Journal of Psycholinguistic Research, Vol. 29, No. 1, 2000

Discourse Before Gender: An Event-Related Brain Potential Study on the Interplay of Semantic and Syntactic Information During Spoken Language Understanding

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

A study is presented on the effects of discourse–semantic and lexical–syntactic information during spoken sentence processing. Event-related brain potentials (ERPs) were registered while subjects listened to discourses that ended in a sentence with a temporary syntactic ambiguity. The prior discourse–semantic information biased toward one analysis of the temporary ambiguity, whereas the lexical-syntactic information allowed only for the alternative analysis. The ERP results show that discourse–semantic information can momentarily take precedence over syntactic information, even if this violates grammatical gender agreement rules.

INTRODUCTION

Understanding spoken language is a balancing act between representational complexity and processing speed. The complexity lies in the many and diverse sources of linguistic and nonlinguistic knowledge that are embodied in the spoken signal and its discourse context (cf. Jackendoff, 1999). The speed follows from the speech rate, which is some two to four words per second

We thank Janet Fodor for her comments on the paper, Edith Sjoerdsma for recording and Petra van Alphen for labeling the materials, Valesca Kooijman, Ellen de Bruijn, en Jelle van Dijk for help in acquiring and analyzing the ERP data, and René de Bruin, Inge Doehring, John Nagengast, and Johan Weustink for technical assistance. Part of this work was reported at the 12th Annual CUNY Conference on Human Sentence Processing, New York, March 18–20, 1999. This research was supported by a grant to Jos van Berkum from the Deutsche Forschungsgemeinschaft, and by Grant 400-56-384 from the Netherlands Organization for Scientific Research, to Colin Brown and Peter Hagoort.

[1] Neurocognition of language processing Research Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

[2] Author to whom all correspondence should be sent. E-mail: Colin.Brown@mpi.nl.

0090-6905/00/0100-0053$18.00/0

in normal conversational settings (cf. Levelt, 1999). Both the diversity of knowledge sources and the stimulation rate pose particular challenges to the listener. Within a fraction of a second, the phonological, syntactic, and semantic information associated with each word in the speech stream has to be activated and evaluated in the light of the preceding lexical, sentential, and discourse information. This high-speed and complex evaluation process is a core feature of human language understanding, and the attempt to determine its functional nature is one of the central themes of sentence-processing research.

In this paper, we focus on the real-time evaluation process by investigating the interplay of discourse–semantic and lexical–syntactic information during spoken language comprehension, by means of the registration of event-related brain potentials (ERPs). The experimental approach is to present subjects with spoken discourses that end with a sentence that is temporarily syntactically ambiguous and to pit two opposing forces against each other in terms of how the ambiguity might be resolved: a discourse-semantic and a lexical–syntactic factor. What is at stake is the relative processing impact of the two factors and the time course of this impact. Does one factor take primacy over the other? At what moment(s) in time do the two factors come into play? These questions lie at the heart of a continuing controversy on the basic processing nature of the sentence processor. Broadly speaking, two classes of model can be distinguished. On the one hand, models that assign primacy to structural information and that do not allow discourse–level information to affect the initial parsing process (e.g., Frazier & Rayner, 1982; Frazier & Clifton, 1996; Ferreira & Clifton, 1986; Mitchell, Cuetos, Corley, & Brysbaert, 1995). On the other hand, models that allow, in essence, any kind of information to affect the comprehension process at any time (e.g., Crain & Steedman, 1985; MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995). Note that a critical difference between the two classes of model is the relative moment in time at which they posit processing consequences of different kinds of information. All models agree that all sources of relevant information can be brought to bear on the comprehension process. The contentious issue concerns the temporal dynamics of the process.

The temporary syntactic ambiguity we used is the, by now, classic ambiguity between a complement and a relative clause, as in "*David told the girl that . . .*". This sentence fragment can be continued as either a complement clause (e.g., "*David told the girl that there would be some visitors*"), or a relative clause (e.g., "*David told the girl that had been on the phone to hang up*"). In the critical target sentences for which we will present the ERP data, the ambiguity was always resolved as a complement clause. In Dutch, the language used in the experiment, the complementizer *that* takes a single form, namely *dat*. The same word form can also be used as a relative pro-

noun (but only for a certain class of nouns, see below). It is this lexical ambiguity that gives rise to the complement/relative-clause ambiguity. It has already been established in several languages, including Dutch (cf. van Berkum, Brown, & Hagoort, 1999a), that in the absence of constraining information a complement-clause analysis is preferred over a relative-clause analysis. In the ERP waveform, the processing consequences of this preference difference emerge as a positive shift elicited by the word that necessitates a relative-clause continuation (e.g., the verb *had* in the above example sentence). This shift starts at about 500 ms after the onset of the disambiguating word, has been observed to reach its maximal amplitude at about 600 ms (but the morphology of the component is not always markedly peaked), and spans several hundreds of milliseconds. In the psycholinguistic literature this ERP effect is labeled the P600/SPS (Syntactic Positive Shift) effect, and is claimed to reflect syntactic (re)analysis processes (cf. Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992). Basically the P600/SPS is elicited any time the parser runs into syntactic trouble, either from an outright grammatical error, or from a nonpreferred structure (see Hagoort, Brown, & Osterhout, 1999 for a review of the literature). In the research we report here, we used the P600/SPS as a diagnostic tool with which to assess the impact of semantic and syntactic information on the resolution of the complement/relative-clause ambiguity. By manipulating both discourse-semantic and lexical-syntactic constraints on complement/relative-clause structures, we investigated whether and when these constraints affected the on-line parsing process, as indicated by the presence or absence of the P600/SPS.

In the experiment, discourse information was manipulated by the presence in the prior discourse of either one or two referents for the noun that immediately preceded the point of temporary ambiguity (cf. Altmann & Steedman, 1988; Crain & Steedman, 1985). Lexical-syntactic information was varied by the syntactic gender of this preceding noun, that was either of neuter or common gender. In Dutch this has consequences for the interpretation of the ambiguous word *dat.* Following neuter-gender nouns, *dat* can be both a complementizer and a relative pronoun and, therefore, gives rise to a complement/relative-clause ambiguity. Following common-gender nouns, however, *dat* can only be a complementizer. The following example illustrates both the discourse and the lexical manipulation (the contexts are given for a subsequent neuter-gender target sentence; the alternative common-gender antecedent is shown in brackets):

1-referent context

David had told the boy and the girl (the woman) 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.

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

neuter-gender target sentence
(1) *David vertelde het meisje dat er visite kwam.*
David told the $girl_{NEU}$ that there would be some visitors.

common-gender target sentence
(2) *David vertelde de vrouw dat er visite kwam.*
David told the $woman_{COM}$ that there would be some visitors.

Note that the target sentences only differ in the syntactic gender of the noun. In the neuter-gender target sentence (1), the form ambiguity of *dat*, together with the syntactic gender of its preceding noun (i.e., *meisje*), creates a complement/relative-clause ambiguity. The ambiguity is resolved as a complement clause on the immediately following word (i.e., *er*). In the common-gender target sentence (2), in principle no ambiguity arises because the syntactic gender of the noun (i.e., *vrouw*) rules out a relative-pronoun interpretation of the word *dat.* For common-gender nouns, the relative pronoun takes another form (namely *die*). In other words, looking at the discourse-final sentences in isolation, without taking their prior discourse into account, there is a lexical-syntactic factor that creates a complement/relative-clause ambiguity for sentence (1), and that rules out any ambiguity in sentence (2), where a complement-clause analysis is the only possible structure.

However, now consider the possible impact of the discourse factor. In a 1-referent context, the discourse provides a unique referent for the noun phrase in the target sentence (*het meisje* or *de vrouw*). This context therefore favors a complement-clause continuation, because the noun phrase preceding *dat* does not require further modification by a restrictive relative clause. This is not the case in a 2-referent context. Here there are two plausible referents (the two girls or the two women) for the noun phrase. Therefore, this context favors a relative-clause continuation, in which referential restrictions are provided that can indicate which of the two antecedents is being talked about (cf., Crain & Steedman, 1985). Can this discourse information override the default complement-clause preference and bias for a relative clause? Our previous work, using visual presentation of the target sentences, demonstrates that this can indeed occur (van Berkum *et al.*, 1999a). In this research, we showed that a word that disambiguates for a complement-clause reading [such as the word *er* in sentence (1) above], elicited a P600/SPS when preceded by a 2-referent context, and not when preceded by a 1-referent con-

text. This finding implies that the 2-referent context lured the parser down a garden path. What happens when the discourse and the lexical information are in conflict? This is the case in the 2-referent context followed by sentence (2). When the word *dat* is encountered, the syntactic agreement rules of Dutch exclude a relative-pronoun reading, because of the preceding common-gender noun. Nevertheless, in our previous work, we observed a P600/SPS elicited by the word *dat* in sentence (2) in the 2-referent context, clearly indicating that initially the parser did pursue a relative-clause analysis, momentarily ignoring the grammatical rules of Dutch (van Berkum, Brown, & Hagoort, 1999b).

The present study further investigates the two basic processing effects that emerged from the visual study. The main purpose is to extend the findings from the visual domain into the auditory domain. This is important for several reasons. The first is primarily methodological and relates to concerns about whether the manner of visual stimulation in ERP experiments might induce artifactual effects. These concerns arise because sentences are presented one word at a time in the center of fixation, to minimize contamination of the ERP signal by the electrical activity associated with eye movements. Moreover, words are usually presented at a quite slow rate of some 500 to 600 ms per word to avoid analysis problems because of the temporal overlap of the brain potentials elicited by individual words. Although there are already indications in the literature that these presentation parameters do not lead to artifactual effects (e.g., Brown & Hagoort, 1999; Hagoort & Brown, 1999; Kutas, 1987), the present study provides a strong test of possible modality and presentation effects by investigating whether established visual P600/SPS effects will also obtain with spoken language stimulation. A second reason for using spoken language is that speech is clearly the primary mode for comprehension. Any sentence processing theory needs to be put to the test in the auditory modality. Nevertheless, to date the overwhelming majority of studies has used, and continues to use, written sentences. Third, some of the most critical differences between competing parsing models relate to the temporal dynamics of the comprehension process. Given that the millisecond-by-millisecond time course of visual lexical processing is as yet unknown, the clearly defined temporal nature of the speech signal might present a better testing ground. Any such endeavor, however, is predicated on the successful elicitation during spoken language stimulation of standard language-related ERP effects, such as the P600/SPS. An interesting additional issue arises here in relation to the time course of language-related brain potentials. Will this be comparable between the modalities? More in general, to what extent are the temporal dynamics of the operations of the language system dependent upon the input signal?

METHOD

Apart from the use of fully spoken materials and different subjects, the method of this experiment is identical to that of its written-language counterpart (van Berkum *et al.*, 1999a, b).

Subjects

We recruited 24 right-handed native speakers of Dutch (19 female, mean age 23, range 19–36 years) from our local subject pool. None had any neurological impairment, had experienced any neurological trauma, or had used neuroleptics.

Material

Every subject listened to 240 three-sentence ministories. We describe the essential features of these stories below and refer to van Berkum *et al.* (1999a) for details. The full set of Dutch materials is available from the authors.

The 120 critical stories ended with a sentence that contained a complement clause headed by *dat* (*that*), such as *. . . dat er visite kwam* in examples (1) and (2) above. In 60 neuter-gender sentences, the noun that immediately preceded this complement clause had neuter-syntactic gender, e.g., *meisje* (*girl*$_{NEU}$). In combination with the form ambiguity of *dat,* this created a complement/relative-clause ambiguity. In all of the neuter-gender sentences, this temporary syntactic ambiguity was resolved in favor of the complement-clause reading by the word immediately following *dat,* which was either the expletive pronoun *er* (as in the example above), or a personal pronoun alternative (see van Berkum *et al.*, 1999a, for details on the latter).

In 60 common-gender sentences, the preceding noun had common syntactic gender, e.g., *vrouw* (*woman*$_{COM}$). Because the correct relative pronoun for a common-gender noun is *die,* not *dat,* this common-gender antecedent formally ruled out the relative-clause reading at the word *dat.*

The remaining 120 stories had a final sentence containing a relative clause, so that the materials would not introduce a complement-clause bias. Half of these relative clauses modified a neuter-noun antecedent (e.g., *David told the girl that had been on the phone to hang up*), the other half modified a common-noun antecedent (e.g., *David told the woman that had been on the phone to hang up*).

For each of the target sentences we created two discourse contexts, which only differed in whether they introduced 1 or 2 referents for the singular noun in the target sentence (e.g., a girl and a boy, or two girls, preceding *het meisje*). In formulating these contexts, we took great care to avoid

foregrounding one candidate referent at the expense of the other one (see van Berkum *et al.*, 1999a, for details).

The discourse contexts had already been recorded for the preceding written-language experiment (where only the target sentences had been presented in written form) and we used these recordings again. They had been read with a normal speaking rate and intonation by a female native speaker. For every item, the two alternative versions of a discourse context had been recorded consecutively in a random order, after which they were sampled at 16 kHz mono for digital speech file storage.

The target sentences were recorded in a separate session, using the same speaker and recording conditions. They were read with a normal speaking rate, and the speaker was instructed to use a natural intonation that was neutral between a complement-clause or relative-clause continuation in the ambiguous region of the sentence (Dutch prosody does allow for such a neutral intonation). A native listener monitored all sentences for biasing prosody and problematic recordings were redone. To time-lock relevant portions of the EEG signal to the exact occurrence of the critical disambiguating words within their carrier sentences (*er* in neuter-gender sentences, *dat* in common-gender sentences), a trained phonetician identified the onset latency of each of these critical words (CWs), using standard speech processing software.

We used two different trial lists, identical to those used in the written-language study. For the first, half of the neuter-gender and half of the common-gender target sentences (30 sentences each) were paired with a 1-referent discourse context; the remaining sentences were paired with a 2-referent context. The resulting 120 critical stories were randomly mixed with 120 comparable stories in which the final sentence contained a relative clause. The second trial list was derived from the first by exchanging 1- and 2-referent discourse contexts only. Half of the subjects were tested with the first list, the remaining were tested with the second list.

Procedure

After electrode application, subjects sat in a sound-attenuating booth and listened to the stimuli over headphones. They were told that, "for technical reasons," EEG recording would only occur as they heard the last sentence of a story and that during recording they should avoid all movement and fixate on an asterisk displayed on the screen before them. Subjects were asked to process each story for comprehension. No additional task demands were imposed.

Each trial consisted of a 300 ms auditory warning tone, followed by 700 ms of silence, the spoken discourse context, 1000 ms of silence, and the

spoken final sentence. The 1000 ms separating the final sentence from its context did not perceptually break the story into two parts and approximated natural pause times measured between sentences within a context. To inform subjects when to fixate and sit still for EEG recording, an asterisk was displayed from 1000 ms before onset of the target sentence to 1000 ms after its offset. After a short practice, the trials were presented in five blocks of 15 min, separated by rest periods.

EEG Recording

The EEG was recorded from 13 tin electrodes in an electrode cap, each referred to the left mastoid. Three electrodes were placed according to the international 10–20 system over midline sites at Fz, Cz, and Pz locations. Ten electrodes were placed laterally over symmetrical positions: left and right frontal (F7, F8), anterior temporal (LAT, RAT, halfway between F7-T3 and F8-T4, respectively), temporal (LT, RT, laterally to Cz, at 33% of the interaural distance), temporoparietal (LTP, RTP, posterior to Cz by 13% of the nasion–inion distance, and laterally by 30% of the interaural distance each), and occipital (LO, RO, halfway between T5-O1 and T6-O2, respectively). Additional leads were used to monitor for eye movements, blinks, and activity on the reference electrode (see van Berkum *et al.*, 1999a, for details). The EEG and EOG recordings were amplified and filtered with Nihon Kohden AB-601G bioelectric amplifiers (hi-cut 30 Hz, time constant 8 s), digitized on-line with a sampling frequency of 200 Hz, and screened off-line for eye movements, muscle artifacts, electrode drifting, and amplifier blocking in a critical window that ranged from 150 ms before onset of the target sentence noun to 1200 ms after onset of the critical word. Trials containing such artifacts were rejected (12.2%).

RESULTS

Prior to statistical analysis, the waveforms were normalized on the basis of the averaged activity 150 ms preceding the critical disambiguating word. Statistical analyses were done separately for the neuter-gender and the common-gender sentences. Figure 1A displays the average event-related brain potentials time-locked to the complement-disambiguating word *er* (or a personal pronoun alternative) in neuter-gender sentences presented in a 1- or 2-referent context, for three representative electrode sites along the midline of the scalp.

Inherent to the use of fully connected speech input, the N1 and other early ERP components that are observable with isolated auditory (or visual)

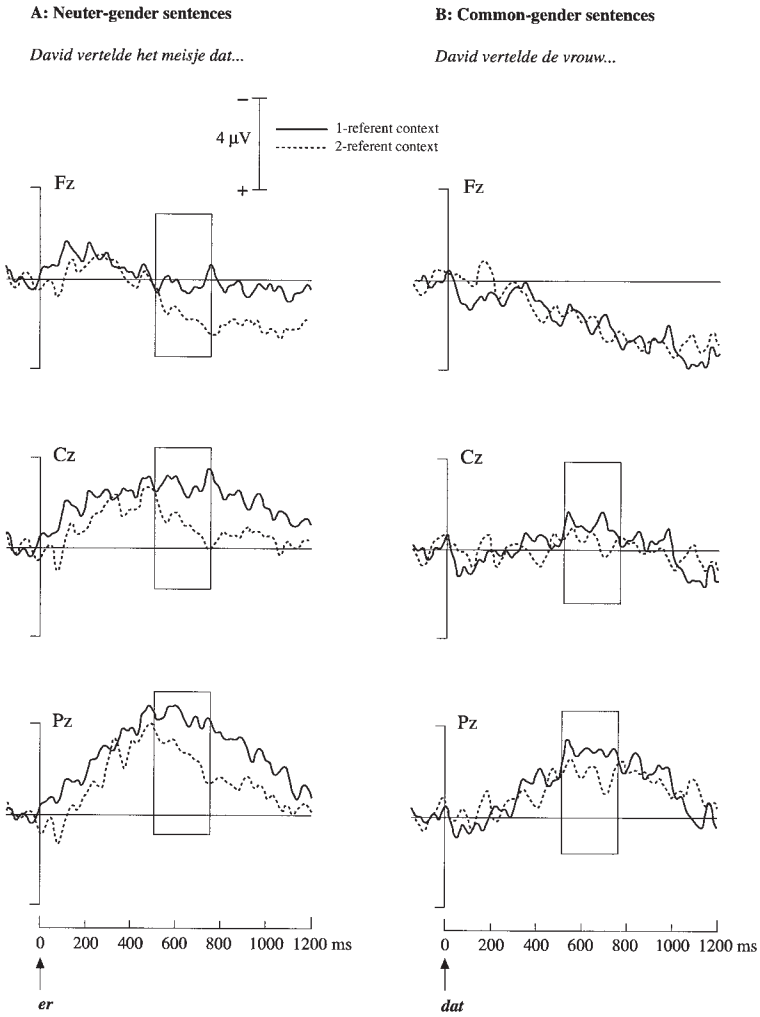

Fig. 1. In both panels of the figure the onset of the disambiguating word is at 0 ms, negative polarity is plotted upward, and the area for the statistical analysis of the P600/SPS is marked by a rectangle. Both panels show data from a frontal (Fz), central (Cz), and parietal (Pz) midline electrode site. (A) Average ERPs elicited by complement-clause disambiguation in neuter-noun target sentences (*er*) presented in a complement-biasing 1-referent context (solid line) and a relative-biasing 2-referent context (dotted line). (B) Average ERPs elicited by complement-clause disambiguation in common-noun target sentences (*dat*) presented in a complement-biasing 1-referent context (solid line) and a relative-biasing 2-referent context (dotted line).

stimuli cannot be distinguished in these waveforms. A first divergence in the waveform occurs at a very early moment, before the onset of the critical word. We will discuss this divergence below, in relation to a likewise early divergence in the waveforms for the common-gender sentences.

As in the prior written-language study the critical word *er* in the neuter-gender sentences elicits a clear P600/SPS in the 2-referent context, relative to the 1-referent context. This can be most easily seen at anterior sites (e.g., Fz), where it emerges at about 500 to 550 ms after CW onset and lasts for at least a few hundred milliseconds, both typical characteristics for this type of effect. It also emerges, at about the same time, at more posterior locations, albeit superimposed on a positive trend with a much earlier onset, reflecting the early divergence we mentioned above. An ANOVA on mean amplitudes in the 500–750 ms interval after onset of the CW revealed a significant main effect of Referential Context [F (1, 23) = 9.64, MSE = 18.9, p = .005), which did not interact with Electrode Site (F (12, 276) = 0.58, MSE = 2.1, Greenhouse-Geisser corrected p = .563], nor, in separate topographical analyses, with Hemisphere, Anteriority across the 3 midline or the combined 2×5 lateral sites, or with Hemisphere by Anteriority (all Fs < 1). Consistent with our assumption that in neuter-gender sentences the preceding word *dat* would generate the syntactic ambiguity but not yet resolve it, we did not observe a context-induced P600/SPS effect in the relevant latency range of this word.

The average ERPs time-locked to the complement-disambiguating word *dat* in common-gender sentences are shown in Fig. 1B, for the three midline electrode sites Fz, Cz, and Pz. Here again, a very early divergence is visible, which we shall discuss below. Like its written counterpart (and like the spoken word *er* in the neuter-gender sentences above), the spoken critical word *dat* elicits a P600/SPS in the 2-referent context, relative to the 1-referent context. The effect is small and has a focal distribution over the midline and right-posterior electrodes Cz, Pz, RT, RTP, and RO, where it emerges at about 500 to 550 ms after CW onset and lasts for approximately 200–250 ms. The ANOVA based on all 13 electrode sites on mean amplitudes in the 500–750 ms interval after CW onset did not reveal a main Referential Context effect [F(1, 23) = 1.64, MSE = 11.1, p = .213), or an interaction with Electrode Site [F(12, 276) = 1.27, MSE = 1.4, Greenhouse– Geisser corrected p = .291]. Referential Context also did not interact with Electrode Site in any topographical subanalysis. However, in line with the focal topography of the effect, a separate analysis over the five centroparietal and right-posterior sites Cz, Pz, RT, RTP, and RO did reveal a clearly significant Referential Context effect in that region [F(1, 23) = 4.77, MSE = 7.2, p = .039]. In line with our assumption that syntactic disambiguation in the common-gender sentences would be achieved by (the gender of) the word *dat,* the subsequent word *er* did not elicit a context-induced P600/SPS effect.

For both the common-gender and the neuter-gender sentences, we also observed a very early divergence in the waveform, well before the onset of the predicted P600/SPS effects. In Fig. 1B, the waveforms obtained with common-gender sentences can be seen to diverge right at the onset of the word *dat* (i.e., at 0 ms). Closer analysis suggests that this early divergence is not elicited by the word *dat*, but originates at the preceding noun and reflects the processing consequences of the referential ambiguity of this noun in the 2-referent context. In our earlier work in the visual modality we had already observed a referential ambiguity effect that was clearly time-locked to the noun (see van Berkum *et al.*, 1999a, for details). This effect was replicated in the present experiment: in the spoken common-gender sentences a referentially ambiguous noun elicits a more negative ERP waveform than a referentially successful one, starting at approximately 400 ms after noun onset (which closely parallels the effect previously obtained for the written common-gender sentences). As can be seen in Fig. 1B, this referential effect also shows up in the waveforms that are time-locked to the next word *dat.* This is because the referential ERP effect begins at about 400 ms after noun onset, which happens to coincide with the average onset of the word *dat* in our spoken stimuli. As can be seen in Fig. 1A, the neuter-gender sentence waveforms diverge as well immediately upon onset of the critical word, although, in this case, the 2-referent condition yields a more positive ERP signal. Again, detailed analysis suggests that this is an indirect consequence of the referential ambiguity effect elicited by the noun. The reason why there is an early negative divergence in Fig. 1B and an early positive divergence in Fig. 1A relates to the relative timing of the critical words *dat* and *er* with respect to the preceding noun: the normalization interval for the word *dat* falls within the early ascending flank of the processing effect of the noun, whereas the normalization interval for *er* falls within the later descending flank. In summary, the very early separation between the waveforms in both Fig. 1A and B is a reflection of the processing effect of the referential ambiguity of the noun that precedes the critical words *dat* and *er*. Because this effect is unrelated to the parsing issues that are the focus of the present paper, we will not further discuss it here.

DISCUSSION

A first finding to emerge from this study is that we obtained significant P600/SPS effects with naturally produced connected speech stimulation. This demonstrates that despite the intrinsically variable and continuous nature of the speech signal, it is possible to measure language-related brain potentials during real-time sentence processing. Our findings are in line with

a still limited set of studies on ERPs and spoken-sentence processing (e.g., Steinhauer, Alter, & Friederici, 1998; Brown & Hagoort, 1999; Connolly & Phillips, 1994; Hagoort & Brown, 1999; Müller, King, & Kutas, 1997), indicating that ERP research on higher-order processes during spoken language comprehension is now a viable option. Note that in our experiment subjects did not have to perform any additional, extraneous task. The instruction was simply to listen attentively for comprehension. Under these conditions we obtained reliable ERP effects. This demonstrates a major advantage of the ERP method for spoken language research, namely, that potentially confounding and contaminating effects from the demands of additional tasks (such as overt grammaticality judgments) can be entirely avoided, while at the same time obtaining an uninterrupted measure of processing activity with millisecond temporal resolution.

The second finding is that there is a significant and early impact of discourse-semantic information on the resolution of the complement/relative–clause ambiguity. In neuter-gender sentences, where the target sentence continues with the by default preferred complement clause, a P600/SPS was observed when the prior discourse biased toward a relative-clause analysis (i.e., in the 2-referent condition). The P600/SPS was elicited by the first word that necessitated a complement-clause continuation. This indicates that discourse information can very rapidly influence the initial analysis of a structural ambiguity, which is in accordance with the claims of context-sensitive parsing models and is problematic for syntax-first models.

The third finding to emerge from the present results demonstrates that discourse information can momentarily take precedence over local syntactic agreement rules. In the common-gender sentences, where the agreement rules of Dutch dictate that a complement-clause analysis is the only possible structural continuation, we again obtained a P600/SPS in the 2-referent condition. This effect was elicited at the earliest possible moment in the speech stream, namely on the lexically ambiguous word *dat.* Despite the fact that this very same word conveyed the local syntactic-gender information that forced a complement-clause analysis, the P600/SPS effect reveals that subjects were momentarily lured by the discourse information into entertaining a relative-clause analysis. This presents a very strong case for an immediate impact of discourse-level information on syntactic processing. In addition, it provides information on the relative weighting of different kinds of (non-)linguistic information during real-time discourse processing. It appears that, at least under the processing conditions of our experiment, discourse-semantic information can take precedence over lexical-syntactic information.

There is one alternative explanation of the data for the common-gender sentences that we should mention here. It could be claimed that the critical word *dat* was sometimes (that is, in at least a significant proportion of trials)

accessed as a complementizer. This, in turn, would result in a perceived pragmatic clash with the discourse in which the relative-clause interpretation was favored. There are several reasons why we can rule out this hypothesis. First, there is no equivalent processing effect on *dat* in the neuter-gender sentences, which cannot be explained by the above account. Second, the literature on language-related ERP components has demonstrated that pragmatic and/or semantic processing difficulty during language comprehension is manifest in modulations of the N400 component [see Brown and Hagoort (1999), Kutas and Van Petten (1994), and Osterhout and Holcomb (1995) for recent reviews]. In the processing window of the critical word *dat,* no N400 effects were observed, while in the same experiment the subjects did produce strong N400 effects to discourse–semantic violations [for purposes of brevity, these data have not been reported in this paper; see van Berkum, Hagoort, and Brown (1999) for details]. Finally, and related to the previous point, the fact that a P600/SPS was elicited by the word *dat* strongly indicates that we are looking at the processing consequences of a failed *syntactic* analysis, reflecting a syntactic commitment during on-line sentence processing [see Hagoort *et al.* (1999) for a review of the functional nature of the P600/SPS].

One of the main purposes of this study was to extend ERP findings of visual sentence processing research into the essentially uncharted domain of spoken discourse understanding. This endeavor has been successful. The ERP processing effects that we previously reported for written language comprehension (van Berkum *et al.*, 1999a, b) are replicated here.[3] One implication of the present results is, therefore, that the concerns that we raised earlier about possible artifactual effects in our visual studies due to the manner of stimulus presentation, can be laid to rest. Within the context of language comprehension, the P600/SPS is a reliable, modality-independent index of syntactic processing, and is not dependent on the manner of visual stimulus presentation.

A second implication of the ERP data concerns the temporal dynamics of the language processing system. It is intriguing that the time course of the P600/SPS in the spoken and the written modality is, to all intents and purposes, identical. Despite the very different timing conditions with which the physical stimulus impinges on the ear or eye, the ERP effect is manifest at the same moment (at about 500 ms) after the onset of the critical word. In other work, using sentences with violations of grammatical agreement, we have also found that the onset of the P600/SPS was at 500 ms, irrespective of

[3] There are, however, some differences in the topographical distribution. Whereas the P600/SPS for neuter-gender sentences had a similar distribution in both modalities, the topography for the common-gender sentences differed between modalities. In the visual study the P600/SPS had a bilateral frontal distribution, in the auditory study the effect is present over central and right-posterior sites.

whether the sentences were spoken, visually presented word-by-word at a slow presentation rate of 1.7 words per second, or at a fast rate of 4 words per second (Hagoort & Brown, 1999). Furthermore, in ongoing work we have observed similar between-modality time invariances in the ERPs elicited by referential ambiguity and by discourse- and lexical-semantic processing.

Although, at present, we lack a clear understanding of the functional significance of the commonalities and divergencies in the electrophysiological profiles of visual and auditory language comprehension, the close correspondence in the time course of the auditorily and visually elicited P600/SPS (and other language-related brain potentials) is not an incidental finding and warrants further research. It appears that at least some part of the sentence processing system operates under its own timing conditions, independent of the temporal parameters of the input.[4]

CONCLUSION

To our knowledge, the present study is the first electrophysiological investigation of the interplay between discourse-semantic and lexical-syntactic information during spoken language understanding. The combined results clearly demonstrate an early effect of high-level semantic knowledge on syntactic analyses: syntactically based preferences can be put aside on the basis of discourse-semantic information. In addition, the results provide insights into the relative weighting of discourse-semantic and lexical-syntactic information during real-time sentence processing. The data strongly indicate that discourse-semantic information can momentarily take precedence over lexical-syntactic information, even when this goes against grammatical gender agreement rules. Taken together, the results support architectures of the language comprehension system in which high-level semantic information can directly affect syntactic analyses.

REFERENCES

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

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

[4] In this respect it is interesting to note that there also is a close temporal similarity between ERP manifestations of lexical processing elicited by written American-English and by American Sign Language (cf. Neville, Coffey, Lawson, Fischer, Emmorey, & Bellugi, 1997), languages that clearly operate under very different temporal conditions.

Jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 213–237). Cambridge: Cambridge University Press.

Connolly, J. F., & Phillips, N. A. (1994). Event-related brain potential components reflect phonological and semantic processing of the terminal words of spoken sentences. *Journal of Cognitive Neuroscience, 6,* 256–266.

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.

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

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.

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.

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

Hagoort, P., & Brown, C. M. (1999). Semantic and syntactic ERP effects of speech compared to reading. Submitted for publication.

Jackendoff, R. (1999). The representational structures of the language faculty and their interactions. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 37–79). Oxford: Oxford University Press.

Kutas, M. (1987). Event-related brain potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. In R. Johnson Jr., J. W. Rohrbaugh, & R. Parasuraman (Eds.), *Current trends in event-related potential research (EEG Suppl. 40)* (pp. 406–411). Amsterdam: Elsevier.

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). San Diego: Academic Press.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). Oxford: Oxford University Press.

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

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,* 469–488.

Müller, H. M., King, J. W., & Kutas, M. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research, 5,* 193–203.

Neville, H. J., Coffey, S. A., Lawson, D. S., Fischer, A., Emmorey, K., & Bellugi, U. (1997). Neural systems mediating American Sign Language: Effects of sensory experience and age of acquisition. *Brain and Language, 57,* 285–308.

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. (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.

Steinhauer, K., Alter, K., & Friederici, A. D. (1998). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2,* 191–196.

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

van Berkum J. J. A., Brown, C. M., & Hagoort, P. (1999b). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research, 28,* 555–571.

van Berkum J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657–671.