The Influence of Semantic and Syntactic Context Constraints on Lexical Selection and Integration in Spoken-Word Comprehension as Revealed by ERPs

Daniëlle van den Brink and Peter Hagoort

Abstract

■ An event-related brain potential experiment was carried out to investigate the influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension. Subjects were presented with constraining spoken sentences that contained a critical word that was either (a) congruent, (b) semantically and syntactically incongruent, but beginning with the same initial phonemes as the congruent critical word, or (c) semantically and syntactically incongruent, beginning with phonemes that differed from the congruent critical word. Relative to the congruent condition, an N200 effect reflecting difficulty in the lexical selection process was obtained in the semantically and syntactically incongruent condition where word onset differed from that of the congruent critical word. Both incongruent conditions elicited a large N400 followed by a left anterior negativity (LAN) time-locked to the moment of word category violation and a P600 effect. These results would best fit within a cascaded model of spoken-word processing, proclaiming an optimal use of contextual information during spoken-word identification by allowing for semantic and syntactic processing to take place in parallel after bottom-up activation of a set of candidates, and lexical integration to proceed with a limited number of candidates that still match the acoustic input. ■

INTRODUCTION

During a conversation, we are able to analyze the acoustic information we receive incrementally, select the proper words by mapping the sensory input onto stored lexical knowledge containing phonological, semantic, and syntactic information, extract the meaning of those words, and integrate them into the ongoing sentential or discourse context to come to an overall interpretation of the utterance. The question of how these processes interact remains a matter of debate. In general, the spoken-word recognition process, which lies at the heart of spoken language processing, is assumed to consist of three basic subprocesses: lexical access, lexical selection, and lexical integration. The functional architectures of many spoken-word recognition models describing these subprocesses differ, but both Marslen-Wilsons's cohort model and Norris's short-list model, for instance, pose that during lexical access multiple lexical candidates that share word onset are activated in parallel on the basis of an analysis of the initial phoneme(s) of a word (Norris, 1994; Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978). This is a purely autonomous or bottom-up process. As pronunci-

ation of a word progresses over time, lexical candidates are dropped or become less activated as soon as they no longer correspond to the incoming acoustic signal. Selection of the proper candidate is said to take place when only one candidate is left that matches the acoustic signal best. Word recognition in sentences or discourse additionally requires that the selected word is integrated into a higher order meaning representation of the preceding context.

The key issues relate to how bottom-up activation interacts with contextual information to establish an interpretation of the utterance. At some point during auditory language comprehension, semantic and syntactic properties of a word are evaluated with respect to constraints provided by the preceding context. An important question that needs to be answered pertains to the time course of the use of different sources of higher level context information in relation to lexical information during sentence processing: Does semantic and syntactic processing occur in parallel or are they serial in nature? At the level of auditory sentence comprehension, this dichotomy is represented by two kinds of models. ''Syntax-first'' models, for instance, claim that syntactic processing (i.e., syntactic structure building) takes place before semantic information from the context is used (Friederici, 1995, 2002; Frazier, 1987). These syntax-first or serial models can be contrasted with

Max Planck Institute for Psycholinguistics, F.C. Donders Centre for Cognitive Neuroimaging

constraint-based or interactive models that claim that all sources of information exert their influence directly and in parallel on auditory sentence processing and that none of them have priority over another (cf. Trueswell & Tanenhaus, 1994). A parallel can be drawn at the level of lexical processing: Is semantic and syntactic information from the context used in parallel during spoken-word processing, or does specific syntactic information have priority over semantic information?

Electrophysiology of Language Processing

The present study was designed to investigate the influence of semantic and syntactic context constraints on lexical selection and integration processes. A technique that seems especially suitable for studying brain processes in real time is the recording of event-related potentials (ERPs). ERPs reflect the sum of simultaneous postsynaptic activity of a large population of neurons recorded at the scalp as small voltage fluctuations in the electroencephalogram. Because of its high temporal (millisecond) resolution this technique has proven to be a sensitive measure of real-time language processing. Furthermore, previous studies have shown that ERPs provide qualitatively distinct correlates of semantic and syntactic processes.

The most salient electrophysiological correlate of semantic processing is the well-documented N400 (Kutas & Hillyard, 1980). The N400 is a negative-polarity potential that starts at about 200 to 250 msec, peaks at 400 msec after stimulus onset, and usually has a posterior distribution across the scalp. In sentence or discourse contexts, the N400 amplitude indexes the relative ease with which the meaning of a word is integrated in the preceding context. The better the semantic fit of a word within the sentence frame or larger discourse, the more reduced the amplitude of the N400 (Van Berkum, Zwitserlood, Hagoort, & Brown, 2003; Hagoort & Brown, 2000; Brown & Hagoort, 1993; Holcomb, 1993). The most prominent effect related to syntactic processing is a positive shift occurring in the latency window of 500 to 800 msec after stimulus onset, the P600. Several syntactic violations, such as violations in agreement of number, gender, and case, phrase-structure violations and subjacency violations, but also violation of structural preference elicit a P600 (cf. Friederici, 2002; Hagoort, Brown, & Osterhout, 1999). In addition, problems in syntactic processing have been found to elicit a negative shift that occurs in a relatively early latency window of 100 to 500 msec after stimulus onset, and usually has a left-lateralized anterior distribution across the scalp. The exact functional interpretation and domain specificity of previously mentioned ERPs in general, and the left anterior negativity (LAN) in particular, is under debate. For example, some researchers contribute the LAN to processes of verbal working memory, rather than claiming it to be specifically syn-

tactic in nature (Coulson, King, & Kutas, 1998; Kluender & Kutas, 1993; cf. Hagoort et al., 1999). Nonetheless, in the context of language comprehension the LAN has been observed to be elicited in situations involving several types of syntactic violations, such as morphosyntactic and word category violations (cf. Hagoort et al., 1999; Hahne & Friederici, 1999). Recently, a further subdivision has been proposed between two negativities within this early time window because of a difference in latency and possible functional roles: an early left anterior negativity (ELAN) between 100 and 200 msec after stimulus onset, related to the processing of word category violations, and a somewhat later LAN between 300 and 500 msec after stimulus onset, related to the processing of morphosyntactic violations (Friederici, 2002; Hahne & Jescheniak, 2001).

Serial versus Interactive Lexical Processing

A number of ERP studies that investigated the interaction of syntax and semantics during lexical processing obtained results that point to their time course being serial in nature rather than parallel. A key finding was that words containing a word category violation as well as a semantic violation elicited an anterior negative effect in the range of 100 to 300 msec followed by a P600, but did not elicit an N400 (Hahne & Jescheniak, 2001; Friederici, Steinhauer, & Frisch, 1999, but see Ainsworth-Darnell, Shulman, & Boland, 1998; Friederici, Pfeifer, & Hahne, 1993). The absence of an N400 in these results were interpreted as providing evidence for the claim that certain syntactic aspects can influence semantic lexical integration processes: A word category violation as reflected by the ELAN appeared to block attempts to integrate a lexical candidate into the preceding context.

A recent serial model that incorporates these findings is a neurocognitive model of sentence processing proposed by Friederici (Friederici, 2002; Friederici et al., 1999). Although presented as a serial or syntax-first model of auditory sentence processing, it makes specific claims at the level of spoken-word processing. An attractive feature of this model is that it specifies the temporal structure of lexical processing in terms of three phases. During the first phase, from 100 to 300 msec, the initial syntactic structure is said to be formed on the basis of word category information. In the second phase, from 300 to 500 msec, thematic role assignment takes place on the basis of lexical-semantic and morphosyntactic processes. In the third phase, from 500 to 1000 msec, integration of the different types of information occurs. The claims are that syntactic-phrase structure building (on the basis of word category information only) is entirely autonomous and precedes semantic processing, and that these processes interact only in the third phase.

However, the temporal characteristics of Friederici's model are primarily based on studies where ELANs were obtained in critical words that were presented auditorily

and that contained clear morphosyntactic markers at the beginning of the critical word indicating the word category. For instance, Hahne and Jescheniak (2001) and Friederici et al. (1993) used auditorily presented sentences like "Die Birne wurde im *gepflückt*" ("The pear was being in-the *plucked*") or "Die Freund wurde im *besucht*" ("The friend was being in-the *visited*") were the morphosyntactic markers *ge-* and *be-* in combination with the preceding auxiliary "wurde" are far more likely to lead to a past participle than a noun. One can question the validity of the temporal parameters of Friederici's model, and possibly even of the claim that word category information is processed earlier than semantic information (Phase I vs. Phase II) if the specifications of the model are primarily based on studies in which word category information was always encoded in the prefix, and thus became available sooner than lexical-semantic information, encoded in the stem.

In most languages information about the word category often is encapsulated in the suffix of a word or in some cases even remains ambiguous. Contrary to many studies giving evidence to the claim that semantic processing acts on a small part of the perceived word, syntax-first models in general, would in these cases predict that semantic processing is postponed until after the word has been heard completely and information about the word category has become available. However, because of its specification of temporal characteristics of lexical processing, Friederici's three-phase neurocognitive model appears to have more difficulty predicting the time course of lexical processing in these cases: Here, word category information embedded in the suffix does not become available during Phase I with temporal parameters of 100 to 300 msec, and according to the model, Phase II cannot be initiated if information about the word category has not been processed.

Now, consider the following account of an interactive model of spoken-word recognition (cf. Zwitserlood, 1989; Marslen-Wilson, 1987): After bottom-up activation of a number of lexical candidates on the basis of an analysis of the initial phoneme(s) of the spoken word (100 to 150 msec of acoustic signal) top-down context information starts to exert its influence. At this moment any information available (i.e., phonological, semantic, and syntactic) is used to assess which candidate is most compatible with both form and content constraints. This candidate is then integrated in the sentence context. If, however, in case of an anomalous word none of the lexical candidates fit the context well, selection of the proper candidate is difficult and can only be done on the basis of the acoustic information. Both this interactive model of spoken word processing and syntax-first models in general, predict that auditorily presented words that violate word category constraints, but lack clear morphosyntactic markers as initial phonemes, would elicit a LAN following the late moment of violation detection. This also holds for the syntax-first model

proposed by Friederici, provided violation detection can take place during the first 300 msec after word onset. However, contrary to syntax-first models, interactive models such as the one described earlier would predict that in the case of a word that violates semantic as well as syntactic constraints set by the preceding sentence context, prior to the LAN, an N400 would possibly be elicited due to semantic information already being available and processed before information about the word category has become available.

The Present Study

In the present study the time course of phonological, semantic, and syntactic processing in spoken-word recognition was evaluated with ERPs. We used semantically constraining Dutch sentences with critical words that differed across three conditions. In the fully congruent (FC) condition, the critical word was a non-sentence-final, highest cloze probability word.[1] In a second condition called the fully incongruent (FI) condition, the critical word was replaced by a verb in the past tense, that apart from syntactic constraints also violated semantic constraints and had initial phonemes that differed from the highest cloze probability word. In addition to these two conditions we were also interested in the ERP waveform to a condition where the onset of the word category violation resembles that of a semantically and syntactically congruent word. This third condition, therefore, contained a verb in the past tense that began with the same initial phonemes as the highest cloze probability word in the FC condition. This condition was called the initially congruent (IC) condition. An example of the experimental materials is given in Table 1. In both examples of target words in FI and IC conditions, the earliest moment at which the listener is able to detect that the word is actually a verb instead of a noun, from here on referred to as the category violation point (CVP), lies at the onset of the word-final syllable *de*. For instance, at the onset of the word-final syllable *de* in *kliederde*, in Dutch the word *kliederboel* (*mess*) is still a possible noun continuation.

Two previous studies using experimental materials that manipulated semantic processing have shown that presenting semantically incongruent words with onsets that, given a semantically constraining preceding sentence frame, differ from those of the highest cloze probability words, elicit a negative effect between 150 to 250 msec after stimulus onset, related to phonological processing. This N200 effect is assumed to reflect the lexical selection process, where word-form information resulting from an initial phonological analysis and content information derived from the context interact (Van den Brink, Brown, & Hagoort, 2001; Van den Brink, Hagoort, & Brown, submitted). Note that in these previous studies all critical words elicited an N200,

Table 1. Examples of the Stimulus Materials

Condition	Sentence
Fully congruent (phon +, sem +, syn +)	Het vrouwtje veegde de vloer met een oude **bezem** gemaakt van twijgen
	(The woman swept the floor with an old **broom** made of twigs)
Initially congruent (phon +, sem −, syn −)	Het vrouwtje veegde de vloer met een oude **bedelde** gemaakt van twijgen
	(The woman swept the floor with an old **begged** made of twigs)
Fully incongruent (phon −, sem −, syn −)	Het vrouwtje veegde de vloer met een oude **kliederde** gemaakt van twijgen
	(The woman swept the floor with an old **messed** made of twigs)

Note. The critical words are typed in **boldface**. The terms "phon", "sem", and "syn" stand for phonological, semantic, and syntactic. The "+"and "−" refer to congruent and incongruent in relation to the highest cloze probability word for that sentence. For example, "phon+" means that the onset phonemes of the critical word are congruent, and "sem−" means that the critical word is semantically incongruent.

irrespective of whether the initial phonemes were identical to those of a congruent word or not. Analogous to the functional interpretation of the N400, the amplitude of the N200 is believed to be indicative of whether the initial assessment of form-based activated lexical candidates reveals the presence of a lexical candidate that on the basis of its semantic and syntactic characteristics fits the contextual specifications. A small N200 is elicited when the set contains a candidate that fits the sentence context, a large N200 indicates that the set does not contain a lexical candidate that fits the preceding sentence frame well. Adding the third condition would enable us to investigate the influence of compatible initial phonemes of semantically and syntactically incongruent words on spoken-word recognition.

Following the interactive account of spoken-word processing we hypothesized that all three conditions should elicit an N200 component, with an amplitude that is largest in the FI condition and does not differ between the FC and IC conditions. In the two semantically and syntactically anomalous conditions this N200 component would be followed by an N400, a LAN, and a P600. The waveforms in the N400, the LAN, and P600 intervals would not be expected to differ between the IC and FI conditions, because both are equally anomalous, both in the semantic as well as in the syntactic sense.
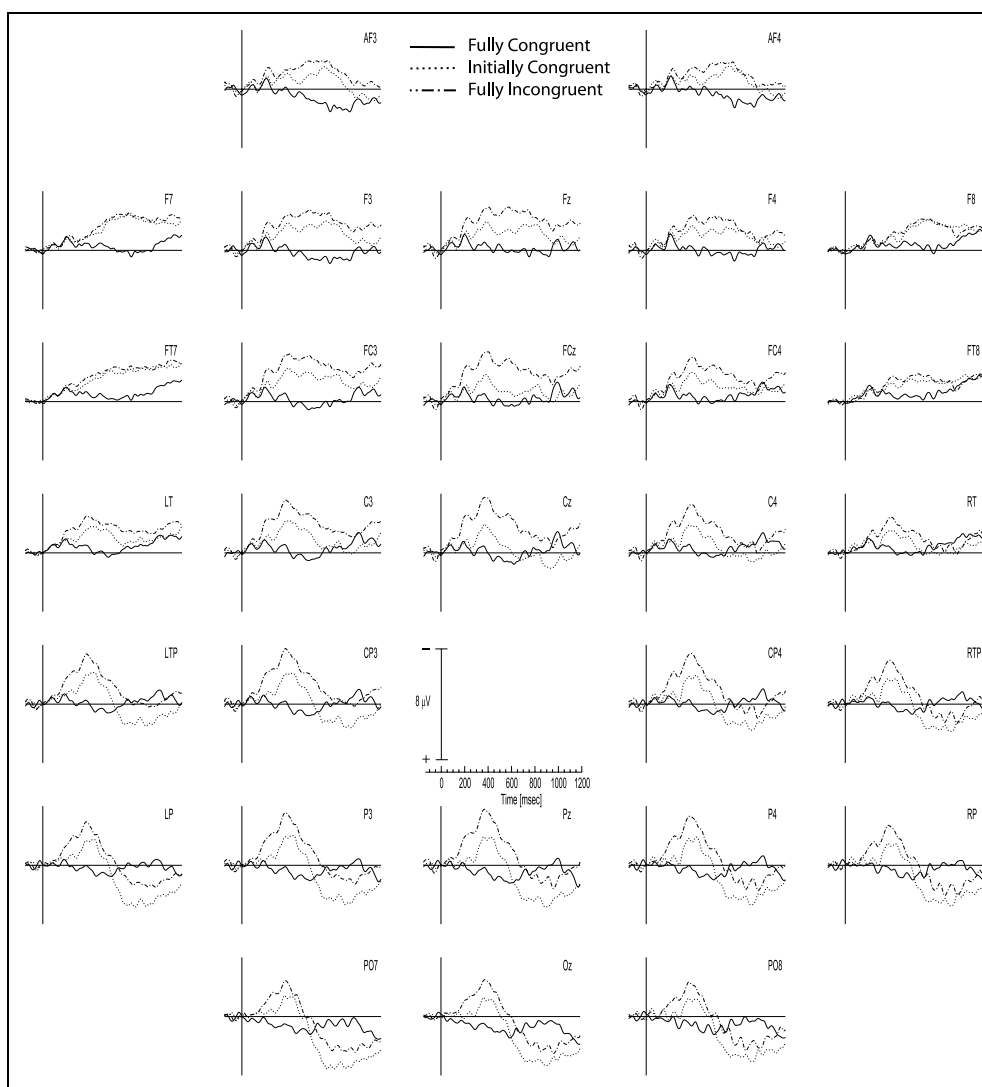
RESULTS

Time-Locking to Word Onset

Figure 1 displays the grand average waveforms by electrode site time-locked to the onset of the critical word. There are several things to note. First, as found in previous studies of ours and a recent study by Sanders and Neville using connected speech, the critical words elicited an N100 component in all three conditions (Sanders & Neville, 2003; Van den Brink et al., 2001, submitted). Second, a negativity at approximately 200 msec is visible in the waveforms. This negativity is

apparent in all three conditions, but is largest in the FI condition and is most visible over the frontal to central sites (see also enlarged FC4 in Figure 2). Third, the syntactically and semantically anomalous critical words in the IC and FI conditions elicit a broad negativity peaking at approximately 400 msec, which is more negative than the ERP elicited in the FC condition. In turn, this broad negativity is larger in the FI condition than in the IC condition. Its latency characteristics and morphology are similar to previously reported N400 effects. Finally, Figure 1 shows a late positivity between 700 and 1000 msec over parietal sites in the anomalous conditions. This positivity reaches maximal amplitude at around 800 to 900 msec and is present in both the IC and the FI conditions, but is largest in the IC condition (see also enlarged P4 in Figure 2). Its morphology is similar to previously reported P600 effects, related to difficulty in syntactic processing, but its latency characteristics differ. However, it is possible that because information about word category does not become available until at a relatively late moment in time (on average 330 msec after word onset), the P600 component may have been delayed.

The N200 Latency Window: 150–250 msec

Table 2 displays the results of the mean ERP amplitude analyses of variance (ANOVAs) in the 150- to 250-msec latency range. The omnibus ANOVA for the N200 latency window resulted in a significant main effect of congruity. The a priori pairwise comparisons revealed that the FI completions elicited a larger N200 than both the FC and the IC completions (corresponding to effects of 0.95 and 0.74 μV, respectively). In both of these comparisons the interaction of congruity with electrodes reached significance. Further topographical analyses revealed that the effects of the FI condition in relation to the FC and the IC conditions were larger over posterior than over anterior sites. The ERPs in the FC and IC conditions did not significantly differ in amplitude (a differ-

Figure 1. Connected speech. Grand average ERPs from 29 scalp sites, to critical words that were congruent (solid line), semantically and syntactically incongruent but shared initial phonemes with congruent completions (dotted line), and semantically and syntactically incongruent and did not share initial phonemes with congruent completions (alternating dash/dot line), after baseline correction in the 150-msec prestimulus interval. Time 0 is the onset of critical words. The time axis is in milliseconds. Note that negative polarity is plotted upwards in this and subsequent figures.

ence of 0.21 µV). No hemispheric differences were observed.

The N400 Latency Window: 300–500 msec

Table 3 displays the results of the mean ERP amplitude ANOVAs in the 300- to 500-msec latency range. The omnibus ANOVA indicated that the basic congruity effect in the N400 latency window was significant. Further analysis revealed that both the FI and the IC completions elicited a larger N400 than the FC completions (corresponding to effects of 2.51 and 1.49 µV, respectively). In both of these comparisons, interactions of congruity with electrodes reached significance. In addition, the FI and the IC conditions differed significantly from each other in the N400 latency window (corresponding to an effect of 1.02 µV), and an interaction with electrodes was found as well. Topographical analyses showed that all three congruity effects were significantly

larger over posterior than anterior regions of the scalp. No hemispheric differences were observed.

To establish whether the difference in ERP amplitude between the FI and IC conditions might be due to an early differential effect in the N200 latency window, we performed additional analyses in the N400 latency window after applying baseline correction in the 250- to 300-msec interval (see also Van den Brink et al., 2001). This time interval corresponds to a period just after the N200 component latency window. Figure 3 shows the grand average waveforms after this alternative baseline correction procedure. The omnibus ANOVA indicated that the basic congruity effect in the N400 latency window was still significant: $F(2,40) = 5.88$, $SEM = 23.00$, $p < .006$. The pairwise comparisons revealed that both the FI and the IC completions elicited a larger N400 than the FC completions, corresponding to effects of 0.72 µV with $F(1,20) = 7.83$, $SEM = 19.92$, $p = .01$ and 0.89 µV with $F(1,20) = 10.96$, $SEM = 21.94$, $p < .01$, respectively. However, the

Figure 2. Connected speech. Grand average ERPs from FC4 and P4, to critical words that were congruent (solid line), semantically and syntactically incongruent but shared initial phonemes with congruent completions (dotted line), and semantically and syntactically incongruent and did not share initial phonemes with congruent completions (alternating dash/dot line), after baseline correction in the 150-msec prestimulus interval. Time 0 is the onset of critical words. The time axis is in milliseconds.
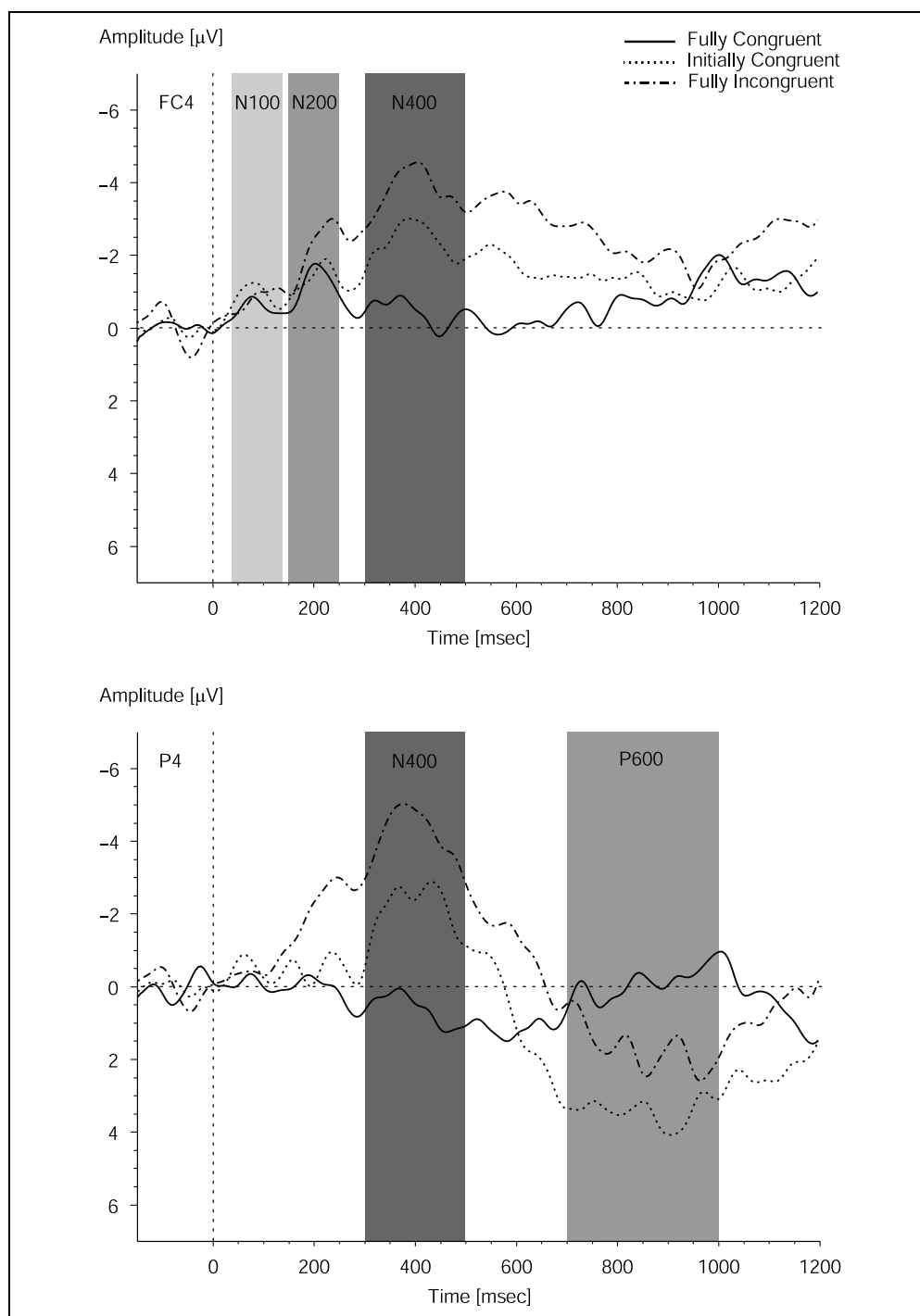
FC and the IC conditions did not significantly differ from each other: $F(1,20) < 1$.

The P600 Latency Window: 700–1000 msec

Table 4 displays the results of the mean ERP amplitude omnibus ANOVAs in the 700- to 1000-msec latency range. The omnibus ANOVA on mean amplitudes did not result in a significant main effect of congruity. Pairwise analyses revealed that only the ERP amplitude

in the FI condition differed significantly from the ERP amplitude in the IC condition. Again, because of the possibility that amplitude effects in this latency window were affected by earlier differential effects in the N200 latency window, we performed further analyses after baseline correction in the 250- to 300-msec interval.

Table 5 displays the results of the mean ERP amplitude omnibus ANOVAs in the 700- to 1000-msec latency range after baseline correction in the 250- to 300-msec interval (Figure 3). The omnibus ANOVA on mean

Table 2. ANOVA on Mean ERP Amplitude in the 150- to 250-msec Latency Range (N200)

	Source	df	F	SEM	p
Omnibus ANOVA (29 electrodes)					
Overall	Con	2, 40	13.54	11.20	.000***
	Con × El	56, 1120	4.36	0.44	.001**
FC vs. FI	Con	1, 20	17.85	15.44	.000***
	Con × El	28, 560	8.96	0.38	.000***
FC vs. IC	Con	1, 20	1.42	9.91	.248
	Con × El	28, 560	0.63	0.50	.591
IC vs. FI	Con	1, 20	20.04	8.24	.000***
	Con × El	28, 560	4.64	0.44	.004**
Anterior vs. posterior (2 × 12 electrodes)					
FC vs. FI	Con × Site	1, 20	13.39	3.37	.002**
IC vs. FI	Con × Site	1, 20	4.72	4.66	.042*
Anterior left vs. anterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	3.22	1.03	.088
IC vs. FI	Con × Site	1, 20	1.49	1.03	.236
Posterior left vs. posterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	0.40	1.63	.533
IC vs. FI	Con × Site	1, 20	0.06	1.83	.806

Note. Con = congruity type; El = electrode.

*p < .05.

**p < .01.

***p < .001.

amplitudes now did result in a significant main effect of congruity. Further analysis revealed that both the FI and the IC completions elicited a larger P600 than the FC completions (corresponding to effects of 1.37 and 1.06 μV, respectively). In both of these comparisons, interactions of congruity with electrodes reached significance. Topographical analyses revealed that the effects in relation to the FC condition (baseline) had a posterior distribution: In the anterior/posterior analyses, the interaction of the factors congruity by site reached significance both in the FC versus FI comparison and the FC versus IC comparison. The ERPs in the FI and the IC conditions did not significantly differ from each other. Again, no hemispheric differences were observed.

Onset Latencies Relative to Word Onset

To investigate the exact onset of the congruity effects we performed additional analyses. Onset latencies were

estimated by first separately computing the mean amplitude values for all 29 electrodes in 20-msec latency ranges (bins) that shifted in steps of 10 msec from target onset until 500 msec after target onset (e.g., 0–20, 10–30, etc.). The values for the latency bins were submitted to ANOVAs that tested against the null hypothesis of zero difference between the a priori selected conditions FC versus FI and FC versus IC. To compensate for multiple comparisons, we considered a congruity effect significant when at least five consecutive bins reached a statistical significance level of $p < .05$.

The onset latency analyses for the FC versus FI comparison on the amplitudes of the two waveforms first revealed a significant congruity effect in the 170- to 190-msec latency bin ($p < .01$) and this effect remained

Table 3. ANOVA on Mean ERP Amplitude in the 300- to 500-msec Latency Range (N400)

|	Source	df	F	SEM	p
Omnibus ANOVA (29 electrodes)					
Overall	Con	2, 40	27.17	35.67	.000***
	Con × El	56, 1120	7.74	0.58	.000***
FC vs. FI	Con	1, 20	48.88	39.22	.000***
	Con × El	28, 560	12.69	0.67	.000***
FC vs. IC	Con	1, 20	13.80	48.68	.001**
	Con × El	28, 560	4.04	0.65	.006*
IC vs. FI	Con	1, 20	16.69	19.12	.001**
	Con × El	28, 560	5.49	0.40	.001**
Anterior vs. posterior (2 × 12 electrodes)					
FC vs. FI	Con × Site	1, 20	17.47	6.83	.000***
FC vs. IC	Con × Site	1, 20	6.61	6.33	.018*
IC vs. FI	Con × Site	1, 20	4.73	4.19	.042*
Anterior left vs. anterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	0.71	2.38	.410
FC vs. IC	Con × Site	1, 20	1.30	1.80	.267
IC vs. FI	Con × Site	1, 20	0.05	1.00	.819
Posterior left vs. posterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	1.32	1.49	.264
FC vs. IC	Con × Site	1, 20	1.56	1.44	.225
IC vs. FI	Con × Site	1, 20	0.01	0.81	.915

Note. Con = congruity type; El = electrode.

*p < .05.

**p < .01.

***p < .001.

Figure 3. Connected speech. Grand average ERPs from 29 scalp sites, to critical words that were congruent (solid line), semantically and syntactically incongruent but shared initial phonemes with congruent completions (dotted line), and semantically and syntactically incongruent and did not share initial phonemes with congruent completions (alternating dash/dot line), after baseline correction in the 250- to 300-msec prestimulus interval. Time 0 is the onset of critical words. The time axis is in milliseconds.
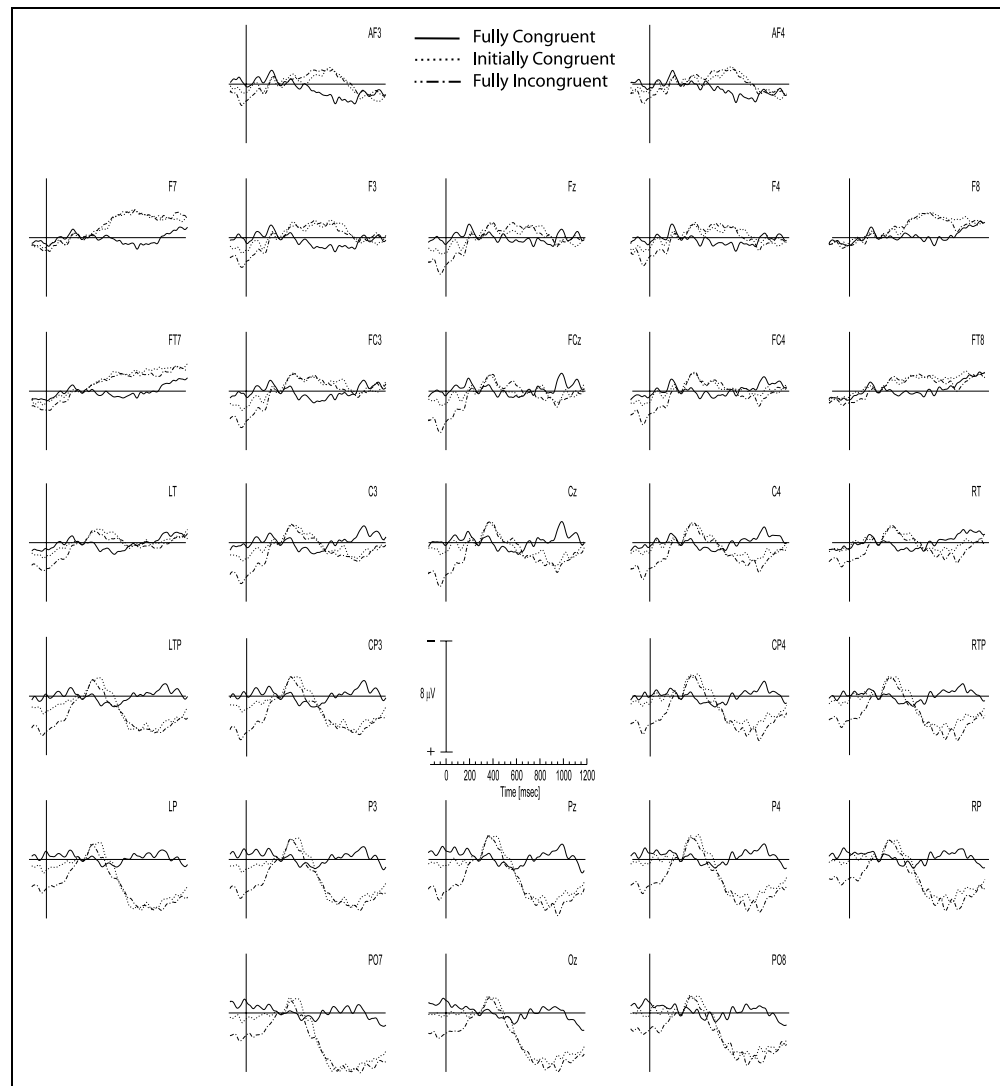
significant over the entire test region (500 msec). In the FC versus IC comparison the congruity effect did not start until the 310- to 330-msec latency bin. So, the FI waveform deviated from the FC waveform approximately 140 msec earlier than the waveform of the IC words (i.e., 170 vs. 310 msec after word onset).

Time-Locking to the Word Category Violation Point

Figure 1 did not reveal the clear presence of a LAN in our data. Consider, however, that a LAN is believed to be elicited upon detection of a word category violation. Because the moment of detection differs between the critical words within a condition, the LAN might be smeared when the waveforms are time-locked to word onset. When time-locked to the moment of detection of a word category violation, the LAN, if present, should become visible. Figure 4 displays the grand average

waveforms for electrode sites F7, F8, Pz, P3, and P4 time-locked on a trial-by-trial basis to the moment at which the critical words in the IC and FI conditions could be identified as verbs only and a word category violation was detected. This word CVP was assessed on the basis of phonetic transcriptions of the critical words. More specifically, on the basis of a Dutch word corpus it was assessed at what moment the critical word was no longer a possible noun. The time period preceding the CVP corresponds to the mean time period preceding CVP between the two conditions, which was 330 msec (336 msec in IC and 324 msec in FI, respectively). Therefore, in the alignment in Figure 4, the zero points of IC and FI correspond to word onset only approximately. Because information about the exact onset of each critical word is lost in this time-locking procedure, baseline correction cannot take place on the basis of information related to word onset. Therefore, baseline correction was applied on the basis of the averaged activity of 150 msec preceding CVP. Remember that the

Table 4. ANOVA on Mean ERP Amplitude in the 700- to 1000-msec Latency Range (P600)

	Source	df	F	SEM	p
Omnibus ANOVA (29 electrodes)					
Overall	Con	2, 40	2.46	48.86	.110
	Con x El	56, 1120	26.20	0.78	.000***
FC vs. FI	Con	1, 20	0.89	62.09	.357
	Con x El	28, 560	24.43	0.95	.000***
FC vs. IC	Con	1, 20	1.11	58.78	.304
	Con x El	28, 560	39.10	0.90	.000***
IC vs. FI	Con	1, 20	9.36	25.73	.006**
	Con x El	28, 560	6.15	0.50	.000***

Note. Con = congruity type; El = electrode.

**p < .01.

***p < .001.

critical words in the FC condition, plotted in gray in Figure 4, do not contain word CVPs. In time-locking the FC waveform to word onset, and adding a time period of 330 msec before the time-locking point in the incongruent conditions IC and FI (and applying baseline correction in the 180 to 330 msec interval for all three conditions), the time line of information processing between the three conditions is kept more or less constant. When we inspect the centroparietal to occipital electrodes, what can be seen is that in both IC and FI the N400 has set in prior to CVP and is followed by a broad positive shift, peaking at around 600 msec after CVP. However, when we look at the frontal electrodes, again for both IC and FI a broad negativity peaking at around 370 msec after CVP is clearly visible. The anterior negative effect is most prominent at F7 and to a somewhat lesser extent at F8. To investigate the latter anterior negative shift we decided to perform a repeated measures ANOVA on the anterior electrode sites (Fz, FCz, AF3, AF4, F3, F4, F7, F8, FC3, FC4, FT7, and FT8) in the latency window of 600 to 800 msec (corresponding to 270 to 470 msec after CVP for FI and IC).

The LAN Latency Window: 270–470 msec after CVP

Table 6 displays the results of the mean ERP amplitude omnibus ANOVA in the 270- to 470-msec latency range after CVP. Baseline correction was based on the averaged activity of 150 msec immediately preceding CVP. The omnibus ANOVA on mean amplitudes resulted in a significant main effect of congruity. Further analysis revealed that both the FI and the IC completions elicited a larger LAN than the FC completions (corresponding to

effects of 1.79 and 1.62 µV, respectively), and in both of these comparisons an interaction with electrodes was found as well. The FC and the IC conditions did not significantly differ from each other. Topographical analyses revealed that the effects in relation to the FC condition (baseline) were larger over the left hemisphere: In the anterior left/anterior right analyses, the interaction of the factors congruity by site reached significance both in the FC versus FI comparison and the FC versus IC comparison.

Inspection of the waveforms in this latency interval at posterior sites reveals a possible short-lived congruity effect between IC and FI. However, statistical analyses of the parietal and occipital band electrodes (Pz, Oz, P3,

Table 5. ANOVA on Mean ERP Amplitude in the 700- to 1000-msec Latency Range (P600) after Averaging in the 250- to 300-msec Interval

	Source	df	F	SEM	p
Omnibus ANOVA (29 electrodes)					
Overall	Con	2, 40	5.24	59.78	.015*
	Con × El	56, 1120	29.94	0.94	.000***
FC vs. FI	Con	1, 20	7.44	76.46	.013*
	Con × El	28, 560	37.10	1.16	.000***
FC vs. IC	Con	1, 20	4.74	72.16	.042*
	Con × El	28, 560	41.82	0.99	.000***
IC vs. FI	Con	1, 20	0.94	30.71	.345
	Con × El	28, 560	0.39	0.68	.728
Anterior vs. posterior (2 × 12 electrodes)					
FC vs. FI	Con × Site	1, 20	86.36	12.47	.000***
FC vs. IC	Con × Site	1, 20	122.55	8.52	.000***
IC vs. FI	Con × Site	1, 20	0.03	7.49	.856
Anterior left vs. anterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	2.99	3.51	.099
FC vs. IC	Con × Site	1, 20	2.54	2.55	.126
IC vs. FI	Con × Site	1, 20	0.22	2.26	.648
Posterior left vs. posterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	0.07	2.72	.800
FC vs. IC	Con × Site	1, 20	2.60	2.04	.123
IC vs. FI	Con × Site	1, 20	1.88	1.88	.186

Note. Con = congruity type; El = electrode.

*p < .05.

***p < .001.

P4, LP, RP, PO7, and PO8) revealed no main congruity effect for this comparison, $F(1,20) = 2.71$, $SEM = 9.78$, $p = .115$.

Onset Latencies Relative to the Word Category Violation Point

We also performed onset analyses for the LAN at F7 and for the N400 at Pz where the effects are largest. Onset latencies were estimated by first separately computing the mean amplitude values for the F7 in 20-msec latency ranges (bins) that shifted in steps of 10 msec from 100 msec before CVP until 500 msec after CVP (e.g., 0–20, 10–30, etc. after CVP). The values for the latency bins were submitted to ANOVAs that tested against the null hypothesis of zero difference between the a priori selected conditions FC versus FI and FC versus IC.

At F7, the onset latency analyses for the FC versus FI comparison on the amplitudes of the two waveforms first revealed a significant congruity effect in the 90- to 110-msec latency bin after CVP ($p < .01$) and this effect remained significant over the entire test region. In the FC versus IC comparison the congruity effect started between 120 and 140 msec after CVP and also remained significant over the entire test region. So, both the FI and the IC waveforms deviated from the FC waveform at approximately 100 msec after CVP.

At Pz, the onset latency analyses for the FC versus FI comparison on the amplitudes of the two waveforms first revealed a significant congruity effect in the latency bin that starts 10 msec before CVP ($p < .01$) and this

Table 6. ANOVA on Mean ERP Amplitude in the 270- to 470-msec Latency Range after CVP (LAN)

Source		df	F	SEM	p
Omnibus ANOVA (12 electrodes)					
Overall	Con	2, 40	13.29	18.48	.000***
	Con × El	22, 440	5.05	0.47	.001**
FC vs. FI	Con	1, 20	26.23	15.33	.000***
	Con × El	11, 220	5.92	0.47	.001**
FC vs. IC	Con	1, 20	11.90	27.84	.003**
	Con × El	11, 220	6.48	0.65	.001**
IC vs. FI	Con	1, 20	0.28	12.27	.602
	Con × El	11, 220	0.71	0.31	.549
Anterior left vs. anterior right (2 × 5 electrodes)					
FC vs. FI	Con × Site	1, 20	7.64	0.98	.012*
FC vs. IC	Con × Site	1, 20	7.65	1.67	.012*

Note. Con = congruity type; El = electrode.
*$p < .05$.
**$p < .01$.
***$p < .001$.

effect remained significant until the 190- to 210-msec latency bin after CVP. In the FC versus IC comparison the congruity effect started between 30 and 50 msec before CVP and remained significant until the 180- to
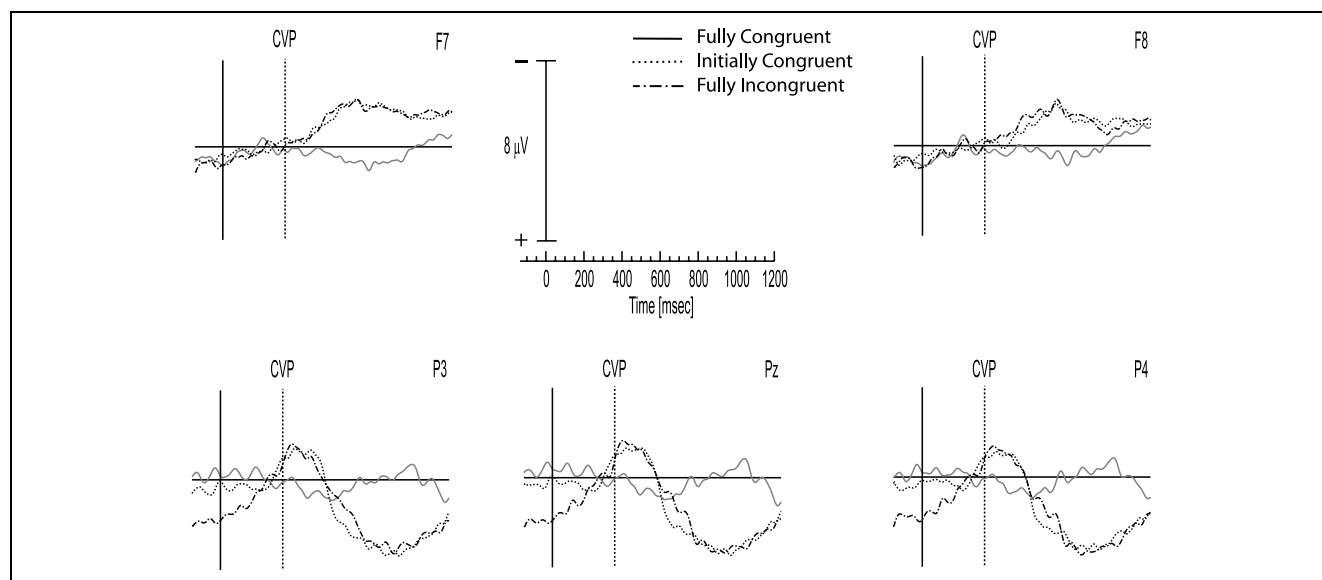

Figure 4. Connected speech. Grand average ERPs from F7, F8, Pz, P3, and P4 sites, to critical words that were semantically and syntactically incongruent but shared initial phonemes with congruent completions (dotted line) and semantically and syntactically incongruent and did not share initial phonemes with congruent completions (alternating dash/dot line), after time locking on a trial-by-trial basis to the moment of word category violation (CVP), and grand average waveforms to congruent critical words (solid line) time-locked to word onset. Baseline correction was applied in the 150-msec interval preceding CVP for FI and IC. The time axis is in milliseconds.
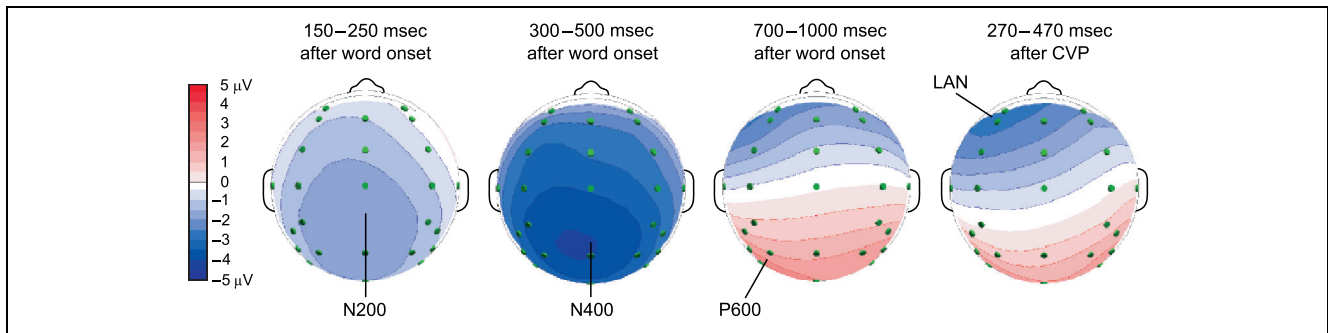
Figure 5. Connected speech. Isopotential voltage maps of the N200, N400, P600, and LAN effects. Maps are based on difference waveforms that resulted from subtracting the mean amplitude in the latency windows of the grand average ERP for the fully congruent critical words from the mean amplitude of the grand average ERP for the semantically and syntactically incongruent critical words that did not share the same initial phonemes as the fully congruent words.

200-msec latency bin after CVP. So, even though we applied a baseline correction procedure in the 150-msec interval prior to CVP, both the FI and the IC waveforms already deviated from the FC waveform at or even before CVP.

Scalp Distribution of Effects

Figure 5 shows the isopotential voltage maps of mean amplitude effects between the FC and FI conditions in the N200, the N400, and the P600 latency window after time-locking to word onset, and in the LAN window after time-locking to CVP. The figure reveals that both the N200 effect and the N400 effect are largest over centro-parietal sites. The LAN effect is largest over left frontal sites, and the P600 effect has a parietal maximum. To establish whether the congruity effects in the N200 and N400 latency windows still have statistically distinct scalp distributions, we performed an additional analysis. First, for every subject difference scores between the FC and FI conditions were computed for every electrode in the N200 latency window and the N400 latency window. Second, a scaling procedure was performed to avoid that differential amplitude effects between the two latency windows would be incorrectly interpreted as distribution effects. In this procedure, the electrode-specific difference scores were z-transformed for each latency window separately (Rösler, Heil, & Glowolla, 1993; see also McCarthy & Wood, 1985). The z-transformed values were entered into an ANOVA that crossed the 2-level latency window factor (N200 and N400) with the 29-level electrode factor. The topographical analysis revealed that the congruity effects found in the 150- to 250- and 300- to 500-msec latency windows do not have different spatial distributions across the scalp: No significant interaction of latency window by electrodes was obtained, $F(28,560) = 1.69$, $SEM = 0.29$, $p = .146$. Although not significant the N400 does seem to have a slightly more posterior distribution than the N200. In

previous studies this difference was found to be significant (Van den Brink et al., 2001, submitted).

DISCUSSION

The present study investigated the influence of semantic and syntactic context constraints on lexical selection and integration processes. More specifically, we used ERPs to investigate the time course of processing of different sources of information such as phonological, semantic, and syntactic, and their impact on lexical selection and integration processes in a situation where all sources of information about the perceived word become incrementally available, namely, in spoken-word comprehension. We used semantically constraining Dutch sentences that contained critical words that were either the highest cloze probability noun for that sentence (fully congruent), a verb in the past tense that apart from syntactic constraints also violated semantic constraints and had initial phonemes that differed from the highest cloze probability noun (fully incongruent), or a verb in the past tense that apart from syntactic constraints also violated semantic constraints but began with the same initial phonemes as the highest cloze probability noun (initially congruent). For the conditions containing the syntactically and semantically incongruent critical words, the moment of word category violation was assessed for each verb and labeled the CVP. The study revealed that in the domain of language comprehension, four different ERP effects can be distinguished that appear to be related to three types of sources of information processing: (a) a phonological effect as reflected by the N200, (b) a semantic effect reflected by the N400, and (c) syntactic effects reflected by the LAN and P600.

Phonological Processing

This study shows that semantically and syntactically incongruent words with initial phonemes that differed

from their congruent counterparts elicited a larger negativity in the 150- to 250-msec range after stimulus onset than both the congruent critical words and the semantically and syntactically incongruent words that began with the same phonemes as the congruent counterparts. The onset analysis reveals that the congruity effect between the FC and the FI critical words sets in at 170 msec. These results are in line with a number of studies using connected speech that revealed that semantically incongruent words with onsets that, given a semantically constraining preceding sentence frame, differ from those of the highest cloze probability words, elicit an early negative shift, related to phonological processing (Van den Brink et al., 2001, submitted; Hagoort & Brown, 2000; Connolly & Phillips, 1994). In these studies, it was assumed that this early negative effect reflects the lexical selection process, where word-form information resulting from an initial phonological analysis and content information derived from the context interact.

In their 1994 study, Connolly and Phillips obtained an early negativity in sentence-final words that were semantically anomalous and had initial phonemes that differed from those of the highest cloze probability word (e.g., ''The dog chased the cat up the *queen*''), comparable to the FC condition in the present study. However, they did not find an early negativity in a condition similar to our IC condition, where sentence-final words were semantically anomalous but had the same initial phonemes the highest cloze probability word (e.g., ''The gambler had a streak of *luggage*''). Connolly and Phillips, therefore, attributed this early negativity to a mismatch of the initial phonemes of the presented word with contextually determined expected phonemes, and labeled it the phonological mismatch negativity (PMN). They proposed that during sentence processing, on the basis of the context, a terminal word is predicted. But only after an initial phonological analysis of the perceived word the top-down context process is claimed to begin to exert its influence. A PMN is elicited if the expectancy of the terminal word on the basis of the context and the phonological analysis of the initial phonemes of the perceived word are not in agreement.

In the present study, as well as our previous studies (Van den Brink et al., 2001, submitted), all critical words elicited an N200, irrespective of whether the initial phonemes were identical to those of a congruent word or not. Therefore, in our previous studies, the process underlying the early negative effect was interpreted somewhat differently from the interpretation laid out by Connolly and Phillips (1994). We proposed that the amplitude of the N200 is indicative of whether the initial assessment of form-based activated lexical candidates reveals the presence of a candidate that fits the semantic and syntactic constraints of the preceding sentence. A small N200 would be elicited when the set contains a candidate that fits the contextual specifications. In the present study, this is the case for the FC and IC conditions, since on the basis of the first phonemes, the highest cloze probability word is activated. A large N200 would indicate that the set does not contain a candidate that fits the preceding sentence frame well, as is the case for the FI condition in the present study.

In the studies mentioned a psychophysiological distinction could be made between the early negative effect and the N400 effect. Failure to do so could be construed as weakening the claim that the two effects are genuinely distinct from one another. Connolly and Phillips (1994) were able to doubly dissociate the PMN and the N400 on the basis of experimental manipulations, finding a PMN in the absence of an N400 in one condition and vice versa in another (e.g., ''The gambler had a streak of *luggage*'' versus ''Don caught the ball with his *glove*,'' where *hand* was the highest cloze probability word). In our previous studies a distinction could be made between the N200 effect and the N400 effect on the basis of their spatial distributions across the scalp (Van den Brink et al., 2001, submitted). Whereas the N200 effect had an equal distribution across the scalp, the N400 effect had a clearly posterior distribution. In neurocognitive research, these statistically different distributions are usually interpreted as reflecting the activity of at least partly distinct neuronal populations and, thus, warrant the claim that the N200 and N400 effects could reflect distinct aspects of the spoken language comprehension process. The present study did not explicitly test the distinction between the N200 and N400 effects, apart from their spatial distributions across the scalp, which both had a posterior maximum, and statistically did not differ from each other. So, even though most aspects of the N200 component in this study are consistent with our previous findings, namely, that it can be obtained in all critical words, that it is largest in semantically incongruent words with initial phonemes that differ from their congruent counterparts, and that its morphology is most prominent over frontal to frontocentral sites, on the basis of the results in the present study, we cannot exclude the possibility that the N200 effect is in fact not distinct from the N400 effect, reflecting overall semantic processing (Van Petten, Coulson, Rubin, Plante, & Parks, 1999).

Semantic Processing

The semantically and syntactically incongruent words elicited an N400, indexing difficulty in lexical integration, irrespective of whether their initial phonemes were similar to those of the congruent word or not. The apparent difference in amplitude between the N400 elicited in the IC and FI conditions can be explained by the differential amplitude in the N200 latency window preceding the effect in the N400 time interval. Figure 3

shows that when we control for the differential effect in the early latency window, the N400 (and P600) differences obtained in the following time intervals vanish. The onset analysis of the congruity effect between the FC and IC critical words is not confounded by this early effect and reveals that the congruity effect sets in at 310 msec after word onset. Important to the current issue, in the present study the N400 has set in before information about the word category has become available. These results are in line with two studies that investigated the temporal relationship between the lexical selection and integration of a word's meaning into the preceding context (Van den Brink et al., submitted; Van Petten et al., 1999). In these studies, information about the moment a word can uniquely be identified on the basis of acoustic information alone, known as the isolation point (IP) of a word, was related to the timing of (the onset of) the N400, to assess whether semantic integration processing begins only after word identification is completed. It was found that in sentences ending with semantically anomalous words (e.g., ''This 14-year-old pianist obviously has a lot of *climate*''), onset of the N400 occurred prior to IP, indicating that semantic integration processing is initiated before a (semantically anomalous) word candidate is selected on the basis of the acoustic signal. In the present study, the N400 sets in even before listeners realize the critical word is actually a verb instead of a noun (i.e., before CVP). Taken together, we conclude that semantic integration processing can start before the acoustic information allows the selection of a unique candidate and, thus, seems to be attempted in parallel for multiple candidates that are still compatible with the bottom-up acoustic input.

Syntactic Processing

When time-locked to word onset, the waveforms of the semantically and syntactically incongruent critical words revealed a syntactic effect in the form of what appeared to be a somewhat delayed P600, which followed the N400. However, we hypothesized that any syntactic effect would be elicited only after detection of the word category violation encapsulated in the critical words. Therefore, we decided to time-lock the waveforms to the earliest moment at which a word category violation can be detected, assessed for each critical word, and labeled the CVP. Figure 4 shows that time-locking the waveforms on a trial-by-trial basis to CVP reveals the presence of an anterior negative shift preceding the P600 (reaching maximal amplitude at 600 msec after CVP) in the semantically and syntactically anomalous conditions IC and FI. This anterior negativity is present at the anterior frontal, frontal, and frontocentral electrodes, and is most prominent at F7. The left-lateralized anterior negativity reaches its maximal amplitude approximately 370 msec after CVP and has an onset at approximately 100 msec after CVP for both anomalous conditions.

It is not entirely clear what the relation is between the LAN obtained in this study and earlier reported ELANs, LANs, or bilateral Anterior Negativities. In some ways our LAN is reminiscent of the other anterior negativities, with respect to its polarity, its (left) frontal distribution and the fact that it is elicited to syntactic violations. However, with respect to timing it seems to be different from these other anterior negativities. Of course, we must keep in mind that our study differs in some respects from the other studies obtaining an anterior negativity. First, whereas many studies that obtained anterior negativities were conducted in the visual modality with syntactic and semantic information becoming almost immediately available, this study was conducted in the auditory domain, where this information becomes incrementally available. Second, in most of the studies using auditorily presented sentences, the critical words contained clear morphosyntactic markers at the beginning of the word indicating the word category and were in sentence-final position (Hahne & Friederici, 1999, 2002; Hahne & Jescheniak, 2001; Friederici et al., 1993, 1999). In the present study, word category information was contained in the suffix of the critical word, and the word CVP was on average 330 msec after onset. Moreover, our critical words were presented in non-sentence-final position, thereby preventing possible pollution of the data from sentence-wrap-up processes (cf. Hagoort, 2003; Osterhout, 1997). In conclusion, the LAN obtained in our study possibly belongs to a class of anterior negativities reflecting problems in syntactic processing. More specifically, this LAN probably reflects the detection of a mismatch between the required word category (in this case a noun) and the word category actually presented (a verb). The latency of this LAN, compared to other anterior negativities that are believed to reflect a syntactic mismatch of word category, is delayed as information about the word category becomes available at a later moment during word processing than was the case in most other studies.

Conclusion

Although our study, unlike any other, not only reveals the presence of both an N400 and a LAN in the same waveform, but also reveals that the LAN has a later onset (and peak) than the N400, we do not presume that during spoken-word processing the processing of lexical-semantic information in general precedes that of syntactic (word category) information. Instead, these results are interpreted as providing strong evidence for the claim of continuous use of any sources of information the moment they become available. These results clearly provide evidence against the claim that building of syntactic-phrase structure, based on word category information, is autonomous and precedes semantic integration processes, and that these syntactic and

semantic processes interact only in a later time interval (Friederici, 2002). This claim was partly based on results from recent auditory studies using critical words that, as in our study, violated both the syntactic and semantic constraints of the preceding sentence context. In these studies an early LAN and a P600 in the absence of an N400 were obtained (Hahne & Jescheniak, 2001; Friederici et al., 1993). However, contrary to the critical words used in the present study, theirs allowed for a rapid detection of a word category violation (due to the morphosyntactic markers *ge-* and *be-*).

The present study testifies to the incremental and cascaded nature of on-line language processing. It shows that the moment a violation of any kind occurs, it is detected by the system. In the case of the FI critical words, the moment the system detected a mismatch between the lexical candidates that were activated on the basis of bottom-up information and the semantic and syntactic constraints imposed by the preceding sentence frame, an N200 effect occurred. For the IC critical words, an N400 was instigated at the moment a semantic violation was detected. It was not until later in the word that a syntactic error became manifest, which resulted in a LAN and a P600.

These data clearly show that contrary to predictions of a serial model, the system does not specifically need syntactic information about word category in order to begin to semantically integrate a perceived word. In other words, semantic processing is not dependent on the result of syntactic processing or vice versa. Instead, the system works with the information it receives and uses it immediately. The results of this study fit with a cascaded model of spoken-word processing that allows for an early influence of contextual information after bottom-up activation of a set of lexical candidates, such as incorporated in the cohort model or shortlist (Norris, 1994; Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978). Although not explicitly specified in these models, our data point to an optimal use of contextual information during spoken-word identification by allowing for semantic and syntactic processing to take place in parallel, and integration to proceed with a limited number of candidates.

METHODS

Subjects

The experiment was conducted with 21 native speakers of Dutch (20 women, mean age 21, range 18–27 years) from the subject pool of the Max Planck Institute for Psycholinguistics. All had normal to corrected-to-normal vision and were right-handed (four subjects reported having a left-handed parent or sibling). None of the subjects had any neurological impairment or had experienced any neurological trauma. Nor had any of the subjects participated in the pretest (see below). The

data of five additional subjects were not analyzed due to too much contamination of artifacts such as blinks, muscle tension and excessive alpha activity. Subjects were paid for their participation.

Materials

The experimental items consisted of a set of 222 triplets of semantically constraining sentences across three experimental conditions. The sentences in each triplet were identical except for the critical word. In the FC condition, the critical word was a non-sentence-final but high cloze probability word, for example, ''Het vrouwtje veegde de vloer met een oude *bezem* gemaakt van twijgen'' (''The woman swept the floor with an old *broom* made of twigs''). In the IC condition, the critical word was a verb in the past tense that had the same initial phonemes (and lexical stress) as the highest cloze probability word, for example, ''Het vrouwtje veegde de vloer met een oude *bedelde* gemaakt van twijgen'' (''The woman swept the floor with an old *begged* made of twigs''). In the third condition the critical word was a verb in the past tense that had initial phonemes that were totally unexpected given the sentential context, e.g., ''Het vrouwtje veegde de vloer met een oude *kliederde* gemaakt van twijgen'' (''The woman swept the floor with an old *messed* made of twigs''). This condition is called the FI condition.

The 222 triplets were derived from a set of sentence materials that had been used in a previous study and which were part of a larger set of sentences that had been submitted to a cloze test. Thirty subjects from the subject pool of the Max Planck Institute for Psycholinguistics were given a list of 350 sentences with the final word omitted. Subjects were asked to carefully read the sentences and to fill in the first word that came to mind at the position of the omitted sentence-final word. On the basis of the results, a set of 222 sentences with a cloze probability of .50 or higher formed the basis for the 222 triplets used in this experiment. To avoid possible sentence-wrap-up effects we added a few words to each sentence (meanwhile making sure that the sentences remained meaningful), so that the critical words were non-sentence-final. Care was taken to avoid adding words that constituted a phrase. This way the prosodic information in the acoustic signal provided a cue that the sentence continued after the critical word. Items for the IC condition were constructed by replacing the critical word by a verb in the past tense, but began with the same phonemes (mean length of overlap was 2.5 phonemes, range 2 to 5, always including a full vowel) and had the same lexical stress as the high cloze probability word, which was used in the FC condition. The items for the FI condition were constructed by replacing the critical word with a verb in the past tense with word-initial phonemes that differed from the

beginning phonemes of the high cloze probability word and other congruent words.

Because the sentences within each triplet were identical except for the critical word, they had the same length (mean length of 13.9 words; range 8–21) and were equally constraining; the critical words in the FC condition had a mean cloze probability of .84 (range: .50–1.0). All critical words were reasonably well-known nouns and verbs selected from a Dutch word corpus called CELEX (Baayen, Piepenbrock, & Van Rijn, 1993) and had a mean duration of 386 msec and a frequency of 43.7 counts per million in the FC condition, 495 msec and 16.2 counts in the IC condition, and 485 msec and 13.6 counts in the FI condition. Each critical word began with a plosive. This provided a clear physical marker that facilitated the alignment of the ERP waveform to the onset of the critical word.

Three lists were constructed to ensure that no subject heard the same sentence or critical word more than once. In addition to the set of 222 experimental items (74 per condition per list), a set of 74 fillers was constructed. These fillers were all semantically congruent, to balance the number of sentences that were anomalous and congruent. Finally, the experimental lists were preceded by a practice list of 20 items, which reflected the experimental materials. The practice list was used to familiarize the subjects with the experimental procedure.

The experimental sentences, fillers, and practice sentences were all spoken by a female speaker, with normal intonation at a normal speaking rate, and were recorded in one recording session. The sentences were recorded in triplets, making sure intonation and speaking rate were kept constant within each triplet. Sentences were spoken in a sound-attenuating booth and recorded on a digital audiotape. The DAT recordings were sampled at 16 kHz mono and stored on computer disk. A speech waveform editing system was used to mark the onset of the critical words. Also, the moment of word category violation was marked for each critical word in IC and FI conditions individually, using the CELEX database for reference. Based on phonetic transcriptions we identified the moment at which the verb in the past tense and a possible noun continuation started to diverge from one another. For instance, the critical FI word *kliederde* (*messed*) has a relatively well-known noun counterpart of *kliederboel* (*mess*). Here, the word category violation occurs at the onset of the word-final syllable *de*, marking for a verb in the past tense, and a marker was placed for alignment purposes. Using this procedure, the word CVP in IC was on average 336 msec ($SD = 76$ msec) and in FI 324 msec ($SD = 70$ msec).

Procedure

Subjects were tested individually in a dimly illuminated sound-attenuating booth. They were seated in a comfortable reclining chair, instructed to move as little as possible, and told that they would hear a number of sentences. Their task was to attentively listen to these sentences and to try to understand them. In addition they were told that several sentences would be anomalous. No additional task demands were imposed.

Each trial began with a 300-msec warning tone, followed after 1200 msec of silence by the spoken sentence. The next trial began 3500 msec after sentence offset. To ensure that subjects would not blink during and shortly after presentation of the sentence, 1000 msec prior to the beginning of the sentence a fixation point (an asterisk) was displayed on the computer screen. The asterisk remained on the screen until 1000 msec after offset of the spoken sentence. Subjects were instructed to fixate on the asterisk during presentation of the sentences, but were free to blink and move their eyes when the asterisk was not displayed on the screen. Subjects listened to the stimuli via closed-ear headphones. After the practice session, the trials were presented in five blocks of approximately 10 minutes.

EEG Recording

The EEG was recorded from 29 Ag/AgCl-sintered electrodes mounted in an elastic cap, each referred to the left mastoid. Five electrodes were placed according to the 10% standard system of the American Electroencephalographic Society over midline sites at Fz, FCz, Cz, Pz, and Oz locations, along with nine lateral pairs of electrodes over standard sites on frontal (AF3, AF4, F3, F4, F7 and F8), frontocentral (FC3 and FC4), frontotemporal (FT7 and FT8), central (C3 and C4), centroparietal (CP3 and CP4), parietal (P3 and P4), and occipital (PO7 and PO8) positions. Three additional pairs were placed laterally over symmetrical positions: (a) a temporal pair (LT and RT) placed laterally to Cz, at 33% of the interaural distance, (b) a temporoparietal pair (LTP and RTP) placed 30% of the interaural distance lateral and 13% of the nasion–inion distance posterior to Cz, and (c) a parietal pair midway between LTP/RTP and PO7/PO8 (LP and RP). Vertical eye movements were monitored via a supra- to suborbital bipolar montage. A right to left canthal bipolar montage was used to monitor for horizontal eye movements. Activity over the right mastoid bone was recorded on an additional channel to determine if there were differential contributions of the experimental variables to the presumably neutral mastoid site. No such differential effects were observed. The EEG and EOG recordings were amplified with a SynAmp Model 5083 EEG amplifier (NeuroScan Inc., Herndon, VA), using a hi-cut of 30 Hz (notch filter 60 Hz) and a time constant of 8 sec (0.02 Hz). Electrode impedances were kept below 3 kΩ for the EEG recording and below 5 kΩ for the EOG recording. The EEG and EOG signals were digitized on-line with a sampling frequency of 200 Hz.

Data Analysis

Data from critical trials were analyzed according to the following procedure. Prior to off-line averaging, all single-trial waveforms were screened for eye movements, electrode drifting, and EMG artifacts in a latency window that ranged from 150 msec before onset of the critical-final word to 2000 msec after onset of the critical-final word. Trials containing artifacts were rejected (11.7% in FC, 13.6% in IC, and 12.5% in FI). For each subject, average waveforms were computed across all remaining trials per condition after baseline correction of the individual trials on the basis of the averaged activity of 150 msec before onset of the critical word.

Statistical analyses of the congruity effects consisted of a number of repeated measures ANOVAs with mean amplitude values computed for each subject and each electrode in different latency windows: (a) 150–250 msec after critical-word onset for the early negativity (N200), (b) 300–500 msec after critical-word onset for the N400, and (c) 700–1000 msec after critical-word onset for the P600. The latency windows were determined after careful visual inspection of the grand average waveforms and corresponded with latency windows used in our previous studies (Van den Brink et al, 2001, submitted). The latency window for the early negativity corresponds to the onset of the ascending flank and the offset of the descending flank. The N400 latency window corresponds to the onset of the ascending flank and 100 msec after its maximal amplitude. This 300- to 500-msec interval is a standard time window for measuring N400 effects. The latency window for the P600 corresponds to a window of 300 msec surrounding its maximal amplitude. For each latency window, the results were first analyzed in an omnibus ANOVA that crossed all three levels of the congruity factor (FC, IC, and FI) with the 29-level electrode factor. In addition to the omnibus ANOVA, a priori pairwise comparisons between the congruity conditions were tested using ANOVAs with a 2-level congruity factor.

Scalp distributions of the congruity effects were subsequently explored in three separate ANOVAs: an ANOVA on anterior versus posterior sites (Fz, FCz, AF3, AF4, F3, F4, F7, F8, FC3, FC4, FT7, FT8/Pz, Oz, CP3, CP4, LTP, RTP, P3, P4, LP, RP, PO7, PO8); anterior left hemisphere versus anterior right hemisphere (AF3, F3, F7, FC3, FT7/AF4, F4, F8, FC4, FT8); posterior left hemisphere versus posterior right hemisphere (CP3, LTP, P3, LP, PO7/CP4, RTP, P4, RP, PO8). Univariate F tests with more than one degree of freedom in the numerator were adjusted by means of the Greenhouse–Geisser/Box's epsilon hat correction (Maxwell & Delaney, 1989). The original degrees of freedom and adjusted p values will be reported.

Note

1. Cloze probability of a word is defined as the percentage of subjects who use that word to continue a particular sentence (see Bloom & Fischler, 1980).

REFERENCES

Ainsworth-Darnell, K., Shulman, H. G., & Boland, J. E. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. *Journal of Memory and Language, 38,* 112–130.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993) *The CELEX lexical database* [CD-ROM]. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory and Cognition, 8,* 631–642.

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

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

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

Frazier, L. (1987). Theories of sentence processing. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 291–307). Cambridge MA: MIT Press.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 50,* 259–281.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6,* 78–84.

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

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

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15,* 883–899.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia, 38,* 1518–1530.

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). New York: Oxford University Press.

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

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13,* 339–356.

Hahne, A., & Jescheniak, J. D. (2001). What's left if the jabberwock gets the semantics? An ERP investigation into semantic and syntactic processes during auditory sentence comprehension. *Cognitive Brain Research, 11,* 199–212.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30,* 47–61.

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

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

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken-word recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 71–102). Cambridge MA: MIT Press.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.

Maxwell, S. E., & Delaney, H. D. (1989). *Designing and analyzing data* (pp. 475–479). Pacific Grove, CA: Brooks/Cole.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203–208.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.

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

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

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research, 15,* 228–240.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.

Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17,* 701–718.

Van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13,* 967–985.

Van den Brink, D., Hagoort, P., & Brown, C. M. The cascaded nature of lexical selection and integration revealed by N200 and N400 effects. Manuscript submitted for publication.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning Memory & Cognition, 25,* 394–417.

Zwitserlood, P. (1989). The locus of effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25–64.