3 Brain Responses to Lexical-Ambiguity Resolution and Parsing

Peter Hagoort
Colin Brown
Max Planck Institute for Psycholinguistics Nijmegen, The Netherlands

We present two sets of event-related potential (ERP) data on separate aspects of sentence processing. The first set focuses on the disambiguation of *biased* ambiguous words by right-context information (biased ambiguous words are words that have a clearly dominant and subordinate meaning). The ERP data provide both converging and contrasting evidence with results obtained in eye-movement studies. The ERP data converge with these studies by showing that when the right-context is consistent with the subordinate meaning, processing effects are observed at the point of disambiguation, whereas no effect is seen when the dominant meaning is confirmed (cf. Dopkins, Morris, & Rayner, 1992; Frazier & Rayner, 1990). The ERP data contrast with eye-movement data (cf. Rayner & Duffy, 1986) by showing that processing effects do obtain at the level of biased ambiguous words preceded by neutral sentential contexts.

The second set of ERP data focuses on neurophysiological manifestations of syntactic processing. These data show an ERP effect to syntactic violations that is qualitatively different from the well-known N400 effect to violations of semantic constraints. Next to the ERP effect elicited by the word that renders the sentence ungrammatical, N400 effects are observed further downstream in sentences that can be interpreted semantically (i.e., normal prose), but not in sentences that are very hard to interpret (i.e., syntactic prose). This shows that syntactic and semantic ERP effects can be dissociated. In addition, these data suggest that a syntactic problem has immediate consequences for the semantic integration of words following the occurrence of a parsing problem. Finally, probably due to the semantic consequences of the lexically specified argument structure of verbs, the pattern of results for subcategorization violations is different from that of the other two syntactic violations in this experiment (i.e., agreement violations and phrase-structure violations).

Before discussing our experimental work, we briefly describe the ERP method, and specify its relevant features for sentence-processing research.

Scalp-recorded ERPs reflect the summation of the synchronous post-synaptic activity of many neurons. ERPs differ from background EEG in that they reflect brain electrical activity time-locked to particular stimulus events. Establishing a reliable ERP trace normally requires averaging over a series of ERP recordings to tokens of the same stimulus type. The resulting average waveform typically includes a number of positive and negative peaks, sometimes referred to as *components*. Usually the peaks in the ERP waveform are labeled according to their polarity (*N* for negative, *P* for positive) and their average latency in milliseconds relative to the onset of stimulus presentation (e.g., *N400*, *P600*). In some cases, the ERP peaks get a functionally defined label (*SPS* for syntactic positive shift; *ERN* for error-related negativity). ERPs are recorded from a number of leads distributed over the scalp, and often have a characteristic distribution, showing larger amplitudes at some leads than at others. These distributional characteristics can be helpful in identifying a certain component.

For the purposes of psycholinguistically oriented ERP research, the most informative ERP peaks belong to the class of the so-called "endogenous" components. Endogenous components are relatively insensitive to variations in physical stimulus parameters (e.g., size, intensity), but highly responsive to the cognitive-processing consequences of the stimulus events. The modulations in amplitude or latency of an endogenous ERP peaks as a consequence of some experimental manipulation, usually form the basis for making inferences about the nature of the underlying cognitive processing events. In terms of experimentation, the ways in which to elicit the relevant experimental effects are not essentially different from other research paradigms in psycholinguistics.

The moment at which the ERP method showed a first glimpse of its potential relevance for sentence-processing research is clearly demarcated in time by the 1980 publication of a paper in *Science* by Kutas and Hillyard. These researchers presented subjects with a variety of sentences either ending in a word that was semantically congruous with the sentence context (e.g., "He shaved off his mustache and beard") or ending in a semantic anomaly (e.g., "I take coffee with cream and dog"). The semantically anomalous words elicited a negative component with a centro-parietal maximum on the scalp, and a latency that peaked around 400 ms. This component has since become known as the *N400*, and the difference between the N400 amplitude in the experimental and the control conditions has become known as the *N400 effect*. It is now clear that N400 effects can be obtained with a variety of paradigms and using a variety of language stimuli, by no means restricted to violations.

Today, the following general characteristics are known to hold for the N400: (a) Each open-class word elicits an N400. (b) The amplitude of the N400 is ent from that of the other two syntactic violations in this experiment (i.e., agreement violations and phrase-structure violations).

Before discussing our experimental work, we briefly describe the ERP method, and specify its relevant features for sentence-processing research.

Scalp-recorded ERPs reflect the summation of the synchronous post-synaptic activity of many neurons. ERPs differ from background EEG in that they reflect brain electrical activity time-locked to particular stimulus events. Establishing a reliable ERP trace normally requires averaging over a series of ERP recordings to tokens of the same stimulus type. The resulting average waveform typically includes a number of positive and negative peaks, sometimes referred to as *components*. Usually the peaks in the ERP waveform are labeled according to their polarity (*N* for negative, *P* for positive) and their average latency in milliseconds relative to the onset of stimulus presentation (e.g., *N400*, *P600*). In some cases, the ERP peaks get a functionally defined label (*SPS* for syntactic positive shift; *ERN* for error-related negativity). ERPs are recorded from a number of leads distributed over the scalp, and often have a characteristic distribution, showing larger amplitudes at some leads than at others. These distributional characteristics can be helpful in identifying a certain component.

For the purposes of psycholinguistically oriented ERP research, the most informative ERP peaks belong to the class of the so-called "endogenous" components. Endogenous components are relatively insensitive to variations in physical stimulus parameters (e.g., size, intensity), but highly responsive to the cognitive-processing consequences of the stimulus events. The modulations in amplitude or latency of an endogenous ERP peaks as a consequence of some experimental manipulation, usually form the basis for making inferences about the nature of the underlying cognitive processing events. In terms of experimentation, the ways in which to elicit the relevant experimental effects are not essentially different from other research paradigms in psycholinguistics.

The moment at which the ERP method showed a first glimpse of its potential relevance for sentence-processing research is clearly demarcated in time by the 1980 publication of a paper in *Science* by Kutas and Hillyard. These researchers presented subjects with a variety of sentences either ending in a word that was semantically congruous with the sentence context (e.g., "He shaved off his mustache and beard") or ending in a semantic anomaly (e.g., "I take coffee with cream and dog"). The semantically anomalous words elicited a negative component with a centro-parietal maximum on the scalp, and a latency that peaked around 400 ms. This component has since become known as the *N400*, and the difference between the N400 amplitude in the experimental and the control conditions has become known as the *N400 effect*. It is now clear that N400 effects can be obtained with a variety of paradigms and using a variety of language stimuli, by no means restricted to violations.

Today, the following general characteristics are known to hold for the N400: (a) Each open-class word elicits an N400. (b) The amplitude of the N400 is inversely related to the cloze probability of a word in sentence context. The better the semantic fit between a word and its context, the more reduced the amplitude of the N400 (Kutas, Lindamood, & Hillyard, 1984). (c) The amplitude of the N400 varies with word position, such that the first content word in a sentence produces a larger negativity than content words in later positions (Kutas, Van Petten, & Besson, 1988). This amplitude reduction is most likely due to the increasing semantic constraints throughout the sentence. (d) N400 effects are obtained in sign language (Kutas, Neville, & Holcomb, 1987; Neville, Mills, & Lawson, 1992), but not with violations of contextual constraints in music (Besson & Macar, 1987; Paller, McCarthy, & Wood, 1992). (e) N400 effects are observed both for written- and spoken-language input, although with slightly different time courses (e.g., Connolly, Stewart, & Phillips, 1990; Holcomb & Neville, 1991). For language comprehension, the processing nature of the N400 has recently been claimed to be related to lexical-semantic integration processes (Brown & Hagoort, 1993). That is, once a word has been accessed in the mental lexicon, its meaning has to be integrated into an overall representation of the current word or sentence context. The easier this integration process is, the smaller the amplitude of the N400 becomes.

Although the past 15 years have seen an increase in ERP research on language processing, most of the research has been dedicated to determining the parameters that modulate language-relevant ERP components such as the N400. These were necessary steps in preparing the ground for ERP research aimed at explicitly testing specific theoretical proposals on different aspects of language processing. Only in recent years have ERP studies directly addressed central issues in sentence-processing research (e.g., Garnsey, Tanenhaus, & Chapman, 1989; Kluender & Kutas, 1993; Osterhout, chapter 2, this volume; Van Petten & Kutas, 1987).

The reason for being optimistic about possible contributions of ERP research to studies of sentence processing arises from some of the characteristics of brain waves as a dependent measure. We discuss two of the most relevant ones for purposes of studying higher order sentence-processing operations.

The first is the *multidimensional* nature of the ERP waveform. ERPs can vary along a number of dimensions: specifically, the latency at which an ERP component occurs relative to stimulus onset, its polarity, its amplitude, and its amplitude distribution over the recording sites. On the basis of these characteristics it is reasonable to assume that different types of ERP peaks are generated by different neural systems. Insofar as the involvement of different neural systems implies qualitatively different processing events, in principle these processing events can show up as qualitatively different in the ERP waveform. This characteristic makes ERPs a useful addition to the recording of unidimensional measures, such as reading times. For instance, if the electrophysiological signatures of semantic integration processes and parsing operations turn out to be qualitatively different, ERPs might provide us with a crucial tool for testing how and at what moments in time the process of assigning a structure to the incoming string of words and interpreting this string semantically, influence each other. Recent evidence suggests that aspects of syntactic processing do indeed elicit ERP responses qualitatively different from those of semantic-integration processes (e.g., Hagoort, Brown, & Groothusen, 1993, Kluender & Kutas, 1993; Münte, Heinze, & Mangun, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992; Rösler, Pütz, Friederici, & Hahne, 1993).

The second important characteristic of ERPs is that they provide a *continuous, real-time* measure. Like speeded reaction-time (RT) measures in the more classical psycholinguistic tasks, such as naming, lexical decision, and word monitoring, ERPs are tightly linked to the temporal organization of ongoing language-processing events. But in contrast to RT measures, ERPs provide a continuous record throughout the total processing epoch and beyond (as is the case with eye-movement registration). Therefore, it is possible to monitor not only the immediate consequences of a particular experimental manipulation (e.g., a syntactic or semantic violation), but also its processing consequences further downstream. As we will show, this feature enabled us to show that the impossibility of assigning the preferred structure to an incoming string of words has consequences for lexical-semantic integration processes further downstream in the sentence (Hagoort et al., 1993).

In the remainder of this chapter, we illustrate the usefulness of these characteristics of ERPs by presenting the two datasets mentioned previously. We consider neither of the experiments decisive for central debates in their respective domains of sentence-processing research. At the same time, both are not entirely without consequences for some of the current positions in these domains.

THE RESOLUTION OF LEXICAL AMBIGUITY BY RIGHT-CONTEXT INFORMATION

In the absence of a well-defined theory of the semantics of natural languages, it is notoriously difficult to develop processing models of the integration of words in message-level representations. Both at the word-meaning level and the message level, it is unclear exactly what these representations are and, hence, what the nature of the eventual end product of the comprehension process is. One way that psycholinguistic researchers have sidestepped some of the problems involved here is to investigate the meaning selection of lexically ambiguous words in sentential contexts. Here at least we have two (or more) clearly distinct meanings for a given lexical form, whatever theory of semantics eventually turns out to hold. With these different meanings in hand, it is possible to investigate the time course of their activation, both as a function of their relative meaning frequency and of the preceding and/or following contextual information. This kind of experimental program is important because it focuses on the integration of words into higher order meaning representations, and thereby, on the interface between the mental lexicon and sentence-processing systems, an interface that lies at the heart of language understanding.

Over the past two decades, a quite stable picture has emerged of the processing of lexically ambiguous words in neutral sentential contexts. This picture has primarily been built up on the basis of RT research with the cross-modal priming paradigm. The data indicate that, initially, all meanings of an ambiguous word are activated, followed by the selection of a single meaning (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Simpson, 1981; Simpson & Krueger, 1991). The time course of meaning selection is influenced by the relative meaning frequency of the distinct meanings of an ambiguous word, and the overall thrust of the results reported in the literature is that dominant meanings are accessed prior to, and remain activated longer than, subordinate meanings (e.g., Simpson & Burgess, 1985; Simpson & Krueger, 1991).

The impact of preceding biasing sentential-context information on ambiguity resolution is still a matter of some debate. There is evidence for multiple meaning activation irrespective of contextual bias, followed by context-sensitive selection (e.g., Duffy, Morris, & Rayner, 1988; Kintsch & Mross, 1985; Onifer & Swinney, 1981; Swinney, 1979). However, contrary findings have been reported. These mainly concern results from studies using biasing contexts for dominant meanings, showing that in these circumstances subordinate meanings are not necessarily activated (e.g., Dopkins et al., 1992; Simpson, 1981, 1984; Simpson & Krueger, 1991; Tabossi, 1988a, 1988b; Tabossi, Colombo, & Job, 1987; Tabossi & Zardon, 1993; Van Petten & Kutas, 1987).

In the present experiment, we focus on the effects of so-called right-context information on meaning selection for ambiguous words with clearly dominant and subordinate meanings (cf. Duffy et al., 1988; Frazier & Rayner, 1990; Rayner & Frazier, 1989). Subjects read sentences that contained words with two distinct meanings. The ambiguities were preceded by neutral sentential information, and were followed (several words downstream) by a word that disambiguated either for the dominant or the subordinate meaning of the preceding ambiguous word. This approach yields two datasets that can contribute to a more detailed characterization of the meaning-selection process as it occurs in real-time. On the one hand, it provides data on the possible processing consequences that derive from multiple meaning representations associated with one lexical form (cf. Rayner & Duffy, 1986). On the other hand, it provides data on possible processing effects at the level of the disambiguating word. Such effects can reveal the extent to which the relative meaning frequency determines the time course of meaning selection and decay; that is, whether the interpretative process is best characterized as one of immediate or delayed meaning selection (cf. Frazier & Rayner, 1990; Rayner & Frazier, 1989). The manipulation of right-context information exploits one of the appealing characteristics of the ERP method that we mentioned earlier; namely, the fact (which it shares with eye-movement registration) that ERPs can be obtained over the entire processing span of a sentence (or sequence of sentences) without having to interrupt the comprehender's ongoing linguistic analysis.

Method

Thirty-six university students read 360 Dutch sentences. Of this set, 120 are test sentences containing an ambiguous word, 120 are control sentences for the ambiguous test sentences, 60 are sentences containing a high-cloze probability word, and 60 are sentences containing a low-cloze probability word. The following sentences exemplify the materials for the ambiguity manipulation (the ambiguous word is in bold roman, the control word is in bold italics, and the disambiguating word is in plain italics):

Concordant with Dominant Meaning	
Ambiguous:	Mijn oom heeft de **as** van zijn *sigaar* snel opgeruimd. (My uncle has the **ash** of his *cigar* quickly [cleared away].)
Unambiguous:	Mijn oom heeft de ***peuk*** van zijn *sigaar* snel opgeruimd. (My uncle has the ***stub*** of his *cigar* quickly [cleared away].)
Concordant with Subordinate Meaning	
Ambiguous:	Mijn oom heeft de **as** van zijn *auto* opnieuw gelast. (My uncle has the **axle** of his *car* again welded.)
Unambiguous:	Mijn oom heeft de ***knalpijp*** van zijn *auto* opnieuw gelast. (My uncle has the ***muffler*** of his *car* again welded.)

The meaning dominance of the ambiguous words was assessed via an association test on 70 university students. The dominant meaning frequency is 81.3%, and the subordinate meaning frequency is 11.7%.

The 60 high-cloze and 60 low-cloze probability sentences were constructed to elicit a standard N400 effect. Given that the present study is the first study to our knowledge to look at brain-potential manifestations of the consequences of right-context information for the processing of lexically ambiguous words, we thought it prudent to establish a standard N400 effect within the same subject population, so as to have a basis for comparing possible morphological differences in the ERP waveforms elicited in the ambiguity conditions.

The high- and low-cloze sentences were identical, with the exception of the high- and low-cloze target words that occurred in sentence-medial position. The cloze probability is 0.58 for the high-cloze words and 0.07 for the low-cloze words. The following two sentences exemplify the two cloze conditions (the target word is in bold):

High cloze: Jenny stopte het snoepje in haar **mond** na afloop van de les.
[Jenny put the sweet in her **mouth** after the lesson.]

Low cloze: Jenny stopte het snoepje in haar **zak** na afloop van de les.
[Jenny put the sweet in her **pocket** after the lesson.]

Two lists of 180 sentences each were made. Each list contains 30 right context dominant-concordant sentences, 30 right-context subordinate-concordant sentences, 60 control sentences, 30 high-cloze sentences, and 30 low-cloze sentences. Within each list, only one version of an ambiguous context occurred (i.e., either dominant or subordinate), with its corresponding control sentence occurring in the same sequential position in the other list. Each list was presented to 18 subjects.

All sentences were displayed word by word on a high-resolution computer screen. Each word replaced the preceding one, and was presented for 300 ms, with an Inter-Simulus-Interval (ISI) of 300 ms.[1] Subjects were instructed to attentively read the sentences for comprehension. No additional task demands were imposed.

Results

For each subject, average waveforms are computed for each condition and electrode site separately, over all trials that are free of artifacts. Mean amplitudes are calculated per subject, per condition, and per electrode for latency windows specified later. These amplitude values are entered into repeated measures analyses of variance (ANOVA). Before turning to the ambiguity data, we briefly present the waveforms for the cloze manipulation. These data establish a standard N400 effect.

ERPs to High- and Low-Cloze Words in Sentence-Medial Position. In Fig. 3.1 and 3.2, 2400-ms epochs are shown for the high- and low-cloze words. This epoch contains three word positions: one word preceding the high- or low-cloze

---

[1]EEG activity was recorded using an Electrocap with seven scalp tin electrodes, each referred to the left mastoid. Three electrodes were placed according to the International 10–20 system (Jasper, 1958), at frontal (Fz), central (Cz), and parietal (Pz) sites. Symmetrical anterior-temporal electrodes were placed halfway between F7 and T3 (anterior left: AL), and F8 and T4 sites (anterior right: AR), respectively. Symmetrical posterior-temporal electrodes were placed lateral (by 30% of the interaural distance) and 12.5% posterior to the vertex (posterior left: PL, posterior right: PR). Vertical eye movements and blinks (EOG) were monitored via a supra- to sub-orbital bipolar montage. A right to left canthal bipolar montage was used to monitor horizontal eye movements. The EEG and EOG recordings were amplified with Nihon Kohden AB-601G bioelectric amplifiers, using a Hi-Cut of 30 Hz and a time constant of 8 sec. The EEG and EOG were digitized on-line with a sampling frequency of 200 Hz. Sampling started 150 ms before the presentation of the first word of each sentence, with a total sampling epoch of 9000 ms.

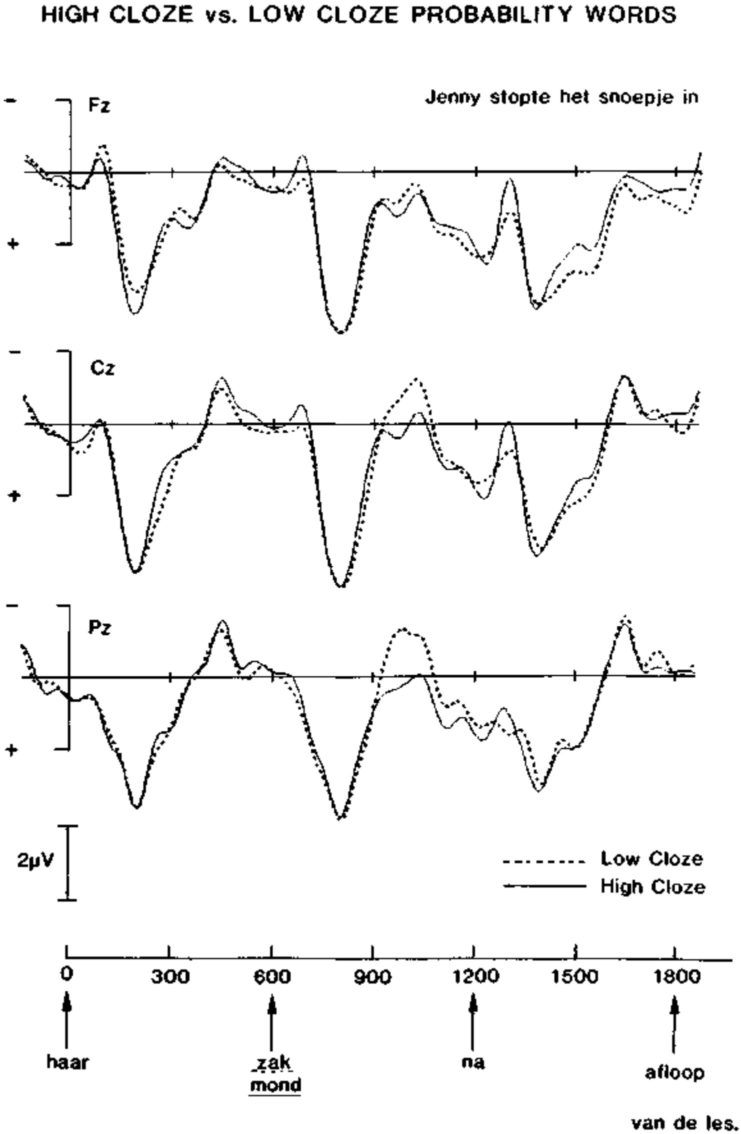

FIG. 3.1. Cloze-probability data. Grand-average waveform for each of the three midline electrode sites, for the high- and the low-cloze words. The cloze target is preceded and followed by one word. The translation of the example sentence is "Jenny put the sweet in her pocket/mouth after the lesson." Negativity is up in this and all subsequent figures.

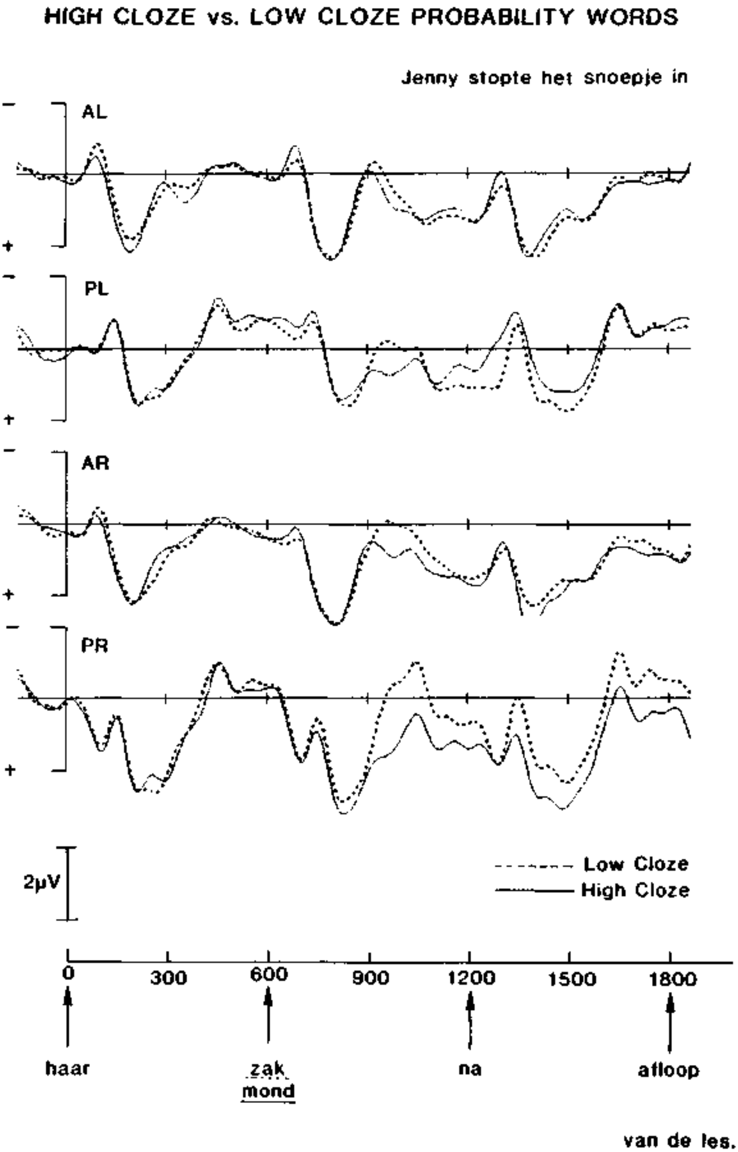

FIG. 3.2. Cloze-probability data. Grand-average waveform for each of the four lateral electrode sites, for the high- and the low-cloze words. The cloze target is preceded and followed by one word. The translation of the example sentence is "Jenny put the sweet in her pocket/mouth after the lesson."

word, the high- or low-cloze word itself, and one following word. Figure 3.1 shows the midline electrode sites, and Fig. 3.2 shows the lateral sites. The waveforms for the high- and low-cloze words show the by-now standard and well-established N400 effect, reflecting the degree to which the words can be readily integrated within the higher order representation of their preceding sentential-semantic context (cf. Brown & Hagoort, 1993). Statistical analyses were done on the standard time window for analysis of N400 effects. The ERP for the low-cloze words shows a significant negative enhancement in the 300–500-ms latency range following word onset, compared with the high-cloze words [$F(1, 33) = 5.60$, MSe = 10.39, $p = .024$]. The distribution of the effect follows the standard topography of the N400 to visual stimulation: largest over centro-parietal electrode sites, larger over posterior than anterior sites, with a slight increase over the right as compared with the left hemisphere. This fits well with previous reports in the literature (e.g., see overviews in Kutas & Van Petten, 1988; Van Petten & Kutas, 1991). No significant differences are predicted to emerge at the preceding word position because at this position the two cloze conditions are identical. None emerge, which demonstrates that the ERP registrations are reliable. Similarly, the waveforms for the position following the ambiguous word do not differ from each other.

ERPs to the Ambiguous Word. For this analysis, the data for the dominant- and subordinate-material sets are collapsed because the factor Dominance is only relevant with respect to the disambiguating word. Figures 3.3 and 3.4 contain 1800-ms epoch waveforms for the ambiguous words and for their controls. Figure 3.3 shows the midline electrode sites, and Fig. 3.4 shows the lateral sites. All words in both conditions elicit an N400, with the standard topography for visual stimulation.

The ERPs elicited in the Ambiguous and Unambiguous conditions at the word position preceding the ambiguity lie on top of each other. This is as it should be because the sentences in the two conditions are identical at this point. Significant differences are observed for the ambiguous word in comparison with its unambiguous control. The overall ANOVA shows a main effect of Ambiguous–Unambiguous [$F(1, 33) = 9.16$, MSe = 4.87, $p = .005$]. However, it should be noted that the distribution of the effect over the scalp does not fit with the standard topography observed fro N400 effects. In particular, the ambiguity effect has a frontal, largely lateralized distribution. There is some separation between the Ambiguous and Unambiguous conditions at Fz, but the largest effects emerge at the anterior left (AL) and anterior right (AR) sites. The posterior left (PL) site shows some effect, whereas at the posterior right (PR) site there is no difference between the two conditions. The most sustained separation holds over the left hemisphere: The effects at AL and PL sites are still present at the onset of the following word. This is not the case for the right lateral sites. Leaving aside imponderables concerning the meaning of this particular topo-

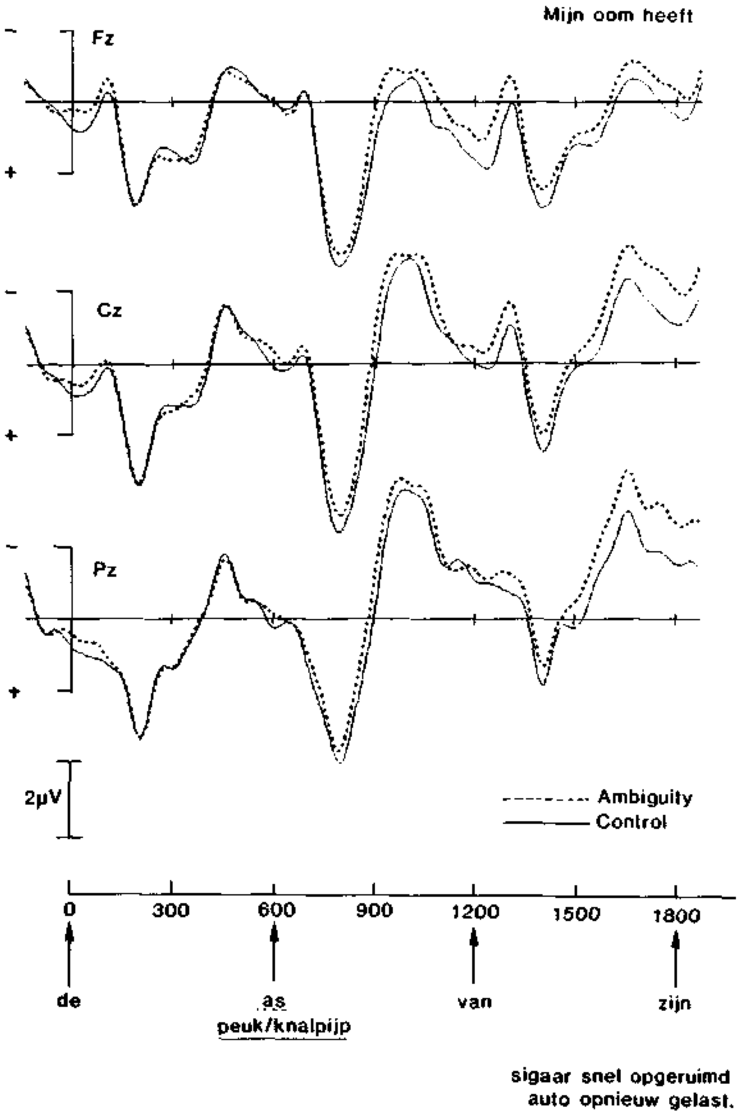

FIG. 3.3. ERPs for the ambiguous words. Grand-average waveform for each of the three midline electrode sites, for the ambiguous words and their controls. These words are preceded and followed by one word. The translation of the example sentence is "My uncle has the (ash/axle)/(stub/muffler) of his (cigar quickly cleared away/car again welded)."

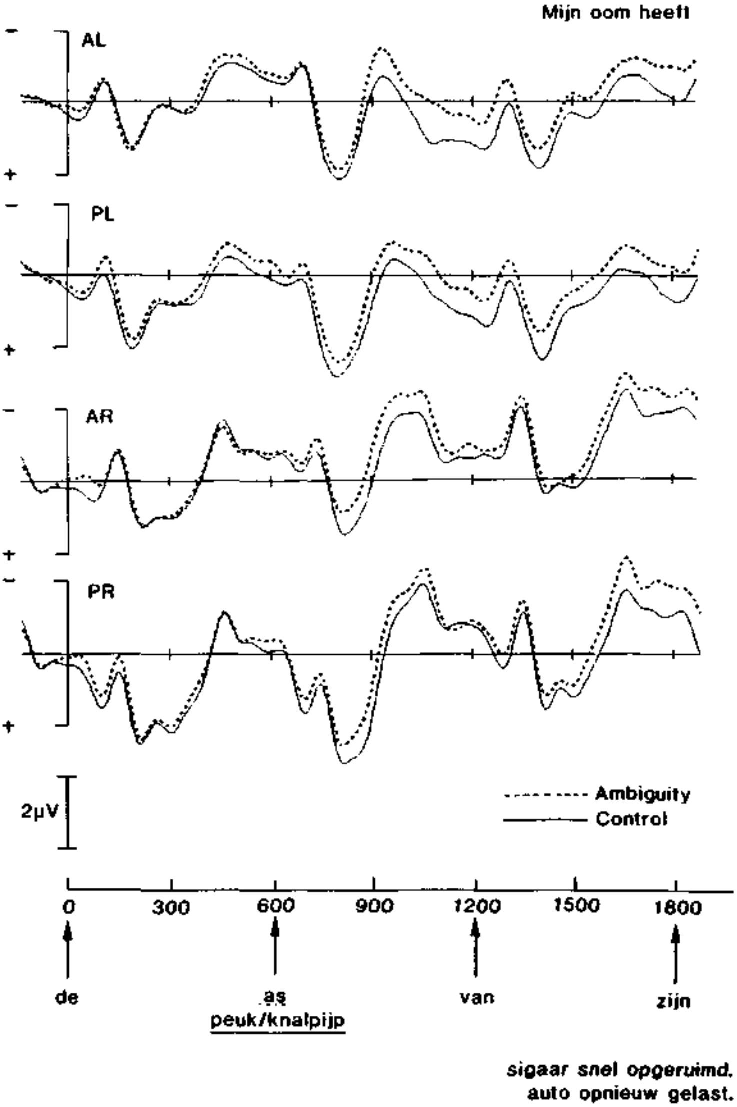

FIG. 3.4. ERPs for the ambiguous words. Grand-average waveform for each of the four lateral electrode sites, for the ambiguous words and their controls. These words are preceded and followed by one word. The translation of the example sentence is "My uncle has the (ash/axle)/(stub/muffler) of his (cigar quickly cleared away/car again welded)."

graphical deviation from the standard distribution of the N400 effect, the overall significant difference between the Ambiguous and Unambiguous conditions indicates that there are processing consequences associated with the reading of biased ambiguous words preceded by neutral sentential contexts.

The processing effects of the ambiguous word carry over into the processing of the word following the ambiguity. Here again, a statistically significant difference exists between the Ambiguous and Unambiguous conditions [$F(1, 33) = 5.45$, $MSe = 8.77$, $p = .026$]. Unlike the effect for the ambiguous word, this effect shows a topography that is more in line with the standard topography for N400 effects.

ERPs to the Word Disambiguating for the Dominant Meaning. Figure 3.5 contains 1200-ms epoch waveforms for the three midline electrode sites for the disambiguating words in the Ambiguous and Unambiguous conditions (i.e., either with an ambiguous or an unambiguous word in the preceding context).[2] As for the previously reported epochs, clear N400s were elicited by each word, and their topography is in line with the standard distribution of the N400 over the scalp.

Although there is a slight separation between the Ambiguous and Unambiguous conditions at the level of the disambiguating word, this does not reach significance. Likewise, for the following word, no significant differences emerge. So, it is clear that when the right-context information as conveyed by the disambiguating word is in accordance with the dominant meaning of a previously encountered ambiguous word, no differential processing consequences are observed. As we shall now see, this does not hold for right-contexts that are in accordance with the subordinate meaning.

ERPs to the Word Disambiguating for the Subordinate Meaning. Figure 3.6 contains 1200-ms epoch waveforms for the three midline electrode sites for the disambiguating words in the Ambiguous and Unambiguous conditions. Again, clear N400s are observed on each word, with the standard topography. At the level of the N400 to the disambiguating word, a statistically reliable negative enhancement emerges for the Ambiguous compared with the Unambiguous condition [$F(1, 33) = 10.65$, $MSe = 7.93$, p .003]. This effect is restricted to the disambiguating word: No significant Ambiguous–Unambiguous differences are present for the following word. The effect at Fz is somewhat larger than usually observed with N400 effects, but overall the topography matches the standard distribution. Here, then, in contrast to the contexts for the dominant meaning of an ambiguous word, an N400 effect is observed when the disambiguating information accords with the subordinate meaning of a previously processed ambiguous word.

[2]In this and the following figure, we do not present the lateral sites because no hemispheric differences emerged in the size of the effects.

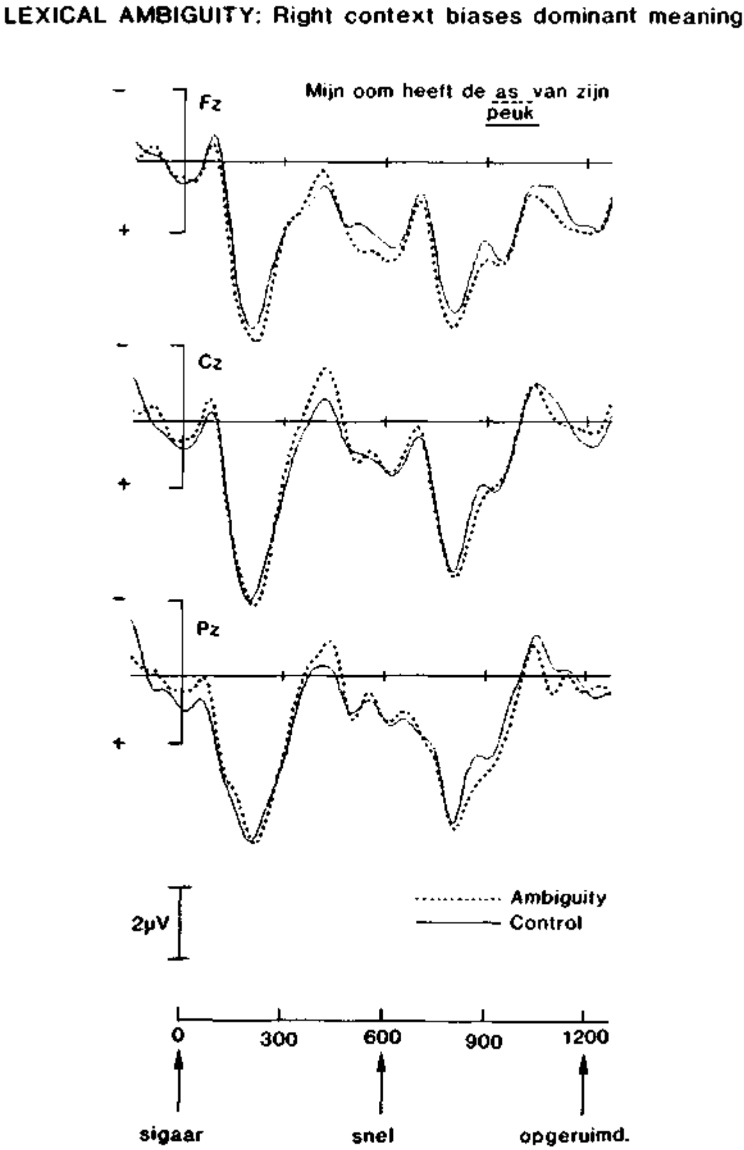

FIG. 3.5. ERPs to the word that disambiguates for the dominant meaning of the ambiguous word. Grand-average waveform for each of the three midline electrode sites. The disambiguating word is followed by one word. The translation of the example sentence is "My uncle has the ash/stub of his cigar quickly (cleared away)."

LEXICAL AMBIGUITY: Right context biases subordinate meaning

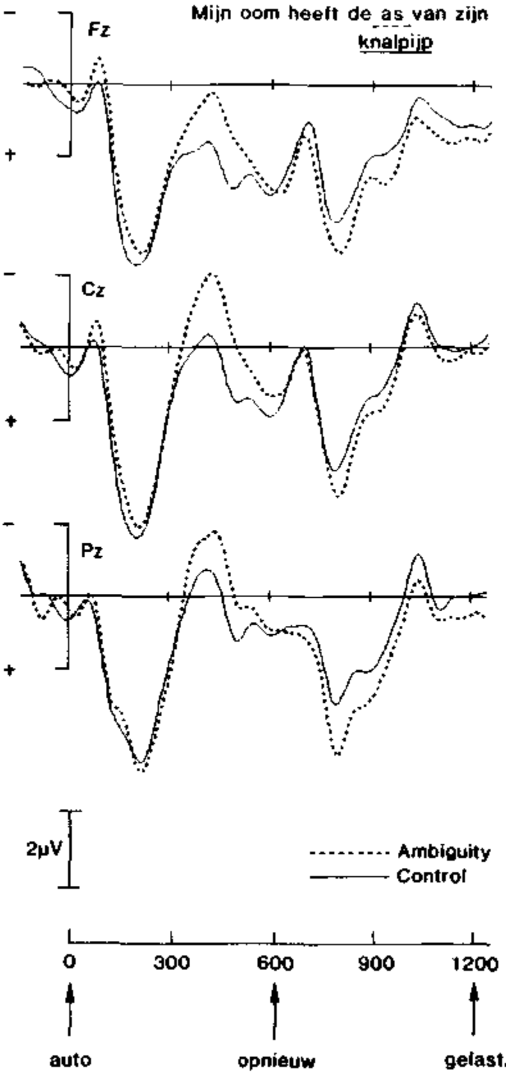

FIG. 3.6. ERPs to the word that disambiguates for the subordinate meaning of the ambiguous word. Grand-average waveform for each of the three midline electrode sites. The disambiguating word is followed by one word. The translation of the example sentence is "My uncle has *the axle/muffler of his car again welded.*"

Discussion

We begin with a discussion of the effects at the level of the ambiguous word, and then turn to the effects of the disambiguation. The significant negative enhancement for the ambiguous words compared with their unambiguous controls indicates that some differential processing is associated with the ambiguous words. This finding is in line with eye-movement data reported by Frazier and Rayner (1990), showing that fixation times are longer on words with multiple meanings compared with words with multiple senses. However, the ERP results are at odds with eye-movement data on balanced versus biased homographs (i.e., homographs with two equally likely meanings, vs. one highly likely and one unlikely meaning). The eye-movement data show that when the preceding context is neutral, readers fixate longer on balanced homographs compared with control words. This fixation difference is not observed for biased homographs (e.g., Duffy et al., 1988; Pacht & Rayner, 1993; Rayner & Duffy, 1986). Contrary to these findings, our ERP results do show a processing effect for biased homographs preceded by neutral sentential contexts. This effect might be due to unknown lexical differences between the ambiguous words and their controls (as could be the case for the pattern of effects obtained in the eye-movement studies). Alternatively, the effect might be reflecting some differential sensitivity between eye-movement and brain-potential data. We did not include balanced homographs in the present experiment, therefore full clarification of this issue awaits further research.

One possible reason for the present ERP effect is that accessing ambiguous words entails accessing multiple meanings, and thereby multiple lexical representations, which is not the case for the control words. Multiple access could be associated with greater processing costs, and this emerges in the ERP waveform. Alternatively, the ERP effect could be reflecting the processing costs associated with computing separate higher order message representations to accommodate the separate meanings of the ambiguity (i.e., multiple integration). Either way, if we accept that the control conditions are appropriate, the implication of the affect is that both meanings of the ambiguous word are activated, and that both are available within a time span of some 300 ms.

Two caveats need to be made here. First, although the ambiguous and the control words are matched on lexical characteristics, it is unclear how to control for, in particular, the lexical frequency of the two classes of words, given the different dominant and subordinate meanings of the ambiguous word (cf. Rayner & Frazier, 1989). However, it is unlikely that differences in lexical frequency underlie the observed effect, since word by Van Petten and Kutas (1990) has shown that the impact of word frequency on the ERP waveform is eliminated after the first two or so words in a sentence. Second, the exact nature of the neurophysiological effect we observed on the ambiguous word is as yet unclear. As we pointed out, the topography of the effect is quite different from the standard distribution of the N400, and this indicates that we are perhaps not dealing with an N400 effect here. Clearly, further research is required before any substantial statements can be made about possible neurophysiological effects related to the processing of ambiguous words preceded by neutral sentential contexts. With these caveats in mind, though, the results do indicate that representational aspects of the mental lexicon—in the present case the lexical complexity of ambiguous words—are manifest in the ERP waveform in the absence of any kind of explicit and interfering task demands (such as those associated with RT tasks).

The pattern of effects at the level of the disambiguating word is much clearer. No processing effects are observed when the dominant meaning of the ambiguous word is in accordance with the disambiguating information, but effects are observed for the subordinate meaning. These results fit well with integration-based accounts of the processing of ambiguous words with clearly dominant and subordinate meanings (cf. Rayner & Frazier, 1989). That is, upon encountering an ambiguous word preceded by a neutral sentential context, the processor integrates the dominant meaning with the prior context. This default assignment occurs either because the dominant meaning is accessed first, or because it receives more activation during a multiple-access process. If subsequent information indicates that the subordinate meaning is in fact the appropriate one, reanalysis processes have to be invoked. Claims about the nature of these reanalysis processes depend on the kinds of assumptions that are made about the time course of the access and integration process. In the integration model proposed by Rayner and Frazier, it is assumed that integration of the dominant meaning with prior context precedes, and if enough of a time lag exists before the subordinate meaning arrives, thereby terminates access to the subordinate meaning. Hence, the reanalysis process will often involve some kind of reaccessing process of the subordinate meaning. The nature of this process is unclear and has not been specified in any of the available models. The results are also compatible with reordered access models (cf. Dopkins et al., 1992). But because we used neutral left-sentential contexts in our experiment, we cannot contrast integration-based accounts with reordered access models. Alternatively, both the dominant and subordinate meanings of an ambiguous word might be accessed in parallel. In this scenario, multiple message-level representations are computed and held in working memory until such a moment in time when disambiguating information comes in. The effect of meaning frequency is then explained by positing differential activation levels for the higher order representations, possibly linked with differential activation-decay functions.

All of the accounts we have described are compatible with the brain-potential data presented in this chapter. In fact, these accounts are, to date, compatible with all of the available data on the effects of right-context information on the processing of ambiguous words that are preceded by neutral sentential contexts. Although the activation and decay functions of the dominant and subordinate meanings in neutral contexts are reasonably well established (e.g., Simpson & Burgess, 1985; Simpson & Krueger, 1991), it is still a matter of debate how these activation and decay functions are modulated by either left- or right-biasing contexts. To determine the nature of the interaction between context and the lexical activation and decay characteristics of the alternative meanings, we need tightly time-locked measures with which we can track on a momentary basis the activational status of the dominant and subordinate meanings over the sentence. Only by using these measures will we be able to determine the time course of the interaction between the activation level(s) of the lexical meaning(s) and the incoming left- or right-context information. Partly due to the relatively slow presentation rate, the present ERP results do not provide sufficient information about this time course, and thereby about the nature of the interaction between context and lexical information. However, what these results clearly demonstrate is that it is possible to pick up on neurophysiological effects that reflect computations at the interface between the mental lexicon and higher order processing systems. Given this demonstration, the challenge now ahead for language researchers using the ERP method is to work out ways of further capitalizing on the on-line, continuous character of the ERP signal, and to attempt to build up a real time processing profile of ambiguity resolution.

THE BRAIN'S RESPONSE TO PARSING

Most ERP studies on the processing of syntactic information have investigated the ERP effects of various types of syntactic violations in visually presented sentence materials. Although ERP studies of syntactic processing are still relatively limited in number, on the whole, the results suggest that the ERP responses to violations of syntactic preferences are qualitatively different from the classical N400 (Hagoort et al., 1993; Kluender & Kutas, 1993; Kutas & Hillyard, 1983; Münte et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992; Rösler et al., 1993).

Existing electrophysiological studies of sentence processing and parsing suggest at least two candidate ERP effects that appear to be related to syntactic analysis: (a) a negative shift that is maximal over left-anterior recording sites (LAN); and (b) a large, broad, symmetric, positive-going shift that has been variously labeled the *P600* (Osterhout & Holcomb, 1992) or the syntactic positive shift (SPS; Hagoort et al., 1993).

Frontal negativities (as well as a small P600) were observed for the first time by Kutas and Hillyard (1983) to words indicating mismatches in number agreement between an adjective and a noun (e.g., "six apple" instead of "six apples") or to illegal tense marking (e.g., "Ice begins to grew" instead of "Ice begins to grow"). Neville et al. (1991) observed LAN effects for violations of phrase-structure constraints, which were realized by changing the obligatory word order of the head noun and a preposition in a noun phrase (e.g., "Ted's about films America"). More complex patterns of results were observed to two additional violation types in the Neville et al. study. Finally, LAN effects were also observed by Kluender and Kutas (1993) in a number of different sentence types, including filler-gap constructions.

Because the extent to which these various LANs are related is unclear, a unifying account of what leads to their elicitation must await further research. It will be especially important to determine what underlies the reported variations in the onset and distributional characteristics of the LAN effects. What is clear, however, is that LAN effects are qualitatively different from the typical N400 effects seen following violations of semantic constraints.

A clearer picture emerges for the SPS/P600 that has been observed to diverse syntactic violations in English and Dutch (Hagoort et al., 1993; Osterhout & Holcomb, 1992). In one of their conditions, Hagoort et al. compared ERPs to Dutch sentences that violated the agreement between the subject NP and the finite verb, as in the following example sentences (literal translation in English between brackets; the word that renders the sentence ungrammatical (the Critical Word [CW]) and its counterpart are italicized):

Het verwende kind *gooit* het speelgoed op de grond.
(The spoiled child *throws* the toys on the floor.)
*Het verwende kind *gooien* het speelgoed op de grond.
(The spoiled child *throw* the toys on the floor.)

The basic pattern of results that we observed is shown in Fig. 3.7 for the posterior midline site (Pz). The CW is preceded by two words and followed by three words.

As can be seen, the ERP waveform to the incorrect CW shows a positive shift in comparison with its correct counterpart. This positive shift is widely distributed over the recording sites and has a centro-parietal maximum. The onset of the positive shift, which we labeled *SPS* (i.e., Syntactic Positive Shift), is at about 500 ms after presentation of the incorrect CW. As can be seen, the SPS is replaced by a negative shift on word positions following the CW. These are N400 effects.

A similar pattern of results is obtained for a completely different syntactic violation. In the phrase-structure violation, the obligatory word order in Dutch of adjective–adverb–noun sequences was violated by changing the order of the adjective and the adverb, as in the following example sentences (the CW is italicized):

De echtgenoot schrikt van de nogal emotionele *reactie* van zijn vrouw.
(The husband [is startled] by the rather emotional *response* of his wife.)

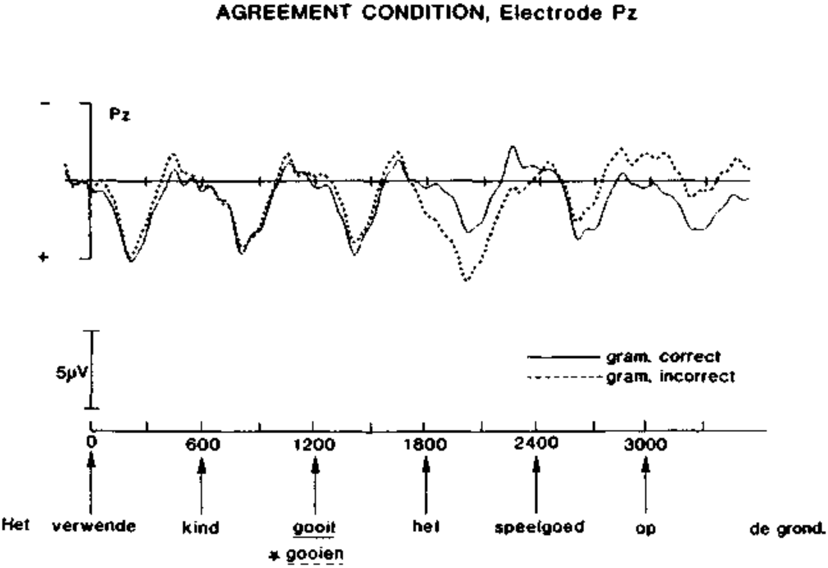

FIG. 3.7. Agreement condition, normal prose. Grand-average waveform for electrode site Pz, for the grammatically correct and incorrect CWs. The CW is preceded by two and followed by three words. The translation of the example sentence is "The spoilt child throws/throw the toy on the ground."

*De echtgenoot schrikt van de emotionele nogal *reactie* van zijn vrouw.)
(The husband [is startled] by the emotional rather *reaction* of his wife.)

For these sentences, again, an SPS was observed, followed by a negative shift to words after the noun that rendered the sentence ungrammatical in the incorrect version. However, there was one major difference with the agreement violation. For the phrase-structure violation, the SPS was already observed to the adverb preceding the noun. At this position, the sentence still could have been continued in a syntactically legal way by adding another adjective (e.g., "the emotional rather aggressive reaction").

This suggests that the positive shift is not an ERP response to violations only. Therefore, we have proposed that the SPS is elicited to the word in the sentence that renders the assignment of the preferred structure (e.g., the less complex, more frequent one) impossible.

To test the independence of SPS and N400 effects, we ran a follow-up study in which sentential-semantic constraints are reduced as much as possible. This was done by constructing syntactic-prose versions of the sentences that were used in the previous study. These sentences are structurally identical to the ones from which they were derived, but are constructed in such a way as to be semantically uninterpretable. In the syntactic-prose version of the experiment, the same syntactic-violation types are used as in the normal-prose version. If SPS and N400 are related to qualitatively different processing events, this experiment should dissociate them because, in the syntactic-prose sentences, we do not expect the syntactic violations to have detectable consequences for (attempts to construct) a semantic interpretation of the sentences. Therefore, in this experiment we expected to see an SPS to the syntactic violations, but no N400 effects for words following the violations.

In addition, this experiment should give us more information about possible reasons for the absence of an SPS to the third kind of syntactic violation we tested in the previous experiment. In this subcategorization violation, we violated the constraint that obligatory intransitive verbs cannot take a noun as direct object. The following sentences are examples of this violation (the CW is italicized):

De zoon van de rijke industrieel leent de *auto* van zijn vader.
(The son of the rich industrialist borrows the *car* of his father.)
*De zoon van de rijke industrieel pocht de *auto* van zijn vader.
(The son of the rich industrialist boasts the *car* of his father.)

For this violation, no effect was observed to the incorrect CW in comparison with the CW in the correct companion sentences. Thus, this violation behaved differently than the other two. However, some consequences of this violation showed up in the N400 effects to words following the CW, reminiscent of the N400 effects for the other two syntactic-violation types.

We speculated that the absence of an SPS might result from the occurrence of an SPS and an N400 effect in the same latency range. Because these are two opposing effects in terms of their electrical polarity, they cancel each other out, with the absence of a significant difference on the CW as the net result. The reason that these opposing effects occur in the same latency range might be due to the intricate relationship between the verb's semantic specifications and its subcategorization frame. Recent empirical evidence suggests that part of the verb's semantic specifications are encoded in its subcategorization frame (Fisher, Gleitman, & Gleitman, 1991). This fits with several linguistic accounts claiming that subcategorization frames are relatively straightforward projections from certain semantic features (Bresnan, 1979; Chomsky, 1981; Jackendoff, 1978).

Exactly how intricate the relationship between subcategorization frame and semantic specifications is could become clear in the syntactic-prose version of this experiment, by removing the semantic consequences of the syntactic violation with respect to the overall interpretation of the sentence. This was another reason for running a syntactic-prose version of the experiment reported in Hagoort et al. (1993).

Method

Three hundred and sixty Dutch sentences were constructed. All sentences were derived from the set of sentences in the normal-prose version of the experiment (Hagoort et al., 1993). For each sentence in the normal-prose version, the lexical items were replaced by other lexical items of the same word class. The replacements were chosen so as to make the sentences semantically uninterpretable. That is, the usual semantic context constraints no longer applied in these syntactic-prose sentences. However, all of the sentences had the same constituent structure as their source sentences in the normal-prose version of this experiment.

As in the normal-prose experiment, half of the sentences are grammatically correct, and half contain a grammatical violation. Each sentence in the violated set is derived from a sentence in the correct set, such that the only difference with the companion correct sentence is the word violating the syntactic constraints. Three kinds of grammatical violations are used: (a) violation of verb–noun number agreement, (b) violation of verb subcategorization, and (c) violation of phrase structure.

The *agreement violations* consist of number violations between verbs and nouns within subject–verb–object (SVO) verb–subject–object (VSO) sentences. The following example gives both the grammatically correct and incorrect version of an SVO agreement violation (literal translation in English between brackets; the CW and its correct counterpart are in italics):

De gekookte gieter *rookt* de telefoon in de poes.
(The boiled watering-can *smokes* the telephone in the cat.)
*De gekookte gieter *roken* de telefoon in de poes.
(The boiled watering-can *smoke* the telephone in the cat.)

The *subcategorization violations* involve obligatory intransitive verbs followed by a noun that has to be assigned the grammatical role of direct object. The correct companion sentence contains a transitive verb at the CW position. For example:

De haargrens in de gewassen boterham leent de *wortel* van zijn krant.
(The hair-line in the washed bread borrows the *root* of his newspaper.)
*De haargrens in de gewassen boterham pocht de *wortel* van zijn krant.
(The hair-line in the washed bread boasts the *root* of his newspaper.)

The *phrase-structure violations* consist of nouns preceded by transpositions of adverbs and adjectives. In Dutch, like English, it is a violation of phrase-structure constraints to have a noun preceded by an adjective–adverb sequence. For example (the CW is italicized):

De hiel valt over de nogal bewoonde *poes* op zijn broekzak.
(The heel tripped over the rather inhabited *cat* on his pocket.)
*De hiel valt over de bewoonde nogal *poes* op zijn broekzak.
(The heel tripped over the inhabited rather *cat* on his pocket.)

Note that the actual violation occurs on the noun following the adverb (i.e., on *poes*) because the adjective–adverb sequence can be part of a larger and grammatically legal adjective–adverb–adjective–noun sequence (e.g., "the inhabited rather talkative cat"). However, in the normal-prose experiment, an SPS was already obtained to the adverb preceding the noun. This is probably because the adverb forces the parser to entertain the possibility of the more complex (less frequent) and, therefore, less preferred adjective–adverb–adjective–noun structure. If this account is correct, also in this experiment the SPS should already be observed to the adverb (i.e., *nogal*) that precedes the noun (i.e., *poes*) in the CW position.

The additional criteria that had to be met in constructing the materials, and the way in which the materials were divided over two lists and three blocks, were exactly the same as in the normal-prose experiment (for details, see Hagoort et al., 1993).

Grammaticality Judgment Pretest

Before running the ERP experiment, the test sentences were pretested in a grammaticality judgment experiment, using a Go/NoGo task, in which subjects were instructed to respond whenever they detected a grammatical violation. The purpose of this pretest was to establish whether subjects were as sensitive to the three types of violations in syntactic-prose sentences as they had been for the normal-prose sentences.

The sentences were displayed word by word in the center of a high-resolution computer screen. Each word was presented for 200 ms, and replaced by the next word in the sentence after a 500-ms blank-screen period. The subjects were told that the sentences they had to read were difficult to understand, but that nevertheless they should try to read each sentence for comprehension. In addition, subjects were instructed to press a button whenever they encountered a grammatical error.

Table 3.1 summarizes the results of the grammaticality pretest for the syntactic-prose sentence and those on a parallel grammaticality pretest for the normal-prose version of these sentences. The percentages indicate the number of times that subjects detected a violation of a certain type on either the CW or the word following it.

In general, compared with the normal-prose sentences, subjects were a little less but still highly accurate in the on-line detection of agreement violations and phrase-structure violations that were embedded in syntactic-prose sentences. In

TABLE 3.1
Performance on the Grammaticality Judgment Task

Condition	*Normal Prose (%)*	*Syntactic Prose (%)*
Agreement Violation	90	73
Subcategorization Violation	74	21
Phrase-Structure Violation	86	73

Note. Percentage of violation detections at the CW position and the following word position for the normal and syntactic prose experiment.

contrast, the performance on the subcategorization-violation sentences decreased dramatically for the syntactic-prose version. This suggests that, unlike with agreement violations and phrase-structure violations, subjects do not recognize the subcategorization violations as purely syntactic in nature. This fits well with the empirically supported claim that part of the verb's semantic specifications are encoded in its subcategorization frame (Fisher et al., 1991). Because the subcategorization violations we created are not only syntactic violations but also semantic violations, they probably are not recognized as different from the other words in the sentences that also violated the standard semantic constraints (such as selectional restrictions).

For the agreement and phrase-structure violations, however, the results of the grammaticality judgment pretest indicate the relatively immediate salience of the ungrammaticalities even in sentences that are difficult to interpret semantically.

The ERP Experiment

Display of the stimuli, including presentation durations of the words, was identical to that of the ambiguity experiment and the normal-prose version of the current experiment. Subjects were informed that they would see sentences that were difficult to understand, but that they nevertheless should read for comprehension. No additional task demands were imposed. Subjects were told that some sentences would be grammatically incorrect, but they were given no information concerning the kinds of grammatical errors that would occur.

Results

For all subjects (N = 40), average waveforms are computed over all artifact-free trials, for the correct and incorrect sentences of the three violation types separately. Mean amplitudes are calculated per subject, per condition, and per electrode for the critical time ranges given later.

To check whether the differences obtained in the critical time ranges might be due to some spurious effects, we also analyzed two word-epochs preceding the critical areas for each of the three violations. Significant differences were not obtained in any of these cases. This further substantiates the claim that the ERP differences between the CWs in the correct and incorrect sentences are real.

Agreement. Figure 3.8 shows the grand-average waveforms for the three midline sites[3] (Fz, Cz, and Pz) in the correct and incorrect agreement conditions. As can be seen in the waveforms, an SPS emerges to the CWs in the incorrect version compared with the CWs in the correct version. The SPS starts at around 500 ms following the onset of the CW, and continues throughout the following word.

To test the SPS to the incorrect CWs, an ANOVA was performed on the mean amplitudes in the 500–1200-ms range following the onset of the CW. This includes the positivity to the incorrect CW itself and its carry-over effect into the processing range of the next word.[4] The analysis yielded a main effect of Grammaticality [$F(1, 39) = 6.82$, MSe = 26.04, $p = .013$].

Although the SPS in the normal-prose sentences was followed by a negative shift toward the end of the sentences, this negative shift was absent in the incorrect agreement sentences in this experiment. Analyses on penultimate and sentence-final words did not result in significant N400 effects. We return to this difference between the normal-prose and the syntactic-prose experiments in the discussion.

Subcategorization. Figure 3.9 shows the grand-average waveforms for the three midline sites (Fz, Cz, and Pz) in the correct and incorrect subcategorization condition. Inspection of the waveforms suggests the absence of any difference between the correct and incorrect conditions, both to the CW and to the words preceding and following it. A statistical analysis on the mean amplitudes in the 500–1200-ms range following the onset of the CW did not result in a significant effect of Grammaticality. Just as in the normal-prose experiment, no SPS was manifest on the noun following the obligatorily intransitive verbs.

Although in the normal-prose experiment increased N400 amplitudes were

---

[3]For all three violation types, the recording sites over the left and right hemisphere showed the same pattern of results as the midline sites. In addition, no hemispheric differences were obtained in the size of the effects.

[4]In the normal-prose version of this experiment, a smaller latency window was used for statistical analysis (see Hagoort et al., 1993). This window went from 500 ms after onset of the CW (or the word preceding the CW in the phrase-structure condition), until 700 ms, which is approximately until the N1 to the following word. Certainly in the current experiment, with the absence of negative shifts following the SPS, the positivity was much more extended than this small 200-ms period. However, we also analyzed the results for the mean amplitude in this reduced-latency window. For this reduced window, the effects of Grammaticality failed to reach significance in the Agreement condition. In the phrase-structure condition, the effect of Grammaticality was significant [$F(1, 39) = 11.28$, MSe = 16.29, $p = .002$].

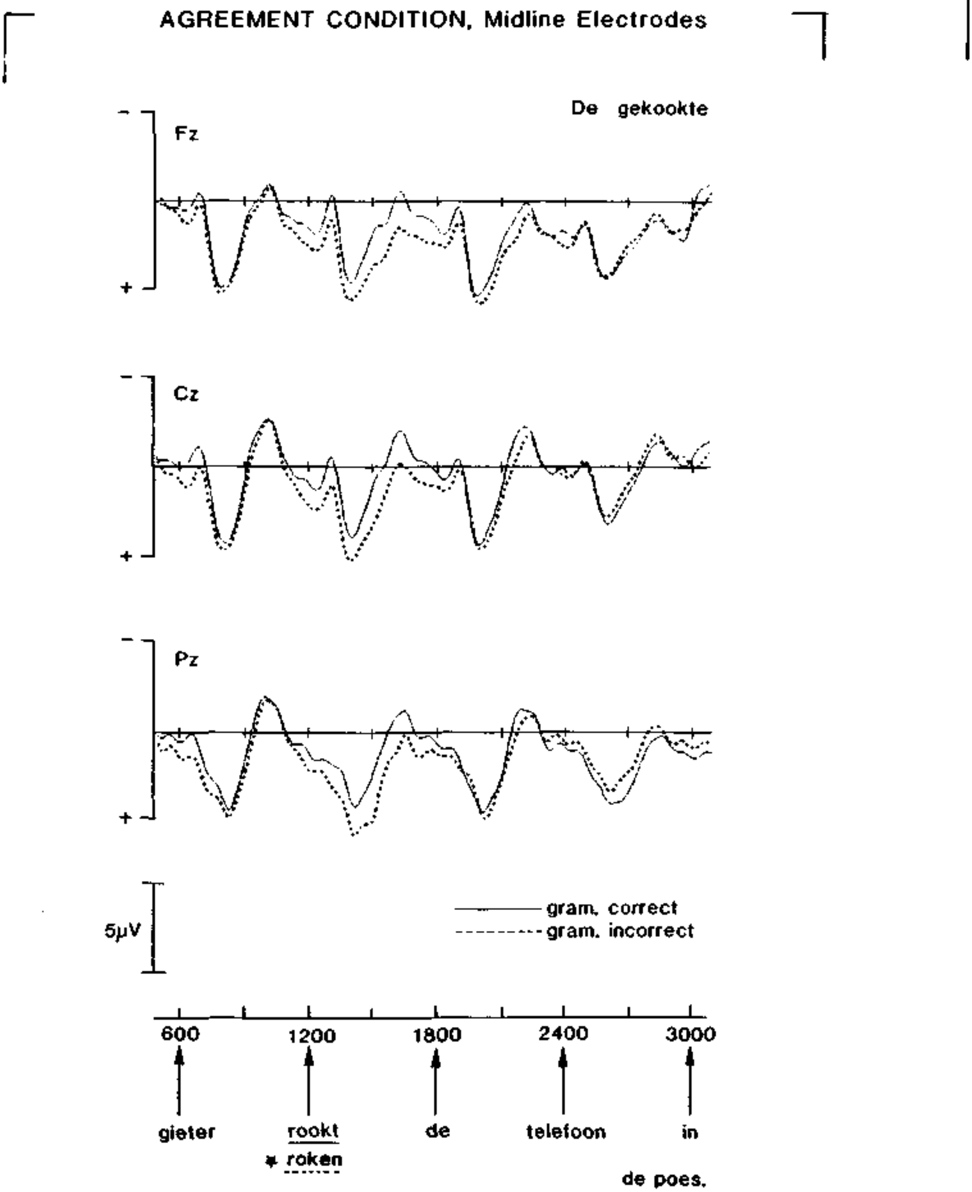

FIG. 3.8. Agreement condition, syntactic prose. Grand-average waveform for each of the three midline electrode sites, for the grammatically correct and incorrect CWs. The CW is preceded by one and followed by two words. The translation of the example sentence is "The boiled watering-can smokes/smoke the telephone in the cat."

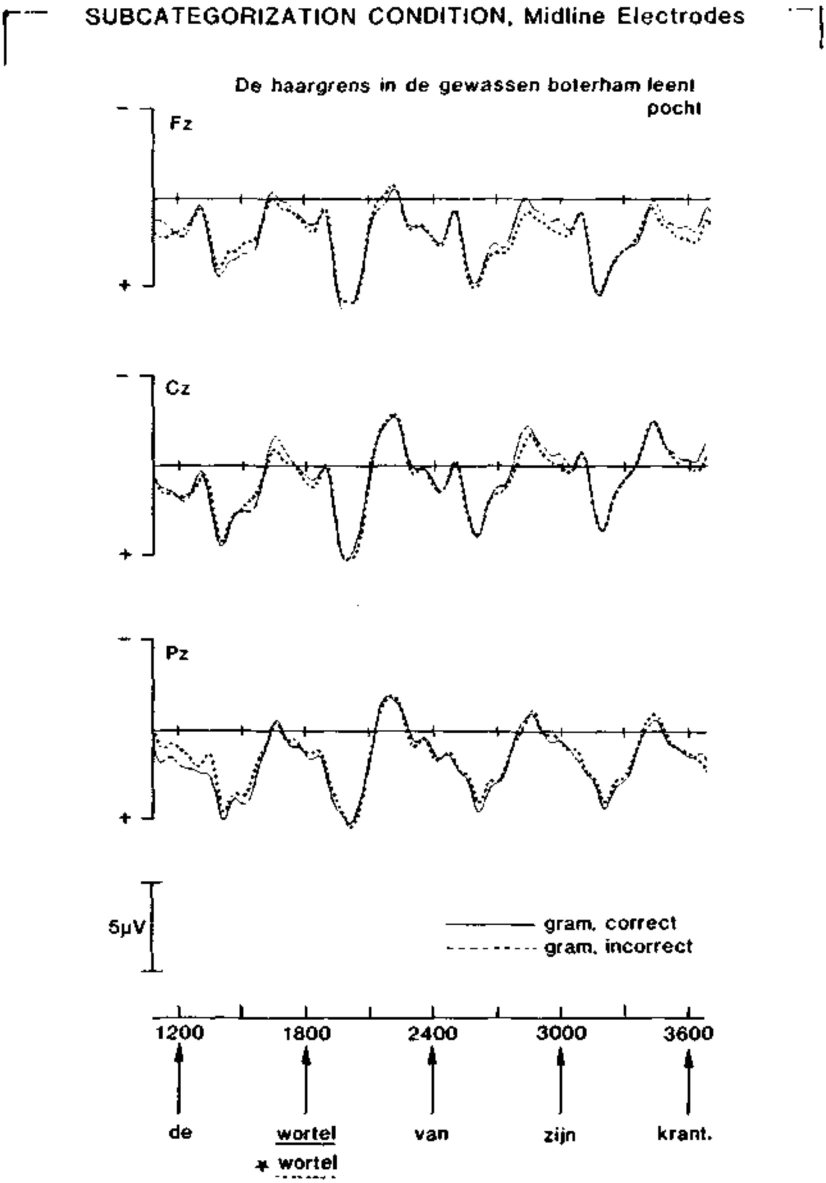

FIG. 3.9. Subcategorization condition, syntactic prose. Grand-average waveform for each of the three midline electrode sites, for the grammatically correct and incorrect (CWs. The CW is preceded by one and followed by two words. The translation of the example sentence is "The hair-line in the washed bread borrows/boasts the carrot of his newspaper."

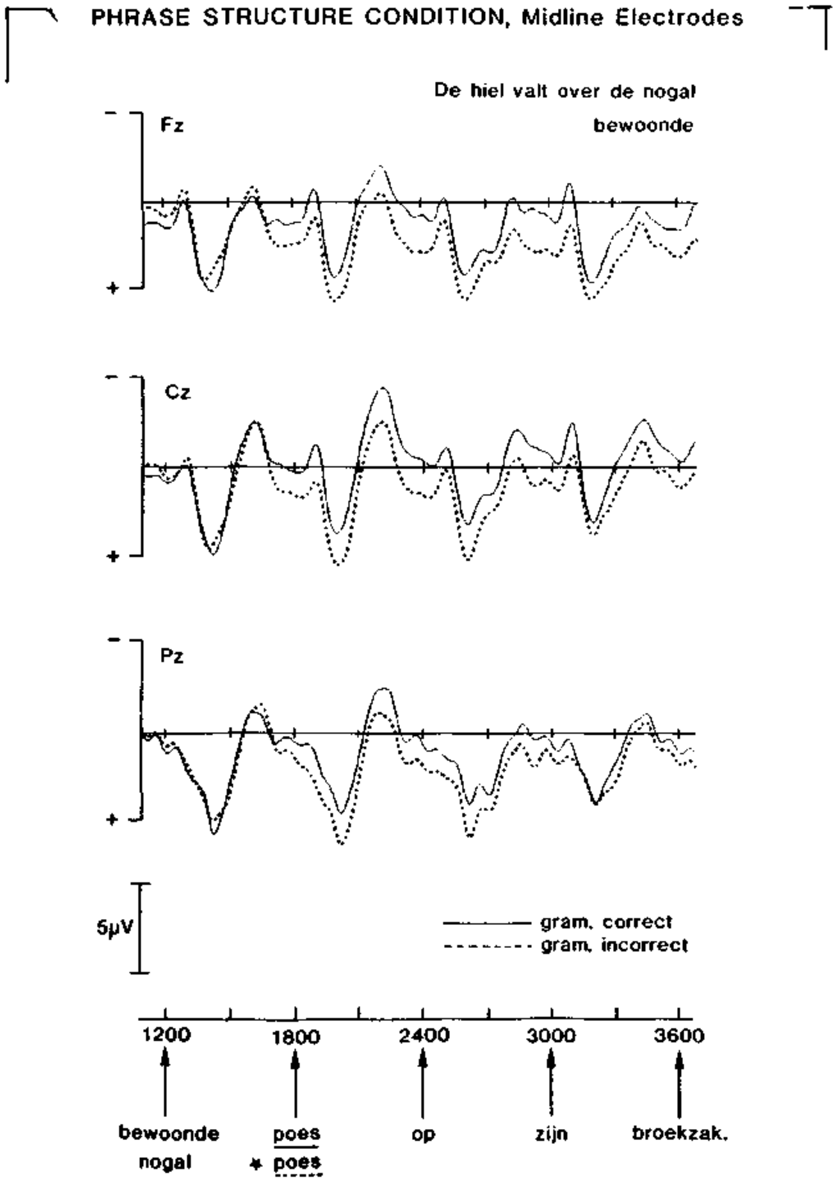

FIG. 3.10. Phrase-structure condition, syntactic prose. Grand-average waveform for each of the three midline electrode sites, for the grammatically correct and incorrect CWs. The CW is preceded by one and followed by two words. The translation of the example sentence is "The heel tripped over the rather inhabited/inhabited rather cat on his pocket."

seen to the penultimate and sentence-final words in the ungrammatical condition, this negative shift was absent in the syntactic-prose sentences. In summary, then, the subcategorization violation did not result in any visible difference between the waveforms of the grammatically correct and incorrect conditions.

Phrase Structure. Figure 3.10 shows the grand-average waveforms for the three midline sites (Fz, Cz, and Pz) in the correct and incorrect phrase-structure conditions. The waveforms show that a positive shift is present in the incorrect phrase-structure sentences compared with their correct counterparts. As in the normal-prose experiment, this positive shift is already elicited by the adverb that precedes the CW. The positivity carries on throughout the epoch of the CW into the following word. An ANOVA was performed for a window that started 500 ms after onset of the adverb preceding the CW and included the CW epoch until the onset of the word following the CW (i.e., 1200 ms after onset of the word that preceded the CW). This analysis results in a highly significant effect for Grammaticality [$F(1, 39) = 10.74$, MSe = 15.83, $p = .002$].

No differences between the two conditions were observed for penultimate and sentence-final words. This contrasts with the negativities at the same word positions in the incorrect normal-prose version of the sentences.

Discussion

The first major result of this study is the widely distributed positivity that is elicited by two of the three types of syntactic violations. This effect is very similar to the SPS that we obtained in the normal-prose version of this experiment, and to the P600 reported by Osterhout and Holcomb (1992).

Figure 3.11 shows the difference waveforms between the grammatically incorrect and the grammatically correct conditions for the agreement and phrase-structure violations. Difference waveforms are presented for the syntactic-prose experiment and for the normal-prose experiment to allow for a comparison of the results in both experiments. The difference waveforms give a straightforward picture of the commonalities and differences in effects between the normal-prose and the syntactic-prose experiment for the two syntactic violations that show an SPS.

A comparison of the two difference waveforms shows two aspects worth mentioning. The first one is that, to the very same word positions for the very same violation types (i.e., the agreement violation and the phrase-structure violation), an SPS is observed in both normal-prose and syntactic-prose sentences. The effect is slightly smaller in the syntactic prose, but has the same onset latency (500 ms after word onset) for the two prose types.

The second aspect is that in the syntactic prose the absence of a negative shift following the SPS near the sentence is striking because a negative shift is so clearly present in the normal-prose difference waveforms. In the normal-prose

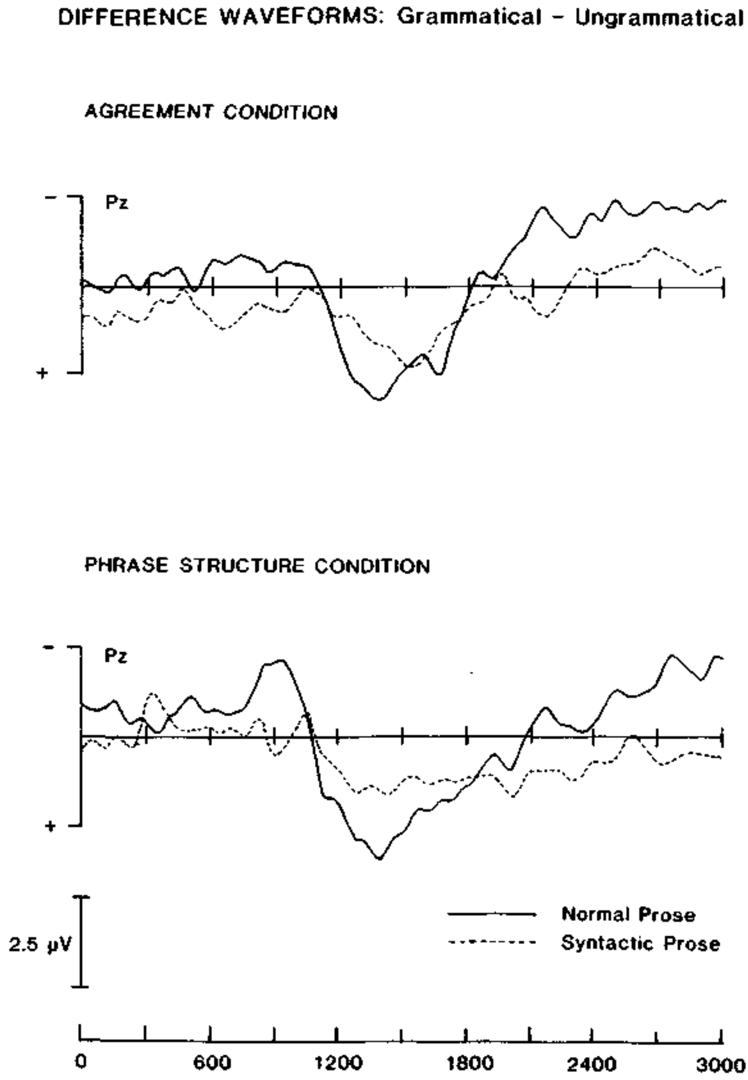

FIG. 3.11. Difference waveforms at electrode site Pz between the grammatically incorrect and correct conditions for the agreement and phrase-structure violations, in the normal-prose and syntactic-prose experiments.

experiment, it was even present for the subcategorization-violation condition, which did not reveal an SPS.

A number of conclusions follow from the results of this experiment, in combination with the results obtained for the normal-prose sentences. First, the SPS that we obtained in this experiment to two different syntactic violations occurred in the absence of N400 effects to words following the syntactic violations. This further substantiates the claim that the SPS is qualitatively different from the negative shift (the N400), which we believe to be especially sensitive to semantic integration processes (Brown & Hagoort, 1993). That is, the processing of syntactic information has a neurophysiological signature that is clearly different from that for the processing of semantic information. This result is difficult to account for in sentence-processing models that deny that qualitatively different constraints (i.e., syntactic and semantic) make qualitatively different contributions to the construction of an interpretation for the whole utterance (e.g., McClelland, St. John, & Taraban, 1989).

Second, the results of the phrase-structure condition show that the presence of a syntactic violation is not a precondition for the SPS to be elicited. In the phrase-structure condition, the SPS was observed to the word that rendered the assignment of a preferred structure (i.e., determiner–adjective–noun) impossible. The adverb following the adjective forces the parser to entertain the possibility of an alternative structure, which is the less frequent and more complex NP structure (i.e. determiner–adjective–adverb–adjective–noun). The result of the phrase-structure condition fits with the proposal that, by default, the parser assigns only one structure to the incoming string of words. This preferred structure is determined on the basis of some computational economy principle (see Frazier, 1987), or on the basis of the frequency of alternative syntactic constructions. The preferred structure gets revised if it is rendered untenable by further incoming words. In general, the SPS seems to arise to the word in the sentence that indicates that the preferred structural assignment is an incorrect syntactic analysis for the incoming string of words.

The presence of N400 effects to words more or less immediately following the syntactic violation in the normal-prose sentences, and their absence in syntactic prose, can be explained in terms of recent proposals about the processing nature of the N400 (Brown & Hagoort, 1993; Osterhout & Holcomb, 1992). According to these proposals, the N400 is especially sensitive to the integration of lexical meaning into an overall representation of the word or sentence context. With the normal sentential-semantic context constraints in place, the syntactic violations seem to have an immediate consequence for the semantic integration of following words into a coherent overall message-level representation of the whole sentence. This integration process becomes more difficult, resulting in an increase of the N400 to words following the syntactic violation. The situation is clearly different for the syntactic-prose sentences. The absence of sentential-semantic constraints probably prevents the construction of a coherent message-level representation. Therefore, the semantic integration process might be extremely difficult, if not impossible. In these circumstances, the syntactic violation probably has no additional disadvantageous consequences for semantic integration processes. Therefore, the syntactic violation does not lead to increased N400 amplitudes to words further downstream.

Finally, the absence of an SPS to subcategorization violations replicates the result for the normal-prose sentences. We explained the absence of an effect for Grammaticality in the normal-prose version of this experiment as resulting from the opposing overlapping effects of an SPS and an N400 in the same latency range. The opposite polarity of these two effects results in the disappearance of both effects in the averaged waveforms. We speculated that, unlike the agreement and phrase-structure violations, in the subcategorization-violation condition the CW renders the sentence ungrammatical via its semantic properties. By removing to a large extent the semantic constraints in the syntactic-prose sentences, we hoped to get a clearer picture of the purely syntactic consequences of subcategorization violations. However, the results of the grammaticality pretest already suggested that this would not work. Subjects are not able to detect subcategorization violations as purely syntactic anomalies. This confirms our earlier suggestion (Hagoort et al., 1993) that verb meaning and the syntactic aspects of the verb that are specified on the subcategorization frames are tightly intertwined (see Fisher et al., 1991, for empirical support). Therefore, the subcategorization violations that we used in both studies are not only syntactic violations, but by necessity also semantic violations. As a result, in syntactic prose they probably get processed in the same way as the semantically anomalous prose in which they are embedded.

In conclusion, the results of the syntactic-prose experiment further substantiate the existence of 'syntactic' ERP components within the domain of language-related ERP effects. The SPS that we observed might either reflect the computation of a separate level of syntactic representation during the process of language understanding, or the initiation of a syntactic reanalysis after a first-pass structural assignment has failed to provide a well-formed structure. Further research is needed to specify the exact processing nature of the SPS within the context of current parsing proposals. However, the clearly syntactic nature of the SPS holds promise for its use in testing more subtle differences between competing parsing theories.

CONCLUDING REMARKS

In this chapter, we have presented ERP data from two experiments that address very different issues in language comprehension research. Both issues, however, have in common that they are related to higher order integration processes. Although in recent years there has been something of an upsurge in experimental work on syntactic and semantic integration processes, on the whole, systems that lie beyond the mental lexicon have not been the focus of on-line investigation, in part because the existing RT techniques pose problems in tracking the comprehension process as it develops across the sentence or discourse. With the advent of eye trackers, a first continuous record was obtained of the normal reading process, in the absence of irrelevant task demands. We hope to have demonstrated that the registration of ERPs presents an additional and insightful tool with which to observe the language comprehension system as it operates in real-time. However, as we pointed out in the general introduction, it is important to emphasize that the present ERP data are but first steps in a psycholinguistic research program on brain manifestations of sentence processing. Before the full potential of the signal characteristics of ERPs can be realized within psycholinguistics, a number of issues need to be addressed, two of which we mention here.

A first issue concerns the relatively slow presentation rates that have been used in the ERP experiments reported here (and in general in the ERP and language literature), which lag far behind the normal reading rate. The main reason for using relatively slow presentation rates is that this minimizes the problem of overlapping components in the waveform. However, this is not a principled problem, as has been shown by Kutas (1987). She registered ERPs to semantically congruous and incongruous words in sentence-final position in sentences presented at a rate of 10 words per second (i.e., about twice as fast as the normal reading rate), and obtained essentially the same N400 effect as when the words of the same sentences were presented once every 700 ms. Furthermore, and clearly contrary to claims about nonlinguistic effects of unnatural rates on language-related ERPs, similar N400 effects have been observed in our laboratory and by others for semantically incongruous words in naturally produced connected speech (Connolly et al., 1990; Holcomb & Neville, 1991). So, the available evidence indicates that rate effects do not severely contaminate the ERP results. Nevertheless, it is clear that researchers using the ERP method will have to move toward more standard presentation rates in reading experiments, certainly when focusing on higher order integration processes. At present, we are running an Rapid Serial Visual Presentation (RSVP) version (one word every 250 ms) of the normal-prose version reported in Hagoort et al. (1993) to ensure that the SPS is also present with more normal reading rates. At the same time, we are running a connected speech version of this experiment to see whether the SPS obtains across modalities, which is to be expected on the basis of work by Osterhout and Holcomb (1993) on ERPs and syntactic processing in connected speech. We believe this kind of simultaneous approach is necessary to test the validity of language-related ERP effects.

A second issue concerns the temporal relationship between the real-time electrophysiological signal and the ongoing linguistic analysis. As we pointed out earlier, an appealing characteristic of ERPs is that they are a real-time signal with which to observe a real-time process. However, before ERPs can be used to obtain a truly on-line processing profile of language comprehension, a better understanding is required of the exact time-locking relationships between the ERP waveform and the presumed underlying comprehension process. This implies that we have to come to grips with the complex problem of the exact moment in time at which a particular component emerges in the ERP waveform. For components like the N400 and the SPS, it is relatively straightforward to determine at what moment after stimulus onset they reach their peak amplitude. However, the latencies of these peak amplitudes clearly overestimate the moment in time at which the components have their onset relative to the onset of the linguistic stimulation that elicits them, and it is exactly these onset moments that provide critical information about the time course of the ongoing comprehension process. This aspect of the time-locking issue poses a real challenge for psycholinguists working with the ERP method.

In conclusion, it is clear that several problems have to be solved before all the promises that ERPs hold for psycholinguistics will be obtained. But we believe that it is equally clear that the ERP method is already a very useful and revealing tool with which to investigate language processes.

ACKNOWLEDGMENTS

We would like to thank Aafke Deckers and Jolanda Groothusen for their assistance in all phases of the work reported here, and Lyn Frazier for her comments on an earlier version of this chapter. This research was supported by a grant from the Volkswagen Foundation (Hannover, Germany) and by the Max Planck Society (München, Germany). The authors contributed equally to this publication, and their order of mention is arbitrary.

REFERENCES

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology, 24*, 14–25.

Bresnan, J. (1979). *Theories of complementation in English syntax*. New York: Garland.

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

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris.

Connolly, J. F., Stewart, S. H., & Phillips, N. A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language, 39*, 302–318.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language, 31*, 461–476.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23,* 331–392.

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

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments. Processing multiple meanings vs. multiple senses. *Journal of Memory and Language, 29,* 181–200.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research, 18,* 51–60.

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.

Holcomb, P. J., & Nevill, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19,* 286–300.

Jackendoff, R. (1978). Grammar as evidence for conceptual structure. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 201–228). Cambridge, MA: MIT Press.

Jasper, H. H. (1958). Report to the committee on methods of clinical examination in electroencephalography: Appendix. The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

Kintsch, W., & Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language, 24,* 336–349.

Kluender, R., & Kutas, M. (1993). The interaction of lexical and syntactic effects in the processing of unbounded dependencies. *Language and Cognitive Processes, 8,* 573–633.

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* (pp. [redacted]–[redacted]). Amsterdam: Elsevier.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition, 11,* 539–550.

Kutas, M., Lindamood, T., & Hillyard, S. A. (1984). Word expectancy and event-related potentials during sentence processing. in S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217–238). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing. In R. J. Ellingson, N. M. F. Murray, & A. M. Halliday (Eds.), *The London Symposium* (pp. 325–330). Amsterdam: Elsevier.

Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 3, pp. 139–187). Greenwich, CT: JAI Press.

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *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.

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

Neville, H., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex, 2,* 244–258.

Neville, H., 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.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9*, 225–236.

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. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8*, 413–437.

Pacht, J. M., & Rayner, K. (1993). The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research, 22*, 257–271.

Paller, K. A., McCarthy, G., & Wood, C. C. (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology, 29*, 202–206.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 779–790.

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.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489–537.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120–136.

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316–340.

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 28–39.

Simpson, G. B., & Krueger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language, 30*, 627–643.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)Consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Tabossi, P. (1988a). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27*, 324–340.

Tabossi, P. (1988b). Sentential context and lexical access. In S. L. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution* (pp. 331–342). San Mateo, CA: Morgan Kaufmann.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research, 49*, 161–167.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language, 32*, 359–372.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related brain potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188–208.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18*, 380–393.

Van Petten, C., & Kutas, M. (1991). Electrophysiological evidence for the flexibility of lexical processing. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 129–174). North-Holland: Elsevier Science Publishers.