Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction

Article *in* Journal of Cognitive Neuroscience · July 2000

DOI: 10.1162/089892900562336 · Source: OAI

3 authors, including:

Thomas C. Gunter
Max Planck Institute for Human Cognitive and Brain Sciences
142 PUBLICATIONS **6,207** CITATIONS

SEE PROFILE

Herbert Schriefers
Radboud University
110 PUBLICATIONS **5,753** CITATIONS

SEE PROFILE

Syntactic Gender and Semantic Expectancy: ERPs Reveal Early Autonomy and Late Interaction

Thomas C. Gunter and Angela D. Friederici
Max-Planck Institute of Cognitive Neuroscience

Herbert Schriefers
Nijmegen University, Netherlands

Abstract

■ This experiment explored the effect of semantic expectancy on the processing of grammatical gender, and vice versa, in German using event-related-potentials (ERPs). Subjects were presented with correct sentences and sentences containing an article-noun gender agreement violation. The cloze probability of the nouns was either high or low. ERPs were measured on the nouns. The low-cloze nouns evoked a larger N400 than the high-cloze nouns. Gender violations elicited a left-anterior negativity (LAN, 300–600 msec) for all nouns. An additional P600 component was found only in high-cloze nouns. The N400 was independent of the gender mismatch variable; the LAN was independent of the semantic variable, whereas an interaction of the two variables was found in the P600. This finding indicates that syntactic and semantic processes are autonomous during an early processing stage, whereas these information types interact during a later processing phase. ■

Whenever we read text or hear speech, words have to be integrated with each other in order to allow an interpretation of a text or an utterance. This process is based on structural (syntactic) and meaning (semantic) features, both necessary for language comprehension. One major debate in psycholinguistics is how and when these syntactic and semantic sources of information are used. Are both sources influencing the comprehension process simultaneously, or is there a strict order with respect to which information is processed first? Besides this time course issue, psycholinguists also want to know whether or not both types of information can be separated and shown to invoke different and independent cognitive processes (i.e., encapsulation; compare, Fodor, 1983, 1988). The present experiment discusses electrophysiological data from a reading experiment, which indicates that semantic and syntactic processes are initially working in parallel, but interact with each other in a later processing stage. Semantic and syntactic processes/information were investigated by manipulating semantic expectancy of a target noun, given a preceding sentential context (high versus low), and gender agreement between a noun and its article. Before turning to the relevant linguistic and electrophysiological details let us first discuss some issues regarding grammatical gender in order to get a feel for what it represents.

If native speakers are aware of gender at all, it is mostly with respect to semantic gender (male versus female) and not with respect to the grammatical gender. The grammatical gender of a word is assigned arbitrarily and, therefore, differs across languages.[1] The word ''house'' for instance is masculine in Russian, feminine in French and neuter in German (Corbett, 1991). For foreign speakers it is typically the case that grammatical gender is a source of inescapable mistakes making languages with an extended gender system more difficult to learn during second language acquisition. There are several theories, which try to explain differences in gender systems or try to explain why grammatical gender exists. It has been argued, for instance, that gender can be used for back tracking references in a complex discourse (Bates & MacWhinney, 1989). Gender may also play a facilitatory role in sentence processing because, (a) gender increases the cohesion of a sentence (Desrochers, 1986); (b) gender cues may facilitate the recognition of words (Bates, Devescovi, Hernandez, & Pizzamiglio, 1996); and (c) gender can sometimes disambiguate homophones (Van Berkum, 1996; e.g., in Dutch: *de* jacht versus *het* jacht [the hunting versus the yacht]). Thus, it is clear that in languages which have grammatical gender, this syntactic feature has a potential role in parsing and comprehension in general.

Linguistic theory suggests that the mental lexicon is organized such that lexical entries contain, besides meaning information, syntactic information about the word's syntactic category. In the case of verbs, the lexical entry also contains the verb's argument structure, and in the case of nouns, it contains gender information. Levelt (1989) views this information as part of the lemma. The availability of the different types of information during comprehension is modeled differently by different approaches. Frazier (1987), for instance, suggests a serial model in which an initial structure is built solely on the basis of word category information. Thematic roles are only assigned in a second step on the basis of other sources of information including semantics. Such a model requires word category information to be available prior to information concerning the meaning, the gender, or the argument structure of a given word. Any model assuming the autonomy of syntactic processes holds that even if the latter types of information are available in the same time domain no interaction between syntactic information and semantic information should occur. Interactive or parallel models (cf., McClelland, St. John, & Taraban, 1989; Marslen-Wilson & Tyler, 1980) suggest that all sources of information interact continuously and predict no sequential advantage of any type of information over the other. For instance, in the interactive approach suggested by Bates et al. (1996) grammatical gender, as provided by the preceding context, can enhance the processing of an upcoming noun by increasing or decreasing the activity of possible lexical candidates (i.e., nouns). According to this view, the processing system can do this before the point at which the actual physical information about the noun is perceived. This model therefore suggests an interaction between gender information and semantic expectancy. Because the interaction between several types of psycholinguistic information is important for distinguishing between different psycholinguistic models of sentence processing, the present paper explores the processing of gender information, and its possible interaction with semantic information.

In two behavioral studies using lexical decision (Schriefers, Friederici, & Rose, 1998; Friederici & Schriefers, 1994), it was shown that in sentences like ''Er trinkt das Bier'' (He drinks the$_{neuter}$ beer$_{neuter}$), the semantic priming between the verb and the noun[2] can be diminished or even removed by a gender mismatch between the article and the noun. Thus, in ''Er trinkt den Bier'' (He drinks the$_{masculine}$ beer$_{neuter}$) there is no facilitation in the recognition of the noun (relative to an unrelated condition), although the verb drink usually enhances the recognition of the word beer. Schriefers et al. (1998) discuss a pre and a postlexical interpretation explaining this elimination of lexical–lexical priming between the verb and the noun. The postlexical interpretation suggests that their findings are due to the negative outcome of a postlexical syntactic procedure which checks whether the gender agreement rule is respected or not. A more prelexical or interactive account of the data would suggest that gender information, like semantic expectancy, is used in order to preselect candidates for the next word to come. Thus, whether gender information comes into play at a prelexical or postlexical level still remains to be solved. That it is difficult to scrutinize this issue might be related to the type of method used to explore it. Typically, reaction time measures are used in these types of studies. It may, however, very well be that these measures are not ''on-line'' enough to validate certain questions, i.e., they do not provide the temporal resolution necessary to distinguish early and late processing phases.

Because of their perfect on-line properties and their sensitivity to semantic- and syntactic-information processing, we used event-related brain potentials (ERPs) in order to explore the interactions of gender agreement violations with semantic expectancy. ERPs are scalp-recorded voltage changes in the electroencephalogram that are time locked to sensory, motor or cognitive events (e.g., Reagan, 1989 or Rugg & Coles, 1995). Electrophysiological psycholinguistics started off in the 1980s exploring semantic and pragmatic violations, which were found to elicit an N400 component (Kutas & Hillyard, 1980; for a review see Van Petten, 1995). This negativity around 400 msec after stimulus onset is typically seen as reflecting the difficulty of integrating a particular word in a preceding linguistic context. It was found for instance that the component was inversely related to word expectancy (i.e., the higher the cloze probability of a word the smaller the component). There are several studies showing that the N400 reflects postlexical processes. Brown and Hagoort (1993), for instance, showed in a priming paradigm that the N400 priming effect on the target is absent when the prime is masked such that it cannot consciously be identified. Using a levels of processing approach, Chwilla, Brown, and Hagoort (1995) showed the N400 in lexical priming to be absent in a shallow task (decision on character size), again suggesting that the N400 reflects lexical integration processes (for similar data but in a sentence context see Gunter & Friederici, 1999).

More recently, researchers started to become interested in how syntax is reflected in ERPs. Generally speaking, two ERP-components have been correlated with the processing of syntactically anomalous sentences. A relatively early negativity (i.e., LAN; with a frontally or left anterior maximum) and a late centro-parietal positivity (i.e., P600 or the so-called syntactic positive shift). The early negativities are typically elicited by elements that turn the sentence into an incorrect one. It is interesting to note that those studies that realized the sentence's incorrectness by a word category error elicited a left-anterior negativity in a very early time window (around 200 msec) (Hahne & Friederici, 1997; Friederici, Pfeifer, & Hahne, 1993; Friederici, Hahne, &

Mecklinger, 1996; Neville, Nicol, Barss, Forster, & Garrett, 1991), whereas studies in which the incorrectness was realized as a morphosyntactic error (i.e., inflectional), evoked a left-anterior negativity between 300 and 500 msec (Vos, Gunter, Kolk, & Mulder, submitted; Coulson, King, & Kutas, 1998; Gunter, Stowe, & Mulder, 1997; Penke, Weyerts, Gross, Zander, Münte, & Clahsen, 1997; Osterhout & Mobley, 1995; Münte, Heinze, & Mangun, 1993; Münte & Heinze, 1994; Münte, Matzke, & Johannes, 1997). A left-anterior negativity in this latter time window was also found for targets violating a verb's argument structure (Rösler, Friederici, Pütz, & Hahne, 1993). This led Friederici (1995) to propose that the different latencies may reflect a temporal hierarchy in the availability or use of the different types of information encoded in the lexical entry, with word category information being available first. P600s are found for both syntactically incorrect as well as syntactically infrequent structures. While the P600 is mostly preceded by an early negativity in the case of incorrect sentences, it is not in the case of correct sentences with a nonpreferred syntactic structure. Thus, a P600 which is often preceded by a negativity, is found for morphosyntactic violations (e.g., subject–verb number disagreement: Vos et al., submitted; Osterhout, McKinnon, Bersick, & Corey, 1996; Hagoort, Brown, & Groothusen, 1993; also a nonword context: Münte et al., 1997; other types of verb inflection violations: Gunter et al., 1997; case disagreement: Coulson et al., 1998). Moreover, this LAN-P600 pattern is also observed with subjacency constraint violations (Neville et al., 1991) and with subcategorisation violations (Osterhout & Holcomb, 1992). A solitary P600 is found in different types of garden-path sentences (Friederici et al., 1999; Osterhout & Holcomb, 1992; Osterhout, Holcomb, & Swinney, 1995; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995).

The functional significance of the syntax-related components is still under debate. Some researchers suggest that the P600 is the most robust syntax related ERP-component (Osterhout et al., 1996; Hagoort et al., 1993) whereas other researchers suggest a specific role of each of the components in the parsing process. Friederici (1995), for instance, suggests that the early left-anterior negativity observed in response to word category errors reflects first pass parsing processes whereas the P600 reflects second-pass parsing processes like reanalysis or repair (cf., Frazier, 1987). Münte et al. (1997) suggested that the LAN reflects the actual detection of a morphosyntactic mismatch whereas the P600 reflects the necessary reprocessing of the sentence in order to make it semantically and syntactically meaningful. This suggestion was based on the observation that the LAN for subject-verb agreement violation is elicited by real and pseudo words whereas the P600 was only present for real words. The debate whether or not the LAN reflects a pure syntactic process or indexes working-memory load (Kluender & Kutas, 1993) seems to be resolved by a recent experiment (Kluender et al., 1998) in which two types of LANs were found. The syntax related LAN only showed up in single word averages whereas the memory related LAN was elicited in both single and across sentence ERPs. Thus, it seems that the LAN can be classified as syntactic in nature.

Although the nature and the underlying mechanism of the P600 is still open for specification, it is not unrealistically to assume we are dealing with a more controlled process. Gunter et al. (1997) suggested that the P600 component reflects a more controlled type of processing because the component was found to be affected by a broad range of task variables such as semantics (Gunter et al., 1997), working-memory load (Vos et al., submitted; Gunter et al., 1997), and probability of occurrence of a violation (Hahne & Friederici, 1999; Coulson et al., 1998; Gunter et al., 1997). This suggestion was also supported by a study in which a levels-of-processing approach showed the P600 to be smaller in a shallow task (Upper-lower character distinction; Gunter & Friederici, 1999). Note that the recent discussion, whether or not the P600 and the P3b originate from the same family (Coulson et al., 1998; Münte, Heinze, Matzke, Wieringa, & Johannes, 1998; Gunter et al., 1997; but see Osterhout, McKinnon, Bersik & Corey, 1996; and Osterhout & Hagoort, 1999), is relevant for this issue because the P3b can also be seen as the reflection of controlled processes.

In the present experiment, we will take the different ERP components, namely N400, LAN, and P600 to reflect underlying processes of a different type. The N400 is viewed to reflect semantic processes and the LAN to reflect morphosyntactic processes. In the present context, we will not commit ourselves to a specific view concerning the P600. As suggested above, from a psycholinguistic perspective it is important to know whether or not semantic and syntactic processes interact with each other and, if they do, in what way. Evidence on this issue comes from a recent experiment (Gunter et al., 1997) in which semantic and syntactic processes were investigated using a 2×2 design crossing a semantic manipulation (normal versus anomalous sentences) and with a morpho-syntactic manipulation (correct versus incorrect inflection of a verb). It was found that the LAN and the N400 reflect independent language processes, which can be related to syntactic (LAN) and semantic (N400) aspects. The P600 appeared to be independent of semantic aspects in its first part. Only in the later part of the P600 were interactions of syntax with semantics significant. As in the scope of the present experiment, gender violations are of particular interest; we will briefly review ERP-effects reported for this type of violation. Two experiments explored the processing of semantic gender in pronouns, and two others the processing of syntactic gender in article-noun relations. One of them is presented in this paper and will be discussed in detail. The other one, conducted by Hagoort and

Brown (1997), can only be discussed on the basis of the conference report as a full report is not yet available. Let us first turn to the semantic gender studies.

The first experiment exploring ERP-effects of gender violations was carried out by Osterhout and Mobley (1995). Subjects were presented with sentences containing a semantic gender violation on a reflective pronoun (e.g., ''The woman congratulated herself/*himself . . .''). A P600 component was found for the gender violation to start at around 400 msec. In the same experiment, sentences containing a pronoun that matched or mismatched the gender of the subject noun also elicited a P600 (e.g., ''The aunt heard that she/he . . .''). Although this last type of sentence is in principle not a real violation, it seems that the parser incorporates semantic gender in its processing the outcome of which is matched with certain structural preferences (relative clause versus complement clause).

Osterhout, Bersick, and McLaughlin (1997) conducted a follow up experiment in which they manipulated semantic gender-related social stereotypes in English (e.g., ''The doctor prepared himself/herself for the interview.''). Although one might expect the less stereotypical reflexive pronoun (in this case ''herself'') to elicit an N400 component because it is a pragmatically less expected word, a P600 was found instead. This P600 was smaller compared to the P600 elicited by definitional gender violations (e.g., ''The woman congratulated herself/*himself . . .''). Because masculine and feminine gender corresponds almost perfectly with the semantic distinction between male and female sex, Osterhout et al. suggest that gender information is also encoded in these gender stereotypes. Thus, the two studies on processing semantic gender consistently show a P600 as a response to a gender violation, be it purely semantic or related to social stereotypes.

Now let us turn to the two recent experiments on the processing of syntactic gender. Hagoort and Brown (1997) explored the interplay of semantic and syntactic processing using a combination of syntactic gender and semantic violations in Dutch noun phrases. Gender violation was realized as a mismatch between a definite article and a noun (e.g., De/*Het kapotte paraplu staat in de garage [The/The$_{\text{incorrect gender}}$ broken umbrella is in the garage]). Semantic violations were realized by a prenominal adjective, which did or did not fit semantically with the following noun (e.g., De kapotte/formele paraplu staat in de garage [The broken/formal umbrella is in the garage]). These two manipulations were fully crossed in a factorial design. Note that both types of violations can only be detected on reading the noun (i.e., paraplu) which therefore constitutes the critical word in this study. It was found that the ERP elicited by the combined semantic and gender violation (i.e., het formele paraplu) showed a larger N400 compared to the pure semantic violation (i.e., de formele paraplu). Most relevant for the present discussion was the finding of a P600 component for the gender violation compared to the correct sentence. Inspection of their Figure 1, in which a prestimulus baseline is used, suggests that a smaller P600 was elicited by the combined semantic and syntactic violation, a finding that replicates a result of Gunter et al. (1997). Whereas the negativity reported by Hagoort and Brown was an N400, usually observed for semantic processes, we found in the present experiment a LAN, which is taken to reflect syntactic processes. We will discuss this difference in more detail in the Discussion section, and now report the experiment we conducted in detail.

In order to explore the interaction of gender disagreement with semantic expectancy 32 native German speakers participated in a 2 × 2 design study in which gender (correct versus incorrect) and semantic expectancy (high versus low) were crossed. The gender (dis)agreement was between a noun and its immediately preceding definite article (das Land versus *den Land). The semantic expectancy was manipulated using the cloze probability (high or low) for a noun as direct object of a given verb. Table 1 gives an example set of a sentence in the four experimental conditions.

During the word-by-word reading of the sentences ERPs were measured on the critical noun (i.e., ''Land''). If, as suggested by Bates et al. (1996), semantic and syntactic information interact early (prelexically) during word recognition, this should be reflected in the influence of semantic expectancy during an early syntactic

Table 1. Examples of the Four Types of Experimental Sentences Used in the Experiment

Cloze %	*Gender*	*Sentence*
High	congruent	(1) Sie bereist <u>das Land</u> auf einem kräftigen Kamel. She travels the$_{\text{neuter}}$ land$_{\text{neuter}}$ on a strong Camel
High	incongruent	(2) Sie bereist <u>den Land</u> auf einem kräftigen Kamel. She travels the$_{\text{masc}}$ land$_{\text{neuter}}$ on a strong Camel
Low	congruent	(3) Sie befährt <u>das Land</u> mit einem alten Wartburg. he drives the$_{\text{neuter}}$ land$_{\text{neuter}}$) with an old Wartburg car
Low	incongruent	(4) Sie befährt <u>den Land</u> mit einem alten Wartburg. She drives the$_{\text{masc}}$ land$_{\text{neuter}}$) with an old Wartburg car

processing phase (i.e., LAN) and the influence of gender violation on the semantic processing phase (i.e., N400). If, as the prelexical suggestion made by Schriefers et al. (1998) is right, and semantic priming is indeed affected by gender violations, one might expect this to be reflected in an N400 effect for the nouns in the high-cloze condition (i.e., a larger N400 for the gender disagreement). The impact of gender disagreement in the low-cloze condition should be minimal. If, however, the influence of syntactic gender violation on semantic priming is due to a postlexical syntactic check, as proposed as an alternative interpretation by Schriefers et al. (1998), then the N400 should remain unaffected and the interaction should only show up at a later processing stage (i.e., P600). If gender violations behave like other types of morpho-syntactic agreement violation such as the one used in Gunter et al. (1997), one would expect a LAN and a P600 as a function of syntactic congruency, with only the P600 being affected by the semantic expectancy and the N400 as a function of semantic expectancy. The N400 should be independent of gender congruency.

RESULTS

Behavioral Data

Reaction time of the probe-detection task showed a main effect of response indicating that the yes-answers have a faster response than the no-answers (679 versus 715 msec; $F(1,31) = 35.09$, $p < .0001$). A gender main effect showed that the responding during the task was faster when the foregoing sentence was syntactically correct than when it as not (691 versus 718 msec; $F(1,31) = 5.8$, $p < .022$). Both factors also had an interaction showing that the effect of gender is larger in the no-answers (6 versus 19 msec; $F(1,31) = 4.23$, $p < .048$).

Percentage correct only showed a main effect of gender indicating that the probe detection task was carried out a bit more accurately when the foregoing sentence had no gender violation (98.1 versus 97.1%; $F(1,31) = 7.83$, $p < .009$).

In summary, the probe-detection task was carried out very well indicating that the subjects were carefully attending the sentences preceding the task. Although there was a small (i.e., 1%) though significant performance difference for the probe detection after the presentation of a correct sentence or a sentence containing a gender violation, it is too small taking it as evidence for different attending of both types of sentences. Particularly since the percentage correct is almost at a ceiling level.

ERPs

Before discussing the results in a statistical sense, let us preview the data. As can be seen in Figure 1A, high-cloze items gave rise to a smaller N400 than low-cloze items. Interestingly, the N400 activity seems to be extended much longer (up to 950 msec) than is typically found for the N400 in sentence processing. The gender violation showed a LAN/P600 pattern (Figure 1B), which was affected by the cloze probability of a particular sentence. Whereas the P600 elicited by the gender violations diminished in the low-cloze sentences (see Figure 2), the LAN was not affected by the cloze probability.

50-msec Interval Analysis

In order to verify the significance of the ERP effects, the data was quantified using MANOVAs carried out on 50-msec latency windows from stimulus onset until 1000 msec after stimulus onset. As can be seen in Table 2, most electrodes showed a main effect of cloze probability. Generally speaking, the effect started around 300 msec and seems to consist of two parts. The first part, between 300 and approx. 450 msec, represents the N400 component. Several electrodes, however, continue to show a significant main effect of cloze probability up to approx. 800 msec. This later part can best be described as a slow negative shift with a similar distribution as the N400. The main effect of gender consists of two effects with different scalp distributions. The first effect, identified as LAN, was indicated by a significant gender main effect between 350 and 450 msec and seems to have a left frontal-scalp distribution. The second effect, identified as P600, was indicated by a significant gender main effect between approx. 550 and the end of the trial (950 msec). Interaction of cloze probability and gender was mainly found in posterior electrodes in the time frame of the P600 (AF8, C3, Tp7, Cp3, P7, P3, Pz). As can be seen in Figure 2 this interaction has to be interpreted as a large reduction of the P600 component in the low-cloze sentences.

On the basis of this interaction we carried out separate analyses for the high- and low-cloze sentences for the more posterior electrodes in the time region of the P600 component. These analyses showed significant main effects of gender such that in the high-cloze sentences the P600 started around 450 msec and lasted up to the end of the trial (e.g., Pz: 450–950 msec). In the low-cloze sentences, there was still a significant P600 difference but it started to become significant between approx. 700 and 900 msec (cf., Pz). Thus, although there was still some P600 activity in the low cloze items, it was clearly delayed, shorter in duration and smaller in amplitude compared to the P600 in the high-cloze sentences.

Scalp Distributions LAN and N400

In order to explore the scalp distribution of the LAN and N400 we calculated an interval measure of 150

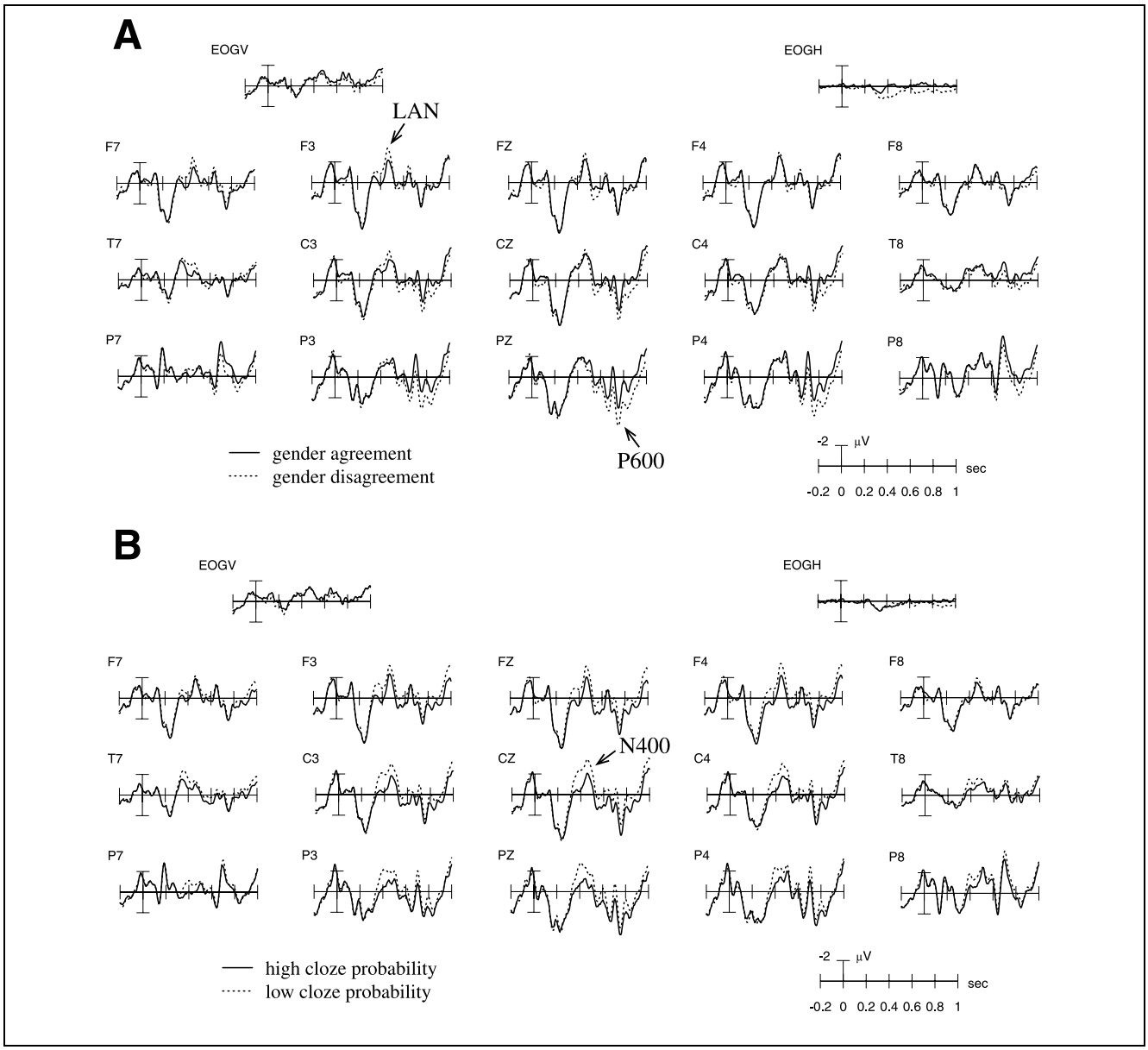


Figure 1. (A) ERPs for gender agreeing (solid line) and gender disagreeing (dashed line) nouns averaged across cloze probability. Gender disagreement elicited a LAN and a P600 component. (B) ERPs for high (solid line) and low (dashed line) cloze probability averaged across gender (dis)agreement. The low cloze nouns show a clear N400 component.

msec long in the latency range between 350 and 450 msec for the LAN and N400 effects. Twenty-three electrode sites were used (i.e., AF7, AF3, AF4, AF8 and Oz were excluded) resulting in an ANOVA with the factors component (LAN versus N400) and electrode (23). The analysis showed significant main effects of component ($F(1,31) = 4.47$, $p < .042$) and electrode ($F(22,682) = 4.06$, $p < .006$, $\varepsilon = .158$). The interaction of component with electrode was significant on the original data and remained significant after applying the McCarthy and Wood (1985) scaling procedure (uncorrected: $F(22,682) = 5.62$, $p < .0001$, $\varepsilon = .24$; corrected: $F(22,682) = 3.54$, $p < .005$, $\varepsilon = .22$). This finding indicated that the LAN and the N400 had different scalp distributions. Visual inspection of the waveforms indicated that the LAN was indeed left anterior distributed whereas the N400 was much more broadly distributed.

DISCUSSION

In this experiment grammatical gender violations elicited a LAN, a component typically found in syntactically incorrect sentences. The effect of cloze probability (i.e., semantic processes) was reflected in the N400. The impact of cloze probability on gender violations was not reflected in the N400 effect but in a P600-effect instead. The P600 elicited by gender disagreement was

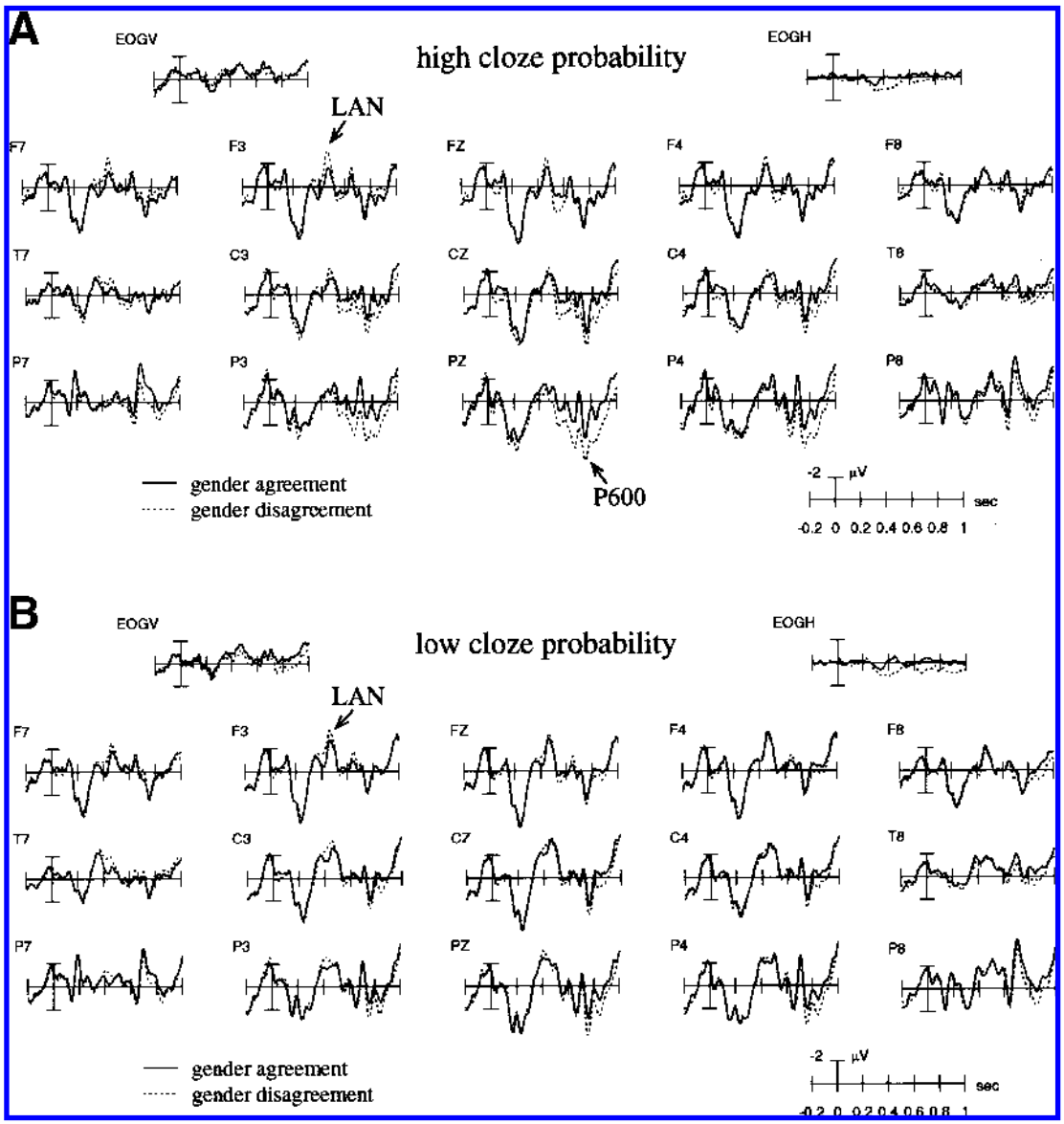
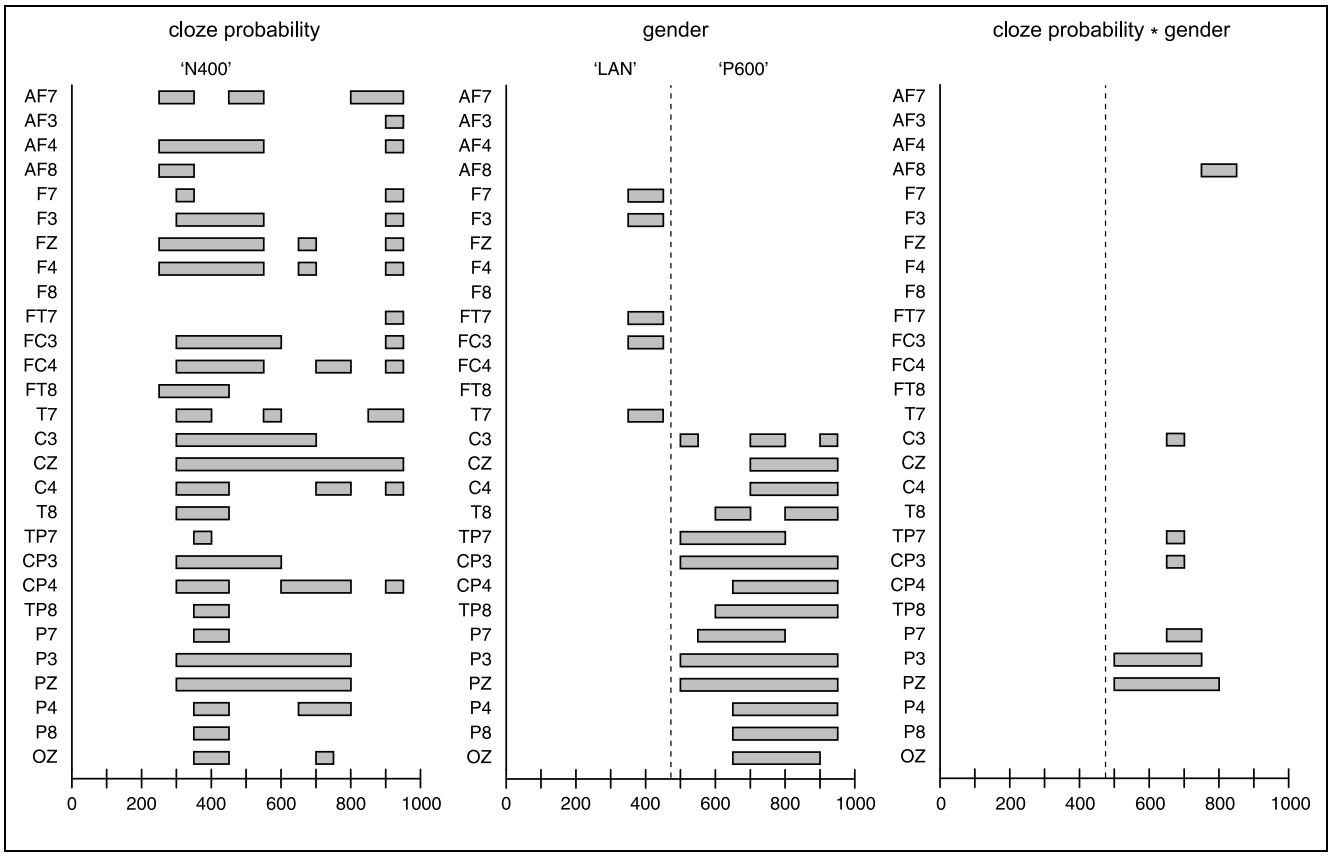


Figure 2. (A) ERPs for gender agreeing (solid line) and gender disagreeing (dashed line) nouns as found for the high cloze probability condition. Gender disagreement elicited a LAN and a P600 component. (B) ERPs for gender agreeing (solid line) and gender disagreeing (dashed line) nouns as found for the low cloze probability condition. Gender disagreement elicited a LAN and a late, small P600 component.

highly reduced and delayed in the condition where semantic expectation was low.[3] Similar findings were reported by Gunter et al. (1997) when crossing verb inflection violations with semantic coherence. From those and the present data, it is clear that during the first part of the comprehension process, semantic and syntactic processes work in parallel and are encapsulated in nature. There is, however, an interaction between both processes in a later time range. Before discussing the general notions suggested by these findings, we will first go into details related to the gender (dis)agreement itself.

The interactive approach (cf., Bates et al., 1996) suggests that gender information given by the article should restrict the activity in a set of potential lexical candidates expected on the basis of the semantic context. When a small set of words are expected, as is the case in a high-cloze condition, the impact of a gender violation on semantic integration processes should be much larger compared to situation where a much larger set of words is selected (low cloze). Thus, the interactive approach would predict N400-effects to be affected by gender disagreement, in particular in the high cloze condition. This is clearly not what happened. Note that these data also seem to contradict the Schriefers et al. (1998) findings, which showed that semantic priming was diminished by gender disagreement. On the basis of their findings

Table 2. Results of the Main Analysis of Variance Tests for Each Electrode



The shaded bars indicate significant ($p > .05$) effects. Effects are shown whenever two or more successive intervals of 50 msec were significant. The left panel shows the main effects of cloze probability, the middle panel shows the main effects of gender (dis)agreement, and the right panel the interaction between both factors.

one would also expect an enlarged N400 elicited by the gender violations in the high-cloze sentences.

The reduction of the P600 in the low-cloze probability situation may seem surprising at first hand. If one considers the P600 to reflect processes of repair or reanalysis, why is the parsing system in this situation not in need of a repair process or, alternatively, does it use a shallower re-analysis process (note that the P600 is not completely absent, there is still some minor activity starting approximately 200 msec later compared to the high-cloze situation). One possible explanation using the mental resource analogy (cf. Kahneman, 1973) is that the comprehension system does not have enough resource capacity to process a difficulty, which is present at both a syntactic and semantic level. This argument seems problematic since one would then expect the LAN to be similarly affected by cloze probability as the P600. One could alternatively argue that the LAN and P600 are energized by different resource pools. It is, however, important to note that caution is in order when using the resource analogy, to make sure it does not become a theoretical ''soup stone'' (cf., Navon, 1984).

A less problematic hypothesis relates to the question of which type of repair process is actually needed to understand the sentences. When the gender-incongruent noun occurs it is clear that in the high-cloze condition, no deep semantic analysis is needed. The word is semantically expected and only a gender agreement repair needs to be done. Thus, meaning is easily present and syntactic gender is repaired. In the low cloze case, however, a more thorough semantic analysis has to be carried out to achieve a semantic integration of the noun with the preceding context. The finding that the P600 is clearly reduced and heavily delayed may indicate that the subjects are analyzing the sentences such that their primary goal is to extract meaning. Why repair a relative minor error when hard work on the semantics is necessary? It is known from several studies that the P600 is under strategical control (Gunter & Friederici, 1999; Hahne & Friederici, 1999). It is therefore possible that what we see here is a strategy of the comprehension system to focus on meaning.

The foregoing discussion presupposes hypotheses that the P600 reflects repair processes (cf., Gunter et al., 1997; Friederici, 1995) and the N400 reflects semantic integration (Brown & Hagoort, 1993). One could, however, go one step beyond these concepts and argue that both the N400 and P600 reflect similar kind of integration processes each done on different type of information, semantic and syntactic information,

respectively.[4] The association of the P600 with integration processes was also put forward by Kaan, Harris, Gibson, and Holcomb (1998) who claim that the P600 reflects difficulty with syntax integration processes in general. They found the P600 to vary as a function of the difficulty of syntactic integration in correct nongarden-path sentences (see also Osterhout, Holcomb, & Swinney, 1995).

The observed LAN-P600 pattern on syntactic gender violations seems partly in accordance with the semantic gender studies of Osterhout et al. (1995, 1997). In these studies, only a stand-alone P600 was reported for semantic gender violations. The explanation for the seemingly contradiction is relatively straightforward. In contrast to syntactic gender, semantic gender needs the meaning of the word before a gender violation can be detected. Thus, a lot of semantic processing is done before detection, making the reanalysis or integration of this violation a relatively late process (i.e., later as the N400 components). The detection of a syntactic gender violation, however, can be accomplished much earlier in time since it is based on information present in the lexicon itself and does not need an elaborated semantic inference process. Thus, when a noun is read, syntactic (gender) information is automatically extracted out of the lexicon making a much earlier impact of the gender violation possible compared to Osterhout's violation where the violation detection was performed via an elaborative semantic analysis.

The observed ERP-pattern is in clear contrast to the Hagoort and Brown (1997) study reporting an N400-enhancement when semantic anomalies are combined with an additional gender violation. Using the N400 window as baseline, they also suggest that the P600 is not additionally affected by the double anomaly. Thus, they claim that the semantic processes are open for syntactic influences whereas the opposite is not true. This suggestion is clearly in conflict with our data which indicate that semantic processes influence syntactic processes (reduced P600 in low cloze) and are in the first instance independent of syntactic processes (LAN versus N400). Note that with respect to the negativity it is very well possible that the enhanced N400 of Hagoort and Brown (1997) is just an overlapping LAN component. Since no information other than the central electrode was given in the report, this possibility cannot be excluded and is even probable if one looks more carefully at the present data. As can be seen in Figure 2, our data also seems to have a slight N400 enhancement (see Cz or Pz). If one looks, however, at the left-frontal electrodes (cf., F3 or F7) it is clear that we are dealing with the differently distributed LAN-component overlapping the N400. With respect to the P600, there seem to be a baseline problem present in the Hagoort and Brown (1997) data. Disregarding this latter problem for the moment, it seems that there are at least three differences between both experiments, which might explain the discrepancies observed.[5] First, in our study we used low cloze items and not outright semantic violations. However, a replication of the present study using outright semantic violations showed a similar pattern of results as found in the present study, i.e., an interaction of semantic and syntax at the P600 level (Gunter et al., in preparation). Therefore, it seems that this factor can be ruled out. The second difference is that the two experiments used different tasks. Whereas the present experiment used a probe detection task, Hagoort and Brown used an acceptability judgment task. It is known that tasks can influence both the N400 and P600 components (cf., N400: Chwilla et al., 1995; N400 and P600: Gunter & Friederici, 1999). It is therefore possible that participants in each experiment had a different strategy for resolving the double violation. The last and probably more interesting difference between the experiments is that the gender violation in the present experiment is more local (article and noun are directly adjecent) than the one of Hagoort and Brown (1997) where an adjective was presented between the article and the noun. The issue of local versus global gender violations is currently under study by our group.

At first glance, the presented data seems to contradict the more serial models of Friederici (1995) and Frazier (1987). These models assume that the first-pass parsing is strictly syntactic in nature. However, both models clearly restrict the first-pass parse to syntactic word category information. Other types of lexical information such as gender or verb-argument information are assumed not to influence the first pass parsing process. The present data show semantic and syntactic processing to run in parallel and not serially. Because the word category of the noun is correct, the first-pass parse has already taken place. It is therefore not surprising to see the electrophysiological reflection of the disagreement for the ''secondary'' gender feature in parallel with the reflection of semantics. In line with this argument, it interesting to note that a double violation consisting of a semantic violation and a word-category violation demonstrated not only serial effects with phrase structure building (early left-anterior negativity) preceding semantic processing (N400) but, moreover, an influence of phrase structure correctness on lexical–semantic integration as reflected by the N400 (Friederici, Steinhauer, & Frisch, 1999; Hahne & Friederici, 1998). The present data indicate that gender information and word meaning are available at the same time, either directly from the lexicon or from an internal buffer, and that these information types are processed in parallel.

To sum up, this experiment has shown that syntactic and semantic processes are encapsulated and run in parallel during a first phase. Interactions between syntax and semantics are present at a later phase. It is clear that the present data are in disagreement with strong interactive models, which would not predict independent early syntactic and semantic processes, but interactions

already early on. Rather they support models, which take syntactic parsing to be autonomous, at least during an early automatic stage, and allow semantic information to come into play only at a later stage. High-temporal resolution of ERP-data seem to provide an explanation for behavioral comprehension studies which report findings in favor of both autonomous models and interactive models. It may well be that the different studies have tapped into different phases of the comprehension process, those which found evidence for autonomy into the early phase and those which found evidence for interaction into the late phase.

METHODS

Participants

Thirty-two native German speaking students (16 female, mean 23.7 years, age range 19 to 28 years) were paid approximately 50 DM for their participation. All participants were right-handed and had normal or corrected-to-normal vision. One of the participants had left-handed relatives.

Stimuli

A total of 160 experimental sentences were constructed containing eight words. The experimental sentences were manipulated in such a way that half of them were low cloze (< 25%, mean: 15.3%; see Taylor, 1953) and the other half were high cloze (> 56%, mean: 74.2%). Half of the experimental sentences had a gender violation (i.e., gender mismatch between article and noun). Thus, a 2 (high/low cloze) × 2 (gender violation/correct) design was used.

Before the actual experimental sentences were chosen, a rating study was carried out in which 25 students had to complete 210 sentences, which contained a personal pronoun, a verb and the neuter article (i.e., Er/Sie trinkt das ... He/She drinks the$_{neuter}$ ...; there were actually 5*25 students who completed a list of 42 items). From this list the 40 sentences were chosen with the highest expectancy for a particular noun. For each of these sentences, a low-cloze counterpart sentence was chosen such that the noun of the high-cloze sentence was semantically acceptable (but not expected in any way). In order to dampen the repetition effect of the noun, several actions were taken. The stimuli were presented in four blocks of 40 experimental sentences each. In each block, a different version of the experimental sentences were used. Blocks 1 and 2 contained either sentence type 1 and 4 or 2 and 3 (see Table 1; type 1: high cloze/congruent gender; type 2: high cloze/inconguent gender; type 3: low cloze/congruent gender; type 4: low cloze/inconguent gender). After at least 1 week, the subjects got blocks 3 and 4 which were the mirror image of block 1 and 2 (if for instance sentence types 2 and 3 were already read, subjects would get sentence types 1 and 4 and the other way around). Thus, no repetition of a particular sentence occurred within one session. The initial type of a particular sentence was pseudo-randomized making sure that all conditions were equally frequent within a block. Four versions were made of each block which permutated all four sentence types. Each of the versions was administered to a total of four subjects. Because Besson and Kutas (1993) have shown that the repetition effect of a word when it is presented in a different sentence context is very small to nonexisting, we can be sure that the influence of repetition is marginal in our design. Because the masculine gender article (den) was always the incorrect one and the neuter article (das) always the correct one in the experimental sentences, filler items of a middle-cloze probability were added in which the mapping between gender and correctness was reversed (i.e., ''das'' is incorrect and ''den'' is correct). In order not to have the violation at the end of the sentence, the critical noun was followed by four words. Thus, each subject received two sessions in which a total of 160 experimental and 160 filler sentences were presented.

All sentences were presented word-by-word (300 msec per word with a blank screen of 200 msec between the words). The first word of a sentence was preceded by a fixation cross of 300 msec. After the presentation of the sentence ending (period included) a blank screen (700 msec) and a probe word (5000 msec maximal) were presented. Subjects were instructed to decide as accurately and as quickly as possible whether the probe word occurred in the preceding sentence or not. On the basis of this response a feedback stimulus was presented for 700 msec immediately after the response (i.e., correct, incorrect, or faster if the response was not within 5000 msec) After the presentation of a blank screen of 300 msec, the next trial started with the presentation of the fixation cross. Word length was always kept within 2° of the visual field. The words, which had a height of approximately .4°, were presented in black on a gray background.

Presentation Sequence

The experiment was conducted in two sessions with a time interval of 1 week between sessions. In each session, subjects received a small training block followed by two blocks of 40 sentences each. Between the blocks, a break of 15 min was given in which subjects were allowed to read a magazine of their choice. None of the sentences used in the training were experimental sentences.

Procedure

A session lasted approximately 3 hr. Participants were seated in a dimly lit room, facing a color video screen

at a distance of 100 cm. When necessary, instructions were given to blink only when the fixation cross-appeared.

Recordings

The EEG was recorded with 28 tin electrodes (electrocap) from AF7, AF3, AF4, AF8, F7, F3, Fz, F4, F8, FT7, FC3, FC4, FT8, T7, C3, Cz, C4, T8, Tp7, CP3, CP4, TP8, P7, P3, Pz, P4, P8, Oz and left mastoid each referred to the right mastoid (nomenclature as proposed by the American Electroencephalographic Society, 1991). Bipolar horizontal EOG was recorded between electrodes at the outer left and right canthus. Bipolar vertical EOG was recorded between electrodes above and below the subject's right eye. Electrode resistance was kept under 5 kΩ. The signals were recorded continuously with a bandpass between DC and 70 Hz and digitized at a rate of 250 Hz.

Data Analysis

Average ERPs, starting 200 msec before and lasting 1000 msec after the presentation of the article and its following noun were computed for each electrode position for each of the four conditions (cloze % [high versus low] and gender [agreement versus disagreement]; see above).

Approximately 6% of the trials were excluded from the averages due to ocular—and amplifier saturation artifacts (EOG rejection +/− 50 μV). Averages were aligned to a 200-msec prestimulus baseline. In order to describe the onsets and length of the ERP effects in reasonable detail, an analysis was carried out in which the data were statistically evaluated using 20 latency-windows, which had a length of 50 msec each. The interval names were chosen in such a way that they indicate the onset of the latency window (e.g., interval ''50'' represents the mean amplitude of sample points 50–100 msec; etc.).

All dependent variables were quantified using repeated measure ANOVAs. Within subject factors were cloze probability (high/low) and gender (correct/violation). In some analyses electrode (23) was also included. The analyses of the 50-msec interval data were carried out for each electrode position separately. Whenever significant effects in a particular range are reported, all 50-msec intervals in this range showed significant effects ($p < .05$). To minimize the danger of false positives due to the large number of comparisons, effects will only be presented when two or more successive 50-msec intervals showed these effects. The Geisser–Greenhouse correction (Vasey & Thayer, 1987; Geisser & Greenhouse, 1959) was always applied when evaluating effects with more than one degree of freedom in the numerator.

The analysis of the behavioral data included the factors cloze probability (high/low), gender (correct/incorrect) and response type (yes/no).

Acknowledgments

We wish to thank Frank Meyer for his technical support; Silke Urban and Suse Wagner for preparing the stimuli and Cornelia Schmidt for the data acquisition. We are indebted to Trevor Penney, Silke Urban and Sandra Vos for helpful comments and corrections on earlier drafts of this manuscript.

Reprint requests should be sent to Thomas C. Gunter at the Max Plank Institute of Cognitive Neuroscience at Stephanstrasse 1a D-04103 Leipzig, Germany, e-mail: gunter@cns.mpg.de.

Notes

1. Note that within a language, the gender information of words can be very transparent. In Spanish, for instance, gender is unambiguously phonologically coded. Although there are certain semantic, phonological and morphological probabilistic regularities which attribute gender to words in the German language (cf., Köpke & Zubin, 1996), gender is typically considered arbitrary (Fodor, 1959).
2. Relative to a semantically unrelated condition like ''Er sieht das Bier'' (He sees the$_{neuter}$ beer$_{neuter}$).
3. One might worry that due to component overlap between N400 and P600 this interaction between semantics and syntax is present. Because the N400 is similar for both the gender-correct and gender-violation condition, this overlap, if one exists, is similarly present in both syntactic condition. We therefore consider this explanation of the data not a realistic one.
4. One could speculate for instance that a word needs a particular amount of integration in a sentence before the sentence can be understood. It does not matter where the problem in integrating is present, if it has its emphasis on semantics one will find more N400 activity, if it is more on syntax one will find more P600 activity.
5. Note that the gender systems of Dutch and German are highly similar. In Dutch, one has two genders: Common gender (''de''; is a merge of the historical masculine and feminine genders) and Neuter gender (''het''). German has three genders: Masculine (''der''), Feminine (''die''), and Neuter (''das''). A clear difference between the two languages is that articles in German contain much more information (syntactic gender, case marking, singular/plural distinction) than articles in Dutch (only syntactic gender and singular/plural distinction). The case marking information as found in German is in Dutch typically represented by word order restrictions. Note, however, that in the stimulus material used in the present experiment only the factor gender was manipulated and none of the other factors!

REFERENCES

American Electroencephalographic Society. (1991). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 8,* 200–202.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception and Psychophysics, 58,* 992–1004.

Bates, E. & MacWhinney, B. (1989). Functionalism and the Competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Cambridge, UK: Cambridge University Press.

Besson, M. & Kutas, M. (1993). The many facets of repetition: A cued-recall and event-related potential analysis of repeating words in same versus different sentence contexts. *Journal of Experimental Psychology: Learning, Memory and Cognition, 5,* 1115–1133.

Brown, C. M. & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5,* 34–44.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32,* 274–285.

Corbett, G. G. (1991). *Gender*. Cambridge, UK: Cambridge University Press.

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

Desrochers, A. (1986). Genre grammatical et classification nominale. *Canadian Journal of Psychology, 40,* 224–250.

Fodor, I. (1959). The origin of grammatical gender. *Lingua, 8,*1–41.

Fodor, J. A. (1983). *Modularity of mind.* Cambridge: MIT Press.

Fodor, J. A. (1988). On modularity in grammar. *Journal of Psycholinguistic Research, 17,* 125–168.

Frazier, L. (1987). Sentence processing. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 49,* 259–281.

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

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Even-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1,* 183–192.

Friederici, A. D. & Schriefers, H. (1994). The nature of semantic and morphosyntactic context effects on word recognition in young healthy and aphasic adults. In D.H. Hillert (Ed.), *Linguistics and cognitive neuroscience: theoretical and empirical studies on language disorders* (pp. 9–32). Opladen: Westdeutscher Verlag.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory and Cognition, 27*, 438–453.

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

Gunter, Th. C. & Friederici, A. D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology, 36,* 126–137.

Gunter, Th. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34,* 660–676.

Hagoort, P. & Brown, C. (1997). *When syntax and semantics meet*. Poster presented at the Fourth Annual Meeting of the Cognitive Neuroscience Society.

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.

Hahne, A. & Friederici, A. D. (1997). Two stages in parsing: Early automatic and late controlled processes. *Experimental Brain Research, 117,* 47.

Hahne, A. & Friederici, A. D. (1998). *ERP-evidence for autonomous first-pass parsing processes in auditory language comprehension*. Poster presented at the Fifth Annual Meeting of the Cognitive Neuroscience Society.

Hahne, A. & Friederici, A. D. (1999). Electrophysiological evidencence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 194–205.

Kaan, E, Harris, A., Gibson, E., & Holcomb, P. (1998). *The P600 as an index of syntactic integration difficulty*. Paper presented at the AMLap-98 Conference.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kluender, R. & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5,* 196–214.

Kluender, R., Münte, T. F., Wind Cowles, H., Szentkuti, A., Walenski, M., & Wieringa, B. M. (1998). *Brain potentials to Englisch and German questions*. Poster presented at the CNS Fifth Annual Meeting.

Köpke, K. M. & Zubin, D. A. (1996). Prinzipien für die Genuszuweisung im Deutschen. In E. Lang & G. Zifonun (Eds.), *Deutsch-typologisch.* Berlin: Walter de Gruyter.

Kutas, M. & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11,* 99–116.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation.* Cambridge: MIT Press.

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

McCarthy, G. & Wood, C. C. (1985). Scalp distribution of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 69,* 218–233.

McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4,* 287–335.

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

Münte, T. F. & Heinze, H. (1994). ERP negativities during syntactic processing of written words. In H. J. Heinze, T. F. Münte & G. R. Mangun (Eds.), *Cognitive electrophysiology* (pp. 211–238). La Jolla, CA: Birkhauser Boston.

Münte, T. F., Heinze, H., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language, *Journal of Cognitive Neuroscience, 5,* 335–344.

Münte, T. F., Heinze, H., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36,* 217–226.

Münte, T. F., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience, 9,* 318–329.

Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review, 91,* 216–234.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3,* 151–165.

Osterhout, L., Bersick, M., & McLaughling, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition, 25,* 273–285.

Osterhout, L. & Hagoort, P. (1999). A superficial resemblance doesn't necessarily mean that you're part of a family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 1,* 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. (1995). 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.

Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience, 8,* 507–526.

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

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T. F., & Clahsen, A. (1997). How the brain processes complex words: An ERP-study of German verb inflections. *Cognitive Brain Research, 6,* 37–52.

Reagan, D. (1989). *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine.* New York: Elsevier.

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

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind: Event-related potentials and cognition*. New York: Oxford University Press.

Schriefers, H., Friederici, A. D., & Rose, U. (1998). Context effects in visual word recognition: Lexical relatedness and syntactic context. *Memory and Cognition, 26,* 1292–1303.

Taylor, W. L. (1953). ''Cloze'' procedure: A new tool for measuring readability. *Journalism Quarterly, 30,* 415–417.

Van Berkum, J. J. A. (1996). *The psycholinguistics of grammatical gender*. Doctoral thesis published in Nijmegen: Nijmegen University Press.

Van Petten, C. (1995). Words and sentences: Event-related brain potential measures. *Psychophysiology, 32,* 511–525.

Vasey, M. W. & Thayer, J. F., (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: A multivariate solution. *Psychophysiology, 4,* 479–486.

Vos, S. H., Gunter, Th. C., Kolk, H. H. J., & Mulder, G. (submitted). *Working memory constraints on syntactic processing: An electrophysiological investigation*.