were performed on midline (Fz, FCz, Cz, CPz, Pz), parasagittal (Fp1/2, F3/4, FC3/4, C3/4, CP3/4, P3/4), and temporal sites (F7/8, FT7/8, T7/8, TP7/8, P7/8), with Sentence Type (Preferred, Nonpreferred, Ungrammatical for the Two NP items), Electrode (five or six levels), and, where applicable, Hemisphere (two levels) as within-subjects factors. In addition, ANOVAs were done to compare the One NP versus Two NP conditions with factors Grammaticality (Grammatical/Ungrammatical) and Number of NP (One NP vs. Two NPs). Electrodes O1, Oz, and O2 were not included in the analyses because of technical problems with these sites in some of the subjects. For effects involving more than one degree of freedom, the Greenhouse–Geisser correction (Greenhouse & Geisser, 1959) was applied, to avoid Type I errors due to unequal variances between the conditions. Since the aim of this study is to investigate differences between the various type of sentences, only the main effects of Sentence Type or interactions of Sentence Type with Electrode Site and/or Hemisphere are reported. Differences in scalp distribution were assessed with z score corrections as described in the Results.

Notes

1. It is still a matter of debate whether the P600 is an instantiation of the P300, a component that occurs for unexpected stimuli in general (Osterhout & Hagoort, 1999; Coulson, King, & Kutas, 1998a). Note, however, that the P600 is not elicited by any unexpected input (semantic violations typically elicit an N400, not a P600). Furthermore, Kaan, Harris, Gibson, and Holcomb (2000) report a P600 for words that are completely expected given the preceding sentence context. We therefore take the stance that, although the P300 may partially overlap with the P600, the P600 is specific to difficulty related to structural (syntactic) aspects of the input.

2. One may raise the concern that an acceptability judgment task induces specific strategies and hence may obscure differences between Nonpreferred and Ungrammatical continuations. However, in a recent study in our lab, we do not find any differences between an acceptability judgment task and a passive reading task on the scalp distribution of the ERPs. Furthermore, qualitatively similar ERP components have been reported in the literature for both judgment (e.g., Osterhout & Holcomb, 1992) and passive reading tasks (Hagoort et al., 1993). We therefore do not think that the task requirements in the present study substantially affected the results.

REFERENCES

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45–93.

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

Coulson, S., King, J. W., & Kutas, M. (1998b). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Cuetos, F., & Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition, 30,* 73–105.

Deevy, P. L. (2000). Agreement checking in comprehension: Evidence from relative clauses. *Journal of Psycholinguistic Research, 29,* 69–79.

Fodor, J. D., & Ferreira, F. (Eds.) (1998). *Reanalysis in sentence processing.* Dordrecht, The Netherlands: Kluwer.

Fodor, J. D., & Inoue, A. (1998). Attach anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 101–141). Dordrecht, The Netherlands: Kluwer.

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., 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., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.

Garrod, S., & Sanford, A. J. (1994). Resolving sentences in a discourse context: How discourse representation affects language understanding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675–698). New York: Academic Press.

Gibson, E. (1991). A computational theory of human linguistic processing: Memory limitations and processing breakdown. Unpublished Ph.D., Carnegie Mellon University, Pittsburg, PA.

Gilboy, E., Sopena, J.-M., Clifton, C., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition, 54,* 131–167.

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

Hagoort, P., & Brown, C. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. J. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 45–80). Hillsdale, NJ: Elrbaum.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–483.

Hagoort, P., Brown, C., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 273–316). Oxford, UK: Oxford University Press.

Hopf, J.-M., Bader, M., Meng, M., & Bayer, J. (in press). Is human sentence parsing serial or parallel? Evidence from event-related potentials. *Cognitive Brain Research,* DOI: 10.1016/S0926-6410(02)00149-0.

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.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language, 59,* 494–522.

Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King, and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 14,* 1–14.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 786–803.

Rösler, F., Heil, M., & Glowalla, U. (1993). Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology, 30,* 170–182.

Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 27–39). Oxford, UK: Oxford University Press.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language, 41,* 147–182.