When Does Gender Constrain Parsing?
Evidence from ERPs

Jos J. A. van Berkum,[1,2] Colin M. Brown[1], and Peter Hagoort[1]

We review the implications of recent ERP evidence for when and how grammatical gender agreement constrains sentence parsing. In some theories of parsing, gender is assumed to immediately and categorically block gender-incongruent phrase structure alternatives from being pursued. In other theories, the parser initially ignores gender altogether. The ERP evidence we discuss suggests an intermediate position, in which grammatical gender does not immediately block gender-incongruent phrase structures from being considered, but is used to dispose of them shortly thereafter.

When we read or listen to language, we have a clear sense of understanding the words as they come in. Psycholinguistic experiments have confirmed this intuition, revealing that semantic interpretation is to a large extent incremental (e.g., Marslen-Wilson & Tyler, 1980; Kutas & Van Petten, 1994). By recording event-related brain potentials (ERPs), eye movements, and other variables related to on-line language processing, psycholinguists have also obtained abundant evidence for incremental syntactic analysis or parsing (for overviews see Mitchell, 1994; Tanenhaus & Trueswell, 1995). For example, the processing consequences of a number agreement error, as in "The spoilt children *throws* the toy on the floor," emerge in the ERP waveforms within some 500 ms after onset of the word at which the problem becomes appar-

ent (Hagoort, Brown, & Groothusen, 1993; Osterhout & Mobley, 1995). Findings like this show that as a sentence unfolds, listeners or readers immediately relate new words to a syntactic analysis of the preceding input.

With natural language input, however, incremental parsing is no trivial task. One major problem is that as a sentence unfolds over time, many words temporarily resist unambiguous syntactic assignment. In "David told the girl that. . .," for example, the word *that* might begin a complement clause, as in "David told the girl that there would be some visitors," but also a relative clause, as in "David told the girl that had been on the phone to hang up." A large body of research on parsing has revealed that the parser provisionally resolves such a local syntactic ambiguity as soon as it arises, by focusing on a single one of the alternatives (Frazier & Clifton, 1996; Mitchell, 1994). We also know that some alternatives are systematically preferred over others.

Much attention has been devoted to exploring the factors that control these parsing preferences. Some have argued that the parser initially always goes for the structurally simplest alternative (e.g., Frazier & Rayner, 1982; Ferreira & Clifton, 1986), or the alternative with the most frequently encountered structure (e.g., Mitchell, Cuetos, Corley, & Brysbaert, 1995). Others have claimed that the parser initially pursues the alternative that is the most plausible one given the current discourse (e.g., Crain & Steedman, 1985) or given a whole range of constraints (including biases based on lexical and structural frequencies, thematic roles, and discourse context; e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995). In the context of this debate, we conducted an ERP study in which we simultaneously explored the role of two possibly relevant factors: discourse context, and grammatical gender.

DOES GRAMMATICAL GENDER INFLUENCE THE PROVISIONAL PARSE?

By definition, the rules of a grammatical gender system constrain the phrase structure analyses that readers or listeners can ultimately arrive at. Just exactly when and how gender information is used to control the parsing process, however, is still an unsolved issue. One possibility is that grammatical gender unequivocally constrains the initial assignment process. Consider the following examples (adapted from Brysbaert & Mitchell, 1996):

(1) The son   of   the actress   who was on the balcony. . .
(2) Het$_{NEU}$ zoontje$_{NEU}$ van de$_{COM}$ actrice$_{COM}$ dat$_{NEU}$ op het balkon zat. . .
    The$_{NEU}$ son$_{NEU}$ of   the$_{COM}$ actress$_{COM}$ who$_{NEU}$ was on the balcony. . .

In (1), a syntactic ambiguity arises at the word *who*, because it is at that point not yet clear whether the following relative clause should be taken to modify the first or the second NP. In the corresponding Dutch example (2), however, this relative clause attachment ambiguity apparently does not arise, because whereas a neuter noun such as *zoontje*$_{NEU}$ agrees with the gender-marked relative pronoun *dat*, a common gender noun such as *actrice*$_{COM}$ requires the alternative relative pronoun form *die* instead.[3]

The idea that in cases like (2) grammatical gender simply precludes syntactic ambiguity from arising seems quite reasonable: it is a syntactic feature, after all, which in this case conspires with the rest of Dutch syntax such that just one well-formed analysis remains. Of course, this linguistic definition does not necessarily bear on actual processing. But in psycholinguistic models of parsing that work with feature unification (e.g., Kempen, 1999; Vosse & Kempen, 1998), for instance, there is a straightforward processing equivalent: phrase structures that require a merge of incompatible gender features simply do not get constructed, and thereby do not enter the set of competing alternative analyses. This means that in this type of processing model, gender can indeed prevent syntactic ambiguity from arising. Within the framework of a radically different model of parsing (*Construal Theory*, Frazier & Clifton, 1996), gender also immediately blocks incongruent phrase structures.

On the other hand, there is some reason to consider the possibility that instead of playing a pivotal role in the construction of well-formed phrase structures, grammatical gender is initially ignored by the parser altogether. Working with sentences like (2), for example, Brysbaert and Mitchell (1996) obtained a pattern of results that seemed to indicate that on their first pass through the sentence, native readers of Dutch are quite happy to attach a neuter relative pronoun to a noun of common gender (and vice versa). In line with their overall account of frequency-driven parsing preferences, Mitchell *et al.* (1995) therefore proposed that the parser always initially pursues the most frequent phrase structure analysis, and at that point does not take gender agreement into account. To explain late effects of gender on pronoun resolution, De Vincenzi (this volume; De Vincenzi & Di Domenico, 1999) likewise suggested that gender is not taken into account during initial parsing.

Thus, the literature suggests at least two plausible hypotheses for when and how grammatical gender constrains the parsing process: gender blocks candidate phrase structures such that the parser does not at any time consider them as viable options, or gender is initially ignored by the parser

---

[3] Dutch nouns are of neuter or common gender, which, among other things, affects the form of the definite article and the relative pronoun (see van Berkum, 1996, Ch. 2).

altogether—for current purposes, we will refer to these as the immediate-blocking and the delayed-impact account, respectively.

Somewhere inbetween these two theoretical options lies a third one: although gender may not block incongruent phrase structures from the very beginning, it might be used to very rapidly detect and dispose of such an analysis, as soon as the parser has begun to pursue it. This so-called rapid-check-and-dispose account differs from the immediate-blocking hypothesis in that gender is used immediately after—rather than before—the parser commits itself to a particular analysis.[4] It differs from the delayed-impact account in that, in the latter, gender is used much later, with delays that are not typically considered to be part of first-pass (initial) parsing.[5]

In the ERP experiment we describe below, we tested these three accounts by looking at the effect of discourse context on syntactic ambiguity resolution in sentence structures that did or did not provide an additional gender constraint. We begin by describing the effects of discourse under conditions where gender is not a constraint (see van Berkum, Brown, & Hagoort, 1999a, for a full account).

ERP EXPERIMENT: THE IMPACT OF DISCOURSE CONTEXT ON THE PROVISIONAL PARSE

In these conditions, we used Dutch target sentences that began like

(3) David vertelde het meisje dat. . . (David told the girl that. . .)

and in which the indirect object NP had neuter gender. Because of the latter (see below), these sentences presented the parser with a syntactic ambiguity at the word *dat*, which could introduce a complement clause, as in "David vertelde het meisje dat er visite kwam" (David told the girl$_{NEU}$ that$_{COMPL}$ there would be some visitors), but also a relative clause, as in "David vertelde het meisje dat had zitten bellen op te hangen" (David told the girl$_{NEU}$ that$_{RELPR(NEU)}$ had been on the phone to hang up).

---

[4] In the parsing literature, this contrast is often referred to in terms of "factor X guides/proposes" versus "filters/disposes."

[5] We are aware of the fact that this formulation does not specify what counts as first-pass parsing, and, hence, what exact processing event the delay is relative to. Following previous empirical operationalizations, we simply take "delayed" to refer to cases where in on-line language processing measures, the processing consequences of some constraint emerge not at the word that provides the constraint but several words downstream (e.g., at the end of the clause or sentence; cf. Brysbaert & Mitchell, 1996; De Vincenzi, this volume).

We placed these temporarily ambiguous target sentences in two types of discourse context, exemplified below:

(4) 1-referent context:

David had de jongen en het meisje gezegd hun kamer voor de lunch op te ruimen. Maar de jongen had de hele ochtend liggen slapen, en het meisje had voortdurend zitten bellen.

(David had told the boy and the girl to clean up their room before lunch time. But the boy had stayed in bed all morning, and the girl had been on the phone all the time.)

(5) 2-referent context:

David had de twee meisjes gezegd hun kamer voor de lunch op te ruimen. Maar het ene meisje had de hele ochtend liggen slapen, en het andere had voortdurend zitten bellen.

(David had told the two girls to clean up their room before lunch time. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time.)

The 2-referent context introduced two plausible referents for the neuter noun phrase in the target sentence (*het meisje*), and as such favored a relative-clause continuation, a construction often used to provide additional referential restrictions (which girl? the girl that had been on the phone; see Crain & Steedman, 1985). The otherwise identical 1-referent context introduced a single unique referent for the NP. Because the latter does not require further modification by a restrictive relative clause (and because written Dutch nonrestrictive relative clauses require a comma after the noun), this context favored a complement-clause continuation.

The issue was whether these contextual biases can affect the parser's provisional resolution of the syntactic ambiguity generated at *dat*. According to context-sensitive theories of parsing, the parser can make immediate use of such information (e.g., Altmann, 1988; Crain & Steedman, 1985; Spivey-Knowlton & Tanenhaus, 1998). According to syntax-first theories of parsing, however, discourse-level information would initially be ignored (e.g., Frazier & Rayner, 1982; Mitchell *et al.*, 1995). To assess how the parser had dealt with the syntactic ambiguity generated by *dat*, we disambiguated the target sentence at the very next word (shown in boldface below):

(6) neuter-noun/complement-clause target:

| David vertelde | het meisje | dat | **er** | visite kwam. |
| David told | the girl$_{NEU}$ | that$_{COMPL}$ | **there** | would be some visitors. |

(7) neuter-noun/relative-clause target:

| David vertelde | het meisje | dat | **had** | zitten bellen op te hangen. |
| David told | the girl$_{NEU}$ | that$_{RELPR(NEU)}$ | **had** | been phoning to hang up. |

We presented the materials to 24 subjects and recorded event-related brain potentials (ERPs) as they read the critical sentence (see van Berkum, Brown, & Hagoort, 1999a, for details). We used ERPs because earlier research had shown that a word at which the parser must abandon its hitherto preferred (or only) syntactic analysis elicits, at about 500 ms after its onset, a distinct positive shift in the ERP, labelled the Syntactic Positive Shift (SPS) or P600 (see Brown & Hagoort, 1999, or Osterhout & Holcomb, 1995, for reviews). This well-established ERP correlate of syntactic garden-pathing allowed us to derive very specific predictions from context-sensitive and syntax-first accounts of parsing. In particular, to the extent that the parser would be more inclined to pursue the relative-clause analysis at the word *dat* in a 2-referent context than in a 1-referent context, subsequent disambiguation as a complement clause (at *er*) should force the parser to abandon this preferred analysis again and should thus elicit an SPS in the 2-referent context. Conversely, to the extent that the parser would be more inclined to pursue the complement-clause analysis at the word *dat* in a 1-referent context than in a 2-referent context, subsequent disambiguation as a *relative* clause (at *had*) should compel it to abandon this analysis, now leading to an SPS in the 1-referent context.

It was exactly this cross-over pattern of results that we observed in the ERP waveforms. As illustrated in the upper half of Fig. 1 for a representative electrode site (Pz; over the midline and back of the scalp), complement-clause disambiguation elicited an SPS in the 2-referent context (relative to the 1-referent one), within about 500 ms after onset of the disambiguating word (*er*). As can be seen in the lower half of this figure, relative-clause disambiguation instead elicited an SPS in the 1-referent context, again within about 500 ms after onset of the disambiguating word (*had*). Furthermore, fully consistent with assumptions about critical aspects of our methodology, no context- induced SPS effects emerged at the preceding word *dat* (which generated the syntactic ambiguity but did not resolve it) in either sentence type. Together with particular control measures discussed elsewhere (van Berkum, Brown, & Hagoort, 1999a), these findings show that discourse context can in principle immediately influence the analysis of a syntactic ambiguity.

ERP EXPERIMENT: THE ADDITIONAL IMPACT OF GRAMMATICAL GENDER

What happens if we add a grammatical gender constraint? To explore this, our experiment also contained target sentences (mixed with those discussed before) in which the referentially manipulated neuter noun was
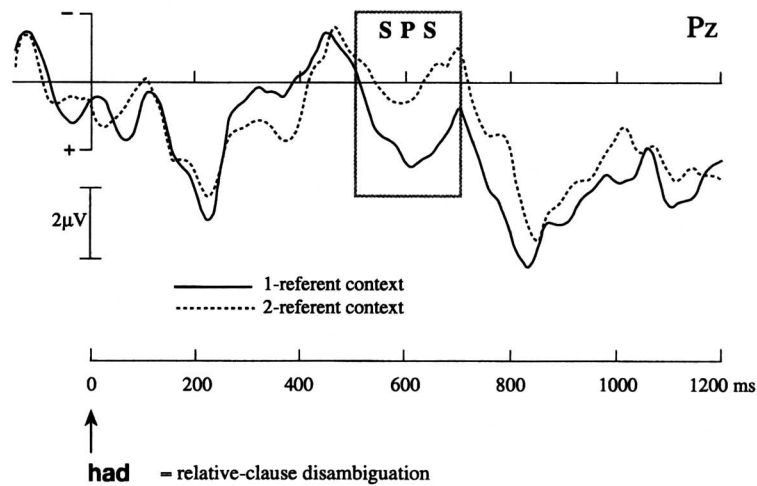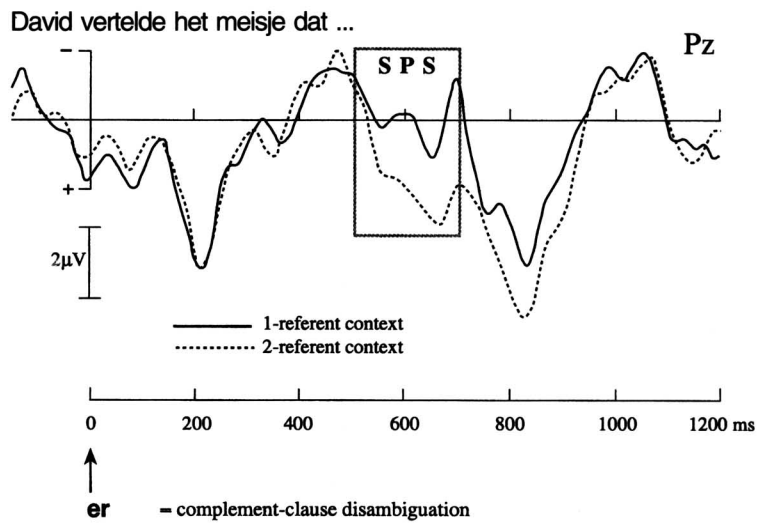
David vertelde het meisje dat ...

S P S

Pz

2μV

——— 1-referent context
·············· 2-referent context

0   200   400   600   800   1000   1200 ms

er   = complement-clause disambiguation

S P S

Pz

2μV

——— 1-referent context
·············· 2-referent context

0   200   400   600   800   1000   1200 ms

had   = relative-clause disambiguation

Fig. 1. Grand average ERP waveforms, at Pz, elicited by complement-clause disambiguation (er, upper half of figure) and by relative-clause disambiguation (had, lower half of figure) in neuter-noun target sentences presented in a complement-biasing 1-referent context (solid line) and a relative-biasing 2-referent context (dotted line). In all figures, the onset of the disambiguating word is at 0 ms, the next word follows at 600 ms, negative polarity is plotted upwards, and the rectangle marks the latency window used to test for SPS effects (500–700 ms after CW-onset).

replaced by a comparable noun of *common* gender. Compare the neuter gender targets (6) and (7) to their common gender variants shown in (8) and (9):

(8) common-noun/complement-clause target:

David vertelde	de vrouw	**dat**	er	visite kwam.
David told	the woman$_{COM}$	**that**$_{COMPL}$	there	would be some visitors.

(9) common-noun/relative-clause target:

David vertelde	de vrouw	**die**	had	zitten bellen op te hangen.
David told	the woman$_{COM}$	**that**$_{RELPR(COM)}$	had	been phoning to hang up.

As was illustrated by the neuter-noun/relative-clause sentence in (7), the form of the relative pronoun for a neuter-gender noun is *dat*. In Dutch, *dat* also happens to be the form of the generic complementizer. It is this lexical ambiguity that gives rise to the associated complement/relative-clause ambiguity. As shown by the common-noun/relative-clause sentence in (9), however, the form of the relative pronoun for a common-gender noun is *die*. Linguistically speaking, the use of a common-gender noun in sentences like (8) and (9) thus completely eliminates the local complement/relative-clause ambiguity, much like grammatical gender eliminated the relative-clause attachment ambiguity in (2) discussed in the beginning of this paper. As before, the issue is what happens in real-time processing.

The critical sentence is the common-noun/complement-clause sentence in (8). If the parser does not construct gender-incongruent phrase structure alternatives, then at the word *dat*, the parser should never consider the relative-clause analysis as a viable option to pursue, whatever the discourse context, and no garden-path effect should arise at complement-clause disambiguation. The immediate-blocking account thus predicts that discourse-induced SPS effects will not be elicited by any of the critical words (*dat* or *er*) in this sentence type.

If the delayed-impact hypothesis is correct, the results for sentence (8) should in fact look exactly like those we obtained for sentence (6). The word *dat* should give rise to the complement/relative-clause ambiguity, a 2-referent context should lure the parser into pursuing the relative-clause alternative, and subsequent complement-clause disambiguation at *er* should force it to abandon this preferred analysis again, resulting in an SPS effect at that latter word.

When we designed the experiment, however, our intuition was that the results would confirm the rapid-check-and-dispose hypothesis. The reason was an informal observation shared by a number of native speakers of Dutch: when reading "David vertelde de vrouw dat . . ." in a discourse context with *two* women, the use of *dat* actually "feels" as if a gender agreement error has been made at that point. Note that this suggests that the word

dat must have been taken as a relative pronoun. We therefore envisaged the following sequence of processing events:

1. When processed as part of "David vertelde de vrouw dat ...," the wordform *dat* makes available two morpholexical entries: the generic complementizer *dat*$_{COMPL}$ and the relative pronoun *dat*$_{RELPR(NEU)}$.
2. The parser momentarily ignores gender and considers two candidate analyses:
   (a) [$_s$ David vertelde [$_{NP}$ de vrouw]   [$_{CC}$ *dat*$_{COMPL}$ ...]]
   (b) [$_s$ David vertelde [$_{NP}$ de vrouw [$_{RC}$ *dat*$_{RELPR}$ ...]] [...   ]]
3. In a 2-referent discourse context, the parser provisionally commits itself to the relative-clause analysis, at least to a larger extent than in a 1-referent context.
4. Before processing the next word, the preferred analysis is checked on gender agreement.
5. To the extent that a 2-referent context has biased the parser to pursue the relative-clause analysis [$_{NP}$ *de vrouw*$_{COM}$ [$_{RC}$ *dat*$_{RELPR(NEU)}$ ...]], the resulting gender agreement violation forces it to abandon its preferred analysis again, leading to an SPS in that context in the ERP waveform to *dat*.

So, although gender does not immediately and categorically block constructions like *[$_{NP}$ de vrouw*$_{COM}$ [$_{RC}$ *dat*$_{RELPR(NEU)}$ ...]], agreement is checked rapidly enough to allow the parser to dispose of a preferred analysis at the word where the provisional commitment was made (rather than several words downstream).

As can be seen in Fig. 2, our findings directly confirmed the above account. Relative to a 1-referent context, the presentation of a common-noun/complement-clause target like "David vertelde de vrouw dat ..." in a 2-referent context elicited an SPS at the word *dat*, within about 500 ms.[6]

---

[6] Whereas the two SPS effects reported before were most clearly visible at parietal electrode sites like Pz, the present SPS effect had a more limited anterior distribution and is therefore illustrated for the frontal electrode site Fz. The exact reason for this topographical variation is as yet unknow, but it may be related to differences in the types of diagnosis and repair required after the parser has garden-pathed at our different syntactic disambiguations. What is critical for current purposes, however, is that the anterior effect clearly falls within the range of SPS effects reported in the literature and can as such be taken to reflect a syntactic dead-end (see van Berkum, Brown, & Hagoort, 1999a, for further discussion). Note, in addition, the small negativity that precedes the SPS in Fig. 2; although not statistically reliable in the 150- to 300-ms latency range (nor between 300 and 450 ms), it does in this early window have a left-anterior distribution (see van Berkum, Brown, & Hagoort, 1999a, Fig. 4b), and as such resembles syntax-related left-anterior negativities reported elsewhere (LAN; e.g., Friederici, Hahne, & Mecklinger, 1996).
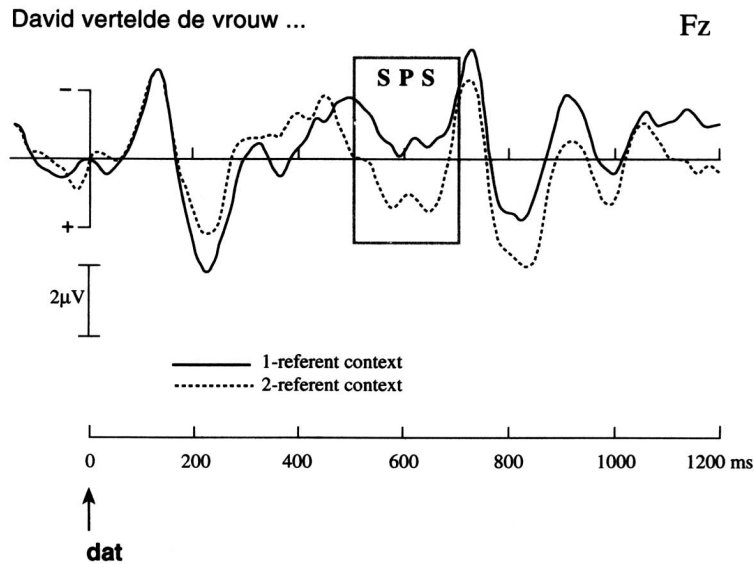
Fig. 2. Grand average ERP waveforms, at Fz, elicited by *dat* in common-noun/complement-clause targets presented in a complement-biasing 1-referent context (solid line) and a relative-biasing 2-referent context (dotted line).

Furthermore, as illustrated in Fig. 3 for the same electrode position, no context-induced SPS effect emerged at the subsequent word *er*. And in line with our assumptions about the possible impact of our referential context manipulation, no discourse-dependent SPS effects were elicited by *die* or *had* in common-noun/relative-clause targets exemplified in (9).

IMPLICATIONS

Let's step back and consider the implications of these results. If, in common-gender/complement-clause sentences such as example (8), grammatical gender had categorically blocked the construction of the relative-clause alternative at the word *dat*, there would be no complement/relative-clause ambiguity, and hence also no provisional parsing decision to be affected by our discourse manipulation. Thus, in a 2-referent context, no discourse-induced syntactic garden-path effect should have been observed at any word, but there clearly was such an effect.
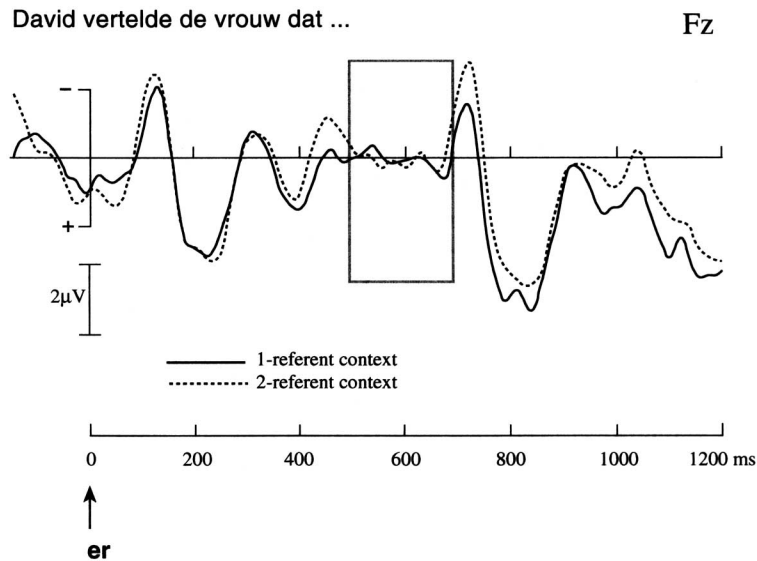
Fig. 3. Grand average ERP waveforms, at Fz, elicited by *er* in common-noun/complement-clause targets presented in a complement-biasing 1-referent context (solid line) and a relative-biasing 2-referent context (dotted line).

The SPS effect at *dat* is also at odds with the delayed-impact account (with "delayed" defined as "several words downstream"; see again footnote 5). If gender had been ignored altogether, the common-noun/complement-clause sentences exemplified in (8) should have been processed in exactly the same manner as their neuter-noun/complement-clause counterparts exemplified in (6). However, in contrast to the word *er* in neuter-noun/ complement-clause sentences, the same word did not elicit any SPS effect in common-noun/complement-clause sentences.

Only the rapid-check-and-dispose account had predicted that in common-noun/complement-clause sentences, a 2-referent context would induce an SPS at the word *dat*. As part of this account, we made two critical assumptions. First, grammatical gender does not immediately and categorically block the relative clause construction to such an extent that a 2-referent context cannot lead the parser to provisionally commit itself to that construction. This explains why we observed a context-induced garden-path effect in the ERP signal. Second, gender agreement is nevertheless checked rapidly enough to allow the parser to dispose of that analysis at the word where the provisional commitment was made. This explains why the context-induced

garden-path effect emerged at *dat*, the very word that also gave rise to the ambiguity in the first place.[7]

How rapid should we take the impact of gender to be? The relevant SPS began at about 450–500 ms after onset of the critical word *dat*. This means that the results of a gender agreement check became available within at most about 450–500 ms after presentation of a relevant agreement target. This is consistent with the results of another ERP study (Hagoort & Brown, this volume), where outright gender agreement violations such as in the Dutch NP *de*$_{COM}$ **meisje*$_{NEU}$ also elicited an SPS effect within about 450–500 ms (see also Osterhout & Mobley, 1995, for antecedent-pronoun agreement violations in English).

Interestingly, the rapid-check-and-dispose account of our results suggests that the impact of referential features of the wider discourse is even faster. Whereas grammatical cues such as gender are typically assumed to be fast-acting (e.g., Frazier & Clifton, 1996; Vosse & Kempen, 1998), discourse-level information is often assumed to bear on processing at a later stage (e.g., Fodor, Ni, Crain, & Schankweiler, 1996; but see van Berkum, Hagoort, & Brown, 1999b). In the current experiment, discourse-level information seems to have affected the parser *before* local gender agreement is checked.

There are at least two possible explanations for this surprising result. One is that the parser is designed to check gender agreement for a single analysis only, so that with two candidate analyses available, the agreement check must wait until other factors have led the parser to commit itself to a provisional parse. A second possibility is that, rather than being scheduled in a principled way, gender agreement information and discourse-level information might for some "accidental" reason exert their respective impacts at different times, even though they can in principle be brought to bear on syntactic ambiguity resolution at the same time. For instance, gender information may have had a somewhat slower "rise time" than the discourse-level information supplied in our experiment, even though gender has a much more categorical impact in the end. Although the reason for this would still have to be established independently, such an account would obviously be in the spirit of recent constraint-

---

[7] We have simplified our account of the processing at *dat* somewhat by ignoring two other analyses that Dutch syntax licenses at this point, one with *dat* as a demonstrative determiner [David vertelde de vrouw dat grapje. (David told the woman that joke)], and one with *dat* as an independent demonstrative [David vertelde de vrouw dat maar niet. (David didn't tell the woman that after all)]. It is not very likely that either of these is pursued given the syntactic context used in our experiment (see Tabor, Juliano & Tanenhaus, 1997, for the impact of a preceding similar "<Agent> said/told . . ." frame on the analysis of a subsequent "that"). Also note that what is critical for our argument is that gender rules out the (discourse-supported) relative-pronoun reading, not whether just one analysis (the complementizer) or several analyses (the complementizer, the demonstratives) remain.

based models of language comprehension (e.g., MacDonald *et al.*, 1994; Spivey-Knowlton & Tanenhaus, 1998; Tabor, Juliano, & Tanenhaus, 1997).[8]

Both accounts may strike one as rather unlikely, because—whatever the exact time-course and reasons for it—they allow noncategorical, probabilistic information to temporarily preempt categorical evidence for ungrammaticality. But there is other evidence to suggest that sufficiently biasing probabilistic information can lead people to momentarily entertain a formally ungrammatical analysis. Tabor *et al.* (1997), for example, observed that structural frequency biases can be strong enough to get people to pursue an analysis that clearly violates available subcategorization restrictions. Readers can sometimes also pursue an analysis that violates subject-verb number agreement (Crocker, 1998). Our present finding may in fact thus not be unique to the use of grammatical gender. Perhaps it is the case that when the parser takes stock of the viable phrase structure options for some current word, *any* grammatical feature other than a word's basic syntactic category (N, V, etc.) can be temporarily overruled by strong contextual biases. In a feature unification architecture (e.g., Vosse & Kempen, 1998), this could be accommodated by allowing imperfect unifications, with their quality indexed by some strength parameter. With its "Attach anyway" principle, the recently elaborated diagnosis model of Fodor and Inoue (1998, 1999) in fact explicitly allows for ungrammatical attachments. What remains surprising under any model of tolerant parsing, however, is that a discourse-induced (and presumably correlation-based) preference outweighs a hard syntactic constraint.

Even though gender does not operate rapidly and categorically enough to block a discourse-supported but gender-incongruent analysis from the start, it does rapidly initiate reanalysis, whatever strong biases are present in the discourse context. The evidence for rapid reanalysis comes from a com-

---

[8] One of the reviewers pointed out that the discourse constraint was supplied well before the gender constraint and might as such have had a processing advantage. This, however, does not necessarily translate to a headstart, because although the referential information is in some sense already in the system, this information still needs to be made relevant to the syntactic ambiguity, a process that can only begin at the word that introduces the ambiguity (as well as the gender marker; *dat*). We by no means wish to rule out the possibility that if we take this additional factor into account, an order-based advantage for the discourse constraint will still remain. But this would be an interesting phenomenon, not an artifact, because discourse context always precedes local syntactic cues.

A related concern involves the relative saliency of our discourse manipulation. By the end of the experiment, most of our subjects had become aware of the referential contrast. We do not believe, however, that this led them to predict an upcoming sentence structure at the point of the critical noun phrase. First, a posttest revealed that most of our subjects were completely unaware of the critical syntactic alternations. Furthermore, a contextual prediction strategy would have been of little use, because we had paired an equal number of complement and relative clause continuations with each type of discourse context.

parison of ERP effects elicted at the subsequent word *er* across neuter-noun/complement-clause sentences and common-noun/complement-clause sentences. Compare once more sentences (6) and (8). In (6), a 2-referent context led the parser to pursue a relative-clause analysis until the word *er* unambiguously indicated that this analysis could no longer be maintained, and as such elicited an SPS effect. In (8), the relative-clause analysis ran into trouble at the preceding word *dat,* which under that analysis did not agree in gender with the antecedent noun, and as such elicited an SPS effect. The absence of a similar SPS effect at the subsequent word *er* can only be explained by assuming that the parser, upon having detected that the relative clause analysis entailed a gender agreement violation, immediately discarded that analysis in favor of the complement-clause alternative. Effectively, our results therefore suggest that, at the word *dat,* gender is first ignored, then checked, and finally used to succesfully revise the analysis, all before the next word is processed.[9]

In the context of the diagnosis model, Fodor (1998) recently proposed that gender agreement violations, although clearly signaling that something is wrong, provide little information to guide reanalysis and are as such relatively ineffective triggers for succesful recovery. The current results suggest that this is not generally the case, at least not to the extent that the correct analysis cannot be found within at most 600 ms (the word-onset asynchrony between *dat* and *er* in our study). Within the diagnosis model, this might be accounted for by assuming that since *dat* has just been read, the alternative complementizer reading dat_{COMPL} is still readily available. In its current formulation, however, the diagnosis parser will resist giving up the relative pronoun reading $dat_{RELPR(NEU)}$ and will initially try to modify some aspect of the partial parse tree to which the latter has been attached. Under this account, the implication of our findings would be that the parse tree to the left has been fully reconsidered within at most 600 ms (the *dat–er* onset asynchrony).

---

[9] In the current study, we used written target sentences, which—due to technical requirements imposed by EEG recording—were presented word by word with a 600-ms word-onset asynchrony (see van Berkum, Brown, & Hagoort, 1999a, for details). In spoken Dutch, however, the two critical words *dat* and *er* have an average onset asynchrony in the order of 100–200 ms (measured on a sample of 120 sentences containing a *dat er* sequence). The findings of a replication study with fully connected spoken versions of the same materials (Brown, van Berkum, & Hagoort, 1999) suggest that even this much shorter word-onset asynchrony leaves the parser with enough time to revise its analysis: in 2-referent contexts, SPS effects were again only elicited by the word *er* in neuter-noun/complement-clause sentences and by the word *dat* in common-noun/complement-clause sentences. This equivalence across input modalities also suggests that the results reported here are not in any way caused by the relatively slow word-by-word visual presentation procedure.

Turning to other findings on the impact of gender agreement violations, it is at this time not clear to us why Brysbaert and Mitchell (1996) did not find grammatical gender effects in parsing. Both studies were in Dutch, and both involved violations of relative pronoun agreement. One difference to consider is that whereas Brysbaert and Mitchell had used various measures of reading time, we recorded ERPs—perhaps the processing consequences of a gender-incongruent relative pronoun do *not* slow down the reading process but nevertheless do affect the ERPs. Another potentially relevant difference is that in our study, the agreement violations always involved the neuter relative pronoun *dat*. Dutch linguists have observed that native speakers of Dutch more and more often use the common relative pronoun *die* to refer to neuter words that denote a human referent (e.g., *meisje, girl*). Critical for our argument is that no such erosion has been observed for the neuter relative pronoun *dat*, which still firmly requires a neuter antecedent noun. To the extent that one of the Dutch relative pronouns is relatively unselective, the mixed use of both in the Brysbaert and Mitchell study may have diluted any effects that hinge on selectivity in agreement.

CONCLUSIONS

Whereas in some theories of parsing, grammatical gender has been assumed to immediately and categorically block gender-incongruent phrase structure alternatives from being pursued (e.g., Frazier & Clifton, 1996; Vosse & Kempen, 1998), others have argued that the parser is initially blind to this feature altogether (e.g., Mitchell *et al.*, 1995; De Vincenzi, this volume). We have discussed recent ERP evidence that suggests an intermediate position. Our findings indicate that in the case of a local syntactic ambiguity, gender-incongruent phrase structures are not blocked to the extent that the parser cannot provisionally pursue them. On the other hand, the ERP results also unequivocally show that the resulting gender agreement violation is very rapidly detected, within at most about 450–500 ms after onset of the problematic word. In the present experiment, that turned out to be rapid enough to allow the parser to succesfully revise its analysis before processing the next word.

Discrete-stage theories of parsing can account for these findings if they assume that the incremental assignment of every word into the unfolding phrase structure is not initially constrained by gender agreement, but is immediately followed by a gender feature check on the resulting, provisionally selected structure—(as an aside, these theories should also allow provisional selection to be affected by discourse context). Within the less discrete framework of constraint-based models of parsing (e.g., MacDonald *et al.*, 1994;

Spivey-Knowlton & Tanenhaus, 1998; Tabor et al., 1997), our findings can be taken to suggest that a gender agreement constraint, although having a categorical impact in the end, can take more time to affect ongoing processing than either core syntactic features such as word category or (perhaps more surprising) pragmatic information about the number of referents in prior discourse. Gender is fast, but apparently not that fast.

REFERENCES

Altmann, G. T. M. (1988). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes, 3(2),* 73–97.

Brown, C. M., van Berkum, J. J. A., & Hagoort, P. (1999). Discourse before gender: An event-related brain potential study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholinguistic Research.* In press.

Brown, C. M., & Hagoort, P. (1999). On the electrophysiology of language comprehension: Implications for the human language system. In M. Crocker, M. Pickering., & C. Clifton Jr. (Eds.), *Architectures and mechanisms for language processing.* Cambridge, UK: Cambridge University Press. (pp. 213–237).

Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology, 49A(3),* 664–695.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. R. Dowty, L. Karttunen, & A. M. N. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge, UK: Cambridge University Press.

Crocker, M. (1998). Statistical mechanisms in language processing. Paper presented at the 4th Conference on Architectures and Mechanisms of Language Processing (AMLaP-98), Freiburg, September 24–26.

De Vincenzi, M. & Di Domenico, E., (1999). A distinction among phi-features: The role of gender and number in the retrieval of pronoun antecedents. In E. Di Domenico & M. De Vincenzi (Eds.), *Gender and number in normal and impaired language processing.* Special Issue of *Rivista di Linguistica,* In press.

Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Fodor, J. D. (1998). Number, case, and argument structure: Triggers for reanalysis. Paper presented at the 4th Conference on Architectures and Mechanisms of Language Processing (AMLaP-98). Freiburg, September 24–26.

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

Fodor, J. D., & Inoue, A. (1999). Garden-path reanalysis: Attach (anyway) and revision as last resort. To appear in Proceedings of AMLaP-1996. In press.

Fodor, J. D., Ni, W., Crain, S., & Schankweiler, D. (1996). Tasks and timing in the perception of linguistic anomaly. *Journal of Psycholinguistic Research, 25(1),* 25–57.

Frazier, L., & Clifton, C., Jr. (1996). *Construal.* Cambridge, MA: MIT Press.

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., 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(5),* 1219–1248.

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

Kempen, G. (1999). Human grammatical coding. Leiden: Leiden University. Submitted.

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). New York: Academic Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101(4),* 676–703.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.

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

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24(6),* 469–488.

Osterhout, L., & Holcomb, P. J. (1995). Event-related potentials and language comprehension. In M. D. Rugg, & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 171–215). Oxford: Oxford University Press.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739–773.

Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24(6),* 1521–1543.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes, 12(2/3),* 211–271.

Tanenhaus, M. K., & Trueswell, C. (1995). Sentence comprehension. In J. L. Miller, & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217–262). San Diego: Academic Press.

van Berkum, J. J. A. (1996). The pyscholinguistics of grammatical gender: Studies in language comprehension and production. Nijmegen, The Netherlands: Nijmegen University Press.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999a). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language,* In press.

van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999b). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience,* In press.

Vosse, T., & Kempen, G. (1998) Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. Leiden: Leiden University. Submitted.