OSU Working Papers in Linguistics 47, 1-19

Discriminating Between Syntactic and Semantic Processing: Evidence from Event-related Potentials

Kim Darnell[1]

Abstract By measuring the event-related brain potentials (ERPs) elicited during a visual word-by-word presentation of sentences containing either a syntactic incongruity, semantic incongruity, or a combined syntactic and semantic incongruity, I investigated whether the N400 and P600 waveforms are discrete components reflective of independent semantic and syntactic processing or simply sub-parts of a larger wave caused by general sentential processing difficulty. Words that were syntactically inconsistent with the sentence structure elicited a P600 potential, while words that were semantically inconsistent elicited an N400 potential. Words that caused both a syntactic and semantic violation of the sentence in which they appeared evoked both a P600 and an N400 waveform. The results support the hypothesis that the N400 and P600 are independent waveforms, suggesting that the brain is capable of responding specifically to anomalies at both the syntactic and semantic levels. These findings are used to evaluate the functionality of three currently popular descriptions of the relationship between the syntactic and semantic levels of the human language processor.

[1]This research was made possible through funding provided by the Ohio State University Center for Cognitive Science, the Department of Psychology, and the Department of Linguistics. My appreciation is extended to my collaborators, Dr. Harvey Shulman and Dr. Julie Boland, for their guidance and support during this project, and to Dr. Peter Culicover for his comments and suggestions on earlier drafts of this paper. Thanks also to Rob Tigner, Christina Pencheff, Karen West, Maria Miles, and Angie Wascovich for their assistance in collecting and analyzing the data. All errors are mine. Questions should be sent via electronic mail to darnellk@ling.ohio-state.edu.

Language comprehension is an intricate process, involving the evolution of a linguistic stimulus from a string of phonological or orthographic code into coherent, hierarchical structure that can be interpreted in terms of stored lexical and pragmatic knowledge. Developing a clear blueprint of the mechanisms that allow this process to take place has been a focus of psycholinguistic research for decades. In particular, there has been substantial debate concerning the relationships between the various levels of processing that compose the language comprehension system. Do the mechanisms that extract the different levels of linguistic representation from the stimulus act in isolation from one another? Or are the processing tiers interconnected, influencing one another during the application of their respective operations?

Essentially, the debate concerning the architecture of the language processor concerns the sequence in which processes from different levels cooperate. According to the serial autonomous view, language comprehension is strictly a bottom-up process, with processes at each level of representation operating in turn. The output of a lower level process serves as the input for the next level of computations. Word recognition processes, for example, generate the mental representations necessary for syntactic rules to be applied. Interactive views, on the other hand, hold that different kinds of linguistic knowledge can be applied simultaneously. Contextual information, then, might be used to guide the development of sentential structure by helping to select the syntactic rules that should be applied to a given string.

The distinction between computationally autonomous, or modular, processes as in Fodor (1983) and non-modular, or interactive, processes has played a key role in this debate. Within the literature on sentence processing, most of the focus has been on the syntactic level and the question of whether it is computationally independent from other parts of the grammar. Syntactic processing is a reasonable candidate for a modular process because the development of sentential structure could take place within a singular and narrowly defined domain. This is how Fodor (1983) defined peripheral processes. Peripheral processes are special purpose routines for inputting and outputting information. They can be distinguished from central processes like problem-solving and reasoning, which are general purpose cognitive processes, and therefore would not be computationally autonomous. The semantic level of processing, where meaning is determined, is seen as more central to human cognition, and thus less likely to be purely linguistic (Fodor, 1983).

Functionally, an autonomous syntax module in the language processor results in an architecture like that shown in Figure 1. Here, hierarchical structure is applied to incoming lexical information without recourse to contextual or pragmatic knowledge; thematic roles and word meanings are considered separately from structural function.

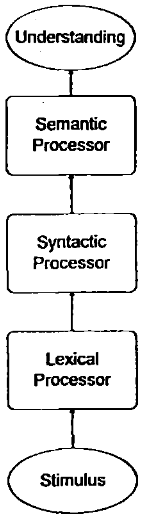

Figure 1 A modular view of the language processor including an autonomous syntax level.

Indeed, when language processing is perceived as the serial application of sets of rules, it is consistent with a modular architecture. But if information from multiple levels of representation influence processing decisions at a given level, the system takes on a more interactive character. Interactive models permit higher processing levels to influence processing decisions together with lower levels (see Boland and Tanenhaus, 1991; Marslen-Wilson and Tyler, 1980; Marslen-Wilson, Tyler, and Seidenberg, 1978; Tyler and Marslen-Wilson, 1977; Taraban and McClelland, 1988). The evaluation of the stimulus still requires an orderly application of grammatical rules, but the ränge of rules that may be applied at a given level is restricted by the outcome of computations at higher levels.

In contrast with their modular counterparts, many interactive models focus on the sharing of information between the syntactic and semantic levels of the processor. In the Marslen-Wilson and Tyler (1980) model, for example, the two levels in question are separate but are able to share information during the sentence comprehension process and guide one another in their respective manipulations of the incoming material (Figure 2). A more extreme model proposed by Bates and MacWhinney (1987), on the other hand, depicts syntactic and semantic information being dealt with on the same processing level; this results in a single mental representation encompassing both syntactic and semantic elements of the current input (Figure 3).

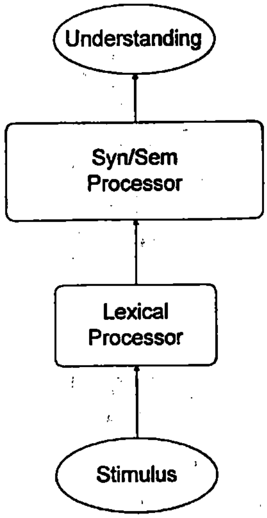

Figure 2 An interactive view of the language processor with parallel processing between the syntax and semantic level.

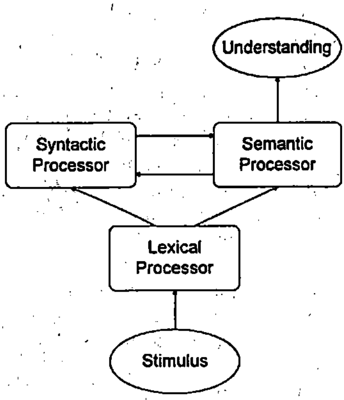

Figure 3 An interactive view of the language processor with a single level for the processing of syntactic and semantic information.

Determining what configuration of the language processor best describes human beings' actual comprehension performance has proven to be an extremely difficult task. Subjects in psycholinguistic experiments, for example, cannot self-report on how syntactic and semantic information come together to produce understanding because the process is normally entirely unconscious. Traditional experimental methodologies, such as reading time and dual task paradigms, have also failed to provide a resolution--a claim supported by the literature, which contains evidence in favor of both the modular and interactive architecture using these sorts of paradigms.

Recent advances in cognitive neuroscience, however, may allow us to investigate the architecture of the linguistic processor in terms of the physiological responses of the brain. If there are physiological differences between syntactic and semantic processing, for example, this would suggest that the two levels are separate. Were this true, there would be cause to prefer models like those in Figures 1 and 2 over the one offered by Bates and MacWhinney (1987). Moreover, physiological evidence that syntactic and semantic responses occur in a fixed order could be used to argue for a modular framework over an interactive one.

One method of measuring brain responses is electroencephalography, or EEG. This involves attaching electrodes to the surface of the head in predetermined locations (Figure 4). The brain's response to an external event is extracted from the EEG by averaging the recordings for several repetitions of the event. These extracted responses appear as waveforms called event-related potentials, or ERPs.

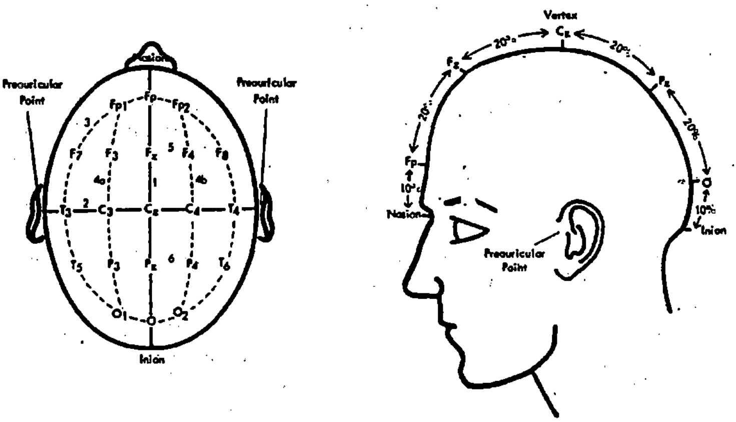

Figure 4 Locations for electrode placement according to the International 10-20 system.

An ERP waveform is composed of positive and negative peaks. These peaks reflect natural changes in the biochemical polarity of the brain that develop as neurons transfer information to one another. To make interpreting ERPs simpler, a standardized system has been devised to label prominent peaks. This includes information about whether a peak is positive or negative, and the average time it takes the peak to maximize after the onset of the stimulus. Thus, a large positive component peaking at approximately 100 milliseconds after a stimulus has been presented is referred to as a P100, while a negative component peaking 100 milliseconds later would be an N200.

Crucially, certain peaks on the ERP waveform seem to be directly correlated to particular kinds of language processing. The N400 wave, for example, has been linked to semantic congruency in two ways: predictability of a word to the context, with larger N400s for less predictable words (Kutas and Hillyard, 1980a, 1983; 1984; Polich, 1985), and the failure of a given word to fulfill thematic constraints (Garnsey, Tanenhaus, and Chapman, 1989).

Examples 1 and 2, below, show sample stimuli from two N400 studies. At the underlined word, the difference in the peak amplitudes of the N400 in sentences like (a) was compared with that in anomalous sentences like (b). The N400 was found to be significantly larger on the reading of the final word in the anomalous condition.

1) Kutas and Hillyard (1980a): N400 as a measure of predictability

a. I take coffee with cream and <u>sugar</u>
b. *I take coffee with cream and <u>mud</u>

2) Garnsey, Tanenhaus, and Chapman (1989): N400 as a measure of thematic constraint violation

a. Which customer did the secretary <u>call</u>?
b. *Which article did the secretary <u>call</u>?

In other work, positive waves known as the P300 and the P600 have been linked with language processing. The earlier component has been shown to reflect the relative importance of certain words to the meaning of the sentence (Friedman et al., 1975), as well as the physical congruency of the orthographic form of particular words in a sentence (Kutas and Hillyard, 1980b). The P600 has been found to appear after various syntactic incongruities in a sentence, such as subcategorization violations, (Hagoort et al., 1993; Osterhout and Holcomb, 1992; Osterhout et al., 1994). Because of some disagreement as to the uniqueness of the P600 as compared to the P300, some refer to the later component simply as the Syntactic Positive Shift (SPS), or, even less dramatically, as the Late Positivity (LP). While making no claims about the validity of P300/P600 contrast, I will refer to the positive component that is sensitive to syntactic incongruity here as the P600.

Example 3 provides sample stimuli from a P600 study by Osterhout and Holcomb (1992). As before, the amplitude of the wave in question was measured at the underlined word in control and anomalous sentence conditions. The P600 to the final word in (b) was significantly larger than that to the word in the same position in (a).

3) Osterhout and Holcomb (1992): P600 as a measure of verb subcategorization violation

a. The broker planned to conceal the <u>transaction</u>.
b. *The broker persuaded to buy the <u>stock</u>.

In short, research has shown that semantic anomalies induce an N400, while syntactic anomalies evoke a P600. While this may suggest a discussion concerning the discreteness of the syntactic and semantic levels of the language processor is moot, there are still several points to consider. First, studies to date investigating the N400 and the P600 have used entirely unrelated stimuli in different experiments to evoke each of the two waveforms, making the results difficult to compare. It is certainly possible that the N400 and P600 are independent phenomena generated in response to the processing of semantic or syntactic incongruencies, respectively. However, it might also be the case that the N400 and P600 are just very salient, measurable sub-parts of a single complex brainwave component that reflects general processing difficulty. Consider, for example, the possibility that there exists some complex waveform that is induced by language processing or some other cognitive process that occurs during the reading of sentences that is initially positive, but becomes negative towards its end. If this waveform were generated in the appropriate time course such that it overlaid an N400-P600 complex, it could conceivably minimize one of the components, making it appear as if only a significant negative or a significant positive shift had been evoked by the anomaly under study.

Thus, in order to get a clearer picture of what information the N400 and the P600 waveforms can truly provide concerning the nature of semantic and syntactic processing and the relationship between them, it is necessary to develop experimental stimuli that permit a comparison across syntactically anomalous and semantically anomalous conditions. In addition, there should be a condition where the critical word reflects both a syntactic and semantic anomaly; this would determine if a doubly incongruous element evokes a waveform in which the N400 and P600 remain distinct.

To this end, stimuli like that described in Example (4) were developed. Each sentence centers around a non-alternating dative verb, or verb that is subcategorized to take both a direct and indirect object, in that order. More specifically, each verb calls for a noun phrase (NP) direct object, to be followed by a prepositional phrase (PP) headed by the preposition 'to' and containing an animate NP indirect object. Every stimulus has four versions, each identical in its acceptability up to the point where the anomaly is introduced (indicated by the underlined word). Version (a), the control, is proper both syntactically and semantically according to the grammar. Version (b) is syntactically incorrect, failing to contain the required preposition, but remains semantically viable because the noun phrases that are present possess the necessary animacy features and can be interpreted meaningfully through application of thematic and real world knowledge.[2] Version (c) is syntactically sound, having the 'to', but semantically unacceptable as a result of the NP in the indirect object position being inanimate and thus violating the verb's thematic constraints. Version (d) is erroneous both syntactically and semantically, having no preposition 'to' and an inanimate indirect object NP.

(4) a. Lee introduced his dog to everyone at the big party.
  b. Lee introduced his dog everyone at the big party.
  c. Lee introduced his dog to entrances at the big party.
  d. Lee introduced his dog entrances at the big party.

Assume that structural issues like subcategorization demands made by the verb are defined as syntactic, while thematic constraints like animacy value are defined as semantic, and that there are different waveform patterns produced when violations occur at each of these levels of linguistic computation. It should be the case, then, that individuals seeing condition (b) should have a significantly larger P600 at the critical word than those seeing condition (a), with no difference between the N400 amplitudes in the two conditions. Those seeing condition (c) should display just the opposite pattern, namely a significantly larger N400 than the control, but not a larger P600. People seeing condition (d) should produce both significantly larger P600s and N400s as compared to in condition (a). Moreover, if the two waveforms in question are truly independent, then the amplitudes of the N400 and P600 in the doubly anomalous condition should be no larger than the same waveforms evoked in the single anomaly. In addition, the waveforms in condition (d) should have similar waveform parameters, such as scalp distribution and waveform topography, as their counterparts in the singly anomalous conditions.

The final concern relates to rate of presentation. It is common for studies of this nature to use time windows around 650 milliseconds for the presentation of each word (e.g., Osterhout and Holcomb, 1992). Considering that one of the components of interest is a broad waveform with a midpoint around 600 milliseconds post stimulus, such a short presentation window could easily result in this late component overlapping

[2]Questionnaires completed by an independent subject group indicate that sentences in the syntactically anomalous condition are interpreted in the manner intended, despite their structural malformation.

with early processing components evoked by the word after the stimulus, making it difficult to measure accurately. Moreover, if a short window of presentation allowed late processing components induced by the word before the critical word to overlap with waves initiated early by that critical word--waves like the N400--it might be difficult to evaluate the actual significance of these waves as compared to those elicited at the same location in the control. To address this issue, the current stimuli were presented at a slower rate, allowing 1000 milliseconds for each word. While this is much slower than average reading speed, it greatly reduces the chance of waveforms from different words overlapping and interfering with the measurements of the components of interest

METHOD

Subjects Thirty-two Ohio State University students from an introductory psychology class participated in this experiment as part of their course requirements. All of the students were native speakers of English, had no reading disabilities, and had normal or corrected-to-normal vision.

Materials Stimuli for this experiment used 20 non-alternating dative verbs identified through an earlier norming study (see Example 4). Two different sentence sets were constructed from each verb, for a total of 40 critical trials. Each sentence had four versions: a control; a syntactically anomalous condition; a semantically anomalous condition; and a doubly anomalous condition that contained both a syntactic and a semantic violation. The validity of the intended interpretation of the anomalous sentences was confirmed through a series of sentence completion questionnaires completed by a separate group of subjects.

There were four experimental lists, each with one version of the 40 critical trials. Conditions were rotated across the four lists so that there were equal numbers of each condition on each list. Each list also contained 40 distracter trials of various syntactic types to prevent subjects from recognizing patters in the critical trials and developing a strategy of response.

Apparatus Surface electrodes were attached to the scalp of each subject at the frontal (Fz), central (Cz), and parietal (Pz) sites located along the midline between the bridge of the nose and the base of the skull. Eye movements were measured by means of additional electrodes above and below the left orbit; jaw movements were measured by electrodes placed on each mastoid. Signal recordings were referenced to the left mastoid. A separate channel recorded the left mastoid referenced to the right. The data were subsequently referenced digitally to the average of the two mastoid electrodes.

Utilizing a Grass Model-12 Neurodata Acquisition system, the EEG recordings were amplified and digitized on-line with a sampling frequency of 100 Hz. Sampling began two words before the presentation of the critical word in each trial (at the third word), with sampling epochs varying according to the length of each sentence. The input data were bandpass filtered with cutoffs of .01 Hz and 30 Hz.

PROCEDURE

The stimuli were presented to each subject in a computer-generated random order. Eight subjects saw each list. Each sentence appeared word-by-word in the center of a computer screen, with the word framed by a white line box 10 centimeters across and eight centimeters high. Each word was presented for 500 milliseconds, followed by a 500 millisecond interval when the box was empty. Subjects were asked to read each word carefully and try to link the words together in their minds to produce a comprehensible sentence.

After a trial was completed, the white box disappeared, leaving the screen blank and alerting the subject that she could blink her eyes. The screen remained blank until the subject pressed a key on a computer keyboard indicating that she was ready for the next trial. At this juncture, one-third of the trials were followed by YES/NO comprehension questions that appeared at the top of the screen; these were included to motivate subjects to be attentive to the sentences. The subject was asked to consider the truth of the question based on the sentence she had just seen and to respond either YES or NO by pressing the Y or N keys on the keyboard. After the subject had given her response, the screen again went blank. Only upon a second keypress was the white box brought back to the center of the screen so that the next trial could begin. If there was no comprehension question, then the first keypress after the subject's "blink break" initiated the onset of the next trial.

Subjects were seated approximately 75 centimeters away from the computer screen, resulting in a horizontal visual angle of at least two degrees and a vertical visual angle of three degrees. Testing took place in a sound-proof cubicle, with the subject seated in a comfortable chair. Before any responses were recorded, each subject ran through 10 practice trials to become familiar with the experimental procedure. Participants were told to expect grammatical problems with some of the sentences, but were not told precisely what types of problems. Including electrode application and removal, each subject's session lasted approximately 90 minutes.

RESULTS

The raw EEG data were digitized and visually evaluated on a trial-by-trial basis for excessive eye or jaw movement that would interfere with standard manipulation of the data for analysis; trials where irreparable amounts of movement occurred were eliminated. Remaining trials were first corrected for the influence of blinking, if necessary, and were then averaged by condition for each subject. These averages were compiled into a grand average across subjects for each condition. The mean N400 and P600 voltages following the critical word were measured and subtracted from a baseline voltage taken from the average of a 50 milliseconds window beginning 250 milliseconds prior to the onset of the critical word.

Responses to the anomalous word in each of the three critical conditions at the Cz and Pz electrodes are contrasted with the control in Figures 5a and 5b, respectively. Each waveform graph shows the grand averaged responses for each pair of conditions from the word prior to the critical word through two words after the critical word. The onset of the critical word is indicated by an arrow on the x-axis timeline. Consistent with previous research in this area, the N400 waveform was found to be most robust at the Cz site, while the P600 was most salient at the Pz site.

The N400 was quantified as the mean voltage in a 40 milliseconds window centered at 400 milliseconds after the onset of the critical word. Because the P600 is a broader, longer lasting potential, it conversely was quantified as the mean voltage in a 400 millisecond window centered 880 milliseconds after the same onset. A bar graph of the average amplitude of the N400 at the Cz and the P600 at Pz is given in Figure 6.

A two-way analysis of variance (ANOVA) was performed with repeated measures on four levels of anomaly type and three levels of electrode position (frontal, central, and parietal). Significance tests were adjusted by the Geisser-Greenhouse correction. This analysis revealed a main effect of electrode and condition for the N400 waveform [$F(1,31) = 11.10$, $P < .01$; $F(3,93) = 4.26$, $P < .01$], and for the P600 waveform [$F(1,31) = 9.68$, $P < .01$; $F(3,93) = 3.01$, $P < .05$]. Paired t-tests by electrode between the mean amplitudes of the N400 waveform in each critical condition with that of the control revealed a reliable difference (2-tail $P < .05$) between the semantically anomalous condition and the control at Cz and Pz. There was also a reliable difference between the doubly anomalous condition and the control at Fz and Cz, with a marginal effect at Pz. The same tests performed for the P600 showed a reliable difference between the syntactically anomalous condition and the control at Pz. At Fz, there was a marginal difference (2-tail $P < .10$) between the control and both the semantically anomalous and doubly anomalous conditions.

DISCUSSION

Our results provide evidence that distinctive waveform patterns are evoked in response to syntactic anomaly, semantic anomaly, and a combination of the two. The syntactic anomaly condition evoked a strong, broad positivity beginning around 600 milliseconds after the presentation of the indirect object, a result that is consistent with previous findings by Osterhout and colleagues (1992, 1994) and Hagoort et al. (1993). The semantic anomaly condition evoked a sharp negativity centering at approximately 400 milliseconds after the presentation of the indirect object, in keeping with reports by Kutas and Hillyard (1980a, 1983, 1984) and Garnsey, Tanenhaus, and Chapman (1989) concerning the N400 component. This pattern of results suggests that the language processor does distinguish between anomalies normally defined as syntactic and semantic within the grammar.

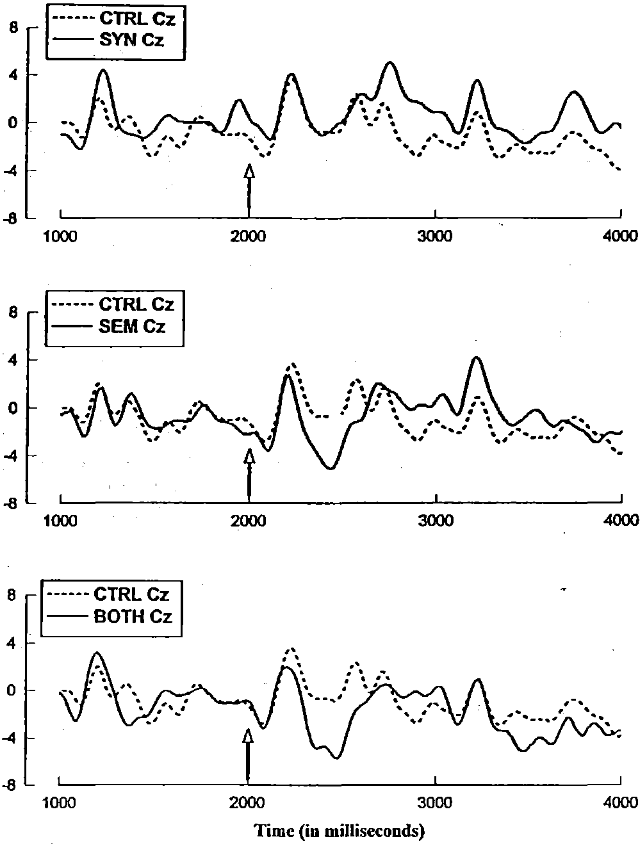

Figure 5a Grand-averaged waveforms at Cz from one prior to one word after the critical word, with each of the anomalous conditions contrasted with the control. The arrow indicates the onset of the critical word. SYN is the syntactically anomalous condition, SEM is the semantically anomalous condition, and BOTH is the doubly anomalous condition.

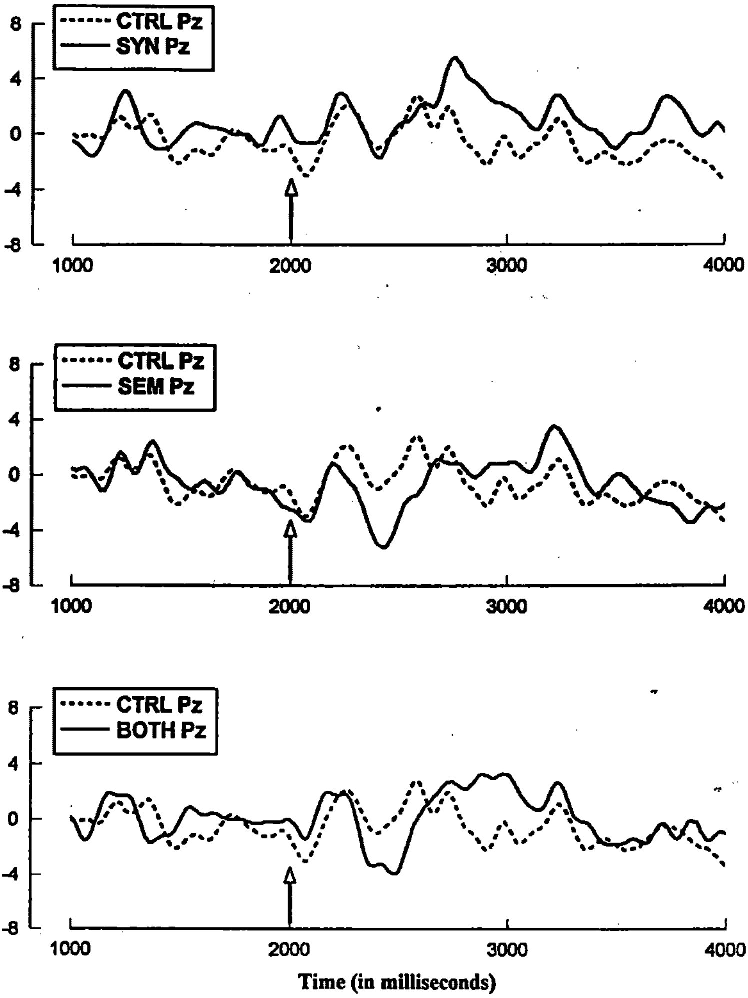

Figure 5b Grand-averaged waveforms at Pz from one prior to one word after the critical word, with each of the anomalous conditions contrasted with the control.

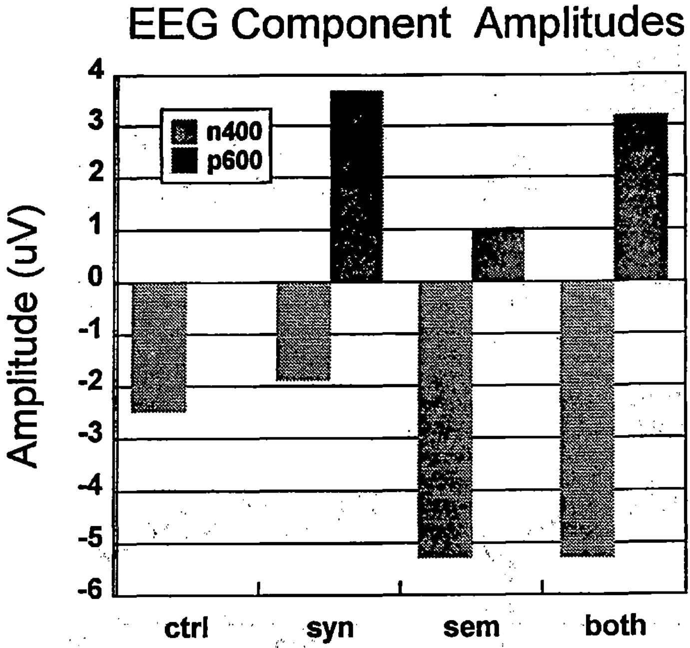

Figure 6 Mean amplitudes in a 40 msec window for the N400 waveform at Cz and in a 400 msec window for the P600 waveform at Pz for each condition.

Although I take this to be evidence that there are distinct cognitive processes associated with syntactic and semantic analysis, it is possible that my anomaly types, and thus my findings, could be interpreted in other ways. For example, it seems intuitively true that the syntactically anomalous condition is less shocking or less inappropriate than the semantically anomalous one. Instead of reflecting semantic processing, then, the N400 could be reflective of major anomaly or a high degree of processing difficulty. The P600, in turn, could reflect a minor anomaly or a low degree of processing difficulty (rather than syntactic processing in particular). From another perspective, our syntactically anomalous sentences are very easy to repair, requiring only the insertion of a highly predictable preposition to regain their grammaticality. The semantically anomalous sentences are much more difficult to fix considering the wide variety of animate nouns a subject could choose from to make them coherent. If ease-of-fix, not anomaly type, was the key factor, it could be that our two waveforms reflect different levels of reparability, not structural and interpretational problems per se.

A closer look at these alternatives, however, reveals that both interpretations may be perfectly confounded with the original predictions. In other words, if the N400 and the P600 reflect degree of processing difficulty or reparability instead of anomaly type, but these factors are directly correlated with what the grammar refers to as semantics and syntax, my line of argumentation is not seriously affected. The current data still offers substantial evidence that the ERP methodology can effectively discriminate between the processing of these two types of sentential information.

Another concern is the purity of the semantically anomalous condition. Recall that the verbs in the current stimuli were chosen because they were non-alternating datives, calling for a NP PP complement. A post-experimental review of these verbs, however, revealed that some of them could also take an infinitive verb phrase (VP^{inf}), such as 'to go' or 'to be', in place of the complement PP. As a result, some subjects could have interpreted the word 'to' after the NP direct object as an infinitive marker, not as the head of a prepositional phrase. Those that perceived the infinitive marker would, in turn, expect the next word to be an infinitive verb form, not an NP. When they did encounter the indirect object NP and attempt to incorporate it into the developing sentential structure, the categorical mismatch of the stimulus and what was expected could easily induce a syntactic processing difficulty.

To address this concern, a post hoc comparison was performed on the sentences in the semantically anomalous condition. Three native English speakers were presented with the stimuli up to the word 'to' and asked to provide both noun and verb completions for each fragment. Stimuli for which one or none of the evaluators were able to think of a completion beginning with a verb were considered "pure" NP PP stimuli; all others were categorized as NP VP^{inf} stimuli. There were 20 sentences in each group. The blink-corrected data for the two groups of stimuli were grand averaged and the resulting waveforms were compared. If the NP VP^{inf} set of stimuli had a larger P600 at the critical word than the NP PP set, there would be evidence that a syntactic anomaly had been introduced. Moreover, if the N400 for the NP VP^{inf} stimuli was smaller than for the NP PP sentences, this would suggest that the intended semantic anomaly had not been perceived by the subjects. Utilizing the same baseline and peak window parameters as before, I examined the amplitudes of the N400 and P600 in both sets of stimuli. The largest difference between the two waveforms was two microvolts, which was clearly not significant. These findings suggest that few subjects, if any, misinterpreted 'to' as an infinitive marker or experienced a syntactic processing difficulty when reading the semantically anomalous sentences.

Lastly, I must address the possibility that our pattern of results may have been influenced by design differences between our study and previous work. In particular, our rate of presentation allowed subjects to view each word for 500 milliseconds (with an additional 500 milliseconds delay between each word) in order to reduce the overlap of brainwave components. A much more common rate is 300 milliseconds display time per word (with an 350 milliseconds delay between words), as found in Osterhout and

Holcomb (1992). Considering that the average person reads text at a rate of 200-300 milliseconds per word, participants are required to read at a substantially slower pace than normal in either case, but the difference was more exaggerated in the present study. It is important to know whether the slower presentation rate could have led to abnormal reading behavior.

To explore this issue, I performed two replications of Osterhout and Holcomb (1992), one using the original 650 millisecond stimulus onset asynchrony (SOA) and one using 1000 SOA. In both cases, I found that the P600 was a significant and identifiable measure of verb subcategorization violation, with the positivity being only slightly attenuated at the slower presentation rate as compared to the faster one. Because these findings are consistent with those published in the source paper, I conclude that for responses as robust as the N400 and P600, increasing the rate of presentation to 1000 milliseconds SOA does not affect the waveform patterns evoked by the type of stimuli in question. Moreover, the slower rate provides the benefit of minimizing component overlap and, thus, facilitating an analysis of the results.

Having dealt with these concerns, I may now evaluate how effectively each of the two interactive configurations of the linguistic processing model predicted our findings. It is almost immediately clear that the configuration in the inclusive model proposed by Bates and MacWhinney (1987), is not supported by our data. If syntactic and semantic processing were taking place by means of the same mechanism, I would not expect to find different waveform components correlated to each of these two levels of linguistic knowledge. The autonomous interactive model, conversely, seems to make the right predictions to account for our data: two levels of processing that have independent manifestations in the realm of electrochemical response.

One might be tempted to suggest that our findings can even reveal which of the two remaining models, namely the Ferreira and Clifton (1986) serial model and the Marslen-Wilson and Tyler (1980) interactive model, provides the most useful description of the relationship between the syntactic and semantic levels of processing. This is because, while both predict that the syntactic and semantic processing functions are discrete, they have different hypotheses about when syntactic and semantic processing take place with respect to each other. The serial processing model maintains that syntactic processing must be complete before semantic processing can begin, whereas the interactive model allows for semantic processing to begin before a final syntactic evaluation of the stimulus has taken place. Even the very labels of the N400 and P600 waveforms tell us that the correlate to semantic processing manifests earlier in the time course of comprehension than does the correlate for syntactic processing. This is consistent with the implication that semantic processing begins before syntactic processing finishes, and thus supports the interactive framework. When making any assumptions about the association between cognitive processes and neurobiological phenomena, however, it is prudent to be cautious. There is no evidence that whatever temporal relationship exists between the output of the syntactic and the semantic levels of processing is preserved in their manifestations as event-related potentials. If these processes are independent, as I am now wont to argue, they could well have completely different functions mapping their cognitive and physical states. This possibility prevents us from drawing too bold a conclusion concerning the limited set of data I have presented here.

This does not, however, mean that it is not possible to use the ERP paradigm to try and determine whether the relationship between the syntactic and semantic levels of the language processor is autonomous or interactive. Indeed, I am currently developing an experiment involving the processing of garden path filler-gap sentences like that in (5) to address this very issue.

(5) a. Which athlete$_i$ did the coach encourage _ _ _ John to watch ____$_i$?
b. Which game$_j$ did the coach encourage _ _ _ John to watch ____$_j$?

Native English speakers have the tendency to assume the fronted NP in (5a), *athlete*, is the direct object of the verb *encourage* (indicated by broken underline), when in fact is actually the direct object of *watch* (indicated by solid underline). When they encounter *John* in the expected gap, speakers realize that they have misanalyzed the sentence and must conduct repairs to make the sentence comprehensible. My colleagues and I are interested in investigating the electrophysiological manifestations of the 'garden path experience' found with these types of sentences. Does a processing difficulty arise because *athlete* and *John* are assigned to the same structural position, giving rise to a syntactic anomaly, or is the problem that the two NPs have been assigned the same thematic role, resulting in a semantic anomaly? Is it both?

A contrastive analysis of the N400 and P600 waveforms produced at the critical gaps in sentences like that in (5a) with that in (5b) should help us find the answer. The serial model predicts that any processing difficulty evoked by the reading of *John* in (5a) should be syntactic and, therefore, marked by a large P600; there should be no N400 at this word position, because, according to this framework, semantic processing of the clause would not yet have taken place. By this same reasoning, the fronted NP *game* in (5b) should evoke a P600 as well, because there is no selectional information, like animacy value, available during the parse to remove this noun phrase as from consideration as a filler for the direct object position.

While the interactive model would also allow for the elicitation of a P600 upon reading the noun phrase *John* in (5a), it differs crucially in its prediction for (5b). Because this model allows for semantic processing to occur in parallel with syntactic processing, it should be possible for the interpreter to identify *game* as an inappropriate theme for the verb *encourage*--due to the concept's lack of animacy--and in turn guide the parser to search for a more appropriate noun phrase to fill the direct object position. The parser should, therefore, have no difficulty when it encounters the noun phrase

John and elects to place it in the direct object slot, and thereby have no cause to generate a P600 waveform.

In sum, I offer evidence that syntactic and semantic processing are indeed independent functions of the language processing system and, thus, that only a model that allows for the separation of these two levels of information can effectively predict the pattern of results displayed in our data. Moreover, I believe this study reinforces the notion that the ERP paradigm is a viable and useful tool for psycholinguistic research.

REFERENCES

Bates, E. and MacWhinney, B. (1987). Competition, variation, and language learning. In MacWhinney, B. (ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum.

Boland, J.E., and Tanenhaus, M.K. (1991). The role of lexical representations in sentence processing. In Simpson, G.B. (ed.), *Understanding Word and Sentence*. North-Holland: Elsevier Science Publishers B.V.

Carroll, J.M., Tanenhaus, M.K., and Bever, T.G. (1978). The perception of relations: the interaction of structural, functional, and contextual factors in the segmentation of sentences. In Levelt, W.J.M. and Flores D'Arcais, G.B. (eds.), *Studies in the Perception of Language*. Chichester: Wiley.

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

Fodor, J.A. (1983). *Modularity of Mind: An Essay in Faculty Psychology*. Cambridge, MA: MIT Press.

Forster, K.I. (1979). Levels of processing and the structure of the language processor. In Cooper, W.E. and Walker, E.C.T. (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.

Friedman, D., Simson, R., Ritter, and Rapin, I. (1975). The late positive component (P300) and information processing in sentences. *Electroencephalography and Clinical Neuropsychology*, **38**, 255-262.

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

Garrett, M.F. (1970). Does ambiguity complicate the perception of sentences? In Flores D'Arcais, G.B. and Levelt, W.J.M. (eds.), *Advances in Psycholinguistics*. North-Holland: Elsevier Science Publishers B.V.

Hagoort, P., Brown, C., and Groothausen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, **8** (4), 439-484.

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

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

Kutas, M., and Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, **11** (5), 539-550.

Kutas, M., and Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, **307**, 161-163.

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

Marslen-Wilson, W.D., Tyler, L.K., and Seidenberg, M. (1978). Sentence processing and the clause boundary. In Levelt, W.J.M. and Flores D'Arcais, G.B. (eds.), *Studies in the Perception of Language*. Chichester: Wiley.

Mitchell, D.C. (1989). Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes*, 4, 123-154.

Osterhout, L., and Holcomb, P.J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language*, **31**, 785-806.

Osterhout, L., Holcomb, P.J., and Swinney, D.A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology*, **20** (4), 786-803.

Polich, J. (1985). Semantic categorization and event-related potentials. *Brain and Language*, **26**, 304-321.

Taraban, R. and McClelland, J.L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, **27**, 597-632.

Tyler, L.K. and Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 683-692.