Bugs, boxes, and binding: ERP indexes and the modularity of the language processor

by
Seana Coulson, Jill Weckerly, and Kathleen Ahrens

Abstract: Modular processes are characterized as having four main attributes: domain specificity, obligatory action, information encapsulation, and automaticity. ERP literature reviewed here concerns the extent to which the defining features of modularity can be said to characterize two paradigmatic examples of modular language processes: *parsing* and *lexical access*. The ERP literature on parsing chiefly concerns the search for a unique electrophysiological index of syntactic processing and has focussed on two ERP components: a left anterior negativity (LAN) which occurs 300-500 msec post-stimulus onset and a late positivity (alternately dubbed the P600 and SPS) with widely varying latency and distribution. Although the LAN and the P600 are elicited in response to syntactic anomalies, their occurrence can also be explained by recourse to domain-general processes such as the operation of working memory and updating hypotheses about the structural characteristics of the linguistic environment.

The ERP literature on lexical access chiefly concerns the extent to which the initial stages of word recognition are influenced by higher-level contextual factors. Empirical tests of this issue have employed the N400 component as a continuous index of semantic priming, testing claims from the behavioral literature on the lexical access of ambiguous words and comparison of lexical access in word pair and sentential contexts. Discussion of these issues involves consideration of the extent to which the N400 component can be considered a valid index of semantic priming, and a reconceptualization of the distinction between lexical and post-lexical processes and their relationship to semantic priming.

Overall, we find the ERP literature on the modularity of the language processor is best described by Cottrell's (1985) phrase *leaky modularity*, and point to the connectionist modelling paradigm as providing an apt metaphor for understanding the ways in which language processes are modular in some respects and non-modular in others.

Acknowledgements: Thanks to Robert Kluender and Marta Kutas for many helpful comments on earlier drafts of this paper. Jill Weckerly and Kathleen Ahrens were supported in part by a grant to M. Kutas from NICHD (HD22614).

Linguistic Notes from La Jolla (1995), Number 18, 82-117

1. Introduction

In this paper we review some of the electrophysiological evidence which bears on the question of whether language processing is modular. In the first section we briefly outline the modularity thesis and how it pertains to language processing. Empirical tests of the modularity thesis have focussed on two proposed modules -- one for lexical access and one for syntactic parsing -- and tend to address what we call the Watergate questions of natural language processing:

i) what do subjects know about the linguistic stimuli they process?

ii) when do they know it?

Under the heading of *What do they know?*, we review the attempts of various investigators to identify an ERP component which indexes syntactic processing and outline how this evidence bears on the modularity thesis. Under the heading of *When do they know it?*, we review some of the electrophysiological literature on the time course of the lexical access process.

1.1 Modularity

In Fodor's (1983) framework, cognition is the interaction of a large number of autonomous, highly-specialized input modules with a general purpose central processor. Input modules can be conceptualized as an array of black boxes, each of which transforms a particular sort of input from the world into a representation which can be handled by the central processor. Input modules perform highly-specialized computations which feed into the central processor that integrates them.

Input systems are referred to as black boxes because they are informationally encapsulated with respect to the central processor, as well as with respect to each other. Each input module takes a particular sort of input, performs a series of transformations on that input, and outputs a representation to the central processor. The central processor has access only to the outputs of the input modules and not to the intervening representations in the modules themselves. Moreover, input modules operate autonomously and thus independently of the modules which surround them.

The role of the central processor is to combine the information which is computed in the input modules, and, when relevant, direct information to other encapsulated modules for further processing. Although the computations done by the individual modules are quite simple, the computation done by the system as a whole may be quite complex. Thus modularity can be seen as a computational form of the divide and conquer strategy: divide up the computational task into its component parts and compute them separately. Each input computes its particular part of the task independently of the others and feeds its results into the central processor for coordination with the output of the other modules.

1.2 Modular Approach to Language Processing

A modular approach to language processing involves the assignment of 'low-level' aspects of processing (such as parsing and word recognition) to informationally encapsulated input modules, while leaving the higher-level aspects such as semantics and pragmatics to the central processor. The difference between modular and non-modular accounts chiefly concerns the time course of processing. In the modular account, lower levels of processing occur autonomously and are integrated later by the central processor. However, on a non-modular account the lower levels of processing are not independent of higher levels but interact with them in the processing of a sentence. While all parties agree that the various levels of linguistic analysis interact, they make different predictions as to exactly when the results of higher-level analyses become available.

Empirical tests of the modularity thesis have tended to focus on whether or not higher-level contextual information can influence processing at lower levels. In the behavioral literature, such tests often involve interrupting the subject in the course of linguistic processing and requiring her to perform a task which requires higher-level information about the stimuli. Especially prominent is the use of measures such as reaction time for lexical decision or naming tasks, both of which are dependent upon the phenomenon of ***semantic priming***, the facilitatory effect observed in the processing of a word which has been preceded by a semantically related word.

However, the use of the reaction time paradigm to address questions pertaining to the availability of high-level information about the stimuli has been criticized because it requires the subject to terminate linguistic processing and to initiate the performance of the experimental task. Ideally one would like an on-line measure of subjects' processing which does not interfere with the linguistic processing itself.

Difficulties associated with behavioral measures of on-line language processing have led many investigators to supplement existing techniques with the use of electrophysiological measures. Especially useful are event-related potentials (ERPs), patterned voltage changes in the ongoing electroencephalogram (EEG) that are time-locked to the onset of a stimulus which requires linguistic processing. ERPs are obtained by recording subjects' EEG and averaging across timelocked events.

Because it can provide a continuous on-line index of the processing which occurs at the advent of a linguistic stimulus, ERPs are well-suited for addressing questions which have to do with what sorts of information subjects are sensitive to and when. In particular, the N400 component, a negative-going component which peaks at approximately 400 msec post-stimulus, has proven to be important in this respect. The amplitude of the N400 component is proportional to the difficulty of integrating a given word into established context (see Kutas and Kluender, 1991 or Van Petten and Kutas, 1991 for review) and can be used as a reliable indicator of semantic priming. Because the N400 is sensitive to the same processes indirectly assessed in the reaction time paradigm, we can view its use in investigations of the modularity thesis as an analogous but more direct version of behavioral measures.

Two linguistic processes which have traditionally been offered as paradigmatic examples of modular processing are the parsing module, which assigns syntactic representations to words, and the lexical access module, which accesses the lexical entry for words such that their core meanings are activated. The modular account dictates that the parser and the lexical access modules each operate quickly, automatically, and independently of any contextual information which might exist elsewhere in the system. Perhaps because investigation of these processes bears so directly on the modularity thesis, their exact nature has been hotly contested in the behavioral literature. Confronting a host of contradictory results, electrophysiologists have entered the fray. The section below -- ***What do they know?*** -- reviews some of the ERP literature addressing the question of the modularity of the parser; the third section -- ***When do they know it?*** -- reviews ERP literature addressing the modularity of lexical access; finally, in the concluding section we summarize our conclusions and explore their relevance to the modularity thesis.

2. What do they know?

On a modular account of language processing, understanding linguistic utterances is dependent upon the assignment of syntactic structure to the string of words. This process is known as *parsing*. Moreover, advocates of modularity (see esp. Fodor, 1983 and Chomsky 1986 for a view opposed to Fodor's) frequently offer parsing as the prototypical example of a modular process. We parse sentences automatically and with great facility -- exactly the type of behavior one expects to elicit from a system whose operations are fast-acting, mandatory, and highly specialized. Further, the impossibility of introspecting upon the actual process of deriving a syntactic representation suggests that it is an encapsulated process.

The main issue with respect to the modularity thesis is whether or not parsing occurs exclusively on the basis of syntactic principles, or whether syntactic, semantic, and pragmatic principles are simultaneously exploited in the course of interpreting utterances. Although proponents of both approaches admit that eventually both sorts of information bear on the interpretation of utterances, they differ in the degree to which the two sorts of information can interact in the initial stages of processing.

2.1 Filler Gap Dependencies

One example of electrophysiological research which has implications for the modularity thesis involves the processing of filler-gap dependencies. In English, creating an information or wh-question usually involves moving the questioned constituent from its usual position, as in (a), to the beginning of the sentence, as in (b):

(a) She told him to attend the meeting.

(b) Who did she tell __ to attend the meeting?

The displaced constituent *who* is referred to as a *filler* while the place it is moved from (in this case, the direct object position) is known as the *gap*.

In order to interpret a wh-question, the parser must assign both a syntactic function and a semantic role to the filler -- a process which involves locating and filling the appropriate gap. If there is a filler but no gap (as in (c)),

(c) What did she put the book on the table?

or a gap with no filler (as in sentence (d)),

(d) Did she put __ on the table?

the sentence will be uninterpretable.

Faced with a temporary ambiguity in the filler-gap construction, the parser has at least two possible strategies. One possibility is the *first resort* strategy in which the parser assigns a filler to the first possible position in the sentence where a gap could conceivably occur, regardless of whether or not that position turns out to be the position where the gap actually occurs. Alternatively, the parser could employ a *last resort* strategy in which it waits until there is unambiguous information about where the actual gap is. A modular parser would probably employ the first resort strategy because it involves a process which is fast, automatic, and operates independently of semantic processing. However, a non-modular parser might also employ the first-resort strategy, assigning a syntactic representation based upon imperfect semantic information.

Garnsey, Tanenhaus, and Chapman (1991)

Testing whether the parser employs a first resort or a last resort strategy, Garnsey et al. constructed sentences with embedded wh-questions in which the filler was either pragmatically plausible or implausible as the direct object of the embedded verb. If the parser employs the first resort strategy it will assign the filler to the first possible gap location -- in this case immediately after the verb *called*. However, use of the last resort strategy would involve waiting until the assignment of the gap is unambiguous, *viz.* until the end of the sentence.

| Plausible

(a) The businessman knew which customer the secretary called __ at home.

Implausible

(b) The businessman knew which article the secretary called __ at home.

If the parser assigns the filler to the first possible gap location (*called*), the subject would be expected to register the presence of a semantic anomaly in the implausible condition, but not in the plausible condition. If, on the other hand, the parser waits until gap assignment is unambiguous, the subject will not register semantic anomaly in the implausible condition until the end of the sentence, where it becomes clear that there is no other possible gap location to which the filler can be assigned.

We can thus utilize the known sensitivity of the N400 component to semantic anomaly to distinguish between a first resort and a last resort strategy. The use of a first resort strategy would result in a larger N400 to *called* in the implausible condition, whereas the use of a last resort strategy would result in N400s of approximately equal amplitude because the subject would not register the anomaly until the end of the sentence.

In fact, Garnsey et al.'s results were consistent with the first resort strategy: while there was no N400 response to *called* in sentences like (a), in which the filler could plausibly be assigned to the gap, there was an N400 response to *called* in sentences like (b),where the filler could *not* plausibly be assigned to the gap. The first resort strategy is consonant with the modularity thesis because it involves a process which is fast, encapsulated, and operates mandatorily. Garnsey's findings can thus be interpreted as supporting the notion of a modular parser.

Although the results reported by Garnsey et al. (1989) go against a completely interactive approach -- because an interactive parser would presumably utilize its semantic/pragmatic knowledge and not assign the filler to the gap in the implausible condition -- the results also go against a completely modular approach. While one would expect a modular parser to attempt filler-gap assignment at the first possible opportunity, one would *not* expect the evaluation of its semantic plausibility until a later stage in processing (Kutas and Kluender, 1991). The elicitation of the N400 component in response to the verb *called* in the implausible condition is an indication that subjects are already sensitive to the semantic incongruity of the first resort filler gap assignment. Moreover, an alternative possibility is that the parser is both meaning-driven *and* uses the first-resort strategy. This would entail a system which constantly tries to make sense of its inputs and assigns syntactic representations according to the best approximation of the meaning at the time of input.

2.2 In search of a syntax wave

Given the assumption that qualitative differences in the ERP waveform reflect the operation of qualitatively different cognitive processes, ERPs can also be used to identify the operation of different cognitive processes as they occur in the interpretation of linguistic stimuli. If syntactic and semantic processing are subserved by independent modules, one might expect the two sorts of processing to be manifested in qualitative differences in the ERP.

Another possible use of the ERP measure, then, would be to identify different components in the waveform which index different levels of processing. This could be brought to bear on the modularity thesis in two ways: first, the very existence of separate components in the waveform which index functionally distinct levels of processing could be seen as implicit support for the modularity thesis. Moreover, once identified, an ERP index of syntactic processing could be examined in various contexts to reveal the extent to which semantic processing either does or does not influence the operation of the parser. An ERP index of syntactic processing would also be useful in evaluating claims about the timecourse of syntactic processing.

Osterhout and Holcomb, 1992

One of the first studies to utilize the ERP measure to address the modular nature of syntactic processing is described in Osterhout and Holcomb (1992). In keeping with the goal of identifying an ERP index of syntactic processing, Osterhout and Holcomb examined subjects' ERPs in response to sentences with syntactic errors, such as the violation of verb subcategorization constraints, as well as violations of phrase structure constraints.

1. Verb Subcategorization Violations

Verb subcategorization constraints involve the specification of the number and type of sentence constituents which must (or may) act as complements of a verb. The lexical entry for verbs contains a specification of the different sorts of constructions which are required or permitted to appear after them. For example, the verb *hoped* requires either an infinitival construction or a prepositional phrase headed by *for*. The verb *persuaded*, on the other hand, requires an object noun phrase, optionally followed by an infinitival complement. A subcategorization violation occurs when a verb is followed by a constituent for which it does not subcategorize.

In order to test the electrophysiological response to violations of verb subcategorization constraints, Osterhout and Holcomb recorded subjects' ERPs as they read sentences such as:

(a) The broker hoped *to* sell the stock.
(b) *The broker persuaded *to* sell the stock.

While sentence (a) above is grammatical, sentence (b) is ungrammatical because the infinitive *to* violates the subcategorization constraints of the verb *persuaded*, which requires a following NP direct object. Osterhout and Holcomb identified a slow, positive-going component with right anterior distribution in response to the infinitival marker *to* in sentences like (b) in contrast to the ERP elicited by *to* in sentences like (a). Although the positive-going component which occurred between 500 and 700 msec post-stimulus did not have a peak, it was dubbed the P600 by Osterhout and Holcomb because its midpoint was about 600 msec post-stimulus onset.

However, interpretation of this result is problematic because there is a possible, albeit improbable, grammatical interpretation of sentence (b) under a passive reading of *persuaded*. For example:

(c) The broker (who was) persuaded to sell the stock was sent to jail.

Strictly speaking, then, sentence (b) does not become ungrammatical until *stock*, when the end of the sentence signals the impossibility of the grammatical reading of the sentence. Moreover, the classification of *persuaded to* as a subcategorization violation depends upon how *persuaded* is interpreted. If it is interpreted on its active reading, *persuaded to* is indeed a subcategorization violation. However, if it is interpreted on its passive reading, *persuaded to* is perfectly grammatical.

Osterhout et al.'s interpretation of subjects' ERP response to *to* in (b) thus depends upon the assumption that *persuaded* is assigned the more probable, active reading. However, given the possible grammatical continuation of (b) at the point of recording, the P600 might have been elicited in response to a reinterpretation of *persuaded* under its passive reading. In that case, the event indexed by the P600 is not a subcategorization violation, but rather a syntactic reinterpretation in a sentence which, as far as the parser is concerned, is perfectly grammatical.

2. Phrase Structure Violations

The second sort of anomaly tested by Osterhout and Holcomb was a violation of phrase structure. This involved a comparison of subjects' ERPs to the auxiliary verb *was* in sentences like (d) as opposed to sentences like (c). Because the verb *hoped* does not allow passivization, sentence (d) with a reduced relative clause is ungrammatical. However, because the verb *persuaded* allows passivization, sentence (c) with a reduced relative clause is perfectly grammatical.

(c) The broker persuaded to sell the stock *was* sent to jail.
(d) *The broker hoped to sell the stock *was* sent to jail.

If the P600 which Osterhout and Holcomb identified in response to the violation of subcategorization constraints was indeed an electrophysiological index of syntactic processing, one might expect to detect a similar response to the violation of phrase structure in sentence (d). Thus ERPs elicited by the auxiliary verb *was* in sentence types (c) and (d) were compared to one another. As in the ERP response to the infinitive *to*, a late positivity was evidenced in the waveform between 500 and 800 msec post-stimulus onset. Moreover, the mean amplitude of ERPs in the 600 to 900 msec window was larger in the ungrammatical condition (d) than it was in the grammatical condition (c). Additionally, they detected an anterior negativity, most prominent in the central frontal regions, 300-500 msec post-stimulus onset, which was larger in the ungrammatical condition.

However, the P600 elicited to *was* in the relevant condition had a different distribution over the head from that elicited in response to *to* . While the P600 to *to* was largest over right anterior regions, the positivity elicited in response to *was* was larger over posterior regions and laterally more symmetric. Besides these differences in the distribution of the waveforms elicited by subcategorization and phrase structure violations, Osterhout and Holcomb also found it necessary to vary the latency window in order to detect statistically significant differences between ungrammatical stimuli and their grammatical counterparts.

In spite of these differences in the latency and distribution of the waveform elicited by subcategorization versus phrase structure violations, Osterhout and Holcomb argue that both are instances of the P600. Proposing the P600 as the syntactic analogue of the N400 (an ERP component often used as an index of semantic processing), the authors suggest the existence of two distinct ERP components for syntactic and semantic processing as indicative of two discrete levels of linguistic processing.

Neville, Nicol, Barss, Forster, and Garrett, 1991

Neville et al. tested the ERP response to a variety of syntactic violations within the framework of Government and Binding Theory (Chomsky, 1981). These included violations of subjacency, phrase structure, and a specificity condition.

1. Subjacency

Subjacency[1] constraints concern the movement of abstract elements in the grammar which ultimately serve to rule out certain configurations. One historically important manifestation of subjacency constraints involves wh-movement, or the movement of wh-phrases which occurs in the formation of wh-questions. Ross (1967) noted an asymmetry in the grammaticality of wh-extraction from inside objects and from inside subjects and adjuncts. While extraction of a wh-phrase from object position results in a grammatical wh-question, extraction from subject or adjunct position results in ungrammatical wh-questions. The ungrammaticality

[1] Although the data upon which it was originally posited are still accepted, subjacency itself has largely been eclipsed by the Empty Category Principle. The interested reader can consult Huang (1982), Lasnik and Saito (1984), Chomsky (1986), Rizzi (1990), and Cinque (1990) for the details of this chapter in the history of syntactic theory.

wh-questions in which the wh-phrase has been extracted from the subject (or adjunct) is an instance of a subjacency violation.

In order to elicit the brain response to subjacency violations, Neville et al. recorded subjects' ERPs to sentences like the following:

(a) Was the proof of the theorem *criticized* by the scientist?
(b) *What was [a proof of__] *criticized* by the scientist?

Note that the wh-phrase in (b) is extracted from inside the subject and its movement violates subjacency constraints. Comparing the ERP response in sentences (a) and (b), Neville et al. report the major effect of the subjacency violation was an enhanced P200 component. However, also of note was a broadly distributed late positivity with onset at 500 msec post-stimulus which was similar to the P600 observed by Osterhout and Holcomb (1992).

However, interpretation of these results is difficult for a number of reasons. The researchers themselves note that this comparison is problematic because the context preceding the comparison point is not well matched. Neville et al. point to the fact that the word immediately preceding the comparison point in (a) is an open class word, while the word preceding the comparison point in (b) is a closed class word.

Unfortunately, this is not the only relevant difference between (a) and (b). Another difference is that (a) is a yes/no question, while (b) is a wh-question. While wh-movement occurs in (b) it does not occur in (a). Consequently, in (a) the comparison point is preceded by a definite noun phrase (*the theorem*), while in (b) the comparison point is preceded by a gap, or wh-trace. Observed differences might be due to any of these factors. Further, the late positivity which resembled the P600 is also difficult to interpret because the onset of the effect was concurrent with the onset of the next word in the sentence. Any putative effect of subjacency, then, is confounded with the effect of stimulus onset and the different lexical items which followed the critical word in each condition.

2. *Phrase Structure*

Neville et al. constructed phrase structure violations by transposing word order in noun phrases with prepositional complements. For example, sentence (c) was contrasted with (d).

(c) The scientist criticized a proof *of* the theorem.
(d) *The scientist criticized Max's *of* proof the theorem.

Because the possessive *Max's* is the specifier of an NP, it should be followed by a noun. Thus the preposition *of* is syntactically anomalous.

Phrase structure violations yielded larger early negativities over the anterior regions of the left hemisphere (N125), a left anterior negativity beginning around 300 msec post-stimulus onset, and a late positivity beginning around 500 msec post-stimulus (coincident with the onset of the next word), which was bilaterally symmetric and largest over occipital regions. However, the authors suggested that the observed ERP effects might not reflect the syntactic violation itself, but rather some compensatory process invoked to recover from the syntactic violation.

3. *Specificity*

In the specificity condition, Neville et al. contrasted ERPs to sentences like (e) with ERPs to sentences like (f).

(e) What did the man admire [a *sketch* of __]?
(f) *What did the man admire [Don's *sketch* of__]?

Sentence (f) is ungrammatical because the wh-phrase has been moved out of an NP with a possessive specifier that is specific in reference (*Max's*). The stimuli in the specificity condition thus constitute true minimal pairs since the hierarchical structure of (e) and (f) is exactly the same. The two sentences differ only in the specificity of the item in specifier position of the final NP. However, Neville et al. point out that this comparison includes both syntax, because it concerns the wh-extraction in specific and definite NPs, and semantics, because the notion of specificity of an NP is semantically based.

The effect of the specificity violation on subjects' ERPS was sustained negativity over anterior regions of the left hemisphere. This potential enhanced the amplitude of the N125 as well as that of the negativity between 300 and 500msec post-stimulus onset which occurred over frontal, anterior temporal, and temporal regions of the left hemisphere. Neville et al. compared and contrasted these results with those in the phrase structure condition as well as a semantic anomaly condition not discussed here, noting that the effect of the specificity violation was more similar to that for phrase structure (read here, syntactic) violations than it was to semantic anomaly. They interpret these results as suggesting that the ERP response indexes the syntactic characteristics of these violations rather than any semantic component they might contain.

Overall, Neville et al. are more concerned with the issue of the modularity of the language processor divided broadly into syntactic and semantic subsystems rather than fine-grained modularity of the syntactic processor itself. Remarking on the overall absence of N400 effects in their data, Neville et al. argue that the presence of distinct ERP effects for syntactic violations clearly suggests a corresponding division of processing mechanisms between those mechanisms sensitive to semantic relations and those sensitive to syntactic relations. Thus Neville et al. find distinct ERP responses for different sorts of syntactic anomalies and interpret those results as indicative of the action of distinct subsystems (*viz.* modules) within the language faculty. Whereas semantic anomaly is indexed by the N400 component, syntactic anomaly is indexed by anterior negativities and/or late positivity. However, they reject the possibility that the P600 is a unique index of syntactic ill-formedness, citing the fact that it was present in only one of the three syntactic violations they tested.

Hagoort, Brown, and Groothusen, 1991

To investigate ERP manifestations of syntactic processing, Hagoort et al. (1991) examined subjects' ERP responses to three sorts of syntactic violations in Dutch: noun-verb number agreement, subcategorization frame, and phrase structure. As in Osterhout and Holcomb (1992) the hope was to identify a distinct component sensitive to syntactic violation.

1. Agreement Violation

The first sort of syntactic anomaly addressed by Hagoort et al. was a simple violation of noun-verb agreement. For example, compare sentence (a) to sentence (b).

(a) Het verwende kind *gooit* het speelgoed op de grond.
(The spoilt child throws-SINGULAR the toys on the floor.)

(b) *Het verwende kind *gooien* het speelgoed op de grond.
(The spoilt child throw-PLURAL the toys on the floor.)

Whereas sentence (a) is perfectly grammatical, the number of the verb in sentence (b) does not agree with that of its noun. Comparing subjects' ERPs to the verbs in the grammatical sentences versus those in the ungrammatical sentences, Hagoort et al. found a broadly distributed positive shift in response to *gooien* (the agreement violation) relative to the ERP to *gooit* (the grammatically correct verb). Distribution was bilaterally symmetric and was largest over parietal sites. The positivity began at approximately 500 msec post-stimulus onset, and continued throughout the following word. Two words after the verb (*speelgoed*) the positive shift in the waveform in the ungrammatical condition was replaced by a broadly distributed negativity which begins at approximately 200 msec post-stimulus onset and continues throughout the following word.

However, the significance of these results is greatly undermined by the finding that the waveforms in the two conditions had actually diverged before the presentation of the verb. The ERP to *verwende* (spoilt) in the ungrammatical condition showed a negative shift (relative to the grammatical condition) which peaked at 450 msec post-stimulus. This difference reveals the noisy nature of these data since, at this point in the sentence, the stimuli are identical in the two conditions.

2. Phrase Structure Violation

The second sort of syntactic anomaly addressed by the Hagoort et al. study was the violation of phrase structure constraints.

(a) Het publiek moet lachen om de *zakkende omlaag broek* van de clown.
The audience must laugh about the falling down trousers of the clown.]
(b) *Het publiek moet lachen om de *omlaag zakkende broek* van de clown.
[The audience must laugh about the down falling trousers of the clown.]

In Dutch, (as in English), the combination of adjective-adverb-noun (*e.g.* comfortable, extremely chair) is unacceptable. Thus while sentence (a) which employs the sequence adverb-adjective-noun (*falling-down trousers*) is grammatical, sentence (b), which employs the sequence adjective-adverb-noun (*down-falling trousers*), is not. However, the subject reading the sentence cannot determine the ungrammaticality of (b) unequivocally until he encounters the noun (*broek*, *viz.* trousers). While the combination of adjective-adverb-noun, (as in *down-falling trousers*), is ungrammatical, the combination of adjective-adverb-adjective-noun is perfectly acceptable, albeit rare (as in, *hard, very uncomfortable chair*).

ERPs elicited by *broek* showed a broadly distributed positive shift in the ungrammatical sentences as compared to the grammatical sentences. Unlike the response to the agreement violations, the positive shift in the ERP to phrase structure violations began almost immediately after the onset of the stimulus. Note also that the waveforms elicited by sentences like (a) and by sentences like (b) began to differ at the beginning of the adjective-adverb-noun sequence -- two words before the sentence could have been determined to be ungrammatical. Two words before the critical word, there was a negative shift in the ungrammatical condition. This was followed by a positive shift on the word immediately preceding the critical word and which continued to occur during the presentation of the critical word itself.

Hagoort et al. (1991) argue that this effect results from the parser entertaining the possibility of the relatively rare (but grammatical) adjective-adverb-adjective-noun construction. While this is a possibility, it is not particularly compatible with their interpretation of the ERP to the critical word as an index of syntactic anomaly because the parsing strategy to which Hagoort et al. appeal would delay (rather than expedite) the detection of ungrammaticality. If the parser were entertaining the possibility of an adjective-adverb-adjective-noun parse in sentence (b), it would not register the ungrammaticality of the sentence until it encountered the noun (*broek*) in the place where the adjective should occur.

3. Subcategorization Violation

The third sort of syntactic anomaly addressed by the authors was violations of subcategorization constraints on obligatorily intransitive verbs. Because obligatorily intransitive verbs cannot take a direct object, placing a direct object after such a verb is a subcategorization violation.

(a) De goed gek lede man gebruikt een *paraplu* tijdens de regenbui.
(The well dressed man uses an umbrella during the shower.)

(b) *De goed gek lede man shuilt een *paraplu* tijdens de regenbui.
(The well dressed man shelters an umbrella during the shower.)

While the Dutch verb *gebruikt* (*uses*) can take a direct object, the Dutch verb *shuilt* (*shelters*) cannot (unlike the English verb *shelter*). The subject will thus realize that (b) is ungrammatical when she reads the object-noun *paraplu*, a difference which ought to be reflected in subjects' ERPs.

However, the subjects' ERPs in response to the grammatical sentences and the ungrammatical sentences actually diverged *before* the point at which they might have determined the grammaticality of sentences like (b). Hagoort et al. speculate that these differences were due to lexical effects of the transitive and intransitive verbs. This part of the study, then, failed to uncover an ERP index of violations of subcategorization constraints. Upon their failure to find a late positive component in this condition, Hagoort et al. argue that the late positivity is masked by an N400. We find this analysis dubious, however, because similar negativity in the phrase structure condition does not similarly preclude the detection of a late positivity in those instances.

Hagoort et al. point to the positivities elicited by subjects in response to phrase structure violations and agreement violations, as well as to the results reported by Osterhout and Holcomb (1992), and label these responses the *syntactic positive shift* (SPS). Further, they argue that the syntactic positive shift elicited in response to syntactic violations is indicative of a distinct level of syntactic processing in accord with the modularity thesis. However, the findings reported by Hagoort et. al. are neither consistent internally nor with those reported by Osterhout and Holcomb (1991). To summarize, Hagoort et al. report a P600 to their agreement violations, a similar positivity in response to phrase structure violations (albeit at time 0), and no positivity at all in the subcategorization condition.

2.3 Is the P600 a syntax component?

Overall, results reported in the literature reviewed here seem to defy a unified explanation in terms of syntactic theory. Osterhout and Holcomb, for example, report a P600 in response to both phrase structure and subcategorization violations. Although the ERPs to both sorts of violations displayed late positivity, the observed differences in the onset and distribution of the two effects would seem to preclude a unified account . Moreover, a similar argument can be made for Hagoort et al.'s syntactic positive shift. The positivity elicited by the phrase structure condition began almost immediately after the onset of the stimulus, whereas the onset of the positivity elicited by the agreement condition was not until 500 msec post-stimulus, and there was no positive shift at all in the subcategorization violation condition.

Further, when we look across studies for a cohesive account of the late positivity elicited in response to syntactic violations, the data are even more mysterious. Phrase structure violations in Dutch elicit a positivity whose onset is almost immediate (Hagoort et al.), whereas similar violations in English elicit a positivity with onset about 500 msec post-stimulus (Neville et al.; Osterhout & Holcomb). However, Neville et al. report a laterally symmetric positivity largest over occipital regions, whereas Osterhout and Holcomb report a positivity with right anterior distribution. Finally, Osterhout and Holcomb report a positivity with symmetric posterior distribution in response to violations of subcategorization constraints in contrast to Hagoort et al.'s null result for subcategorization violations in Dutch.

Whereas differences in the latency of an evoked potential do not necessarily imply qualitative differences in the response, differences in the scalp distribution are important factors in the definition of an ERP component (Donchin, 1979). Given the variety in the scalp distribution of the ERP response to syntactic violations across studies, existing syntactic theory would seem to preclude any coherent mapping of syntactic violations and ERP indices. While some of these discrepancies can be explained away by recourse to models (*e.g.* Bates and MacWhinney's 1987 Competition Model) which predict differences in the neural implementation of different languages, not all of the differences can be attributed to differences between English and Dutch. Especially mysterious is the observation [see chart below] of a preceding negativity whose intermittent presence has yet to be cogently explained. We return to this question in the following section.

Results from the studies reviewed above are listed in the following chart according to the sort of syntactic violation each study addressed and the latency and distribution of the positivity found in the ERP. The last column in the chart tells whether or not the observed positivity was preceded by a negative going component.

Violation Type	Latency	Distribution	Preceding LAN/N400
Agreement			
Hagoort et al.	500ms-next word	Lat Sym; Parietal	None
Phrase Structure			
Osterhout et al.	500-700ms	Right Anterior	LAN
Hagoort et al.	Almost immed.	Broad	None
Neville et al.	500-700ms	Lat Sym; Occipital	LAN
Subcategorization			
Osterhout et al.	600-900ms	Lat Sym; Posterior	None
Hagoort et al.	-----		N400?
Specificity			
Neville et al.	-----	-----	LAN
Subjacency			
Neville et al.	200 ms-next word	Lat Sym; Broad	None

In view of the disparate array of results reported above, we are in accord with the conclusions of Kutas and Kluender (1991) that the variety of ERP components elicited in response to syntactic violations suggests that there is no unique index of syntactic violation which parallels the extent to which the N400 can be used as an index of semantic violation. Further, there is no particular reason to accept the so-called syntactic positive shift as syntax specific. It is likely that the late positivity time-locked to syntactic irregularity is actually a member of the P300 family, a number of positive components with varying onset latency and centro-parietal distribution.

The P300 is known to reflect the resolution of prior uncertainty and the task relevant surprise value of the stimulus. For example, in the auditory oddball paradigm (in which the subject is directed to attend to a series of long beeps periodically interspersed with short beeps, or vice versa) the P300 is elicited by the less frequently occurring beeps. Its amplitude is proportional to the rarity of the target stimulus and its latency varies with the difficulty of the discrimination task (Picton, 1992). Moreover, the P300 does not reflect physical parameters of the eliciting stimuli; nor, is it necessarily an index of the objective probability of the stimulus' occurrence. Rather, the P300 component has been associated with subjective aspects of the stimuli, such as task relevance, salience, and conformity to expectations (see Pritchard, 1981 for review).

One proposal as to the functional role of the cognitive process or processes associated with the P300 component is *context updating*, or updating hypotheses, models, or expectations about the environment (Donchin, 1979 cited in Pritchard, 1981). Considering the fact that so much of our interaction with other people involves language, it is quite plausible that we entertain hypotheses about our linguistic environment just as we do about our physical environment. Moreover, those expectations might not always be borne out, but would require reanalysis and updating.

Take the following sentence used by Osterhout et al. in their subcategorization condition:

*The broker persuaded to sell the stock.

Note that the P600 which Osterhout et al. report in response to *to* cannot index the recognition of an ungrammatical sentence, because, strictly speaking, the sentence does not become ungrammatical until *stock*. However, if we assume that the parser assigns the most common reading of ambiguous words by default, then it is a safe bet that its linguistic expectations would be based upon the active reading of *persuaded* . Upon encountering *to*, then, it would become clear that the active reading is untenable and requires reanalysis. Perhaps the P600 which Osterhout et al. reported in response to *to* is indexing the reanalysis of *persuaded* from the more common active reading to the passive reading.

Further, Osterhout (1993) has demonstrated that the late positivity elicited in response to syntactic violation is in fact a member of the P300 family via a comparison of ERPs to syntactic violations and oversized words. Although the P300 to oversized words was larger in amplitude, it showed the same latency and morphology as that to syntactic anomalies. Because the P300 represents a domain-general index of the information value of a stimulus, its implication in syntactic processing argues against the existence of a modular parser.

However, Osterhout and Holcomb (1992) and Osterhout (1993) argue that an ERP component such as the P600 might *co-occur* with domain general psychological processes associated with the P300 family without necessarily being identical to them. In other words, the detection of syntactic anomalies might co-occur with domain-general processes which underlie the P300 family without being specifically tied to them.

However, in view of the fact that the search for a syntactic component is motivated by the hope of substantiating claims about the existence of autonomous processing systems for syntax and semantics, the hesitation evidenced by both Hagoort (1992) and Osterhout (1992; Osterhout, 1993) to equate the late positive components in their data with a domain-general rather than a domain-specific processing mechanism is suspect.

2.4 What is the LAN indexing?

Besides late positivities, ERP responses to syntactic violations have often included left lateralized anterior negativities (LAN) between 200 and 500 msec post-stimulus. Osterhout and Holcomb (1992) observed a LAN effect in their phrase structure condition. Neville et al. (1992) also observed LAN effects in their phrase structure condition, as well as in their specificity condition. Were it not for the absence of a similar effect in Hagoort et al.'s data, one might speculate that the LAN effect indexes violations of phrase structure. Explanation of this effect in terms of syntactic theory has thus proved elusive.

However, Kluender and Kutas (1994) have suggested that the LAN elicited in response to certain sorts of syntactic violations is actually indexing some aspect of working memory use. They find a consistent LAN effect associated with entering a filler in working memory, storing it, and subsequently retrieving it to assign fillers to gaps. Rather than attempting to account for the LAN effect purely in terms of syntactic theory, Kluender and Kutas address the problem in terms of the computational tasks associated with parsing sentences. For instance, Kluender and Kutas suggest that the LAN effect elicited to *was* in Osterhout and Holcomb's phrase structure condition indexes the parser's search through working memory for a discourse referent to serve as its subject.

Recall that Osterhout and Holcomb contrasted subjects' ERPs to sentences like the following:

(a) The broker persuaded to sell the stock *was* sent to jail.
(b) The broker hoped to sell the stock *was* sent to jail.

The word *was* in (b) elicited both a late positivity and a preceding LAN effect relative to the same word in sentence (a). In processing sentence (a), the parser can employ the passive participial reading of *persuaded* thus allowing *broker* to serve as the subject for *was*; however, in sentence (b) the subcategorization constraints on *hoped* preclude a similar reading, leaving *was* without a subject. Thus Kluender and Kutas argue that the P600 elicited to *was* indexes the parser's realization that (b) is ungrammatical; moreover, the preceding LAN indexes the parser's vain search through working memory for a discourse referent to serve as the subject for *was*.

Kluender and Kutas also propose a working memory account of the LAN effects which Neville et al. observed in their phrase structure and specificity conditions. The authors note that Neville et al. detected a LAN effect in response to *proof* in sentence (d) relative to the same word in (c).

(c) The scientist criticized Max's proof of the theorem.
(d) What did the scientist criticize Max's proof of __?

This effect is easily accounted for by the presence of a filler gap relationship in (d) and the corresponding absence of such a relationship in (c). Thus the LAN indexes the storage of the filler (*what*) into working memory until it can be assigned to its gap at the end of the sentence.

However, Neville et al. also detected a similar LAN effect for sentence (d) (above) relative to sentence (e) (below).

(e) What did the scientist criticize a proof of __?

Because there is a filler gap relationship in both (d) and (e) this LAN effect cannot be explained away so easily. Kluender and Kutas argue that the relevant difference between (d) and (e) is the presence of a unique discourse referent (*Max's*) in (d) with an aspecific referent in (e). Although both (d) and (e) require the parser to store a filler in working memory, there is a greater cost to working memory associated with the activation of the discourse referent *Max's*. Thus there is a larger LAN effect for (d).

While the working memory account of the LAN seems to do a good job of explaining the presence of LAN effects in the data, it also needs to account for the absence of similar effects in Neville et al.'s subjacency condition. In this condition subjects' ERPs to *printed* were compared in the following two sentences:

(f) Was a picture of the accident *printed* by the newspaper?
(g) What was a picture of __ *printed* by the newspaper?

If the LAN indexes working memory in accordance with Kluender and Kutas' hypothesis, we would expect to see a LAN effect in sentence (g) at the point of comparison (*printed*). However, Neville et al. found no such effect. One possible explanation is that the word *printed* is not interpreted by the parser as an (ungrammatical) passive verb, but rather as a grammatical passive adjective as in:

What was a picture of printed obscenities doing __ in the newspaper?

The LAN effect at *printed* is thus absent because the passive adjectival reading does not involve a gap at that point. Noting the prevalence of possible passive adjectival readings for the critical verbs in Neville et al.'s stimuli, Kluender and Kutas argue that the failure to find a LAN effect results from this aspect of the stimuli.

Thus the search for an electrophysiological index of syntactic processing has yielded results somewhat contrary to the modularity thesis. Both of the ERP components associated with syntactic processing -- the late positive component and the LAN effect -- seem to be associated with domain general rather than domain specific processes. Although this fact does not invalidate the modularity thesis, it does somewhat undermine it. Nonetheless, it may be that the failure to date to detect the action of domain-specific modules rests with the insensitivity of electrophysiological measures. However, the suggestion that domain-general processes are sensitive to structural regularities of natural language is itself a fascinating result.

3. When do they know it? The time course of the lexical access process

A great deal of debate about the validity of the modularity thesis has dealt with the question of how sentential context can affect *lexical access*, the process in which a word's core meaning or meanings are activated. This controversy concerns the extent to which the relatively low-level process of lexical access is penetrated by information from higher-level processes. One might ask, for instance, whether lexical access is facilitated by some sentential contexts relative to others. Indeed, is lexical access at all affected by sentential context or does it proceed in the same manner regardless of the context in which a given word is encountered?

Although the influence of context on lexical access has been a highly controversial topic, one way in which context affects lexical access is agreed upon by all participants in the debate. This is semantic priming. A number of behavioral measures reveal that the processing of a single word is facilitated by the prior occurrence of a semantically related word. For instance, the word *cat* is easier to process if it is preceded by a word (such as *dog*) which belongs to the same general category. This facilitation effect is referred to as *semantic priming* and is generally interpreted as indicative of the way in which word representations are organized in the lexicon.

In the semantic priming paradigm, the subject's reaction time for a variety of tasks is used as an index of the degree to which the meaning of any given word has been activated or "primed" by prior context. These measures include naming latency (in which investigators record how long it takes subjects to pronounce a visually presented word), the lexical decision task (in which subjects distinguish words from nonwords), the word monitoring task (in which the subject monitors a set of sentences for the occurrence of a target word), as well as others. The shorter the subject's reaction time in the performance of these tasks, the greater the degree of priming.

However, the use of the semantic priming paradigm to address questions pertaining to the availability of high-level linguistic information has been criticized because it requires the subject to terminate linguistic processing and to initiate the performance of the experimental task. Moreover, because the predictions entailed by the modular account of language processing hinge on the question of which levels of linguistic analysis influence processing *when*, ideally one would like an on-line measure of subjects' processing which does not interfere with the linguistic processing itself.

Electrophysiological methodology provides such measures. In a number of conventional psycholinguistic tasks (e.g. reading, lexical decision), ERPs to primed words are less positive between 200 and 500 msec poststimulus than unprimed words. Beginning at approximately 200 msec, this monophasic negativity is referred to as the N400 semantic priming effect. This affords use of N400 amplitude as an inverse indicator of semantic priming -- the greater the N400 amplitude the less the word has been primed by prior context.

Further, the ERP measure has several advantages over existing techniques in the semantic priming paradigm. First, it provides a realtime continuous signal which can be used to evaluate the time course of processing. Second, it does not necessarily require the subject to interrupt processing of the target word. Thus the N400 semantic priming effect provides a measure of semantic priming which is both more sensitive and more direct than the behavioral measures traditionally employed in investigations of these issues. Reviewed below are two ERP studies which explored the time course of lexical access and ambiguity resolution, as well as a series of studies which compared the time course of the effect of lexical and sentential context on the lexical access process.

3.1 Lexical Access and Ambiguity Resolution

Empirical inquiries concerning the modularity of lexical access have largely focussed on ambiguous words, that is, words with two or more meanings. If lexical access can be influenced by contextual information, one might expect one of the meanings of an ambiguous word to be easier to access than the others. Take the word *ball* in the following sentences:

(a) Hillary's dress was sure to be discussed at the upcoming charity ball.
(b) Hillary heard the crack of the bat and reached out to snag the ball.

If lexical access is encapsulated with respect to contextual information, the lexical access of *ball* should proceed identically in sentences (a) and (b). However, if lexical access is not modular, one might expect to see facilitation of the contextually appropriate meaning. While the modular lexical access system would always activate both meanings of *ball* (leaving the suppression of the inappropriate meaning for a later stage in the processing), a more interactive system might bias the activation of the contextually appropriate meaning.

In a landmark study, Swinney (1979) used the cross-modal lexical priming technique in which the subject first listens to auditorily presented sentences containing ambiguous lexical items, and then performs a lexical decision task on a visually presented probe word. For example, the subject might hear a sentence such as:

The boy dreamed he was being eaten alive by bugs.

in which the target word (*bugs*) has at least two possible meanings (*insects* and *hidden microphones*). This sentence was then followed by the visual presentation of one of three types of probe words:

Contextually Appropriate: ANTS
Contextually Inappropriate: SPY
Control (Unrelated): GLOVE

or by a nonword. Subjects' lexical decision latencies were used as an index of semantic priming for the different meanings of the target words.

Swinney found that the temporal interval between the probe and the target is the critical variable in whether or not the contextually relevant meaning of the ambiguous word is primed relative to the contextually irrelevant meaning. If this interval is extremely short (200 msec or less), both meanings of the ambiguous target show priming; however, at longer intervals (500-1500 msec) only the contextually appropriate meaning is primed (Onifer and Swinney, 1981).

Prima facie, this supports a modular account of word recognition known as *multiple access*. Priming for both of the target's meanings in the short SOA reflects the action of a modular lexical access process. Fast, autonomous, and informationally encapsulated, the lexical access module very quickly retrieves all possible meanings and passes them on to the semantic processor which suppresses the contextually inappropriate meaning. Priming for only the contextually appropriate meaning in the long SOA condition is due to the fact that the contextually inappropriate meaning has been suppressed by this point in the processing.

Alternative Interpretations

However, these results are not unequivocal. As useful as it has been, there are several problems with the semantic priming paradigm which need to be addressed. First, the behavioral measures employed in the semantic priming paradigm are necessarily indirect, because of the absence of a clear delineation between the processing of the target word and the task decision. Moreover, they do not provide a continuous measure with which to evaluate the time course of the processes underlying the priming effect. Such a measure is necessitated by the existence of word recognition models that make fine-grained predictions about the time course of processing in experiments conducted within the semantic priming paradigm.

Further, interpretation of priming effects in the behavioral literature rests on questionable assumptions. Specifically, the assumption that the subject terminates processing of the target upon the presentation of the probe -- regardless of the SOA -- may simply be erroneous. One problem with the traditional interpretation of Swinney (1979) is that it relies upon the assumption that the effect of varying the SOA between the target and the probe is to vary the amount of time which the subject is given to process the target before she is obliged to process (and respond to) the probe. Thus the subject's reaction time in the short SOA would reflect the output of context-blind lexical access, while the reaction time in the long SOA would reflect the output of lexical access as well as subsequent processes more sensitive to context.

However, Van Petten and Kutas (1987) suggest that, in fact, the subjects' behavior in the short SOA reflects the output of a language processor which is obliged to process two words (that is, both the target and the probe) during essentially the same time period. In that case, the results of the short SOA are indicative of a phenomenon known as *backward priming* (see Kiger & Glass, 1983). When backward priming occurs, the subject's reaction time to the probe represents not only the extent to which the preceding context primes the probe, but also the extent to which the probe affects the processing of the preceding context.

Suppose the subject hears the sentence, "The boy dreamed he was being eaten alive by bugs" and subsequently sees the probe word SPY, upon which she performs the lexical decision task. According to the backward priming hypothesis, it is possible for the probe (*spy*) to affect the processing of the target word which precedes it (*bugs*) when the SOA between the two is so small as to cause the two words to be processed concurrently. Moreover, because the word SPY is congruent with the meaning of the contextually inappropriate meaning of the target, the backward priming hypothesis provides an alternate explanation for the observed data in the short SOA condition.

Another alternative to the multiple access model has been dubbed the *biased activation* hypothesis (St. John, 1991). Based on the well-known *interactive activation connectionist* (IAC) model (McClelland, 1987), the biased activation hypothesis predicts an interaction between low-level processing of lexical items and higher level representations of sentential context. Specifically, this involves the feedback of information from the developing interpretation of the sentence to the lexical access process. This feedback bolsters the activation of the contextually relevant meaning of the ambiguous word, thereby biasing it over the contextually irrelevant meaning. Top-down activation thus serves to reinforce the initial activation of the contextually relevant meaning so that its activation level increases faster and is sustained for longer than the contextually irrelevant meaning.

The biased activation hypothesis resembles the multiple access account in that, initially, all meanings of the ambiguous word are activated. However, it differs from the modular account because in the biased activation model, sentential context affects the rise time of lexical access. So, although lexical access is penetrated by sentential context, the influence of context proceeds at a rate which is too slow to influence traditional measures such as naming latency and reaction time in the lexical decision task (McClelland, 1987). In order to test the biased activation hypothesis it is necessary to use a continuous measure of semantic priming such as that afforded by the ERP.

There are several reasons to suspect that one or more of these alternative accounts of the lexical access of ambiguous words is the right one. Although the behavioral literature contains somewhat contradictory findings on the question of whether ambiguous words initially prime both contextually biased and unbiased meanings, most authors report slightly faster reaction times for the former (McClelland, 1987). Moreover, statistical meta-analysis of the ambiguity resolution literature reveals a significant difference between the reaction times for contextually biased and unbiased meanings of the target (St. John, 1991).

Van Petten and Kutas, 1987

Capitalizing on known properties of the N400 component as an inverse index of semantic priming, Van Petten and Kutas (1987) compared the relative time course of activation of contextually biased and unbiased meanings of ambiguous words. Subjects read sentences such as:

The gambler pulled an ace from the bottom of the deck.

Sentences were then followed by a probe word which was either related to the contextually biased meaning of the target (e.g. *cards*), the unbiased meaning of the target (e.g. *ship*), or was unrelated to the target (e.g. *pill*).

At the long SOA (700 msec), there was a clear difference in the N400 component of the ERPs to contextually biased as compared to either the unbiased or the unrelated probes. As in the earlier studies, the N400 priming effect showed a greater degree of priming for the contextually biased probes. Moreover, the N400 amplitude of the unbiased and unrelated probes did not significantly differ from one another. The results in the long SOA condition thus replicate the effects found in the behavioral literature and are in accord with both modular and non-modular accounts.

It is the results of the short SOA which bear most directly on the question of whether multiple access, biased access, or backward priming causes the facilitation of the contextually inappropriate meaning which was evidenced in the behavioral literature. As in the behavioral literature, Van Petten and Kutas observed the N400 priming effect for both biased and unbiased meanings in the short SOA condition. However, the priming effect for the unbiased meaning was by no means identical to that for the biased meaning. First, it was slightly smaller (viz. the amplitude of the N400 to the contextually inappropriate meaning was slightly larger than that of the contextually appropriate meaning). More importantly, it displayed a different time course. The onset of the N400 component of the ERP to the unbiased meaning of the target was slightly later (occurring at 500 msec poststimulus) than the onset of the N400 priming effect for the appropriate meaning (300 msec poststimulus).

Recalling that the multiple access account predicts the initial activation of the contextually biased and unbiased probes to be identical, Van Petten and Kutas argue that the detection of any statistically significant difference whatsoever between the ERP to contextually biased and unbiased probes in the short SOA supports backward priming. Although Van Petten and Kutas findings warrant rejection of the multiple access account, they also support McClelland's (1987) biased activation hypothesis. Both biased activation and backward priming predict a difference in the degree and time course of priming for the biased and the unbiased meanings in the short SOA condition.

However, the time course of the ERP response to contextually unbiased probes provides additional support for the backward priming hypothesis. Initially, the ERP response to the contextually unbiased probes was virtually identical to ERP to the unrelated probes; however, the ERP to the unbiased probes subsequently became less negative in a way which then resembled the ERP to the contextually appropriate targets. This is precisely the pattern one would expect if the unbiased probe (*e.g. spy*) is initially processed as if it were completely unrelated to the contextual meaning of the prime (*e.g. bugs* where the biased meaning is *insects*), and subsequently causes a reinterpretation of the target (*i.e. bugs*) such that the probe is processed as being congruent with the alternative meaning of the target (*i.e.* the *hidden microphone* meaning).

Canseco-Gonzalez et al., in press

The results of Van Petten and Kutas (1987) have been questioned by Canseco-Gonzalez and collaborators (Canseco-Gonzalez, Hickok, Zurif, Prather, and Stern, in press) because of the artificial presentation rate in the former study. The sentences in Van Petten and Kutas were presented serially, one word at a time. Although this form of presentation is referred to as Rapid Serial Visual Presentation (RSVP), the 900 msec SOA is in fact three times slower than the average person's reading rate. Canseco-Gonzalez et al. argue that the RSVP paradigm fosters strategic processing by subjects and that consequently the ERPs recorded by Van Petten and Kutas reflect post-lexical integrative processes rather than lexical access itself.

Arguing that a more natural presentation of the sentential stimuli would yield an ERP response which was better suited for assessing the modularity of lexical access, Canseco-Gonzalez et al. performed a similar ERP study which utilized the cross-modal lexical priming technique. Subjects in the Canseco-Gonzalez study heard sentences like the following:

The little girl was eager to learn math and she couldn't wait until she was able to COUNT [probe] without using her fingers.

200 msec after the 'onset' of the ambiguous target, one of four types of visual probes was presented for lexical decision. Probe types included:

Contextually Biased: NUMBER
Contextually Unbiased: DUKE
Unrelated: PUBLIC

as well as nonwords (such as PLIFF). Subjects' ERPs to the probe words were recorded.

Comparing the N400 amplitude to contextually biased probes with unrelated (control) probes, the authors found a greater degree of priming for the former; moreover, comparison of the N400 semantic priming effect also revealed a greater degree of priming for the contextually inappropriate probes relative to the unrelated (control) probes. Noting that they found evidence of priming for both contextually appropriate and inappropriate meanings of the ambiguous target in spite of the biasing sentence context, Canseco-Gonzalez et al. concluded that their data support the modular multiple access account of the lexical access process.

Modular or Interactive Lexical Access

However, the analysis given by Canseco-Gonzalez et al. of their data is problematic in several respects. First, they distributed the variable of interest (probe type) between rather than within subjects, thereby lessening the power of their design. This is especially important when we consider the fact that the modular hypothesis of multiple access is in fact the null hypothesis[1] (St. John,

[1] The null hypothesis is the assumption that the true (population) means are equivalent in the conditions corresponding to the experimental variable. Statistical tests employed by psycholinguists are usually invoked to reject the null hypothesis and to infer that observed

1991). Recalling that multiple access predicts no difference between the subjects' response to contextually relevant and contextually irrelevant meanings of an ambiguous word, it is clear that the multiple access hypothesis corresponds to the statistical null hypothesis.

Note that the function of the null hypothesis in statistical reasoning is somewhat analogous to the presumption of innocence in the American justice system: the independent variable is presumed to have no real effect on the subjects' response measure unless it can be proven by statistical tests. Moreover, just as the presumption of innocence in the justice system results in a bias for finding people innocent, the logic of statistical tests results in a bias for accepting the null hypothesis. Thus the null result reported in the Canseco-Gonzalez study serves only as weak evidence for the modularity thesis.

Second, although Canseco-Gonzalez et al. admit the possibility of the biased activation hypothesis, the issue of backward priming is never addressed. Consideration of the backward priming hypothesis requires a direct comparison of the ERP response to the contextually biased and unbiased meanings of the target. Although Canseco-Gonzalez et al. actually found that the N400 amplitude was smaller for the biased than for the unbiased meaning of the target (thus indicative of a greater degree of priming for the biased meaning) they discount its significance on the grounds that the contextually biased meanings were also the more frequent meanings.

Finally, we might question Canseco-Gonzalez et al.'s claim that Van Petten and Kutas' (1987) results issue from an unnatural presentation rate and reflect post-lexical integrative processes. Although Canseco-Gonzalez and her collaborators used a more ecologically valid presentation rate than Van Petten and Kutas (1987), both sets of investigators used the N400 as their index of semantic priming. Given (i) the assumption common to ERP research that quantitative differences in the waveform index quantitative differences in processing while qualitative differences in the waveform index qualitative differences in processing; and (ii) the assumption common to psycholinguistics that lexical and post-lexical processing are qualitatively distinct, no single response measure can validly index both processes. Thus it makes little sense to argue that the N400 indexes lexical access in the Canseco-Gonzalez study and that it indexes post-lexical access in the Van Petten study simply on the basis of presentation rate. Either the N400 indexes the lexical access process in both studies or it does so in neither.

Further, data suggest that the presentation rate has little or no effect on the N400 priming effect (Kutas, 1993). Kutas (1993) presented the same set of sentences at various presentation rates including one word every 100, 250, 750, and 1150 msec. Comparison revealed that irrespective of presentation rate, N400 amplitude was smallest for words judged the best completion of the sentence, intermediate for words related to the best completion, and largest for words judged as unrelated endings. The fact that the N400 priming effect was essentially the same for all four presentation rates implies that this effect cannot be dismissed as purely the result of an unnatural presentation rate.

Note that Canseco-Gonzalez et al. are not alone in questioning the relevance of ERP studies such as Van Petten and Kutas (1987) utilizing the N400 semantic priming effect on the grounds that they reflect post-lexical (viz. nonmodular) integrative processes. We return to the question of how to interpret the N400 priming effect in section 3.3.

---

differences between experimental conditions are real, and, further, that they result from the manipulation of the independent variable(s).

3.2 Lexical versus Sentential Context

As we have seen, the modularity thesis -- besides predicting that different sorts of computations are performed in fast-acting, encapsulated modules -- also makes strong predictions about the time course of processing. Modular cognitive models are aimed at delineating when the outputs of particular modules become available. Commitment to the modularity thesis thus requires not only the demonstration that the steps in language processing occur independently of one another, but also that the modules operate in a particular order. With respect to language processing this aspect of the debate has turned on the issue of information derived from lexical access (presumably modular) versus information derived from post-lexical, sentential processes -- processes which are presumably non-modular.

Modularists make a distinction between lexical priming which takes place within the lexicon and sentential priming that arises from the integrative processes that can take place only after the word has been accessed. Thus contextual mechanisms that occur prelexically, such as semantic priming, are necessarily distinct from those mechanisms that are post-access. Sentence-level effects or non-lexical context effects are excluded from the lexicon by definition. The studies reviewed in this section are aimed at examining the differences between lexical and post-lexical context effects -- whether and when such effects occur.

Van Petten and Kutas, 1990

On a modular account, lexical access is driven entirely by low-level (viz. perceptual) characteristics of the word. The word recognition module takes the visual (or auditory) characteristics of the stimulus, matches this input against information in the lexicon, and outputs this information for further processing. The initial process of word recognition is thus encapsulated with respect to sentential context. For example, Forster (1981) and Bradley and Forster (1987) propose a modular model of word recognition in which the initial stage is sensitive only to frequency, remaining oblivious to context until the next step in the processing.

In order to test the predictions of a modular account of word recognition such as Forster's (1981), Van Petten and Kutas (1990) recorded subjects' ERPs as they read sentences. ERP responses to open class words were sorted according to each word's frequency (as assessed in printed norms: Kucera and Francis, 1967), as well as its position in the sentence (beginning, middle, or end). Using the known correlation between the amplitude of the N400 component and semantic expectancy, Van Petten and Kutas (1990) compared the effects of word frequency and position in the sentence to see if the two factors were additive or interactive, according to the assumptions of the Sternberg additive factors model.

The assumptions of the additive model are that if two factors (viz. processes) are independent of one another, they will each influence the response measure autonomously. On the modular account, for example, lexical access and post-lexical integration are said to be independent processes. Thus if both factors are operating in semantic priming their influence on the dependent measure will be additive. On the other hand, if two factors are not independent, the influence of one factor might affect the influence of the other, and the corresponding effect on the dependent measure will contain an interaction.

Because, at least on the modular account, lexical access works purely on the basis of word frequency, it can be argued that the N400 frequency effect indexes the lexical access mechanism. Moreover, because the influence of post-lexical integrative processes will inevitably increase over the course of the sentence as the listener builds up a representation of the meaning of the sentence as a whole, we can assume that the effect of ordinal position reflects the influence of these post-lexical processes. A modular account of lexical access predicts that lexical and post-lexical effects are independent, and therefore,that the effects of frequency and ordinal position will be additive. A non-modular account, on the other hand, predicts the influence of higher level, viz. post-lexical effects on lexical access, and therefore, that there will be an interaction between the effects of frequency and sentence position.

Van Petten and Kutas (1990) found main effects of frequency, position, and an interaction between the two. The effects of word frequency were most prominent in words in the beginning of sentences, somewhat attenuated in sentence intermediate words, and entirely absent in words which occurred at the ends of sentences. Moreover, when Van Petten and Kutas examined data from the sentence-intermediate words on a closer basis -- word-by-word from position 3 through position 10 -- they found a significant downward linear trend in N400 amplitude. The decreasing amplitude of the N400 component in sentence intermediate words seems to be unrelated to differences in the frequency of the words, and suggests that the effects of sentential context increased monotonically across the course of the sentence.

If we assume that the effect of frequency is a manifestation of a word's initial activation in the lexicon, and that the effect of sentence position is a manifestation of higher-level semantic processes, Van Petten and Kutas' finding of an interaction between position and frequency implies that lexical access and post-lexical integrative processes are not independent of one another. Moreover, Van Petten and Kutas argue that this pattern of results suggests a non-modular model of word recognition in which each increment of sentence context influences the recognition of words which follow. The interaction between position and frequency effects, especially the fact that there was NO N400 frequency effect for words occurring in sentence terminal position, suggests that the buildup of contextual knowledge across the course of the sentence influences the initial activation and retrieval of lexical items, as well as some subsequent selection process. This is in contrast to the modular account of word recognition in which the buildup of contextual knowledge across the course of the sentence facilitates the operation of the semantic processor, but does not influence a word's initial activation and retrieval from the lexicon.

Once again, one might question the relevance of the ERP results to the question of modularity of language processing. If the N400 indexes a post-lexical integrative (i.e. non-modular) process, then the position effects (in Van Petten &Kutas, 1990) are orthogonal to the modularity question. An alternative account of the data reported by Van Petten and Kutas involves the assumption that N400 frequency effects do not reflect frequency-driven lexical access, but rather the consequences of integrating low-frequency words into minimal sentence contexts. Just as in Van Petten's non-modular account, the modular alternative would also predict that the integrative effects of frequency would disappear as sentence context begins to play a role in the integrative process.

However, N400 frequency effects are elicited by words presented in lists as well as for words in sentences (Smith and Halgren, 1987). If the N400 does indeed index a post-lexical process and N400 frequency effects arise from the integration of low-frequency words into sentence context, one is hard pressed to explain why it occurs in word lists. Van Petten and Kutas (1991) note the oddity of the employment of a post-lexical integrative process in a task which does not require lexical integration. This is especially peculiar in view of the fact that post-modular processes are putatively under conscious control. Moreover, explanation of the N400 frequency effects which occur at the first open class word in a sentence by recourse to the difference in the difficulty of integrating high and low frequency words into virtually non-existent contexts (e.g. "the squirrel" versus "the rock") is similarly absurd (Van Petten & Kutas, 1991).

Van Petten, 1993

Van Petten (1993) compared lexical and sentential context effects by presenting the same associated pairs of words (for instance, *moon* and *stars*) in both a congruent and an anomalous sentence.

Congruent Sentence/ Associated Words
(a) When the *moon* is full it is hard to see many *stars* of the Milky Way.
Anomalous Sentence/ Associated Words
(b) When the *moon* is rusted it is available to buy many *stars* of the Santa Ana.

Because the words *moon* and *stars* are semantically related, lexical context alone should produce a priming effect of *moon* on *stars*. However, the interesting question concerns whether the congruent sentential context in (a) will produce an effect which interacts with this lexical context effect.

Besides varying the type of sentential context, Van Petten also varied the lexical context by including unassociated word pairs in both congruent and anomalous sentential contexts. Stimuli, then, included sentences such as the following:

Congruent Sentence/ Unassociated Words
(c) When the *insurance* investigators found out that he'd been drinking they *refused* to pay the claim.
Anomalous Sentence/ Unassociated Words
(d) When the *insurance* supplies explained that he'd been complaining they *refused* to speak the keys.

Subjects' ERPs were recorded as they read sentences such as the ones above in order to see if there would be any detectable difference caused by the processing of targets which benefit from lexical as opposed to sentential context. Note that in the Congruent Associated condition, the target word benefits from both lexical and sentential context; in the Anomalous Unassociated condition, it benefits from neither ; in the Congruent Unassociated, the target benefits from sentential context alone; finally, in the Anomalous Associated context, the target benefits from lexical context only.

Although there was no difference between the onset of lexical versus sentential effects, there was a difference in the duration of the two sorts of context effects. The sentential context effect was present in both the 300-500 msec and the 500-700 msec latency windows. By contrast, the lexical context effect was present only in the earlier latency window. Van Petten and Kutas (1991) interpret the difference in duration as indicative of the need for more extended processing of sentential than for lexical context. However, one might also argue that the difference in the duration of the two effects is indicative of the operation of two different underlying mechanisms: one lexical mechanism, indexed by the context effect in the early latency window, and one post-lexical mechanism, indexed by an ERP effect with longer duration.

Van Petten suggests that the similarities in waveshape, distribution, and time course of the ERP to the target words in varying sentence contexts suggests that the two sorts of context effects are produced by a similar underlying mechanism. Since the onset times of the two effects were found to be nearly identical, there is no evidence from this experiment that there is a discrete stage of processing which produces lexical effects and other later stages produce the sentential effects.

Kutas, 1993

Kutas (1993) directly compared the effects of lexical and sentential contexts on priming. The study employed two sorts of stimuli: sentential stimuli at varying levels of contextual constraint, and word pairs derived from the sentential stimuli in the first half of the experiment. During the first half of the experiment, subjects' ERPs were recorded as they read sentences for content. The sentences varied with respect to how predictable the last word of each was given the preceding context. Half of the sentences ended with the most predictable response. For example, sentence (a) below ends with the most likely word. However, sentence (b) ends with a word of low cloze probability.

Best Completion Sentences
(a) Before exercising Jack always stretches his *muscles.*
Low Probability Sentences
(b) Fred put the worm on the *table.*

In the second half of the experiment, subjects' ERPs were recorded as they read word pairs. The second word in every word pair also occurred as the last word from one of the sentences used in the first half of the experiment. The first word in word pairs that were derived from best completion sentences was always a semantic associate of the second word. For example, the word pair derived from sentence (a), above, was BICEPS -- MUSCLES. The first word in word pairs derived from low probability sentences was always the best completion for that sentence. For example, the first word in the word pair derived from sentence (b) was HOOK --TABLE.

This design allowed Kutas to compare ERP responses to the same word in lexical and sentential contexts. Assuming that qualitatively different mechanisms are indexed by qualitative differences in the ERP, we can use the ERP measure to reveal the operation of different mechanisms in contexts which require higher-level processing and those in which only low-level processing is needed. If lexical and sentential context effects are due to modular and nonmodular processes, respectively, there should be a qualitative difference in the ERP response to stimuli in the different context conditions. However, if lexical and sentential context effects are due to the same underlying process, one would expect only quantitative differences in the size and timing of the two effects.

Kutas found that the waveforms for the sentence terminal words were remarkably similar to those for the second word in each word pair. However, the onset of the N400 effect was 35 msec earlier in the sentence than in the word pair condition. Further, the size of the priming effect was much larger in the sentence condition than it was in the word pairs. Nonetheless, the peak latencies of the N400s were nearly identical (379 msec in the sentence condition versus 380 msec in the word pair condition), as were their scalp distributions.

Kutas argues that the differences between the effects of lexical and sentential context on the ERPs are quantitative rather than qualitative in nature. Further, she interprets these data as undermining a distinction between lexical and sentential semantic processes. Citing similarities in the morphology, distribution, and modulation (via semantic relatedness) of the N400 priming effect in lexical and sentential contexts, Kutas argues for the existence of a single process underlying both sorts of context effects.

Once again interpretation of Kutas' findings turns on an accurate understanding of the functional role of the N400 component. If the N400 indexes lexical processing these results undermine the notion of a modular lexical access process. However, if the N400 is indeed indexing some post-lexical integrative process then these data are largely irrelevant to the modularity thesis. In the next section we turn to this question of whether or not the N400 priming effect should be interpreted as indicative of lexical or post-lexical processes.

3.3 What is the N400 indexing?

The literature reviewed in this section points to a common process underlying the N400 semantic priming effect in sentential and lexical contexts. Moreover, a telling series of experiments by Van Petten and her collaborator Kutas point to a process which displays substantial usage of top-down information. First, the differences observed by Van Petten (1993) between the same word pairs in congruent versus anomalous sentential contexts, that is to say, the fact that the second word of a word pair showed a greater priming effect in a congruent sentential context than in an anomalous sentential context argues strongly for the influence of top-down (sentential) information on lexical processing. Second, the decrement in N400 amplitude across the course of the sentence also argues for the influence of sentential context on the processing of each word. Third, the difference in the size of the priming effect in sentential versus lexical contexts in Kutas (1992) supports the idea that lexical processing is affected (and facilitated) by the build-up of sentential context. Perhaps more importantly, the difference in the time course of the sentential and the lexical context conditions in Kutas (1992) -- especially the fact that sentential context effects have an *earlier* onset than lexical effects -- strongly suggests top-down penetration of sentential context on lexical processing.

However, interpretation of the ERP literature discussed above clearly turns on the issue of what process is indexed by the N400 priming effect. The finding that the N400 semantic priming effect is influenced by contextual information is only germane to modularity if we assume that the process underlying that effect is a modular one. If the N400 indexes the lexical access mechanism responsible for semantic priming effects, the findings reviewed above clearly cast doubt on the modular account. However, if the N400 does not index lexical access, but rather post-lexical processes, then the influence of higher-level information should come as no surprise.

One suggestion along these lines is that the N400 reflects the disparity between predicted and actual endings of sentences, perhaps similar to the output of expectancy-based priming mechanism. This suggestion, however, stems from an unfortunate misinterpretation of the N400 as an index of the degree semantic violation. The issue of whether the N400 indexes the violation of expectancy or the congruence of the target with preceding context is addressed by Kutas and Hillyard (1984). In order to tease apart the possibilities, Kutas and Hillyard crossed several levels of *contextual constraint*, the degree to which preceding context constrains the possible completions of a sentence, and *cloze probability*, the probability that a particular word will be predicted from preceding context. If N400 indexes the violation of expectancy, it should be possible to modulate that effect by manipulating the degree of expectancy which is possible for a subject to entertain. However, Kutas and Hillyard (1984) found that, in fact, N400 amplitude is proportional to cloze probability. Thus N400 amplitude is an inverse indicator of the congruence of the actual target with preceding context, rather than a direct indicator of disparity between actual and predicted targets.

Another objection is that the slow presentation rates employed in most ERP studies lead to unnatural processing strategies, strategies not normally invoked in natural language processing. Thus the findings from ERP studies which employ RSVP are suspect on the grounds that they reflect contrived processing schemes occurring only in the laboratory rather than natural language processing. However, this objection is nicely answered by Kutas (1993) comparison of ERPs to sentences presented via RSVP at varying rates. As seen in earlier studies, the amplitude of the N400 was proportional to the congruence of the stimulus with preceding context. Moreover, the amplitude of the N400 component was unaffected by the presentation rate. If slow presentation rates cause subjects to invoke unnatural processing strategies, those strategies are not assessed by the N400 semantic priming effect.

Kutas (1993) found that onset and peak latency of the N400 effect were inversely related to the presentation rate. The differences among the SOA of 250 msec (normal reading rate), 700 msec, and 1150 msec were relatively small, viz. on the order of 30 msec. However, the onset of the N400 effect in the 100 msec condition was larger, viz. delayed by 80-100 msec. The data presented by Kutas (1993) show that slower presentation rates have relatively little impact on the N400 priming effect. In fact, the N400 priming effect was most affected by the presentation rate which was significantly faster than the normal reading rate.

Perhaps not surprisingly, advocates of the modularity thesis pose the question in a way which presupposes the modularity of lexical access and only questions the relevance of the N400 as a dependent measure. Framed in this way, the issue is whether the N400 indexes early, automatic, lexical (read here modular) or later controlled integrative effects. Thus one of the main issues in this debate (and one which has concerned the N400 priming literature) is whether or not the N400 indexes an "automatic" (Posner & Snyder, 1975) or a "controlled" process.

This question has been addressed by testing whether the N400 response can be modulated by factors known to increase subjects' attentional involvement in their task. This includes manipulation of task requirements, the proportion of related to unrelated word pairs, and visual masking of the target. Variations of task requirements (Kutas 1981), and the proportion of related to unrelated pairs (Holcomb, 1988) are both seen to modulate the N400 in some way. Thus larger N400 effects are seen with various experimental manipulations which presumably increase the likelihood of attentional involvement.

It is important to note, however, that the phenomenon of semantic priming probably cannot be explained by a single underlying process. In their review of the semantic priming literature, Neely and Keefe (1989) argue that there are actually three different mechanisms involved in explaining priming effects. The first is *automatic spreading activation* (ASA) in the lexicon. ASA is roughly defined as excitation along pre-established pathways among logogens, or memory representations of lexical items. ASA can account for the facilitation effect found in various word pair paradigms by positing shorter pathways between semantically related items than for between unrelated items. The second is *expectancy-induced priming*, which occurs when subjects generate a set of predictions about upcoming stimuli. If the actual word is congruent with the subject's expectations, its processing is facilitated. However, if the actual word is unexpected, its processing is inhibited. The third priming mechanism is *semantic matching*. This type of priming occurs post-lexically and results from the subject checking the coherence of the stimulus with the preceding context. Coherent stimuli will thus be responded to more quickly than incoherent stimuli -- especially when the response measure involves the lexical decision task.

ASA stands in contrast to the other two priming mechanisms in being automatic. Semantic matching and expectancy induced priming are thought to be controlled processes. It has been argued that the strongest indicator for the presence of an automatic process (both in general, and for the presence of ASA) is finding a priming effect using the visual masked priming technique. Visual masked priming involves the brief presentation of a prime word (for as short as 10 msec or as long as 40 msec); subsequent presentation of a pattern mask (e.g. hash marks); and the presentation of an unmasked target word for lexical decision. Because the masking technique prevents the prime from reaching conscious perception, it is argued that it precludes controlled processing of the prime. The discovery of a significant semantic priming effect under masking thus strongly suggests that there is an automatic component to semantic priming (Marcel, 1983).

Two fairly recent studies have looked at the impact of a masked prime on the N400 in a lexical decision task. Brown and Hagoort (1992) contrasted the effects of masked and unmasked primes on lexical decision latency and N400 amplitude. Although a significant facilitation effect was seen in both conditions in the reaction time measure, no significant N400 effect was seen in the masked condition. In contrast, Neville, Praterelli, and Forster (1989) found a small semantic priming ERP effect, even with the prime masked. However, this effect was earlier and more frontally distributed than the typical N400 effect. Although, existing evidence is not unequivocal, it does seem to support the view that the N400 indexes a controlled process.

However, note that Neely's initial division of the semantic priming phenomena into ASA and lexical integration was based on the need to account for inhibition effects in word-pair priming experiments (*viz.* the finding that the processing of target words following unrelated primes is actually inhibited relative to neutral primes). Although early contributions seemed to indicate that inhibition effects occurred only when the interval between prime and target is greater than 500 msec (Neely, 1977), subsequent researchers have found inhibition effects with shorter inter-stimulus intervals (e.g., Antos, 1979; McClean & Schulman, 1978). Thus it is not necessarily the case that the process underlying inhibition effects is slow-acting, and therefore, a controlled, nonmodular process.

Further, the exclusive division of all cognitive processes into automatic and controlled on the basis of whether or not they can be modulated by attention may not be theoretically useful. There is reason to question this criterion on the grounds that processes best thought of as automatic and unconscious can be modulated by attention. Implicit learning, for instance, can be modulated by attention, despite the learners' lack of conscious awareness of the process (Cohen, Ivry, & Keele, 1990). Moreover, Kutas (unpublished manuscript) has proposed blurring the distinction between automatic and controlled processes. Instead of positing a dichotomous classificatory scheme, one might locate processes on a continuum which runs from automatic to controlled. This would afford us with a means of explaining why some processes are more susceptible to strategic interference than others.

As noted above, empirical consideration of this issue has taken place in an environment which presupposes that fast, automatic processing is necessarily lexical and modular. Similarly, it is assumed that slower, controlled processing is necessarily post-lexical and therefore need not be accounted for on a modular account. Thus any demonstration of automatic processing is taken as *prima facie* evidence for modularity. However, it may be the case that the attributes fast, automatic, lexical, and modular are not perfectly correlated.

Take semantic priming, for example. It has generally been assumed that semantic priming reflects the organization of word representations in the lexicon. Words with similar meanings were stored close to one another. Thus ASA results in facilitating the lexical access of words associated in meaning, presumably on the assumption that associated lexical items are often encountered close to one another. The functional utility of semantic priming, then, is to facilitate processing in sentential contexts by biasing the lexical access of related terms.

Although it is not necessary for every psychological mechanism to have an obvious utility value, there are times at which it behooves us to reexamine the logic behind our assumptions. In the case of semantic priming, the assumptions include:

- semantic priming reflects the organization of lexical items in the lexicon

○ the automatic character of semantic priming is advantageous for sentence processing by facilitating lexical and post-lexical access processes.

However, given empirical findings about the actual nature of ASA priming it is unclear how such a process would facilitate sentence processing. First, ASA primes the semantic associates of a word, but only for a miniscule amount of time. Moreover, ASA priming in word lists disappears if there is even one intervening item between the semantic associates (Neely, 1991). If ASA were the only process operating to produce priming, the processor would be able to capitalize on that priming effect only in instances in which two semantically related words follow upon each other's heels. Although sentences generally contain a few lexical associates, it is quite rare for those associates to be immediately adjacent to one another.

Luckily, priming in sentential contexts displays different characteristics than priming in word-list contexts. Specifically, priming of semantic associates can be sustained over intervening items if they occur in sentential contexts. The different characteristics of priming in word list and sentential contexts suggests two possibilities. The first is that there are two different priming mechanisms: one, presumably modular and of little functional utility, and, another non-modular, but potentially useful in facilitating the representation of sentence meaning.The second possibility is that there one mechanism responsible for priming effects in both word lists and sentential contexts. In this case, the fact that priming has different characteristics in sentential contexts than in word lists suggests the operation of a context-sensitive mechanism in both cases.

Perhaps the question should not be whether the N400 indexes a lexical or a post-lexical process, but rather whether the semantic priming phenomenon itself is generated by a lexical or a post-lexical process. The functional utility of ASA to the language processor is dubious, and its utility to the theory of language processing even more so. Hodgson (1991) argues that little or none of the priming effects reported in the literature are attributable to automatic spreading activation in the lexicon. Thus semantic priming effects are not to be interpreted as a function of lexical access, but rather from a post-lexical integrative process which is automatic in character.

While we dispute the suggestion that priming effects are automatic in the sense that they are not affected by attentional manipulations, we agree that they result form an obligatorily acting process which integrates lexical items into pre-existing context. Hodgson speculates that such a process might be thought of as an early, pre-conscious component of semantic interpretation. Note that the sort of mechanism we are proposing to underlie semantic priming effects is neither clearly modular nor is it clearly interactive. It resembles modular lexical access in that it is a fast-acting process which acts obligatorily on every lexical item. However, it resembles an interactive process in that it involves high-level inferential information of the sort traditionally associated with attentional processing.

We suggest that the N400 indexes the post-lexical semantic integration process responsible for the bulk of semantic priming effects reported in the literature. Clearly the N400 indexes a fast-acting process which acts obligatorily on every lexical item. Because such a process is invoked automatically, one would expect it to operate in word-list contexts as well as sentential contexts. This, too, is borne out by the N400 data: both Van Petten (1993) and Kutas (1993) report qualitative similarities between the waveform underlying the N400 effect in word-pairs and the N400 effect in sentences. Van Petten and Kutas' (1990) report of a downward linear trend in N400 amplitude over the course of a sentence suggests that the process being indexed [as well as Hillyard & Kutas; and a host of others] is sensitive to the ease with which a lexical item can be integrated into preexisting context. Thus one and the same process is sensitive to low-level properties of a lexical item (e.g. frequency) and to higher-level contextual information.

What then of lexical access? Given the skeptical eye we have cast on the explanatory value of the concept of ASA in the lexicon and the failure of Kutas (1993) and Van Petten (1993) to uncover qualitative differences in the processing of word lists and sentences, the best move would be to jettison the whole concept of lexical access. Data reviewed above suggest that even the initial processing of a lexical item is inextricably bound up with the incorporation of that item into the existing context. It is this process which is responsible for the bulk of semantic priming effects and it is this process which is indexed by the N400.

4. Conclusion

We have reviewed electrophysiological literature bearing on the modularity of language processing in the brain. Under the heading of *What do they know?* we reviewed the literature dealing with the modularity of the parser. Much of this section addressed the search for an ERP index of syntactic violation. Under the heading of *When do they know it?* we reviewed the literature dealing with the modularity of lexical access.In this section, we briefly discuss the ERP literature on parsing and lexical access in order to assess the extent to which those processes display the defining characteristics of modularity: domain specificity, obligatory action, information encapsulation, and automaticity. Overall, we find the ERP literature on the modularity of the language processor is best described by Cottrell's (1985) term *leaky modularity* and point to the connectionist modelling paradigm as providing an apt metaphor for understanding the ways in which language processes are modular in some respects and non-modular in others.

4.1 Domain Specificity

The issue of domain specificity chiefly involves the first of our watergate questions -- what do subjects know about the linguistic stimuli they process? On a modular account, the parser is a highly specialized mechanism which knows only syntactic facts and performs only syntactic computations. The idea behind domain specificity is to avoid the pitfalls associated with being a jack of all trades, master of none. Moreover, the hypothesis that the parser is a domain specific process predicts the existence of discrete levels of processing represented in the brain. Although the modularity thesis does not strongly predict a detectable index of syntactic processing, the detection of a unique index of syntactic processing would strongly substantiate the claims of modularists. In the second section we reviewed the search for an electrophysiological index of syntactic processing. Researchers have identified two ERP components associated with syntactic processing: a late positivity occurring 500 to 700 msec post-stimulus in response to grammatical errors and a left anterior negativity occurring 300-500 msec post-stimulus in response to both grammatical and ungrammatical stimuli.

Overall, we found no evidence for a syntax-specific ERP component which is equivalent to the extent to which the N400 can be considered an index of semantic processing. In fact, existing evidence favors the hypothesis that ERP effects elicited by syntactic violations reflect the action of domain general rather than domain specific processes. The P600 and the SPS are probably late positive components of the P3 family (Osterhout, 1993). Thus the late positivity associated with syntactic processing is probably a manifestation of domain-general computational tasks indexed by the P300 family. Moreover, the LAN probably indexes a process associated with storing items in working memory (Kluender & Kutas, 1993a, 1993b).

4.2 Obligatory Action

Modular systems embody a computational trade-off between speed and accuracy. Obligatory action is one property of modular systems which is involved in this trade-off. Modular processes, because they fire obligatorily, will necessarily fire in cases in which it is less than optimal to do so. Thus modular systems will occasionally compromise accuracy for speed. However, the advocate of modularity will remind us that the system is designed so that the central executive can remedy the occasional errors in the outputs of encapsulated modules. Moreover, this property provides us with testable hypotheses regarding the outputs of modular systems.

For example, the obligatory action of the first resort strategy of gap filling discussed in section 2.1 will often generate seemingly irrational filler-gap assignments. Garnsey, et al. relied upon the hypothesized property of obligatory action to demonstrate the first resort strategy. Recall that the first resort strategy involves assigning a filler to the first possible gap -- regardless of the semantic and/or pragmatic plausibility of that assignment. Thus Garnsey, et al. were able to interpret the larger N400 response to the verb in their implausible condition relative to their plausible condition as consistent with the results of the obligatorily acting first resort strategy and inconsistent with the more rational (albeit slower) last resort strategy.

Further, the lexical access process also seems to be characterized by obligatory action. The results of Van Petten and Kutas (1989) support both the biased activation hypothesis of lexical access and the backward priming account of the results in the short SOA condition. Both of these accounts of semantic priming phenomena involve the notion of obligatory action. According to the biased activation hypothesis, both meanings of an ambiguous word are activated regardless of the context (viz. obligatorily). The effect of context is more subtle -- that of biasing the more contextually appropriate meaning -- and pertains more to the information encapsulation property than to the property of obligatory action.

4.3 Information Encapsulation

Results reviewed here go against the thesis that the processes of parsing and lexical access are completely encapsulated with respect to information contained elsewhere in the system. First, Garnsey's results show that the parser is at least encapsulated enough to semantic and pragmatic contextual information to allow the assignment of fillers to gaps in conditions which could have been excluded on the basis of contextual information. Moreover, the lexical access process seems to be penetrated by context to a certain extent because processing of contextually biased meanings of ambiguous words are facilitated with respect to the processing of their unbiased meanings.

4.4 Automatic

There is a tendency in the arguments put forth by proponents of the modularity thesis to want to dismiss any findings which seem to support context effects (and hence undermine information encapsulation) as indicative not of low level processes like lexical access but, rather of later, post-lexical integrative processes. This argument can be seen in its extreme in Brown and Hagoort's (1992) adoption of the division of semantic priming effects into three components (Neely), only one of which is automatic. Because only processes which are automatic are deemed germane to the issue of modularity, divisions such as these are used to dismiss the relevance of data which show context effects.

Unfortunately, in doing so they systematically reduce the pool of data which modular models do account for. For example, following the logic of Brown and Hagoort , the only portion of semantic priming effects which is automatic (and hence germane to modularity) is ASA. Moreover, the only dependent measure of ASA is reaction time in a masked priming paradigm. However, note that the size of the facilitatory effect produced by masked priming is a mere 15 msec (Hagoort & Brown).

4.5 Integration and Dissociation

In this section we explore the utility of dissociating the different properties of modular processors. For example, certain processes might act obligatorily yet nonetheless be penetrable by contextual considerations. Moreover, the automatic/attentional dichotomy might be an inadequate characterization of neurolinguistic processes.

While the idea of the existence of autonomous subcomponents which render complex computations involved in language processing fast and efficient has much to offer, the claim that subcomponents are informationally encapsulated does not add much to the appeal of the aforementioned parts. The advantage of encapsulation (*qua* Fodor, 1983) is that it allows modules to avoid the computational problems involved in determining relevance. Modules have a finite set of possibilities to consider because they consider all and only those possibilities each time they operate. Because they consider the same set of possibilities regardless of context, the question of contextual relevance never arises.

However, the computational advantages accruing to a modular cognitive system can be retained in the absence of information encapsulation. If only we relax the requirement of strict information encapsulation, it is possible to maintain the computational advantages of a modular system while nonetheless letting the various sorts of information interact. One example of a computational system which retains the advantages of the divide and conquer strategy while relaxing the encapsulation requirement is provided by Cottrell (1985).

Cottrell (1985; Cottrell & Small, 1983) describes a connectionist model of lexical access in which the process of disambiguation is explained as competition between activations of the features of each of the meanings of an ambiguous word. Designed to simultaneously exploit syntactic and thematic role constraints on a word's meaning, the model disambiguates by choosing the "best fit" from the competing alternatives.

Cottrell (1985) characterizes the extent to which the components in his model are encapsulated with the phrase *leaky modularity*. In such a model, the task is split up into component parts (divide and conquer strategy) while allowing information from the computation of the various components of the task to interact with and mutually constrain one another. Dell and O'Seagda (1991) propose a similar concept with respect to spreading activation models of word production in the phrase "globally modular yet locally interactive." Interaction between the various levels of computation construed in this manner does not serve to confuse the processor with information irrelevant to the task of the subprocessor, but rather to constrain the possibilities.

Although the ultimate validity of the modularity thesis is not yet determinable, the evidence to date supports a view of language processing amenable to the leaky modularity characterization. The extent to which syntactic processing seems to perform some functions mandatorily and irrespective of good semantic and pragmatic reasons to avoid such strategies indicates that the parser has these modular characteristics. On the other hand, the semantic processing of words seems to occur in a non-modular fashion, interacting with information from higher levels of linguistic analysis.

References

Antos, S. 1979. Processing facilitation in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 527-545.

Bates, E. and MacWhinney, B. 1987. Competition, variation and leanguage learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bradley, D. and Forster, K. 1987. A reader's view of listening. *Cognition, 25*, 103-134.

Brown, C. and Hagoort, P. 1992. The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5, 34-44.

Canseco-Gonzalez, E., Hickok, G., Zurif, E., Prather, P., and Stern, C. 1993. Electrophysiological analysis of lexical access.

Chomsky, N. 1986. *Barriers*.

Chomsky, N. 1986. *Lectures on government and binding*. Dordecht: Foris.

Cohen, A., Ivry, R., and Keele, S. 1990. Attention and structure in sequence learning. Journal of Experimental Psycholgoy: Learning, Memory, and Cognition, 16, 17-30.

Cottrell, G. 1985. *A Model of Lexical Access of Ambiguous Words*. Unpublished Ph.D. dissertation.

Cottrell, G. and Small, S. 1983. A connectionist scheme for modeling word sedisambiguation. *Cognition and Brain Theory* 6:89-120.

Dell, G. and O'Seagda, P. 1991. Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98* (4), 604-614.

Donchin, E. 1979. Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), *Evoked brain potentials and behavior*. New York: Plenum Press, 1979.

Fodor, Jerry. 1983. *The Modularity of Mind*. Cambridge, MA: The MIT Press.

Forster, K. 1981. Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465-495.

Garnsey, S. Tanenhaus, M., and Chapman, R.. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18 (1), 51-59.

Hagoort, P., Brown, C., and Groothusen, Y. 1992. The syntactic positive shift as an ERP-measure of syntactic processing.

Halgren, E. and Smith, M. 1987. Cognitive evoked potentials as modulatory processes in human memory formation and retrieval. *Human Neurobiology, 6*, 129-139.

Hodgson, J. 1991. Information constraints on pre-lexical priming. *Language and Cognitive Processes, 6* (3), 169-205.

Kiger, J. and Glass, A. 1983. The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition, 11*, 356-365.

Kutas, M. 1993. In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8* (4) 533-572.

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

Kutas, M. and Kluender, R. 1992a. What is who violating? CRL Newsletter.

Kutas, M. and Kluender, R. 1992b. What is who violating? A reconsideration of linguistic violations in light of event-related potentials. In H.J. Heinze, T. Munte, and G.R. Mangun (Eds.), *Cognitive Electrophysiology*. Boston: Birkhauser.

McClelland, J. 1987. The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Neely, J. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibition-less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.

Neely, J. 1991. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), Basic processes in reading: Visual word recognition, 264-336. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Neely, J. and Keefe, D. 1989. Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24). New York: Academic Press.

Neville, H., Nicol, J., Barss, A., Forster, K., and Garrett, M. 1991. Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3 (2), 151-165.

Neville, H., Praterelli, M, and Forster, K. 1989. Distinct neural systems for lexical and episodic representations of words. *Society for Neuroscience*, Abstract.

Onifer, W. and Swinney, D. 1981. Accessing lexical ambiguities during sentence comprehension: Effects of frequency on meaning and contextual bias. *Memory and Cognition, 9*, 225-236.

Osterhout, L. 1993. Presented at CUNY Sentence Processing Conference.

Osterhout, L. and Holcomb, P. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*.

Picton, T. 1992. The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology 9* (4): 456-479.

Posner, M. and Snyder, C. 1975. Attention and cognitive control. In R.L. Solso (Ed.), Information processing and cognition: The Loyola Symposium (pp.55-85). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Pritchard, W. 1981. Psychophysiology of P300. *Psychological Bulletin, 89* (3), 506-540.

St. John, M. 1991. Hitting the right pitch: A meta-analysis of sentence context on lexical access. *Center for Research in Language Newsletter*, Vol. 5, No. 6.

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

Van Petten, C. 1993. A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8* (4) 485-531.

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

Van Petten, C. and Kutas, M. 1990. Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, 18 (4), 380-393.

Van Petten, C. and Kutas, M. 1991. Electrophysiological evidence for the flexibility of lexical processing. In G. Simpson (Ed.), *Understanding word and sentence*, 129-184. Amsterdam: North-Holland.

Van Petten, C. and Kutas, M. 1991. Influences of semantic and syntactic context on open and closed class words. *Memory and Cognition*, 19, 95-112.