Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage in Reading

Jonathan W. King and Marta Kutas
University of California, San Diego

Abstract

■ ERPs were recorded from 24 undergraduates as they read sentences known to differ in syntactic complexity and working memory requirements, namely Object and Subject Relative sentences. Both the single-word and multiword analyses revealed significant differences due to sentence type, while multiword ERPs also showed that sentence type effects differed for Good and Poor comprehenders. At the single-word level, ERPs to both verbs in Object Relative sentences showed a left anterior negativity between 300 and 500 msec postword-onset relative to those to Subject Relative verbs. At the multiword level, a slow frontal positivity characterized Subject Relative sentences, but was absent for Object Relatives. This slow positivity appears to index ease of processing or integration, and was more robust in Good than in Poor comprehenders. ■

INTRODUCTION

Studies using ERP measures of sentence processing have usually focused on the processing of single words rather than larger linguistic units (for reviews see Kutas & Van Petten, 1988; Fischler & Raney, 1991). This study augments the single word approach by also analyzing multiword ERPs that span whole syntactic constituents. In so doing, we hope to demonstrate how patterns of slower cortical potentials can help to elucidate aspects of language processing that extend across the sentence. The use of longer recording epochs should be particularly valuable in cases where expected differences are not due primarily to lexical factors, but rather to differences in the processing and use of information presumably stored in working memory (just & Carpenter, 1992). This is because language comprehension is a process that must necessarily unfold over an extended time-course, as single words rarely if ever convey the complete meaning of an utterance. To take the maximum advantage of both the multiword and single-word approaches, we investigate two sentence types that are syntactically similar, share enough surface features to make multiword comparisons informative, and yet differ noticeably in parsing difficulty for reasons universally attributed to differences in their working memory requirements. In addition, we focus on how differential success in comprehension can be tied to particular ERP components that index cognitive processes including working memory.

WORKING MEMORY AS A COGNITIVE PROCESS

As most cognitive functions are not simply mental reflexes, but depend on previous events and future expectations, some form of working memory is clearly involved in processes as complex as language processing. For many years, working memory models were essentially "register-based" short-term memory (STM) models (e.g., Atkinson & Shiffrin, 1968), but more recently the most influential models of working memory have abandoned the register/slot architecture for activation-based network architectures. In these kinds of models, items in working memory are either the most highly activated nodes (e.g., Anderson, 1983) or implicitly defined by patterns of activation across several nodes in a network (e.g., St. John & McClelland, 1990). The process of storage is accomplished via the spreading of activation; forgetting can result from the spontaneous decay of activation over time, and/or the redistribution of activation that occurs as the result of cognitive processing.

The realization that the operation of working memory encompasses both storage and processing functions led to a greater appreciation of the degree to which the processes acting on different "basic" information types were distinct. Baddeley (1986) proposed a more segmented memory system that consisted of a Central Executive and a number of slave memory systems, such as the "articulatory loop," the "phonological store," and the "visuospatial scratchpad." Since then, researchers have found evidence supporting the articulatory loop in humans (Vallar & Baddeley, 1984; Paulesu, Frith, & Frackowiak, 1993) as well as similarly specialized memory subsystems in nonhuman primates (Goldman-Rakic, 1992). While the extent to which such specialization encapsulates behavior is still unclear, the idea that various parts of the brain compute specific functions has become more generally accepted. A relevant example in light of the visuospatial scratchpad is the visual system, which is now thought to include more than 20 distinct brain regions whose interarea connections total almost 40% of those needed for complete interconnection (Felleman & Van Essen, 1991). A representative model of this more complex view of working memory is the CAPS 2 system, a production system that can simulate processing architectures ranging from micromodular to lushly interactive, but wherein all processing and storage functions share a dynamically allocated but limited pool of cognitive resources (Just & Carpenter, 1992). In essence, the architecture of the CAPS model involves three different kinds of productions: The first implements the storage and maintenance functions of working memory, in possibly both modality-specific and modality-independent formats. The second includes processes that operate on the contents of working memory with the implicit goal of achieving greater integration of items in working memory, thereby reducing the resources required for maintenance. The third includes the processes responsible for the allocation of resources between other simultaneously active storage and processing processes, both from moment-to-moment and strategically over longer processing episodes. (This third kind of production includes some of the functionality of Baddeley's central executive, albeit perhaps in a more distributed and less centralized form.)

Within this framework, working memory operates as a parallel production system, with all processing performed by altering the activation levels of various items contained in working memory on each processing cycle. Thus, the essential limitation on working memory is on the total amount of activation that can be propagated throughout the system at a given time, with the implication that resource-intensive computations will take more production cycles to complete when working memory capacity is most heavily taxed. In addition, items are assumed to decay out of working memory at a slow rate, so that cognitive resources must be spent to maintain them. Processes can also use WMC to suppress the activation levels of working memory elements, which is one way to implement competition between items that represent conflicting propositions (e.g., whether a given noun phrase is the agent or patient of a given verb, or which meaning of an ambiguous word will be chosen in context).

BASIC PRINCIPLES OF WORKING MEMORY AND PARSING

An investigation of the role of working memory in parsing produced one of the earliest collaborations in cognitive science (Miller & Chomsky, 1963). In this early framework, working memory capacity was believed to affect parsing by limiting the number of "registers" available for the application of syntactic transformations in the grammar to the input sentence, producing an output deep structure representation (cf. Savin & Perchonok, 1965). But as modern theories of parsing have turned from a reliance on specific, often ad hoc rules to general parsing principles, the role of working memory has also changed, and, in particular, it has assumed a more predictive role in processing. Modern parsing principles are designed to be flexible, on-line, and interactive in the sense that the set of such principles mutually constrains the syntactic form of grammatical sentences. Various authors have proposed limitations on what structures are built during parsing, and especially on what kinds of structures are preferred. This trend began over 20 years ago, when Kimball (1973) suggested how a parsing mechanism might reduce its working memory requirements by adopting a more heuristic approach to the analyses attempted. Frazier (1979) distilled some of these suggestions into the principle of Minimal Attachment, which stated that the syntactic processor entertaining only one analysis at a time for a given input string, namely the one with the fewest nonterminal nodes in its tree structure. Since then, much research has been aimed at verifying if or to what degree the principle of Minimal Attachment applies, and the answer remains controversial (e.g., Ferreira & Clifton, 1986; Altmann & Steedman, 1988; Clifton & Ferreira, 1989; Steedman & Altmann, 1989). Frazier's proposal is important in that it was based at least indirectly on working memory considerations, and has inspired work on other similarly motivated principles of parsing.

A recent proposal by Gibson (1990) relates parsing performance not to the number of nodes in a sentence's tree-structure, but to a more direct measure of how thematically integrated a given possible reading is. Thematic role assignment is the process of mapping thematic roles such as "agent," "patient," "theme," and "goal" onto syntactic constituents according to the argument structure of thematic role assigners such as verbs and prepositions. Modern syntactic theories such as Chomsky's (1981) Government and Binding assume that every (argument) noun phrase (NP) in a sentence must have a thematic role, and that all thematic roles that can be assigned must be assigned. Gibson has hypothesized that each NP that is momentarily without a thematic role and each unassigned thematic role imposes a burden on working memory. In the absence of any processing strategy, these processing loads accumulate until the correct thematic role assignments are made or until total working memory capacity is exhausted, in which case parsing grinds to a halt. Gibson used this principle to predict garden path effects in some sentences with temporary syntactic ambiguities. Specifically, if one reading of the ambiguity taxes working memory more than the other, the parser can drop it from consideration even though it might eventually be the correct reading. While Gibson originally made his proposal to account for garden path phenomena that were essentially self-evident, we can adopt it to make predictions about local processing difficulty in unambiguous sentences where we can assess the importance of working memory to the processing of sentences that do not usually result in total parsing failure.

ERP COMPONENTS IN LANGUAGE PROCESSING AND WORKING MEMORY USE

The working memory-based approach to the dynamics of cognitive processing, when combined with information about task-specific computations and their costs, yields testable predictions about the time-course of processing. The usual assumption is that resource-constrained computations take longer than "easier" computations, as indexed by increased reaction times in on-line reading tasks. However, at least two caveats must be expressed about this proposal. The first is that resource usage is a function of both the memory load and the time over which it must be maintained; therefore subjects faced with a substantial memory load may reduce resource usage by storing memory load items for less time (essentially a form of speed-accuracy trade-off). The second is that while reaction time data can provide a good estimate of how long a processing event took, they offer only an indirect reflection of what distinct processes were involved and the time-course of the processing that led to a single reaction time. ERP data provide a complementary source of information about WM use, difficulty in parsing, and the relation between these two. Reliable differences in ERPs between conditions provide strong evidence that processing differed, and to the extent that ERP effects can be identified with specific cognitive processes, they provide some evidence of how processing differed. Moreover, such differences may be evident even when RT data show no effect.

ERP methodologies have promised both greater temporal resolution and the possibility of distinguishing between cognitive processes based on other dimensions such as the morphology and scalp distribution of the potentials. Among the ERP components that have been implicated in sentence processing are the N400 (Kutas & Hillyard, 1980), the Left Anterior Negativity [LAN] (Kluender & Kutas, 1993), and the N400-700 (Neville, Mills, & Lawson, 1992), as well as a variety of components posited to index various kinds of grammatical violations (Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992). A brief review of ERP components important in language processing follows.

The N400 is an ERP component that is reliably elicited by semantically anomalous words in context (Kutas & Hillyard, 1980). But the N400 has been found to be more broadly influenced by the degree to which a word in context builds on prior semantic expectancies (Kutas & Hillyard, 1984; Van Petten & Kutas, 1991). The N400 has also been used to test indirectly the effect of WMC on more strictly syntactic aspects of parsing. For example, Garnsey and her colleagues used the N400 to detect semantic garden-pathing apparently caused by their subjects' processing choices in sentences with a temporary syntactic ambiguity (Garnsey, Tanenhaus, & Chapman, 1989). In the Garnsey study, readers were often misled about the syntactic structure of the sentence, and thus had gone down a "garden path" only to reach an apparent semantic dead end, at which point a large N400 developed. In particular, Garnsey examined sentences with embedded "wh-questions" wherein the filler was either plausible or implausible:

(la) The businessmen knew which
customer the secretary *called* _____ at home ...
(lb) The businessmen knew which
article the secretary *called* _____ at home ...

They observed an N400 effect at the word "called" in (lb), due to the implausibility of "article" as a filler for the gap following "called." Insofar as garden-pathing results from limitations in working memory capacity, this use of the N400 might also help us index working memory demands at the syntactic level in a more abstract way; that is, we might expect N400s to be elicited by words introducing unexpected syntactic continuations.

The existence of ERP components with a more direct relation to loading WM is suggested by at least one study on retention of information in WM (Ruchkin, Johnson, Canoune, & Ritter, 1990). Moderate WM loads were associated with greater negativity between 250 and 600 msec poststimulus, particularly at left hemisphere central electrode sites, when compared to smaller WM loads. This negative component was followed by a positive slow wave during the retention interval at other scalp locations. These components, obtained in a nonlanguage processing task, are remarkably similar to those observed by Neville et al. (1991) in response to local phrase structure (i.e., word order) violations. The similarity is intriguing because such local phrase structure violations are likely to induce a temporary WM load when the current word cannot be integrated into the current sentence representation.

Another ERP component possibly related to Ruchkin's finding is the left anterior negativity (LAN) described by Kluender and Kutas (1993) in their study of sentences containing long distance dependencies. This component was elicited by at least two kinds of words, those that were being processed while a working memory load was being carried, and those at and immediately following a sentence location where thematic role assignments were being made. (And therefore where the working memory load could be integrated into the sentence representation being constructed.) Kluender and Kutas only observed the LAN embedded in the ERP to closed class words, but there is every expectation that it would be observed in response to open class words as well.

Another potentially relevant ERP component is the N400-700 described by Neville, Mills, and Lawson (1992). The N400-700 is an anterior, slightly left-lateralized component that is reliably evoked by function words in sentential contexts. An important feature of most function words in English is that they introduce new syntactic constituents (e.g., noun phrases often begin with an article such as "the"; prepositional phrases begin with a preposition, etc.). To the extent that the function word sets up an anticipation for a specific syntactic constituent, the N400-700 might be seen as a language-specific version of anticipatory potentials such as the contingent negative variation (CNV). Another possibility, however, is that a function word per se induces a greater working memory load and thus evokes the N400-700 for this reason rather than for any expectation it sets up.

The existence of ERPs that are sensitive to these different aspects of language processing related to working memory is helpful in that it allows us to design an experiment that might be more sensitive to working memory-related effects than a reading time experiment. Further, we can have some idea of how particular ERP components (e.g., N400, LAN, slower cortical potentials) would change throughout the course of processing.

AN ERP EXPERIMENT WITH ENGLISH RELATIVE CLAUSES

To study how working memory capacity is deployed in language processing, one could manipulate the processing requirements of the linguistic structures being processed, the processing capacity of the subjects in the study, or both. King and Just (1991) did both by comparing the word-by-word reading times of High- and Low-capacity readers (as measured by the reading span test, Daneman & Carpenter, 1980) for sentences differing in the processing demands posed by their syntactic structure; high-capacity subjects should be less affected than low-capacity subjects in processing more demanding syntactic structures. The predicted patterns of results were found with sentences similar to those in (2a) and (2b):

(2a) The reporter who harshly attacked the senator admitted the error.

(2b) The reporter who the senator harshly attacked admitted the error.

Both sentences (2a) and (2b) contain a relative clause modifying the subject of the sentence, but differ in the role that the main subject NP ("the reporter") plays in the relative clause; in (2a), the main-clause subject is also the subject (and agent) of the verb in the relative clause, while in (2b), it is the object (and patient). For this reason, sentences like (2a) are known as subject-subject relative (SS) sentences, while those like (2b) are known as subject-object relative (SO) sentences.

By the thematic role assignment principles discussed above, processing an SO sentence (2b) should induce a greater WM load within the relative clause. Further, there should be a greater working memory load at and just following the relative clause verb of SO sentences, where two separate thematic role assignments are being carried out; that is, where readers encounter the first verb of the sentence and have to determine which noun phrase is indeed the subject. Note that, with these materials, neither semantic nor pragmatic information can be used to choose which noun is the subject. As anticipated, both groups showed similar reading times on SS sentences like (2a), but their reading times diverged on SO sentences like (2b), with high-processing capacity subjects requiring less time to read the relative clause verb and the word immediately following than low-processing capacity subjects. Surprisingly, there were no clear reading time differences between SO and SS sentences earlier in the relative clause, as might have been expected from the different loads in working memory. But, as discussed previously, RT measures may not be sensitive to the effect of simply carrying a temporary memory load, or may reflect a trade-off between the intensity of processing and its duration.

By recording ERPs to SO and SS sentences (like 2a and 2b), we can examine an inherently multidimensional measure of processing and also obtain a continuous recording of the brain activity resulting from the processing of two sentence types that differ in syntactic structure. This approach uncovers ERP effects that apparently reflect differences in working memory use during parsing, and that also distinguish the readers in our sample who are either better or worse comprehenders.

RESULTS

Behavioral Data

Comprehension accuracy was calculated separately for SS and SO probes. Furthermore, subjects were classified as "Good" or "Poor" comprehenders according to a median split on their total comprehension scores. The data were analyzed using an ANOVA with gender (male vs. female) and comprehension ability (good vs. poor) as between subjects variables, and Syntactic Structure (SO vs. SS) as a within-subjects variable. (This ANOVA verified that our so-called Good comprehenders did indeed perform significantly better than Poor comprehenders [$F(1,20) = 45.55$, $p < 0.0011$, with an overall comprehension rate of 87% correct for the Good compared to 68% correct for the Poor.)

Regarding the effect of Sentence Type, we find that subjects in this sample were not reliably better at comprehending SS sentences (79%) than SO sentences (77%), [$F(1,20) = 1.02$, n.s.]. This contrasts with findings from previous studies as discussed below. The main effect of Gender was not reliable, but the interaction of Gender and Sentence Type was [$F(1,20) = 5.91$, $p < 0.05$]. In this experiment, male subjects had comprehension scores that were 10% worse on SO sentences than female subjects (72 vs. 82%), but 6% better on SS sentences (82 vs. 76%). The explanation for this last interaction is unclear, although Neville et al. (1991) hinted that gender had an effect on some aspects of their data that involved ERPs to syntactic violations. No other main effects or interactions were significant.

This lack of an effect of Sentence Type with these materials is slightly surprising given the history of finding either reliable differences (e.g., Foss & Cairns, 1970; Larkin & Burns, 1977) or stronger trends (King & Just, 1991). There were, however, at least three differences in the current study that could partially explain these divergent results. First, the materials differed in that an adverb was inserted into the relative clause for both sentence types; this lengthened the distance between the subject of the main clause and its verb and may have affected comprehension of the main clause probes (which were comprehended less well in SS sentences than in SO sentences). Further, in our procedure we tested comprehension only on a random subset of the critical materials, which may have had an effect on either subject vigilance or strategic processes in sentence comprehension that would have decreased performance on SS trials. On the other hand, the lack of a strong Sentence Type main effect in the comprehension data makes it somewhat more difficult to ascribe differences in the ERPs for these two sentence types to completely uninteresting causes (e.g., SO sentences are just impossible to process, subjects were paying less attention to them, etc.).

Multiword ERPs

For the purposes of analysis, these multiword ERPs were divided into nonoverlapping two- and three-word regions that reflect the syntactic structure of the two sentence types, as indicated in Table 1. The grand average multiword ERPs for both SO and SS sentences themselves appear in Figure 1. [Actual analyses were based on multiword ERPs beginning after the word "who" so that we could capture effects throughout the main verb phrase (MVP) region; the ERPs in the figure show that processing did not diverge before the relative clause.] The mean amplitudes of the ERP were measured for each of these regions, which were then separately submitted to an ANOVA with the within-subjects variables of sentence type (SO vs. SS), Electrode Site (6 pairs of lateral sites) and Hemisphere (Left or Right), and the between-subjects variable of Comprehension (Good or Poor). In all the results reported below, the Huyn-Feldt correction was applied where sphericity assumptions were violated; in these cases the uncorrected degrees of freedom are reported with the corrected probability levels.

In the Early Relative Clause (ERC) region (SO: article-noun2; SS adverb-verb), there was a significant effect of Electrodes, indicating more absolute negativity over posterior electrode sites [$F(5,100) = 12.53$, $p < 0.0001$]. In addition, there was a significant Electrodes x Sentence Type interaction due to the development of a relative frontal negativity for SO sentences compared to SS sentences [$F(5,110) = 13.41$, $p < 0.001$]. As Figure 1 suggests, this relative frontal negativity was both bilateral and sustained throughout the ERC region, and occurred precisely where the syntactic structure of the SO and SS sentences diverges. In the SO case, the ERC region spans the second noun phrase (NP2) of the sentences, which must be maintained in working memory until the corresponding verb is encountered; in the SS sentence, the ERC region spans the adverb and verb of the relative clause, which should trigger the assignment of the initial noun phrase in the sentence as the subject of the relative clause.

In the Late Relative Clause (LRC) region (SO: adverb-verbl; SS: article-noun2), the only reliable effect was a strong main effect of Electrodes, with posterior sites showing more negativity than frontal sites [$F(5,110) = 46.22$, $p < 0.001$], and a maximum negativity at Posterior Temporal sites. Note that the lexical items being contrasted in the LRC region are identical to those contrasted in the ERC region except that the sentence types in which they occur are reversed. In the LRC region, however, we do not find the interaction between Sentence Type and Electrode Sites that occurred in the ERC region. If lexical factors were the sole cause of the Sentence Type effects in the ERC region, then we would not expect this asymmetry. This pattern of results is

Table 1. Definition of Sentence Regions within the Relative Clause of SO and SS Sentences

	Sentence Regions			
Sentence Type	*Pre-relative Clause*	*Early Relative Clause*	*Late Relative Clause*	*Main Verb Phase*
SO	The reporter who	the senator	harshly attacked	admitted the error
SS	The reporter who	harshly attacked	the senator	admitted the error

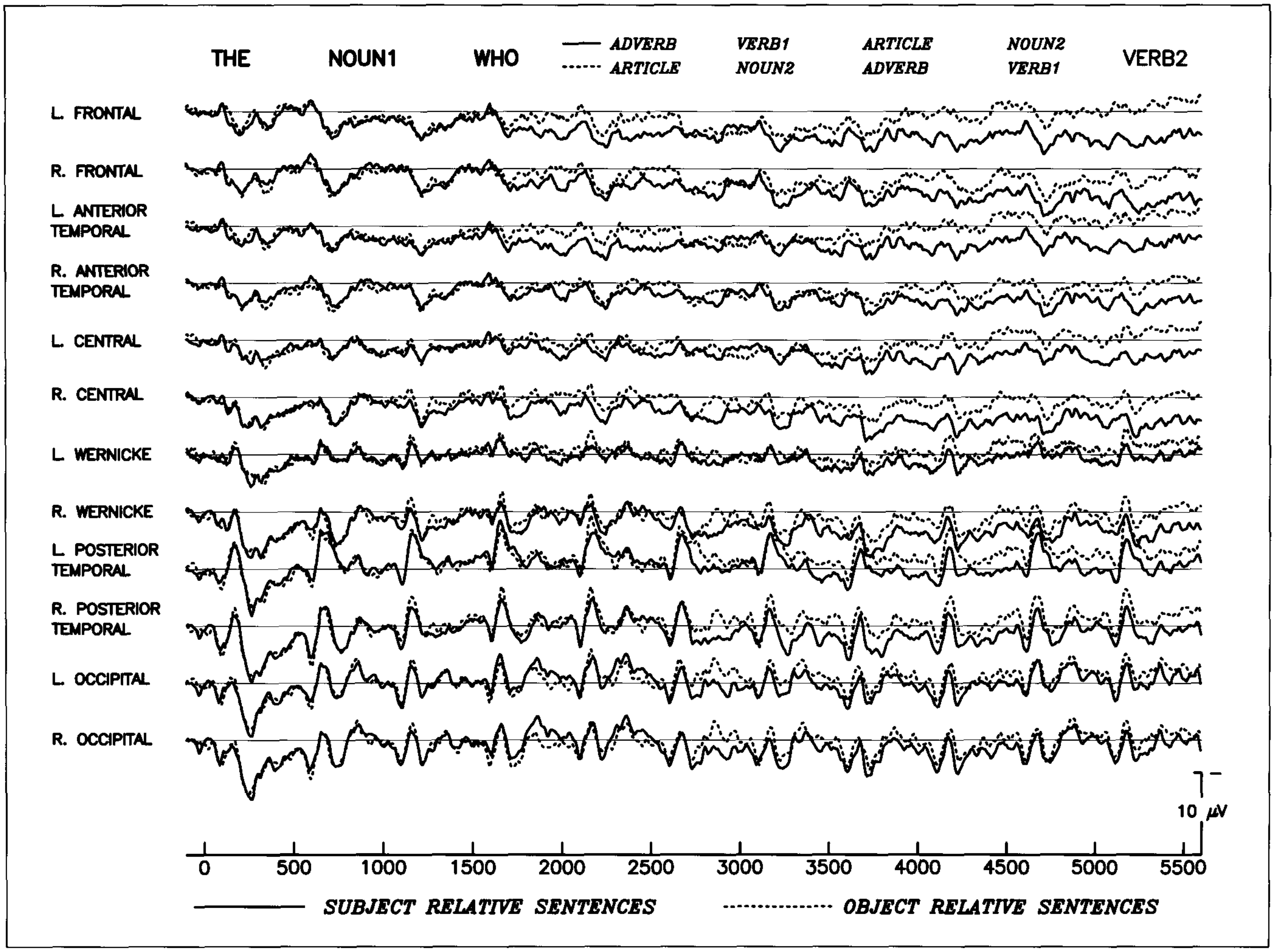


Figure 1. Grand average ERPs (n=24) from 12 Electrode Sites for the multiword waveform evoked by the initial eight words of both Object Relative (SO) and Subject Relative (SS) sentences.

instead consistent with various forms of the WM-loading hypothesis, since that would be a substantial load in the ERC region of SO sentences, but not in the LRC region of SS sentences.

In the Main Verb Phrase (MVP) region (both SO and SS: verb2-article-noun3), ERPs to SO sentences show a widespread negativity relative to the Subject Relative sentences. This leads to a reliable main effect of Sentence Type in the ANOVA [$F(1,22) = 5.71$, $p < 0.05$]. There is also a reliable effect of Electrode Sites [$F(5,110) = 4.50$, $p < 0.01$] with a distribution of anterior and posterior negativity similar to that seen in the ERC region. While Figure 1 suggests that the greater negativity for SO sentences in the MVP region may be more prominent anteriorly, there is no reliable interaction between Sentence Type and Electrodes [$F(5,110) = 1.04$, $p > 0.3$] due to substantial variability between subjects. This variability is due to different patterns of frontal negativity in Good and Poor comprehenders, as Figure 2 shows.

While both Good and Poor comprehenders show a slow positive drift in their ERPs as the whole sentence progresses, this pattern is more pronounced in the Good Comprehenders. Further, Good comprehenders show much more (relative) anterior negativity to SO sentences in the MVP region than do Poor comprehenders, who show only small differences between sentence types at Frontal and Anterior Temporal sites. Thus, when we consider the factor of Comprehension in the ANOVA of the MVP region, we find a reliable three-way interaction of Comprehension (Good vs. Poor) x Sentence Type x Electrodes [$F(5,110) = 3.42$, $p < 0.05$].

In summary, we have documented the existence of at least three scalp potentials effects associated with sentence processing. The first is a slow, frontal, positive-going wave that appears to index successful integration in SS sentences. While not commented on previously, this type of positivity has appeared in previous work (e.g., Kutas, Van Petten, & Besson, 1988) that used more syntactically simple sentences. In our study, this positivity was far more prominent in Good than in Poor comprehenders. The second and third were slow negative modulations of this pattern associated with both the addition of a load to working memory and the aftermath of thematic role assignment in SO sentences, especially for

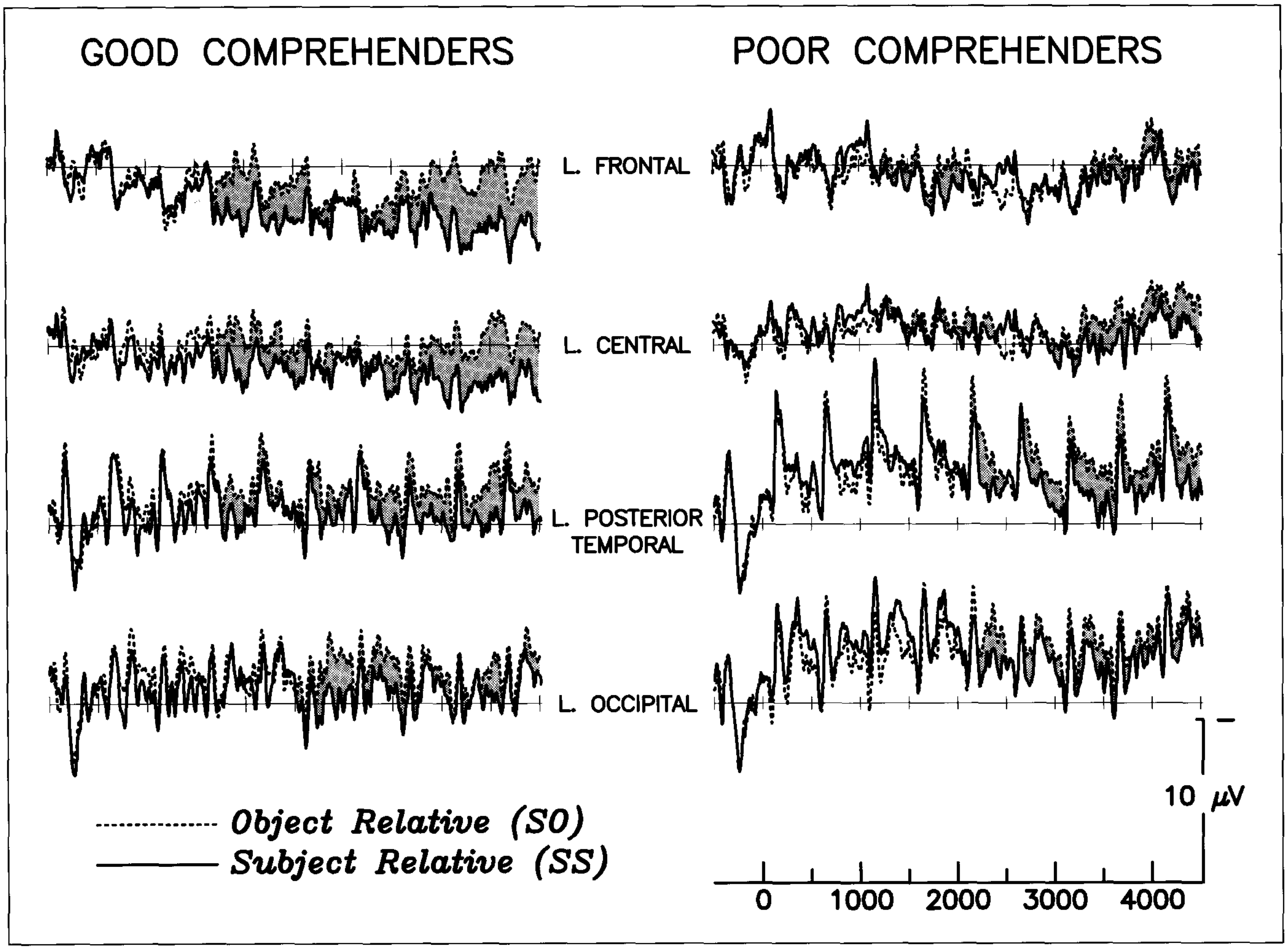


Figure 2. Average ERPs from representative left Hemisphere Sites from Good (n = 12) and Poor (n = 12) Comprehenders for Object Relative (SO) and Subject Relative (SS) sentences. Waveforms are aligned on the first word of each sentence type and include the response to words up to and including the final word of the main verb phrase.

Good Comprehenders. It is also clear from the multiword ERPs, however, that there are numerous ERP effects that appear to be associated specifically with single words, which we will discuss next.

Single Word ERPs

To examine some of the processing differences noted above on a shorter time scale, we also studied the ERPs to the single words that constitute the whole clauses. While these are essentially the same data, a comparison of Figure 1 and Figure 3 shows that the change in time scale and rebaselining greatly alters their visual appearance, and the kinds of inferences one is likely to draw. Separate ANOVAs were conducted on the mean amplitude between 300 and 500 msec post-onset for the specific lexical items being investigated. This window was chosen to examine not only possible differences in the N400, but also features of the Left Anterior Negativity (Kluender & Kutas, 1993), which may be related to aspects of parsing. We present results for the single words in the serial order they appear in the SO sentences (top line of Table 1), followed by some comparisons that unconfound word class, thereby clarifying the role of structural factors that may not be attributed to lexical factors alone.

The first comparison in the ERC region contrasts a definite article in the SO sentence with an adverb in the SS sentence. In this case we would expect to find either a main or interaction effect of Sentence Type because it mirrors a word class difference known to produce an anterior negativity to the closed class item (here "the'). The ANOVA confirms this result, as Sentence Type and Electrodes interact robustly [$F(5,110) = 34.23$, $p <$ 0.0011. The ERP to the definite article shows more anterior negativity than the ERP to the adverb, while the pattern is reversed at Occipital electrode sites. This pattern seems to reflect the larger N400 component observed for open class words compared with closed class items (Kutas, Van Petten, & Besson, 1988). There is also an effect of Electrode Site, with greater overall negativity at posterior electrode locations [$F(5,110) = 24.07$, $p <$

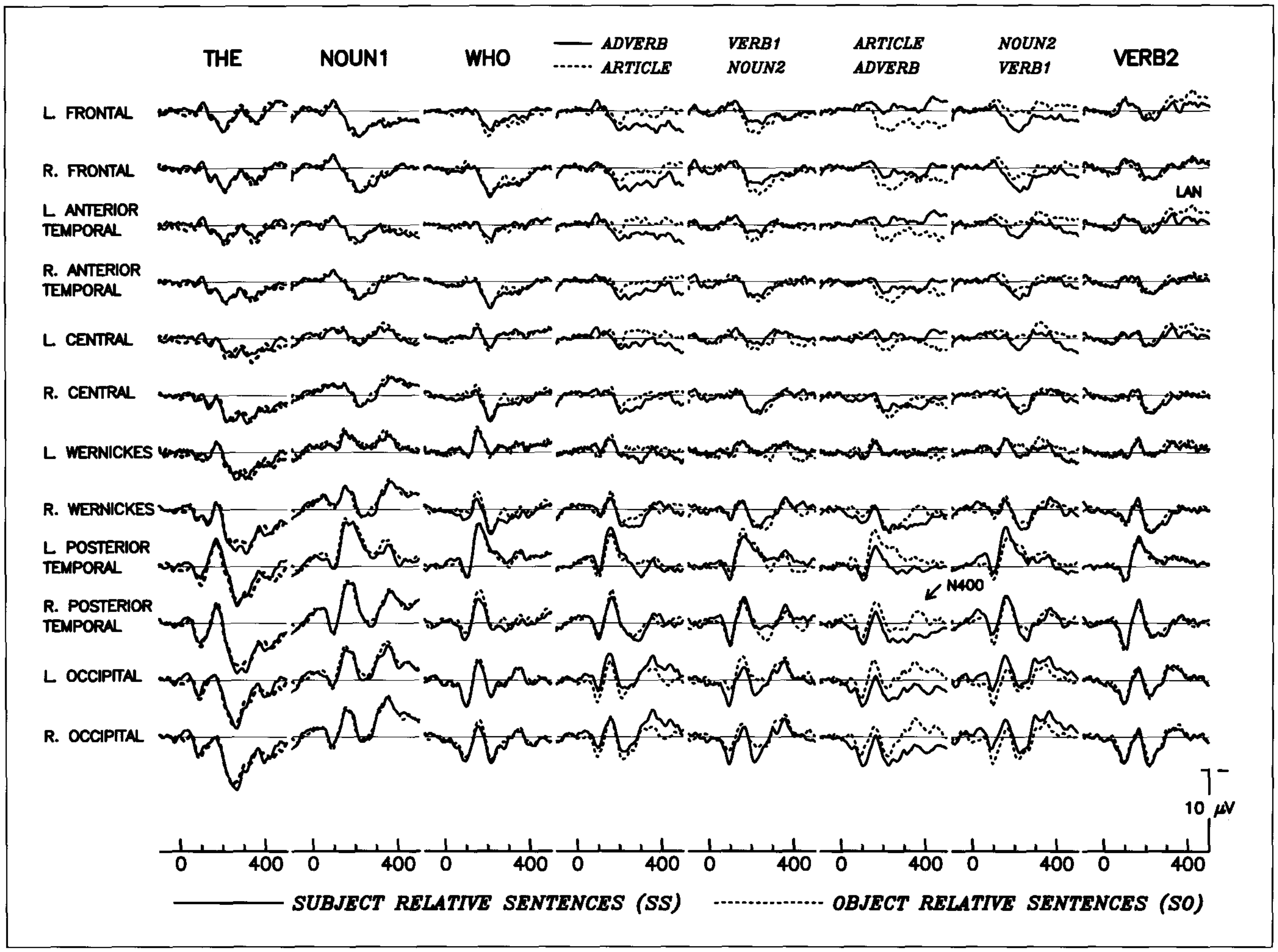


Figure 3. Grand average ERPs (n = 24) from 12 Electrode Sites for each of the initial eight words of both object Relative (SO) and Subject Relative (SS) sentences.

0.001]. There were no reliable main effects of either Sentence Type, or interactions with either Hemisphere [$F(1,22) = 2.91$, $p = 0.10$] or Hemisphere x Electrode [$F(5,110) = 0.80$, n. s.].

So far, these results just mirror those found for the ERC region as a whole, but the picture becomes more complete when we consider the additional effect of Comprehension Skill. Comprehension Skill does not generate a reliable main effect [$F(1,22) = 2.72$, $p = 0.11$], but does interact significantly with Electrode Site [$F(5,110) = 3.6$, $p < 0.05$] as well as participating in a three-way interaction of Comprehension Skill, Sentence Type, and Electrode Site [$F(5,110) = 6.99$, $p < 0.01$].The Comprehension Skill x Electrode interaction is best characterized as a reduction in posterior negativity for Good compared to Poor Comprehenders. This seems in turn to reflect a noticeable difference in the size of the N400s generated by Good and Poor comprehenders, as can also be seen in the multiword ERPs in Figure 2. More specifically, the three-way interaction involving Comprehension Skill shows that this difference is due to a larger N400 elicited by the closed class item "the" in the SO sentence for Poor than for Good comprehenders. While N400 effects to closed class items are unusual, they have been observed in other contexts (e.g., Kluender & Kutas, 1993). Similarly, differences in N400 effects due to individual differences have also been observed (Kutas et al., 1988). We delay further consideration of these issues, however, until we have presented data from other sentence locations. In comparison to this multitude of effects early in the ERC region, the contrast of the relative clause noun in SO sentences with the relative cause verb in SS sentences produces only two effects, one of which is marginal. The significant effect of Electrode site [$F(5,110) = 10.82$, $p < 0.001$] reflects the expected greater posterior negativity typical of open class items.

The next sentence position, which begins the LRC region, is a contrast between the same lexical items as in the first ERC word comparison, except that the effect of Sentence Type is reversed: the SO sentence now provides the adverb while the SS sentence provides the definite article. The results of this simple reversal are, however, quite different from those observed at the earlier sentence location. Whereas the analysis of the early ERC yielded a very robust main effect of Electrode Site, the same analysis here generates only a marginally reliable one [$F(5,110) = 2.58$, $p = 0.09$]. Also, even though there is a reliable Sentence Type x Electrode interaction here [$F(5,110) = 30.05$, $p < 0.0001$] as in the ERC, the form of the interaction is different. In the early ERC, the definite article generated fairly widespread negativity. Here, the exact reverse holds: the definite article (now in the SS sentence) generates only a small frontal negativity and a widespread positivity at posterior sites. This is the pattern usually associated with closed class items, as opposed to the N400-like activity observed in the ERP to the definite article in the ERC. These changes in the ERPs to the definite article contrast with the similarity of the ERPs to the adverb at both sentence locations.

The other differences between the ERPs to these lexical items in the ERC and LRC corroborate the notion that the article in the ERC was generating an atypical N400-like response, but in the LRC an N400-700 response. In contrast with the results in the ERC region, at this sentence location there is both a reliable main effect of Hemisphere [$F(1,22) = 15.89$, $p < 0.001$] and a reliable interaction of Hemisphere with Sentence Type [$F(1,22) = 13.68$, $p < 0.01$]. The effect of Hemisphere results from slightly less positivity over the left scalp than the right, while the interaction results from a clear (and usual) left scalp negativity in the ERPs to the definite article, compared with relatively equal positivities to the adverb. The reliable three-way interaction of Sentence Type, Hemisphere, and Electrode [$F(5,110) = 3.50$, $p < 0.05$] further indicates that this left hemisphere negativity is largest at Frontal and Anterior Temporal electrode sites. Finally, Comprehension Skill had no main effect at this sentence location, and did not interact significantly with any other factors at any level (all ps > 0.3). Overall, it is clear that the ERP effects present at the beginning of the ERC are quite different from those seen at this location; the differences are related to the processing of the definite article in two substantially different contexts, one where the syntactic structure induces a greater working memory load (the ERC portion of the SO sentence) and the other where the syntactic structure does not (the LRC portion of the SS sentence).

A clear indication that this difference results in an N400 response to the article in the SO sentence can be seen in the left column of Figure 4, where the single-word ERPs to the two articles are compared. The early differences in these two waveforms probably result from the differences in word class of the words immediately

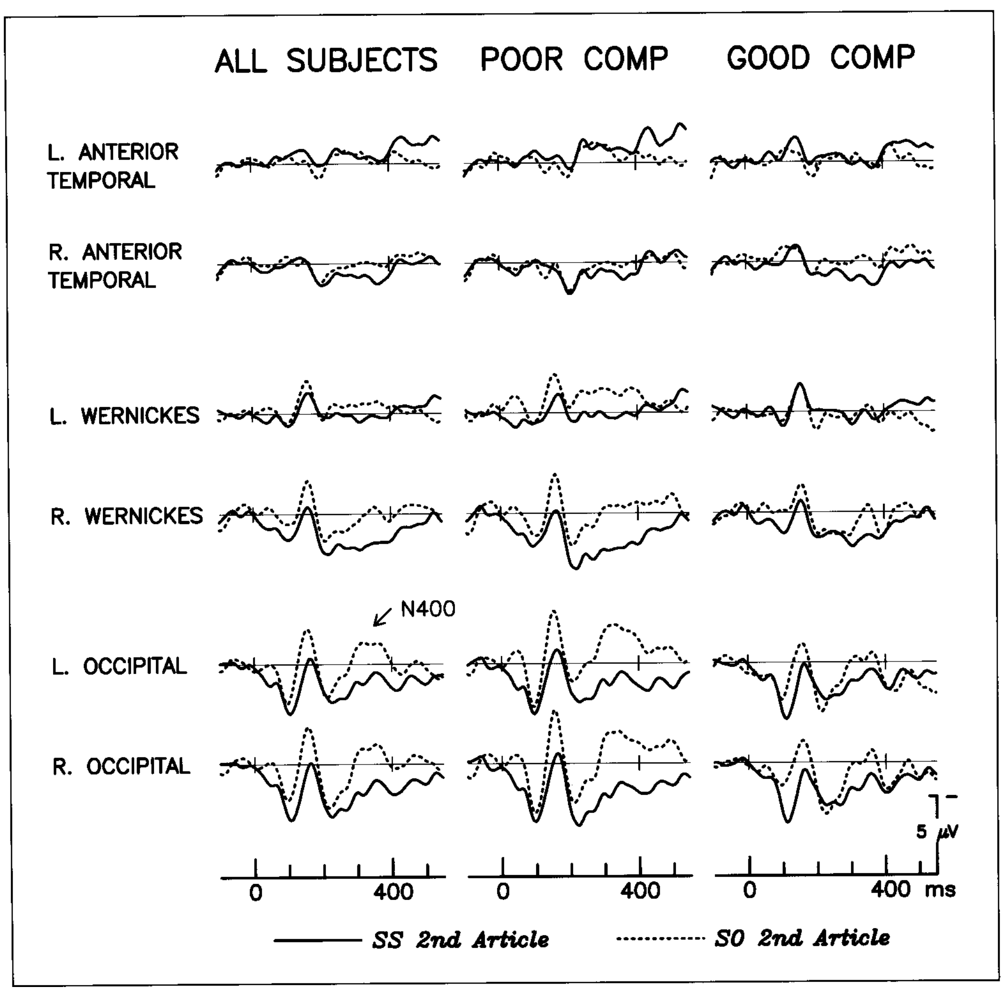


Figure 4. Average ERPs to the second definite article in Subject Relative (SS) and Object Relative (SO) sentences. The left column presents the grand average ($n = 24$) ERPs, while the center and right columns present averages separately for Poor ($n = 12$) and Good ($n = 12$) comprehenders, respectively.

preceding (cf. Van Petten & Kutas, 1991). The difference after about 200 msec, however, shows the anterior-posterior and left-right distribution of the classical N400, here peaking at about 350 msec post-onset over Right Parietal (Wernicke's) scalp.

An ANOVA done on the ERPs to these two definite articles in the 300-500 msec post-onset window provides some confirmation of this hypothesis. There are reliable main effects of both Hemisphere (left more negative than right) and Electrode Site (greatest negativity at anterior locations) as well as a reliable interaction of the two (scalp negativity is greatest at left anterior sites) [Hemisphere: $F(1,22) = 27.59$, $p < 0.001$; Electrode: $F(5,110) = 4.72$, $p < 0.05$; Hemisphere x Electrode: $F(5,110) = 4.53$, $p < 0.01$]. These effects merely indicate that we are looking at closed class items with a substantial N400-700 component, just as one would expect. Additionally, however, there are reliable interactions of Sentence Type with Hemisphere (SS but not SO sentences show more left hemisphere negativity), and with Electrode Site (SS shows more anterior negativity, SO shows more posterior negativity) [Sentence Type x Hemisphere: $F(1,22) = 12.94$, $p < 0.01$; Sentence Type x Electrodes: $F(5,110) = 5.94$, $p < 0.05$]. These results are all consistent with the hypothesis that the definite article in the SO sentence is eliciting an N400 in addition to the expected N400-700.

As suggested previously, the amplitude of this N400 appears to differ for Poor and Good comprehenders, as shown in the center and right columns of Figure 4. Thus, in the same ANOVA, Comprehension Skill participates in a reliable three-way interaction with Sentence Type and Electrode, due to Poor Comprehenders showing much more posterior negativity to the article when it occurs in the SO sentences [$F(5,110) = 7.00$, $p < 0.05$].

The last word of the LRC region ends the relative clause for both SO and SS sentences, but the syntactic roles played by these words are very different. In the SO sentence, the LRC ends with a verb that must be simultaneously associated with both its subject and its object, the latter having been maintained in working memory without a thematic role throughout the relative clause. In the SS sentence, the LRC ends with a noun that is quite easy to assign as the object of the relative clause verb. We expect the additional processing required in the SO sentence to be manifested in a (possibly lateralized) anterior negativity when compared with the noun in the SS sentence, given the findings of Kluender and Kutas (1993).

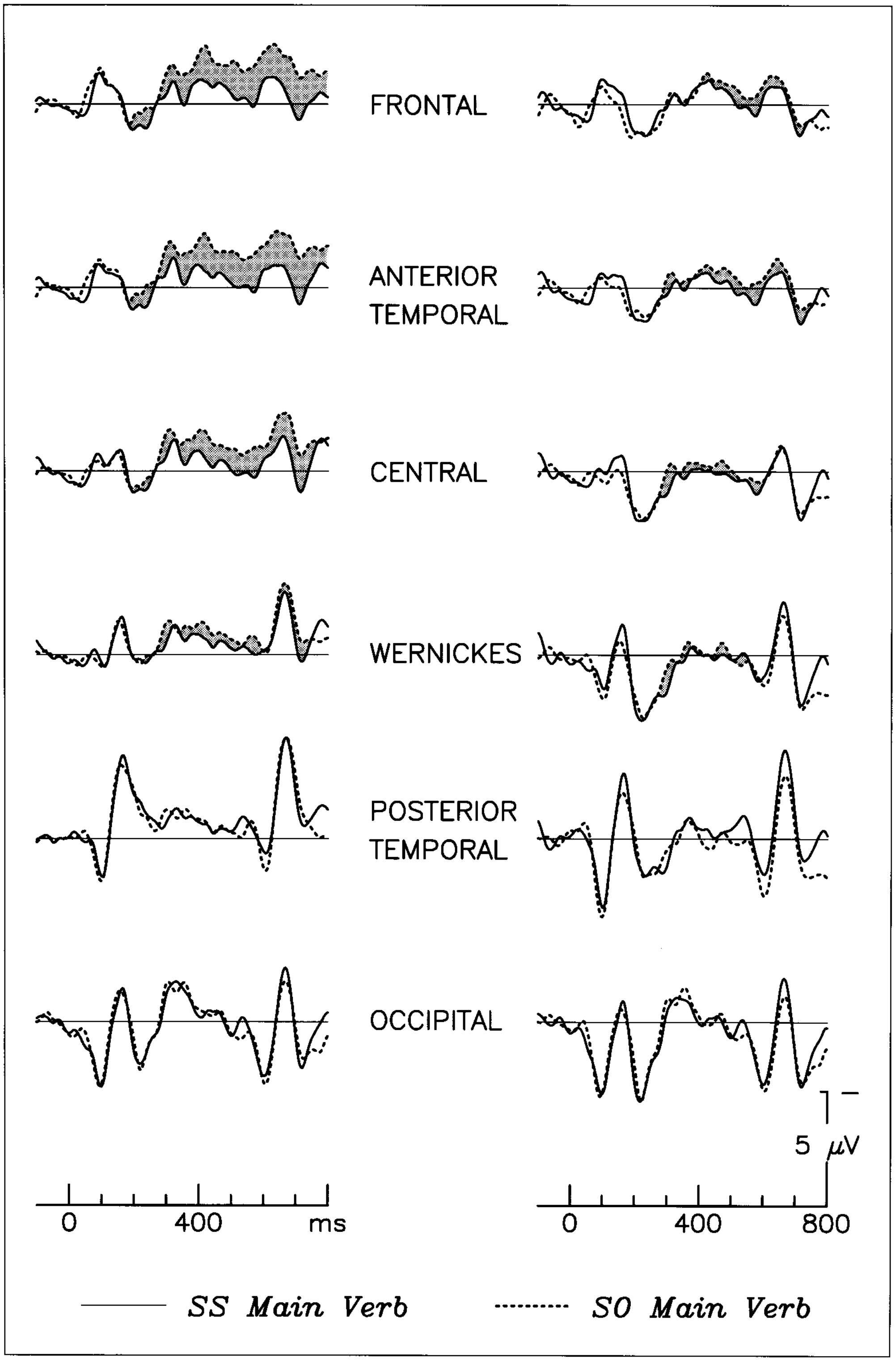
This expectation is fulfilled by the reliable interaction of Sentence Type and Electrode Sites, which indicates greater formal negativity in this window for the verb in the SO sentence relative to SS sentences [$F(5,110) = 5.86$, $p < 0.05$]. The only other reliable effect is that for Electrodes, which shows a greater overall negativity at posterior electrode sites [$F(5,110) = 7.81$, $p < 0.01$]; an interaction of Sentence Type and Hemisphere (indicating greater left negativity for the SO sentence verb) just misses statistical reliability [$F(1,220) = 4.13$, $p = 0.054$].

While these findings are consistent with the notion of anterior negativity as a reflection of processing complexity and working memory operations, these effects may also reflect ERP differences to items of different lexical categories. Moreover, the LAN effect itself has been shown only for words of the same lexical category. To investigate this possibility, we directly compared the two relative clause verbs in an ANOVA with the same factors as those above. If the LAN depends only on raw processing load, we would expect significant interactions involving Sentence Type and either Electrodes, Hemisphere, or ideally both. The results were suggestive, but the interaction of Sentence Type and Electrodes and the three-way interaction of Sentence Type, Electrodes, and Hemisphere were both only marginally reliable [Sentence Type x Electrodes: $F(5,110) = 3.04$, $p = 0.09$; Sentence Type x Electrodes x Hemisphere: $F(5, 110) = 2.40$, $p = 0.07$]. Another idea is that the LAN could depend on the effective processing load, which might vary from reader to reader depending on their verbal abilities. This hypothesis found greater support, as Comprehension Skill did enter into a reliable three-way interaction involving Sentence Type and Hemisphere [$F(1,22) = 5.14$, $p < 0.05$]. This interaction indicates that Poor comprehenders showed greater left hemisphere negativity for the SO relative clause verb than did Good comprehenders, who showed approximately equal potentials over both Hemispheres and Sentence Types. One additional difference between Good and Poor comprehenders was in the lateralization of the anterior negativity elicited by both verb types; Poor comprehenders showed greater anterior negativity only over the left hemisphere, while Good comprehenders showed a clear anterior-posterior gradient over both hemispheres. This resulted in a reliable interaction of Comprehension Skill, Hemisphere, and Electrode Site [$F(5,110) = 2.67$, $p < 0.05$].

When we move our analysis into the MVP region, matters are simplified by the complete absence of lexical confounds, which will help clarify the effects of Sentence Type. In particular, the contrast of the two main clause verbs occurs just after the "gap" in the SO relative clause, at a sentence location where the greatest behavioral differences caused by the two sentence types are generally found (e.g., Ford, 1983; King & Just, 1991). This difference in processing load just after the relative clause should lead to a substantial LAN in favor of the SO sentences, and this is, in fact, what is found, as shown in Figure 5. A reliable three-way interaction of Sentence Type, Electrode Site, and Hemisphere demonstrates that the difference in negativity between SO and SS main clause verbs is greatest at left anterior sites, which is the strict definition of the LAN [$F(5,110) = 3.53$, $p < 0.01$].

In addition to evidence for the specific LAN elicited by SO main clause verbs, Figure 5 suggests there is a more general negativity over the left hemisphere for

Figure 5. Grand average ERPs (n = 24) from 12 Electrode Sites for the main verb and following article in both Subject Relative (SS) and Object Relative (SO) sentences. The left and right columns in the figure present waveforms for left and right hemisphere Electrode Sites. The left anterior negtivity (LAN) is shaded.



both SO and SS main clause verbs, especially in the temporocentral regions. In fact, the ANOVA shows both a reliable main effect of Hemisphere, and an interaction of Hemisphere and Electrode [Hemisphere: $F(1,22) = 14.39$, $p < 0.01$; Hemisphere x Electrode: $F(5,110) = 3.39$, $p = 0.05$]. This last result is interesting because the maximum overall hemispheric difference was not at Frontal or Anterior Temporal sites but over Central sites. It is tempting to suggest that this different peak may be related to the processing task in common between the two sentence types: the reactivation of the (distant) subject of main clause verb. This temptation becomes stronger given that the scalp location of the hypothesized effect is consistent with the work of others who have studied verbal tasks that may have taxed this aspect of working memory (e.g., Lang, Lang, Uhl, Kornhuber, Deecke, & Kornhuber, 1987). These effects and related studies will be further elaborated in the Discussion below.

Moving past the main verb to the following article, the general negativity over the left hemisphere, which is more properly termed left anterior continues, as indicated by the reliable main effect of Hemisphere and the interaction of Hemisphere and Electrode, respectively [Hemisphere: $F(1,22) = 23.71$, $p < 0.001$; Hemisphere x Electrode: $F(5,110) = 3.45$, $p < 0.05$]. As with the last article comparison, these effects would seem to be due to the development of the N400-700 for the closed class items involved. There is also a reliable interaction of Comprehension Skill x Sentence Type in this window [$F(1,22) = 5.22$, $p < 0.05$], caused by Good comprehenders showing more overall negativity to the article in the SO sentence, and Poor Comprehenders showing the opposite pattern. Taken by itself, the single-word pattern here is puzzling, but the multiword ERPs suggest that the effect is due to the previously noted difference in the slower cortical potentials underlying the single-word ERPs rather than a local one.

These single word results give us a local view of the processing that happens during reading, which in turn emphasizes the processes by which single words are categorized, processed semantically, and integrated into the ongoing context. The process of categorization in this case is probably best seen in the differences between open- and closed-class items in their generation of lexically determined (left) frontal negativity, while both the N400 and the LAN seem to index processes related to semantic and thematic integration, respectively. Further, it is in these more integrative processes that we see the greatest differences between Good and Poor comprehenders, with Good comprehenders being the more successful integrators as defined by, for example, N400 effects to individual words.

DISCUSSION

One of the most striking aspects of this study was the degree to which the ERP effects in single-word and multiple-word comparisons differed from each other, demonstrating that a complete understanding of sentence-processing phenomena requires analyses at multiple time scales. Results from the single-word analysis would generally suggest that, like reading time data, the differences in processing SO and SS sentences were essentially localized to the verbs, with a smaller effect at the beginning of the relative clause. Specifically, the ERPs to the verbs in SO sentences showed more prolonged negativity over left anterior regions of the scalp than those to ERPs to verbs in the SS sentences; this effect was similar to the left anterior negativity (LAN) reported by Kluender and Kutas (1993) as a sign of working memory load. A potentially informative feature of this LAN effect is that it was significantly larger for Poor than Good comprehenders at the relative clause verb, but not at the main clause verb. Indeed, Poor comprehenders appeared to have special difficulties with the processing of the relative clause in the SO sentences. In addition to showing a larger LAN effect at the relative clause verb, they also showed a distinct N400-like response to the definite article beginning the SO relative clause, which was essentially absent in Good Comprehenders, possibly because Poor comprehenders were discounting the possibility that the SO structure could occur.

Although most of the single-word ERP results might lead one to believe that the overall processing of these sentence types was similar except at those locations where linguistic theory suggests they should differ, our multiword ERP data show that this is definitely not the complete picture, at least for Good comprehenders. Overall, ERPs to both SS sentences and simple declarative fillers showed a very slow positive drift over frontal electrode locations. This pattern contrasts with that for SO sentences, which were characterized by a greater relative negativity at frontal Electrode Sites. Not coincidentally, the point of electrophysiological divergence (i.e., a greater positivity for SS than for SO sentences) is the point at which the relative clause subject is added to working memory. At this time scale, in contrast to the single-word results, it is the Good comprehenders who show reliable and robust ERP differences between the two sentence types.

In the remainder of this discussion, we show how this pattern of effects is consistent with the predictions of the working memory-based model outlined in the introduction, and with the results of previous experiments.

Working Memory and Single Word Effects

Although ERPs to single words are unlikely to show the full extent of working memory involvement in sentence processing, we begin at this level because it is the common point of departure for both ERP and conventional reading time data. Indeed, at this timescale, ERP data and reading time data often mirror each other closely. Thus, for example, most of our ERP effects occur at the same sentence locations where reading time effects are found (Ford, 1983; King & Just, 1991). Single word ERPs, like reading times, are relatively poor indicators of storage effects (but see Wanner & Maratsos, 1978).

Perhaps the most likely site of working memory effects in single-word data is in the responses to words that naturally set up expectations for future storage based on lexical or (local) syntactic factors. Most function words (e.g., articles, prepositions, and complementizers) in English fulfill just these qualifications, introducing new syntactic units whose processing may require additional temporary storage (e.g., Frazier & Fodor, 1978). Thus, to the extent that this expectation of a working memory load has any electrophysiological consequences, ERPs to function words should differ from those to content words (e.g., nouns, verbs, adjectives, and -ly adverbs). And, in fact, they do; ERPs to function words tend to have a slow, (left) frontal negativity that starts at roughly 400 msec postword onset and lasts at least 300 msec (e.g., Neville, Mills, & Lawson, 1992). We too observed this effect in the present experiment when comparing the ERPs to articles versus those to adverbs. In some respects, these differences in the N400-700 resemble those found in more classical (and nonlinguistic) contingent negative variation (CNV) paradigms discussed below, except that here their size is smaller and their time-course more compressed. The possible connection between this late "function word effect" and similar (although slower) nonlinguistic potentials cannot be proved at this time (see Van Petten & Kutas, 1991); however, one study has found that the amplitude of the CNV in matching tasks using language materials varied with reading skill in poorer (but not in good) readers (Segalowitz, Wagner, & Menna, 1992).

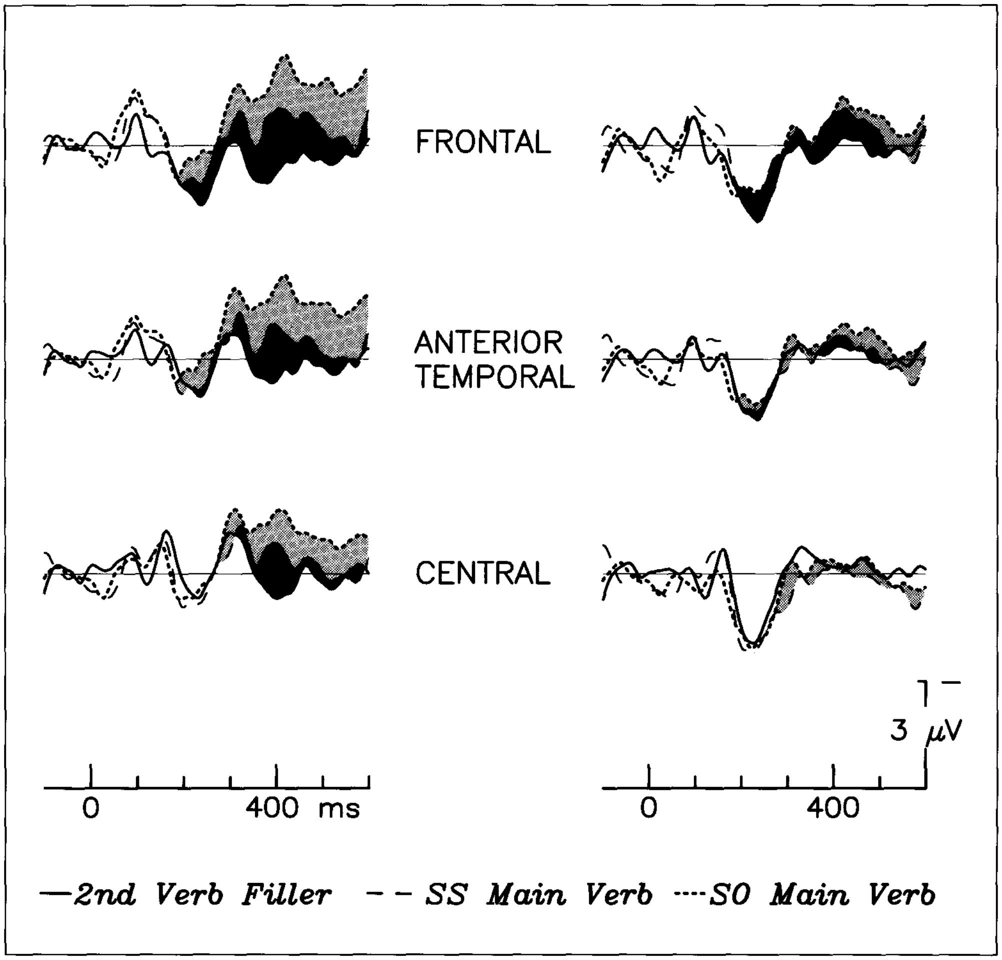
Examination of the computational and integrative aspects of working memory in single word ERP data is more promising. Unlike the process of storage, which is almost necessarily spread out in time, some aspects of language processing such as the assignment of thematic roles to actors intuitively seem to have a more punctual aspect to them. (Certainly, when a garden path sentence leads to a failure in role assignment, the point of failure is usually unambiguous and the "suddenness" of the failure seems compelling.) Our ERP data suggest that it is possible to see the effects of such thematic assignments, especially when the process of assignment is difficult in and of itself (such as when multiple assignments must be made) or when the assignment process must compete with the storage functions of working memory for cognitive resources. Thus, Figure 5 shows a greater LAN evoked by the main clause verbs in SO sentences relative to SS sentences. Both the distribution and timing of this LAN effect are similar to those reported by Kluender and Kutas (1993), even though in the present case they were elicited by open class words; this rules out explanations for the LAN that depend solely on word class. Moreover, these LAN effects were obtained during word-by-word reading for comprehension of a series of isolated, meaningful sentences, which suggests that the LAN effect does not rely on "abnormal" stimuli, unusual task demands, or extralinguistic monitoring processes.

One process that is common to both SS and SO sentences is the need to "reactivate" the head noun when the main verb is reached and requires a subject (and agent). It has been shown that some partially activated representation of the head noun is stored throughout the processing of the relative clause, and then activated fully when it can be assigned its thematic role (see, e.g., the 1989 review by Fodor). Thus, if the LAN indexes load in some fashion, the main verb in both SO and SS sentences should both show a LAN relative to some appropriate control. While such a control condition was not explicitly constructed in this study, a reasonable candidate among our filler items is the second transitive verb in sentences with two verbs. (In this experiment, such sentences involved simple forms of temporal subordination, coordination, or so-called purpose clauses, which did not induce large working memory loads.) Figure 6 shows that the ERP to the main clause verb in both SO and SS sentences was, in fact, characterized by a LAN relative to the control verbs as predicted.

Figure 6 also shows that both SS and SO main clause verbs also produce notably left-lateralized negativities relative to fillers over central sites as well. We previously suggested that this may reflect the fact that the agent in SO and SS sentence types is more distant from the site of its probable "reactivation" at the verb than in the filler case, and is, therefore, yet another working memory effect. While it may be dangerous to assume that ERP effects are generated by directly underlying cortex (Nunez, 1981), it may not be coincidental that these ERP effects are what would have been expected from a simple extrapolation of the data from two recent regional cerebral blood flow (rCBF) studies (Paulesu, Frith, & Frackowiak, 1993; Petrides, Alivisatos, Meyer, & Evans, 1993). Paulesu et al. (1993) suggested that Broca's area and the left supramarginal gyrus were involved in the “articulatory loop" component of Baddeley's working memory model, localizations that are only slightly posterior to where we observed the greatest negativities at the scalp. Our frontal maximum, however, is even closer to the site of greatest blood flow differences seen by Petrides et al. (1993) during the performance of two tasks requiring subjects to maintain an ordered list of numbers in working memory. Whether the list of numbers was ordered by the subject or the experimenter, increased blood flow was observed in mid-dorsolateral frontal cortex including Brodmann's area 46, which is heavily implicated in the functioning of working memory in nonhuman primates (Goldman-Rakic, 1992). These two results suggest that a PET study designed to examine the extent to which linguistic information guided the serially ordered regeneration of verbal information in working memory might find rCBF effects that would coincide with the "obvious" localization of both our LAN and the slightly more posterior activity. Of course, this is highly speculative, and requires more rigorous source localization of these ERP effects (which in turn requires more electrodes than we used in this study).

Whereas the LAN effect may reflect resource competition between integrative and reactivation processes, it may not tell us much about the outcome of this integrative effect. Traditionally, the N400 component has been taken as an index of the degree to which the current word is consistent with the contextual constraint provided by preceding words as a representation of the whole sentence is computed (e.g., Kutas, Van Petten, & Besson, 1988). According to this view, our materials

Figure 6. Grand Average ERPs (n = 24) from six most anterior Electrode Sites for the second verb in Filler, Subject Relative (SS), and Object Relative (SO) sentences. Dark shading emphasizes the differences between Filler verbs and verbs from the two experimental conditions, while fight shading indicates the relative left anterior negativity of SO compared to SS main clause verbs.

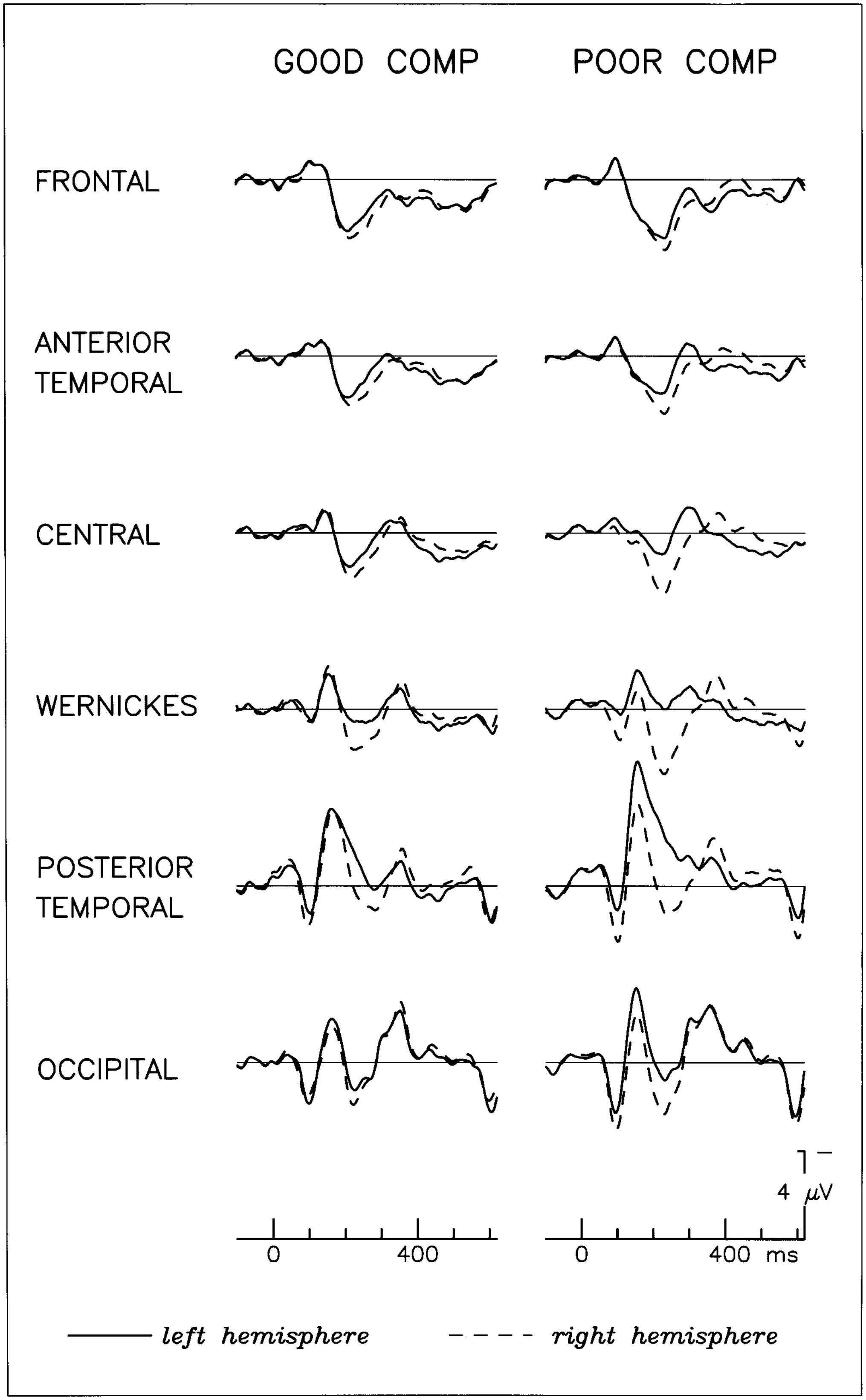


should have generated relatively small N400s inasmuch as successive words were consistent with and could be integrated into previous context. In fact, there was not even a hint that the N400 was larger in SO than SS sentences at any open class word position. Thus, there is no evidence in our data that subjects were led down a syntactic garden path in the SO sentences. Had this been the case, we would have expected to find a larger N400 for the SO word at some critical sentence location (probably the main clause verb) as was observed by Garnsey et al. (1989).

Whereas there was no syntactic garden path, our sentence types clearly differed in cloze probability at least at one point. The SO continuation of a sentence fragment like "The reporter who" is less frequent than an SS continuation (Fox & Thompson, 1990), and an N400 effect observed here might reflect this difference in cloze probability. The fact that the N400 to the definite article in SO sentences is larger for Poor comprehenders suggests either that their `baseline" expectation for this structure is lower, or that they discount the possibility of such a continuation. As discussed in the Introduction, one reason for discounting this continuation would be a limit in the available resources in working memory. However, if a reader completely rejects the SO possibility, then the resulting sentence fragment would be ungrammatical. Based on several studies, this ungrammaticality would be manifest in a P600 (which we did not observe) rather than an N400 effect (e.g., Osterhout & Holcomb, 1992; Hagoort, Brown, & Groothusen, 1993). Finally, note that the larger N400 in this case develops although the eliciting (closed class) word adds no semantic content of its own, but instead merely implies a continuation that is unexpected in context.

Another aspect of working memory to consider briefly is resource usage and allocation, although such effects may be more likely to develop over the course of several words. A prominent ERP component that has been linked to changes in resource allocation (in the form of shifting attention) is the N1-P2 complex. In general, N1 and sometimes P2 amplitudes are greater for attended stimuli than for unattended stimuli in both the auditory (Hillyard, 1985) and visual (Mangun & Hillyard, 1991) modalities, and the former appears to index the amount of resources devoted to processing an information channel that can be selected on the basis of spatial information. Figure 7 shows that there are substantial variations in the size of the N1-P2 complex linked to overall levels of comprehension-Poor comprehenders show larger posterior N1-P2 amplitudes than Good

Figure 7. Average ERPs to mid-sentence filler nouns for Good (n = 12) and Poor (n = 12) comprehenders at six different electrode locations over both the left and right hemispheres.



comprehenders in most cases.[1] This could be interpreted as consistent with a fairly conventional hypothesis (Hunt, Lunneborg, & Lewis, 1975; Perfetti & Lesgold, 1977) that Poor comprehenders allocate more attention to visual processing or word encoding than do Good comprehenders. This kind of N1-P2 effect has not been widely investigated among normal readers, although Raney (1993) did show that N1-P2 amplitude to tones in secondary task (tone detection) could be used as an indirect measure of cognitive load during reading. Specifically, he found that the Nl-P2 evoked by the tones increased upon a second reading of the same passages, consistent with the view that more cognitive resources were available to be devoted to the secondary tone detection task as the processing demands of word recognition decreased with rereading. While probe tasks such as this are a traditional way to obtain transient measures of continuous processing, in this work we have access to more direct indices of sustained processing in our slow potential data, to which we now turn.

Working Memory and Slow Cortical Potentials

The connection between scalp-negative slow cortical potentials and working memory seems to have been made first in monkeys performing simple delayed response tasks (reviewed by, e.g., Goldman-Rakic, 1987; Fuster, 1989). Slow negative shifts over the frontal cortex were related to the activity of single cells in dorsolateral prefrontal cortex of the macaque that maintained their firing as long as information was held in working memory, as well as to scalp-negative slow potentials that fill the delay interval in delayed response tasks. Although our current paradigm is radically different, we find some evidence for an analogous frontal negativity reflected in the ERP to the first two words of SO relative clauses. This negativity has a duration of over 1 sec and is bilaterally distributed over frontal and central sites, consistent with the localization of verbal working memory processes based on rCBF measures (Petrides et al., 1993; Paulesu et al., 1993). Here, the difference could be reflecting the fact that the head noun must be stored in working memory, in contrast to the SS sentences, wherein the head noun is almost immediately assigned its appropriate thematic roles.

More striking are the noticeable differences between Good and Poor comprehenders in the size and distribution of slow effects after this point. Good comprehenders show large effects of sentence type especially at frontal sites. By contrast, Poor comprehenders show a rather small frontal effect of sentence type but a prominent posterior standing negativity during the processing of both SS and SO sentences. In classical delayed response paradigms, the amplitude of the relative negativity is also correlated with performance. Moreover, the overall pattern of slow potentials for Poor comprehenders is also reminiscent of that reported by Lang, Starr, Lang, Lindinger, and Deecke (1992) on the effects of stimulus modality and mnemonic strategy on negative DC potentials in a short-term memory task. In the study of Lang et al., auditory memory stimuli elicited larger and more sustained frontal negativities than did visual memory stimuli, while the visual stimuli elicited larger posterior negativities. Moreover, subjects who used subvocalization strategies showed larger auditory stimulus effects while subjects who used imagery strategies showed larger visual stimulus effects; both strategies also altered the ERPs to stimuli in the "other" modality. We have previously suggested that Poor comprehenders might allocate more resources to visual word decoding, which would explain their standing posterior negativities for both sentence types. But, it is also often assumed that poorer readers are more likely to subvocalize while reading, which could be reflected in the larger sustained frontal negativities observed for both sentence types in these subjects. The allocation of resources to both these lower-level processes, naturally, would leave poorer readers with fewer resources to devote to higher order aspects of the task and result in reduced comprehension. The latter effect is another potential source of the large posterior N400s recorded in poorer readers (cf. Neville et al., 1993).

While the sustained negativity of SO sentences relative to SS sentences may in part be related to differences in storage requirements (along with resource allocation), it is the variations in this standing negativity that indicate where and when computationally intensive processes such as multiple thematic role assignment occur. The grand average ERPs (see Fig. 1) indicate that the anterior negativity, which we have linked to working memory storage, disappears over frontal sites just before the SO relative clause gap is reached; the resulting positivity is slightly larger over the left hemisphere. Apparently, there is a substantial left anterior positivity at this site that then develops into the LAN enhancement seen on the following words.

Given its timing, localization, and larger amplitude in Good comprehenders it is tempting to claim this slow positive wave indexes a particular computational function. Indeed, slow waves (SW) are generally taken to reflect additional processing instigated by perceptually or conceptually difficult operations (Ruchkin, Johnson, Mahaffey, & Sutton, 1988), although reports of frontally-positive SWs are rare, and none are lateralized. This positive frontal SW bears some resemblance to one observed by Karis, Fabiani, and Donchin (1984) for subjects attaining superior free recall scores presumably due to the use of elaborative encoding strategies. While the similarity between this elaborative mnemonic process and thematic role assignment may be accidental, both involve the processing of serial order relations between list items. It is this selfsame process that Petrides et al. (1993) presumed to be the cause of a significant increase in rCBF in left anterior cingulate cortex. Because the cingulate cortex is on the mesial side of the hemisphere, this localization raises the possibility of "paradoxical lateralization" wherein a negativity at the cortical surface appears over the right hemisphere while a complementary positivity appears over the left.

This hypothesis is also consistent with the results of Grossman, Crino, Reivich, Stern, and Hurtig (1992), who assessed sentence processing deficits in patients with Parkinson's disease (PD). They found that not only did PD patients do poorly in a series of linguistic tasks (including the parsing of center-embedded sentences similar to the SO stimuli in the present study), but that what little success they did have correlated strongly with rCBF in the left anterior cingulate. These language deficits in PD apparently stem from disease-related changes in the dopaminergic innervation of cerebral cortex. Dopamine has long been hypothesized to play a key role in the functioning of working memory circuits in (monkey) prefrontal cortex (Sawaguchi & Goldman-Rakic, 1991). Further, the sheer duration of the slower cortical potentials observed in the present experiment is better-matched by the effective lifetime of modulatory neurotransmitters such as dopamine in cortex than they are by the short-lived postsynaptic potentials believed to be responsible for most ERP effects (for information on modulatory neurotransmitters, see Foote & Morrison, 1987). Other factors such as the effects of continued neuronal firing, intracellular potassium, or glial polarization also may contribute (Rokstroh et al., 1989).

Thus far, we have not yet addressed one of the more prominent features of our data, namely the very slow positive drift over frontal electrode sites. This positivity is correlated with the ease of processing and is quite pronounced in Good comprehenders and virtually absent in Poor comprehenders (for our critical materials). Again, the SCP literature does not describe many frontal positivities, but those that exist may be relevant. For example, paralleling the distinction between SO and SS sentences in our study, Uhl et al. (1990) found that a difficult (proactive interference) paired associate task was associated with greater bilateral negativity over frontal sites than was an easier ("release") task. Another potentially relevant study is that of Asenbaum, Lang, Egkher, Lindinger, and Deecke (1992), who recorded a slow frontal drift with a slight right hemisphere dominance during an auditory selective attention task. Landwehrmeyer, Gerling, and Wallesch (1990) showed more overall right hemisphere negativity for dyslexics compared to left hemisphere negativity in controls in a slow brain potential study using various language tasks.

This section has tried to show how patterns of slower cortical potentials elicited during language processing may be explained in part by processes involved with more general cognitive processes such as working memory. To illustrate this point, we have referred to several SCP studies of cognition, very few of which were explicitly designed to study linguistic processes. Clearly, it would be useful to have more SCP studies of language processing to relate our findings to.

FUTURE DIRECTIONS AND CONCLUSIONS

The most obvious extension of this research involves the investigation not only of other structures that are difficult to process but also structures that are relatively easy to process. Moreover, the patterns of individual differences seen in this and other studies strongly suggest that we have much to learn about how fast and slow cortical potentials interact in populations more widely varied than college sophomores, such as children, normal elderly adults, and those suffering from diseases believed to have indirect effects on language processing, including Parkinson's disease (e.g., Grossman et al., 1992) and senile dementia of the Alzheimer's type (e.g., Murdoch, Chenery, Wilks, & Boyle, 1987).

In this study, we have attempted to explicate a number of ERP effects on sentence processing in terms of basic cognitive processes involving working memory. Further, we have shown that the ERPs of Good and Poor comprehenders differ in potentially revealing ways, and indicated how these could arise from differences in more basic cognitive processes, which evolve over time. While we can make substantial progress in this research program working at a fairly "cognitive" level, eventually it will be necessary to understand patterns of fast and slow ERP effects, and the factors that modulate their functioning, in terms of their neural generators. This, of course, is an immense task, but one that may help us understand what besides language makes us uniquely human.

METHOD

Subjects

Twenty-four UCSD students (12 women) between 18 and 27 years of age participated in the study, receiving $5.00 an hour for their time. All subjects were right-handed monolingual native English speakers who had no history of reading difficulties or neurological disorders.

Materials

The critical materials were 36 examples each of SS and SO sentences illustrated above, combined with 216 filler trials of various syntactic structures. In the construction of the materials, it was necessary to use the same 36 verbs in the relative clause of both the SS and SO sentences due to a shortage of verbs in English that were appropriate both for these syntactic structures and for use in ERP research. Each verb thus occurred once in an SS and once in an SO sentence. Critical trials were pseudorandomly mixed with the fillers with the constraint that no two critical items occurred consecutively.

True-false comprehension probes were constructed for 16 of the 36 stimulus sentences in the two critical conditions by using all plausible combinations of nouns and verbs in the stimulus sentences. Thus, for the SO sentence (2a), the possible probes (and correct answers) were

The reporter attacked the senator. (False)
The senator attacked the reporter. (True)
The reporter admitted the error. (True)
The senator admitted the error. (False)

For SS sentences like (2b), the probes were identical, while the correct responses to the first two probes were reversed. Four examples of each of these four probe types were used for both SO and SS sentences. A wider variety of true-false probes were constructed to follow half of the filler trials so that subjects would be less likely to use artificial comprehension strategies intended to improve performance on the critical materials. As detailed below, a median split on comprehension scores for critical items was used to categorize individual subjects as either "Good" or "Poor" comprehenders so that we could investigate what patterns of ERP components best predicted eventual comprehension.

Experimental Procedure

All 288 stimulus sentences were presented to subjects one word at a time in the center of a CRT while their electroencephalogram (EEG) was being recorded. Words were presented for a duration of 200 msec with a stimulus-onset asynchrony of 500 msec. Subjects were instructed to read the sentences normally, and warned that they would have to respond to a true-false comprehension question after slightly less than half of the sentences. The comprehension probe appeared centered on the screen 1500 sec after the onset of the final word. Subjects responded TRUE or FALSE by pressing one of two buttons held in either hand (the assignment of buttons to hands was counterbalanced across subjects). Following trials with no real comprehension probe, the direction "Press either button to continue" appeared on the screen. Subjects were directed to favor accuracy over speed in their responses. There was a total of 6.8 sec between experimental sentences including the time required for responses to the comprehension probes.

EEG Recording Parameters

ERPs were recorded from six pairs of lateral electrodes on an Electro-Cap and the left and right mastoids. All electrodes were referenced to a noncephalic lead derived from a pair of electrodes placed at the stemo-clavicular junction and on top of the seventh cervical vertebra; these leads were fed through a potentiometer adjusted to eliminate cardiac artifact. The Electrode Sites used included both standard 10-20 sites (F7, F8, T5, T6, 01, 02), and three pairs of electrodes approximately over Broca's area, Wernicke's area, and primary auditory cortex. [These six pairs of leads will be referred to as Frontal (F7 and F8), Anterior Temporal (Broca), Central (Area 41), Parietal (Wernicke's), Posterior Temporal (T5 and T6), and Occipital (O1 and 02) henceforth.]

Subjects” EEG was digitized on-line with a sampling rate of 250 Hz and stored for analysis. The amplifiers were set with half-amplitude cutoffs of 0.01 and 100 Hz, yielding a time constant of approximately 8 sec. Horizontal eye movement and blink artifacts were detected using electrodes placed at the outer canthi and under the right eye, respectively; trials with artifacts were rejected off-line prior to averaging. Overall, approximately 25% of all trials were discarded in forming the multiple word averages, while between 5 and 10% of the trials were discarded in forming single word ERPs.

Recording Epochs

Artifact-free EEG was averaged both over a sentence-length epoch of the critical materials and over individual words as described below. Sentence-length ERPs lasted 5000 msec with the ERP to the first noun in the critical sentence being used as a 500 msec "prestimulus" baseline. (The first noun was chosen as the baseline so that the most linguistically crucial parts of the sentences lay entirely within the window of analysis while minimizing the amount of EEG that had to be rejected due to artifacts.) Data for the sentence-length ERPs were decimated to an effective sampling rate of 125 Hz to facilitate averaging, but this poses no limitation to our analysis of slower activity. ERPs for each word of the critical sentences were also obtained, with the 100 msec preceding stimulus word onset serving as the baseline.

Acknowledgments

Work reported in this paper was supported by NICHD Grant HD22614 and by NIA Grant AG08313 to Marta Kutas. Jonathan W King was also supported in part by a McDonnell-Pew postdoctoral fellowship from the San Diego Center for Cognitive Neuroscience, funds from the Center for Human Information Processing (CHIP; NIMH Grant MH-14268), and funds from the Center for Research on Language (CRL; NIH Grant T32 D000041-01) at UCSD. The authors wish to thank Robert Kluender for useful discussions of earlier drafts of this work.

Reprint requests should be sent to Jonathan W King, Department of Cognitive Science 0515, University of California, San Diego, 9500 Gilman Drive. La Jolla, CA 92093-0515, or via E-mail to: king@cogsci.ucsd.edu

Note

1. Midsentence nouns were chosen for this comparison because they are numerous and not associated with significant late effects in this study as are, e.g., verbs.

REFERENCES

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

Anderson, J. R. (1983). The Architecture of Cognition. Cambridge, MA: Harvard University Press.

Asenbaum, S., Lang, W, Egkher, A., Lindinger, G., & Deecke, L. (1992). Frontal DC potentials in auditory selective attention. *Electroencephalography and Clinical Neurophysiology, 82,* 469-476.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In J. T. Spence (Ed.), *The psychology of learning and motivation Advances in research and theory*. New York: Academic Press.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

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

Clifton, C. Jr., & Ferreira, E (1989). Ambiguity in context. *Language and Cognitive Processes,* 4, S177-103.

Daneman, M., & Carpenter, P A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning* & Verbal *Behavior,* 19, 450-466.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1-47.

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

Fischler, L, & Raney, G. E. (1991). Language by eye: Behavioral and psychophysiological approaches to reading. In M. G. H. Coles (Ed.), *Handbook of cognitive psychophysiology: Central and autonomic nervous system approaches.* New York: Wiley.

Foote, S. L., & Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annual Review of Neuroscience,* 10, 67-95.

Ford, M. (1983). A method of obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior,* 22, 203-218.

Foss, D. J., & Cairns, H. S. (1970). Some effects of memory limitation upon sentence comprehension and recall. *Journal of Verbal Learning and Verbal Behavior,* 9, 541-547.

Fox, B. A., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language,* 66, 297-316.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Ph.D. dissertation, distributed by the Indiana Linguistics Club.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition,* 6, 291-325.

Fuster, J. M. (1989). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe.* New York: Raven Press.

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

Gibson, E. (1990). Recency preference and garden-path effects. In *Program of the twelfth annual conference of the Cognitive Science Society (pp.* 372-379). Hillsdale, NJ: Lawrence Erlbaum.

Goldman-Rakic, E S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In V Mountcastle (Ed.), *Handbook of physiology: The nervous system. (pp.* 373-417). Bethesda, MD: American Physiological Society.

Goldman-Rakic, E S. (1992). Working memory and the mind. *Scientific American,* 267, 111-117.

Grossman, M., Crino, E, Reivich, M., Stern, M. B., & Hurtig, H. I. (1992). Attention and sentence processing deficits in Parkinson's disease: The role of anterior cingulate cortex. *Cerebral Cortex,* 2, 513-525.

Hagoort, P, Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439-483.

Hillyard, S. A. (1985). Electrophysiology of human selective attention. *Trends in Neurosciences,* 8, 400-405.

Hunt, E. B., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology,* 2, 194-227.

Just, M. A., & Carpenter, E A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Karis, D., Fabiani, M., & Donchin, E. (1984). P300 and memory: Individual differences and the Von Restorff effect. *Cognitive Psychology,* 16, 177-216.

Kimball, J. P (1973). Seven principles of surface structure parsing in natural language. *Cognition,* 2, 15-47.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing. The role of working memory. *Journal of Memory and Language,* 30, 580-602.

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

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

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

Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. In M. G. H. Coles (Ed.), *Advances in psychophysiology*. Greenwich, CT: JAI Press.

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology,* 69, 218-233.

Landwehrmeyer, B., Gerling, J., & Wallesch, C. W (1990). Patterns of task-related slow brain potentials in dyslexia. *Archives of Neurology,* 47, 791-797.

Lang, M., Land, W, Uhl, E, Kornhuber, A., Deecke, L., & Kornhunber, H. H. (1987). Slow negative potential shifts indicating verbal cognitive learning in a concept formation task. *Human Neurobiology,* 6 183-190.

Lang, W, Starr, A., Lang, V, Lindinger, G., & Deecke, L. (1992). Cortical DC potential shifts in accompanying auditory and visual short-term memory. *Electroencephalography and Clinical Neurophysiology, 82,* 285-295.

Larkin, W, & Burns, D. (1977). Sentence comprehension and memory for embedded structure. *Memory and Cognition, 5,* 17-22.

Mangun, G. R., & Hillyard, S. A. (1991). Modulation of sensory-evoked brain potentials indicate changes in perceptual processing during spatial priming. *Journal of Experimental Psychology: Human Perception and Performance,* 17, 1057-1074.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology. (Vol.* 2, pp. 419-492). New York: John Wiley

Murdoch, B. E., Chenery, H. J., Wilks, V, & Boyle, R. S. (1987). Language disorders in dementia of the Alzheimer type. *Brain and Language,* 31, 122-137.

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

Neville, H. J., Coffey, S. A., Holcomb, P. J., & Tallal, P. (1993). The neurobiology of sensory and language processing in language-impaired children. *Journal of Cognitive Neuroscience, 5, 235-253.*

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

Nunez, P L. (1981). *Electric fields of the brain.* New York: Oxford University Press.

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

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature (London), 362, 342-345.*

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P A. Carpenter (Eds.), *Cognitive processes in comprehension. (pp. 141-184).* Hillsdale, NJ: Erlbaum.

Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences USA.*, 90, 878-882.

Raney, G. E. (1993). Monitoring changes in cognitive load during reading: An event-related brain potential and reaction time analysis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19, 51-69.*

Rockstroh, B., Elbert, T, Canavan, A., Lutzenberger, W, and Birbaumer, N. (1989). *Slow Cortical Potentials and Behavior.* Baltimore, MD: Urban and Schwartzenberg.

Ruchkin, D. S., Johnson, R., Canoune, H., & Ritter, W (1990). Short-term memory storage and retention: An event related brain potential study. *Electroencepbalography and Clinical Neurophysiology*, 76, 419-439.

Ruchkin, D. S., Johnson, R., Mahaffey, D., & Sutton, S. (1988). Towards a functional categorization of slow waves. *Psychophysiology, 25, 339-353.*

Savin, H. B., & Perchonok, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 4, 348-353.

Sawaguchi, T, & Goldman-Rakic, P. S. (1991). D1 dopamine receptors in prefrontal cortex: Involvement in working memory. *Science, 251, 947-950.*

Segalowitz, S. J., Wagner, W J., & Menna, R. (1992). Lateral versus frontal ERP predictors of reading skill. *Brain and Language*, 20, 80-103.

Steedman, M., & Altmann, G. (1989). Ambiguity in context: A reply. *Language and Cognitive Processes*, 4, 105-122.

St. John, M. E, & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 97, 332-361.

Uhl, E, Franzen, P., Series, W, Lang, W, Lindinger, G., & Deecke, L. (1990). Anterior frontal cortex and the effect of proactive interference in paired associate learning: A DC potential study. *Journal of Cognitive Neuroscience*, 2, 373-382.

Vallar, G., & Baddeley, A. D. *(1984).* Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, 23, 151-161.

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

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In G. A. Miller (Ed.), *Linguistic theory and psychological reality*. (pp. 119-160). Cambridge, MA: MIT Press.