

Sidestepping Garden Paths: Assessing the Contributions of Syntax, Semantics and Plausibility in Resolving Ambiguities

Weijia Ni

Haskins Laboratories, New Haven, Connecticut, USA

Stephen Crain

*Haskins Laboratories, New Haven, Connecticut and University of
Maryland, College Park, Maryland, USA*

Donald Shankweiler

*Haskins Laboratories, New Haven, Connecticut and University of
Connecticut, Storrs, Connecticut, USA*

A central issue in the study of sentence processing is the manner in which various sources of information are used in resolving structural ambiguities. According to one proposal, the garden path model (e.g. Frazier & Rayner, 1982), perceivers are initially guided by strategies based solely on the structural properties of sentences. Another class of models, constraint satisfaction models, emphasise the influence of lexical properties in decisions among the alternative analyses of an ambiguous sentence fragment (e.g. Tanenhaus, Garnsey, & Boland, 1991). In this paper, we explore the prediction of an alternative model, the referential theory (e.g. Crain & Steedman, 1985).

Requests for reprints should be addressed to Weijia Ni, Haskins Laboratories, 270 Crown Street, New Haven, CT 06510, USA. E-mail: NI@HASKINS.YALE.EDU.

The research reported in this paper was supported in part by a Project Program HD-01994 from the National Institute of Child and Human Development to Haskins Laboratories and in part from a grant to Stephen Crain and Donald Shankweiler from the University of Connecticut Research Foundation. Portions of this research were presented at meetings of the Annual CUNY Conference on Sentence Processing: Rochester, NY, 1991; Amherst, MA, 1993; and New York, NY, 1994. We would like to express our gratitude to Michael Tanenhaus and John Trueswell for providing us with operating software and instructions in conducting eye movement research. Thanks are also due to Slavoljub Milekić for his assistance in setting up the eye movement laboratory, and to Meredith Daneman for her courtesy in providing the materials for the memory-span test. Finally, we are indebted to Gerry Altmann, Don Mitchell and an anonymous reviewer for many insightful comments on earlier drafts of the manuscript. Any errors that remain are our own.

The referential theory maintains that the relative complexity of discourse representations plays a key role in determining the perceiver's immediate parsing preferences. We present four experiments designed to weigh the influence of semantic/referential complexity and general world knowledge in the on-line resolution of two kinds of structurally ambiguous sentences. In each experiment, we examined pairs of sentences that were identical except for the alternation between the definite determiner THE and the focus operator ONLY. Two techniques were used to assess ambiguity resolution: word-by-word reading and eye movement recording. The results indicate that semantic/referential principles are applied immediately in on-line ambiguity resolution and that these principles pre-empt general world knowledge. The use of world knowledge was found to depend on working memory capacity, whereas the resolution of ambiguity by means of semantic/referential principles appeared to be independent of memory resource. Taken together, the findings are interpreted as support for the referential theory of ambiguity resolution.

INTRODUCTION

One of the central goals of psycholinguistic research is to provide a systematic account of how people interpret structurally ambiguous sentences. This paper presents the results of an interlocking set of experiments that shed further light on the bases of ambiguity resolution. Two experiments focus on the kind of structural ambiguity exhibited by Bever's (1970) well-known "garden path" sentence *The horse raced past the barn fell*. In this sentence, the verb *raced* is morphologically ambiguous: it can be analysed as a simple past tense main verb or as a past-participle. It is striking that most people find the sentence extremely difficult to process at the verb *fell*, suggesting an overwhelming preference to analyse *raced* as a main verb: this decision leads to a "garden path effect". Two additional experiments investigate a more subtle garden path effect,¹ the momentary semantic anomaly that is manifested in sentences like *The spy saw the cop with a revolver*. As Rayner, Carlson and Frazier (1983) have demonstrated, the anomaly arises because people prefer to analyse the prepositional phrase *with ...* to modify the verb *saw*, rather than the immediately preceding noun phrase *the cop*. Evidence of a preference for verbal attachment in constructing a structural representation was found by comparing eye movement patterns during reading in sentences such as *The spy saw the cop with a revolver* and sentences such as *The spy saw the cop with binoculars*. Apparently, an elevation in eye fixation durations occurs when people

¹We use the term "garden path effect" quite broadly, to refer to cases in which the attachment of a linguistic item in an on-going structural analysis results in a structure that is incompatible with later input, and also as a description of cases in which an attachment requires reinterpretation.

encounter the noun phrase *a revolver*. Because the prepositional phrase is attached to the verb, this noun phrase makes the sentence anomalous. By contrast, the noun phrase *binoculars* is a plausible continuation of the sentence fragment *The spy saw the cop with . . .*

The existence of garden path effects and semantic anomalies is evidence that the human sentence processing mechanism (the parser) makes rapid decisions about which alternative, grammatically well-formed structural representation to adopt when the input is ambiguous. The factors that influence the decision making of the parser and the manner in which the parser is influenced are matters of controversy, however. One account of ambiguity resolution is known as the "garden path model" (Clifton & Ferreira, 1989; Ferreira & Clifton, 1986; Frazier, 1979; Frazier & Rayner, 1982; Rayner et al., 1983). The garden path model maintains that the parser is a serial processing device. The model contends that the parser's initial analysis of ambiguity is based solely on structural properties of the linguistic input. One structurally based parsing strategy, Minimal Attachment, instructs the parser to pursue the analysis that postulates the fewest non-terminal nodes in constructing the phrase structure representations of a sentence. The model predicts that in the sentence *The horse raced past the barn fell*, the word *raced* is initially analysed as the main verb because this analysis is structurally simpler than the alternative analysis on which the verb *raced* is analysed as a past-participle. Consequently, the "real" main verb *fell*, which comes later, cannot be readily incorporated into the analysis. The main verb analysis thus leads the parser down a "garden path". Although the parser uses only structural information in making its initial decisions, according to the garden path model other sources of information contribute to reanalysis if the initial analysis turns out to be incompatible with subsequent linguistic material. For example, Ferreira and Clifton (1986) propose that a second stage "thematic processor" operates on the output of the syntactic component of parsing. When triggered by "an error signal in the disambiguating region" (Ferreira & Clifton, 1986, p. 366; our emphasis), the thematic processor supplies alternative thematic representations that may prove useful in structural reanalysis.

A recent version of the garden path model, by Mitchell, Corley and Garnham (1992), maintains that the parser may begin to revise a misresolved ambiguity even before the point of disambiguation. According to this account, the effects of syntactic parsing strategies such as Minimal Attachment persist only briefly, perhaps for no longer than a word or two. This allows the parser sufficient time to begin processing semantic and discourse information, which may then be used in the reanalysis of the sentence, if that is needed. Mitchell et al. (1992) point out that most of the early research on the influence of discourse representations on ambiguity resolution test for syntactic commitments two or more words after the onset

of ambiguity. According to Mitchell et al., such test points arrive too late to reveal the effects of parsing strategies because discourse factors may have had sufficient time to override the parser's initial syntactic commitment. It is therefore necessary to test for the parser's initial syntactic commitment at the "earliest feasible point" in the unfolding structural analysis assigned by the parser.² It is important to note that on all versions of the garden path model, non-structural sources of information can only serve as evidence confirming or disconfirming the parser's initial structurally based decision. Recent research motivated by this theoretical framework has therefore been concerned with the costs incurred by reanalysis, and with the diagnostics used by the parser in recovering from ambiguities that have been misresolved (Frazier, 1994; Frazier & Clifton, 1996).

While a number of studies provide support for one or another version of the garden path model (e.g. Britt, Perfetti, Garrod, & Rayner, 1992; Murray & Liversedge, 1994; Rayner, Garrod, & Perfetti, 1992), much recent work has led to alternative accounts of ambiguity resolution that explore the possibility that on-line decisions of the parser are affected by a range of non-structural factors. Adopting the terminology used in a recent review article by Tanenhaus and Trueswell (1995), we refer to one general line of research as the "constraint satisfaction model". Proponents of this approach have investigated the on-line influence of lexically based factors such as verb frequency, information from argument structure and conceptual-semantic information. Each of these factors is assumed to play a role in evaluating the alternative structural representations of ambiguous sentences (e.g. Boland, Tanenhaus, & Garnsey, 1990; Juliano & Tanenhaus, 1993, 1994; MacDonald, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994; Merlo, 1994; Spivey-Knowlton & Sedivy, 1995; Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1994; Taraban & McClelland, 1990; Trueswell & Tanenhaus, 1994; Trueswell, Tanenhaus, & Garnsey, 1994). Tanenhaus and Trueswell (1995) conclude, "When these [lexically based] factors are quantified and combined ... there is no need for either an initial category-based parsing stage or a separate revision stage". In a similar vein, other relevant factors have been identified, including intonation and prosody (e.g. Beach, 1991; Marslen-Wilson et al., 1992; Nagel & Shapiro, 1994; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Speer, Crowder, & Thomas, 1993)

²The empirical basis for Mitchell and co-workers' (1992) claims has been challenged by Altmann et al. (1994), who questioned both the effectiveness of the test materials and the proper interpretation of the data. Regardless of the outcome of this debate, however, it is important to heed the general point made by Mitchell et al., who emphasise the importance of testing for syntactic biases at the earliest point possible, so as to detect subtle, possibly unconscious garden path effects, if these exist. One way to accomplish this is to obtain a precise record of the parser's on-line operations, using measures that are sufficiently sensitive. We return to this point in the discussion of Experiment 2.

and the memory costs associated with the processing of various syntactic constructions (e.g. Gibson, in press; Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992).

So far we have identified the garden path model and the constraint satisfaction model. In this paper, we pursue the predictions of a third model. On this model, yet another class of factors is viewed as essential in the resolution of structural ambiguities, namely the referential properties of sentences. We call this model the "referential theory" (e.g. Altmann & Steedman, 1988; Crain & Steedman, 1985; Ni & Crain, 1990). According to the referential theory, the complexity of the alternative discourse representations (corresponding to the alternative structural analyses) is often crucial in the resolution of structural ambiguities. A wide variety of parsing preferences that have often been attributed to structural properties of sentences are viewed by the referential theory as consequences of the application of semantic/referential principles.

Both the constraint satisfaction model and the referential theory can be contrasted with the garden path model in certain respects. Each of the former typically maintains that the parser computes multiple (partial) structural analyses of an ambiguous phrase; the parser is regarded as a *parallel processing* mechanism. In addition, both models take the position that real-world knowledge is invoked to resolve some structural ambiguities. Crain and Steedman (1985, p. 330) have explicitly proposed and supported the claim that parsing decisions are influenced by considerations of general knowledge of the world: "If a reading is more plausible in terms either of general knowledge about the world or of specific knowledge about the universe of discourse, then, other things being equal, it will be favored over one that is not". While the use of information about the *a priori* plausibility of the alternative readings of an ambiguous sentence fragment is acknowledged, the referential theory maintains that principles of discourse pre-empt *a priori* plausibility. As Crain and Steedman put it: "in case of a conflict between general and specific knowledge, the latter must clearly take precedence" (op. cit.).

The research presented in this paper is designed to test these two specific tenets of the referential theory; that is, the claim that there is immediate application of semantic/referential principles in the resolution of ambiguity, and the claim that these principles pre-empt general knowledge of the world. A further goal of the present study was to examine the manner in which these two different sources of information are used by subjects with varying working memory capacities to resolve structural ambiguities and to recover from garden paths. To accomplish these goals, two experiments investigated main-verb/reduced-relative-clause ambiguity, and two investigated ambiguities involving the attachment of prepositional phrases. The experiments were conducted using two experimental techniques: (1)

self-paced, word-by-word reading and (2) eye movement recording.³ In the two experiments on main-verb/reduced-relative-clause ambiguities, the structural content of the test sentences informed the subjects whether or not they had been led down a garden path. In contrast, in the two experiments testing the preferences for attachment of prepositional phrases, the subjects were informed by *a priori* plausibility when an incorrect analysis had been pursued. The results of the latter experiments were analysed by dividing the subjects into groups according to individual differences in memory span (Daneman & Carpenter, 1980). Between-group comparisons enabled us to distinguish properties of sentence processing that are relatively undemanding of memory resources from properties that are highly sensitive to limitations in memory. The view has been put forward that people with high memory spans are better able to maintain parallel alternative structural analyses of an ambiguous sentence, and are better able to use a variety of sources of information in resolving ambiguities (MacDonald et al., 1992; Pearlmutter & MacDonald, 1995). The findings of our research are consistent with these claims. All our subjects appeared to be capable of applying semantic/referential principles on-line to resolve local ambiguities, but those with greater memory capacity also proved able to rapidly access real-world knowledge to recover from a misanalysis, whereas subjects with less memory capacity typically re-read portions of ambiguous sentences when the use of real-world knowledge was required to recover from a misanalysis.

The findings bear directly on the issues of how and when ambiguities are resolved and the costs that are incurred. As discussed, the referential theory leads to the expectation that decisions concerning specific discourse representations should be made earlier than decisions that are based on general world knowledge. In a resource-limited system, such as the human sentence processing mechanism, disruptions should occur most often at later stages of processing; if principles used to construct discourse representations are accessed earlier than information about the real world, then the former should be less likely than the latter to exhaust memory resources. We appeal to this assumption of the referential theory to explain why individuals with higher memory capacity appear to be better able to access real-world knowledge on-line in order to recover from a misanalysis.

Since this research was motivated by the referential theory, we think it appropriate to spell out its operating principles in greater detail in the next section. Then we describe the semantic properties of the focus operator ONLY, and explain the rationale for alternating ONLY and THE as

³Although the main findings of the two techniques are almost entirely complementary, the results from eye movement recording provide a more fine-grained view of processing difficulties within the ambiguous region of a garden path sentence.

pre-nominal modifiers in our experimental manipulations. We then report the experimental findings and, finally, provide a general discussion of the findings.

THE REFERENTIAL THEORY

The referential theory contends that primary responsibility for resolving structural ambiguities rests with the immediate, word-by-word evaluation of alternative structural analyses by the semantic/discourse processor. On this view, no particular structural configurations are intrinsically prone to elicit garden paths but, instead, certain discourse contexts either promote or deter garden path effects. The theory assumes a “weak” interaction between components of the language processing system. The syntactic processor putatively computes multiple (partial) structural analyses when it encounters an ambiguous sentence fragment. The alternative analyses are shunted to the semantic/discourse processor, which chooses among them. The following is a list of the basic tenets of the theory:

1. All permissible structural analyses of an ambiguous sentence are computed in parallel by the syntax. They are presented to the semantic/discourse processor for adjudication.
2. Semantic evaluation is carried out incrementally, more or less word by word.
3. The semantic/discourse processor evaluates and chooses among the alternative syntactic analyses on the basis of their fit to the conversational context.
4. If no decision is rendered by the semantic/discourse processor, then factors such as general knowledge of the world may be used to decide on the analysis to pursue.

Ambiguity resolution is ordinarily achieved within some discourse context. According to the referential theory, decisions by the parser follow what Crain and Steedman (1985, p. 331) have called the “principle of referential success”: “If there is a reading that succeeds in referring to entities already established in the perceiver’s mental model of the domain of discourse, then it is favored over one that does not”. Another version of this principle was formulated by Altmann and Steedman (1988), as the “principle of referential support”: “An NP analysis which is referentially supported will be favored over one that is not”.

The referential theory also explains how ambiguities are resolved in the absence of context. When processing a sentence in the so-called null context, the perceiver actively attempts to construct a mental representation of a situation that is consistent with the sentence. In addition to the characters and events *asserted* in a sentence, the construction of a mental model of the

situation sometimes requires the perceiver to represent information that a sentence *presupposes*. The process of augmenting one's mental model to represent the presuppositional content of sentences has been called "the accommodation of presuppositional failure" by Lewis (1979), "extending the context" by Stalnaker (1974) and Karttunen (1974), and the "addition of presuppositions to the conversational context of an utterance" by Soames (1982).

According to the referential theory, the accommodation of presuppositional failure plays a critical role in explaining how ambiguous sentences are interpreted outside of context. The parser attempts to construct all permissible discourse representations of a sentence but, due to limited computational resources, it settles on the analysis that requires the fewest modifications in establishing a coherent representation. Crain and Steedman (1985, p. 333) call this the "principle of parsimony": "If there is a reading that carries fewer unsatisfied but consistent presuppositions than any other, then that reading will be adopted and the presuppositions in question will be incorporated in the perceiver's mental model".

The principle of parsimony can explain why the sentence *The horse raced past the barn fell* produces a garden path effect in the absence of context: At the onset of the ambiguous phrase (i.e. at the verb *raced*), the parser must actively create a mental model of a discourse in which the sentence could felicitously occur. According to the principle of parsimony, when there is a choice between alternative analyses, the analysis that requires the fewest extensions to the mental model is favoured. The noun phrase *the horse* leads the parser to assume that a particular horse is in the domain of discourse. To make felicitous the alternative reduced relative clause analysis, the parser would have to establish a representation in which there is more than a single horse, with one of the horses being raced by someone. Because nothing in the fragment *The horse raced ...* demands such additions to the mental model of the discourse, the main verb analysis of *raced* is favoured, and the parser is led down the garden path.

Empirical support for the referential theory has come chiefly from studies that show the on-line influence of linguistic context on the resolution of structural ambiguities (Altmann & Steedman, 1988; Altmann, Garnham, & Dennis, 1992; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Spivey-Knowlton & Tanenhaus, 1994).⁴ However, much debate has centred on the

⁴Crain and Ni (1991) show how the referential theory applies to purely semantic ambiguities; that is, ambiguities that could not, in principle, be explained by structurally based criteria. Based on the observation that the principles of discourse and reference are independently motivated, and cover a range of phenomena not handled by structurally based models, Crain and Ni contend that the referential theory has an edge on these other models to the extent that the referential theory can also explain the garden path effects that occur in structural ambiguities of the sort discussed in the literature.

immediacy of such influences; that is, whether or not contextual information is available early enough to be effective in resolving local ambiguities (e.g. Altmann, Garnham, & Henstra, 1994; Crain & Steedman, 1985; Ferreira & Clifton, 1986; Fodor, Ni, Crain, & Shankweiler, 1996; Mitchell et al., 1992). It is apparent that the information represented in the mental model of an extended discourse may not be accessed immediately, and that the effectiveness of context on parsing may be differentially affected by characteristics of the contextual manipulations. A more rigorous comparison between the different accounts of ambiguity resolution can be made, therefore, if garden path effects are manipulated without providing an explicit discourse context. The experiments we report in this paper follow this research strategy. The test sentences differ only in a single respect: one version of each sentence contains the focus operator ONLY, and another version contains the definite determiner THE (e.g. *Only horses raced past the barn fell* versus *The horses raced past the barn fell*). Because the experiments involve *minimal pairs* of sentences, the influence of the referential properties of sentences is investigated in isolation from other factors: The experimental manipulations vary only the referential content of the test materials, while holding constant the effects of verb frequency, argument structure, semantic/conceptual information, and so on.

To sum up, by manipulating referential properties sentence-internally, the present experiments circumvent problems that have sometimes plagued studies attempting to investigate referential effects by manipulating discourse context. Because the referential contributions of the focus operator ONLY, and of the definite determiner THE, are essential to incremental semantic interpretation, the influence of these pre-nominal modifiers in the resolution of ambiguity is likely to be felt immediately. Indeed, we will demonstrate that garden path effects can either be instigated or deterred by substituting one pre-nominal modifier for another. Thus, with all other factors held constant, the present studies demonstrate robust effects of changes in referential content (which arise by substituting ONLY for THE in the experimental materials) on the decisions made by the parser when an ambiguity is encountered. These findings are predicted by the referential theory. In the following section, the rationale for using the focus operator ONLY in the experimental manipulations is explained more fully.

THE FOCUS OPERATOR ONLY

The semantic function of the focus operator ONLY is to signal that the denotation of a linguistic constituent, which we call the focus element, is being contrasted with a set of alternatives. Consider (1):

1. *In New Haven, only Willoughby's coffee is really good.*

It is appropriate to use the sentence in (1) only if coffee from Willoughby's is being compared with coffee from other shops in New Haven. If the speaker had sampled coffee from Willoughby's, but nowhere else, it would be infelicitous to utter (1). Note, however, that the use of sentence (1) does not *assert* that a comparison among coffee shops has been made; rather, this comparison is *presupposed* to have occurred prior to the utterance of (1). The presupposition that coffee from other shops has been sampled is triggered by the focus operator ONLY.

The semantic representation of sentences with the focus operator ONLY can be partitioned into three parts. One part represents background information, a second represents the element in focus and the third represents a contrast set—the alternatives to the focus element. The contrast set is not mentioned explicitly in the sentence; instead, it is presupposed to exist. Two conditions must be met for sentences with ONLY to be true: first, the information in the background must apply to the element in focus; second, the background information must *not* apply to any members of the contrast set. That is, the background must apply *uniquely* to the focus element.⁵

Based on semantic properties associated with ONLY, the referential theory predicts that sentences like (2) will not evoke garden path effects, but that ones like (3) will. These differences are expected despite the fact that (2) and (3) are identical following the initial noun phrase; in particular, the sentences are identical at the point of disambiguation and thereafter.

⁵Formally, the semantic value of the focus operator ONLY is captured by the following rule (adapted from Krifka, 1991; also see Jackendoff, 1972; Rooth, 1985):

MEANING RULE FOR ONLY:

$$B(F) \ \& \ \forall X[\{ X \in \text{CON}(F) \ \& \ B(X) \} \rightarrow X = F]$$

Where X is a variable of type F, and CON(F) is a set of contextually determined alternatives to F.

The first conjunct of the meaning rule, B(F), states that the background must apply to the focus element. The second conjunct is the statement of uniqueness: $\forall X[\{ X \in \text{CON}(F) \ \& \ B(X) \} \rightarrow X = F]$. Here, the universal quantifier ranges over a *metavariable*, X. By replacing the metavariable X with actual variables of different types, different interpretations may be derived, depending on the nature of the entities that are being contrasted with the focus element. This provides the flexibility to cope with alternative interpretations for sentences with ONLY. If the element in focus is an *individual*, then the contrast set contains individuals. In this case, the metavariable is replaced by an individual variable: x, y, etc. By contrast, if the focus element is a *property* of individuals, then the contrast set consists of sets of properties of individuals, rather than individuals themselves. In such cases, the metavariable is replaced by a variable of this type: P, Q, etc. The meaning rule ends by guaranteeing the *uniqueness* of the focus element—for each member of the contrast set, if the background applies to it, then that member is the focus element itself.

2. *Only businessmen loaned money at low interest were told to record their expenses.*
3. *The businessmen loaned money at low interest were told to record their expenses.*

According to the referential theory, the subject NP *only businessmen* in (2) causes the parser to establish a discourse representation (a mental model) of the conversational context in which a set of businessmen is represented. The pre-nominal modifier ONLY in the initial NP prompts the parser to search for a contrast set. If a contrast set has not been previously established in the discourse, the parser has two options. First, it could attempt to construct a contrast set "from scratch"; that is, the parser could conjure up some set of individuals to be contrasted with businessmen. There is a second option, however. Since the verb *loaned* is ambiguous, the parser could choose to satisfy the presupposition associated with ONLY by adopting the reduced relative clause analysis of the verb phrase. Pursuing this second option requires the parser to partition the set of businessmen already admitted into the mental model, rather than adding new entities. According to the principle of parsimony, the second option should therefore be preferred.

If a decision is made to analyse the ambiguous fragment as a reduced relative clause, no garden path effect will occur when the main verb (*were told*) is encountered. According to the referential theory, then, sentences like (2), which begin with the focus operator ONLY, should tend to pattern like sentences with an unambiguous verb, such as (4):

4. *The vans stolen from the parking lot were found in a back alley.*

There is a further prediction of the referential theory: If a contrast set is established *before* the ambiguity is encountered, then garden path effects should tend to emerge in sentences with ONLY. This is illustrated in (5):

5. *Only wealthy businessmen loaned money at low interest were told to record their expenses.*

The phrase *wealthy businessmen* in (5) satisfies the requirement of setting up a contrast set: the set of businessmen who are not wealthy. Having established the contrast set in advance of the ambiguity, the main verb analysis is more highly favoured (by the principle of parsimony, as discussed earlier). Adopting the main verb analysis results in a garden path effect, however, when the real main verb, *were told*, is encountered.

In the following sections, we report the results of four experimental studies that test the predictions of the referential theory, using both the self-paced, word-by-word reading paradigm and the technique of monitoring subjects' eye movements during reading. The former technique is used to provide continuity with past research and the latter to gain greater naturalness and finer temporal resolution. Experiments 1 and 2 test

sentences with main-verb/reduced-relative-clause ambiguities. Experiments 3 and 4 test sentences with ambiguous attachment sites for prepositional phrases.

EXPERIMENT 1

The aim of Experiment 1 was to examine whether the parser's on-line decisions are affected by manipulations of the referential properties in sentences containing main-verb/reduced-relative-clause ambiguities. A single substitution of either the focus operator ONLY or the definite determiner THE was made within the ambiguous test sentences. This substitution does not alter the syntactic structure at the point of ambiguity (i.e. at the ambiguous verb), but it does alter the referential content of the initial NP. According to the referential theory, sentences in which the word ONLY precedes a noun that is followed by an ambiguous verb should not induce garden path effects at the main verb, in contrast to their counterparts that substitute the word THE. In addition, the ambiguous test sentences were manipulated by including or excluding an adjective in the noun phrase that contained either THE or ONLY. Garden path effects are expected to occur for both these sentences, following the referential theory.

Method

Subjects. Thirty-two undergraduate students participated in the experiment, all of whom were native speakers of English and were naive about the purpose of the experiment.

Materials. Thirty-two ambiguous test sentences and 16 unambiguous controls were constructed for the experiment.⁶ There were four versions of each of the test and control sentences. One version of the test and control sentences contained the definite determiner THE in the initial noun phrase ("The-amb" and "The-unamb"); one version of each contained the word ONLY in the initial noun phrase ("Only-amb" and "Only-unamb"); one version of each contained THE and an adjective in the initial NP ("The-adj-amb" and "The-adj-unamb"); and, finally, one version of the test and control sentences contained the word ONLY followed by an adjective in the initial NP ("Only-adj-amb" and "Only-adj-unamb"). A full list of test

⁶We used fewer control sentences than experimental sentences in an effort to minimise subject fatigue.

and control sentences can be found in Appendix 1. The following is an example of one complete set of test and control sentences:

Ambiguous test sentences

- The-amb *The businessmen loaned money at low interest were told to record their expenses.*
- Only-amb *Only businessmen loaned money at low interest were told to record their expenses.*
- The-adj-amb *The wealthy businessmen loaned money at low interest were told to record their expenses.*
- Only-adj-amb *Only wealthy businessmen loaned money at low interest were told to record their expenses.*

Unambiguous control sentences

- The-unamb *The vans stolen from the parking lot were found in a back alley.*
- Only-unamb *Only vans stolen from the parking lot were found in a back alley.*
- The-adj-unamb *The new vans stolen from the parking lot were found in a back alley.*
- Only-adj-unamb *Only new vans stolen from the parking lot were found in a back alley.*

The test and control sentences were intermixed in the experiment. A counterbalanced design yielded four lists of stimuli, such that no more than a single version of any particular test or control sentence was present in each list. Each of the four lists contained 32 ambiguous test sentences (with 8 tokens of each of the 4 versions) and 16 unambiguous control sentences (with 4 tokens of each of the 4 versions). These sentences were interspersed among 92 filler sentences. The fillers included a variety of structures, half of which were grammatically ill-formed. Eight subjects were randomly assigned to be tested on one of the four stimulus lists. All versions of the test and control sentences were shown to each subject; therefore, the substitutions between THE and ONLY and the presence and absence of an adjective in the initial noun phrase are within-subject (and within-item) variables, while ambiguity is a within-subject but between-item variable.

Procedure. The experiment used a grammaticality judgement task embedded in the self-paced, word-by-word reading paradigm (Ford, 1983; Kennedy & Murray, 1984). Each consecutive word in a test sentence appeared on the computer screen, from left to right, at the request of the subject by way of a key press. The words accumulated on the screen. The subject's task was to press a key marked "YES" if each newly appearing word was a grammatical continuation of the previous material. The subjects

were instructed to press a key marked "NO" whenever a sentence fragment stopped being grammatical. The "NO" key was used by subjects thereafter to finish displaying the remainder of the sentence. The computer recorded the time elapsed, in milliseconds, between the onset of each new word and the next key press. The subjects' responses ("YES" or "NO" to each word) were also recorded by computer. A short pre-test was conducted to familiarise the subjects with the task.

Results and Discussion

Two-way ANOVAs were performed separately for the ambiguous test sentences and for the unambiguous control sentences. These ANOVAs examined the two types of pre-nominal word (THE/ONLY) and the presence or absence of an adjective (ADJ/NOADJ) intervening between the word THE or ONLY and the head noun. Planned comparisons by subject were performed between ambiguous test sentences and unambiguous controls. The dependent variables were mean reaction times and percent of errors. The reaction time data included the time subjects took at each word to correctly judge that the word was a grammatical continuation of the ongoing sentence fragment. A "NO" response to any word in a test or control sentence was counted as an error, and reaction times on any sentence in which an error occurred were excluded from the reaction time analyses.

For comparisons of reaction times and error rates, the test and control sentences were divided into six regions. Region 1 contained the subject NP [*The/Only (wealthy) businessmen*]. Region 2 included the first verb [*loaned*], which is morphologically ambiguous in the test sentences. Region 3 [*money at low*] contained the remainder of the first verb phrase except the last word. The sole content of Region 4 was the last word in the first verb phrase [*interest*]. Region 5 was the region of focal interest in this experiment. It contained two words: either an auxiliary verb and the main verb, or the main verb and the following word [*were told*]. These words either confirmed a correct analysis or corrected a misanalysis. The final region included the remainder of the sentence minus the terminating words [*to record their*]. The final word was excluded from the analysis to avoid the distorting influence of end-of-sentence wrap-up effects (see Just & Carpenter, 1980).⁷

⁷Just and Carpenter (1980) observe that there are extra processing requirements associated with the end of a sentence. Subjects tend to pause longer at the word or phrase that terminates a sentence for several reasons. First, they may be searching for references that have not been assigned; second, they may be constructing interclause relations (contextual integration); and, finally, they must handle any inconsistencies that could not be resolved within the sentence. In a word-by-word reading task, the maintenance of verbatim information in working memory may also contribute to the sentence wrap-up effect.

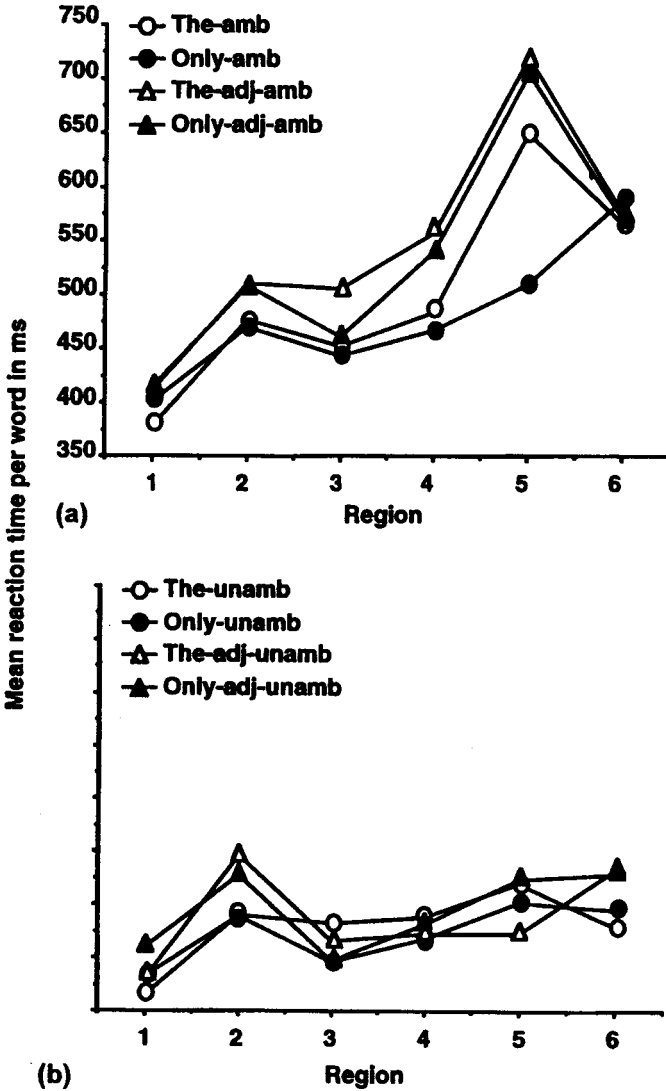


FIG. 1. Experiment 1: Mean reaction time per word at each region for (a) ambiguous test sentences and (b) unambiguous control sentences.

Figure 1 depicts the mean reaction time per word at each of the six regions for the ambiguous test sentences and for the unambiguous control sentences. The ambiguous sentences and unambiguous controls were analysed separately. At Regions 1-4, there was no significant effect of the substitution between THE and ONLY (hereafter THE/ONLY), or between

the presence and absence of an adjective in the initial noun phrase (hereafter ADJ/NOADJ). However, average reaction times were greater for ambiguous sentences than for unambiguous controls at Region 2 [$F(1,31) = 4.83, P < 0.04$], Region 3 [$F(1,31) = 18.21, P < 0.01$] and Region 4 [$F(1,31) = 17.89, P < 0.01$].⁸

At Region 5, a two-way analysis of variance revealed a main effect of THE/ONLY for the ambiguous test sentences [$F_1(1,31) = 10.31, P < 0.01$; $F_2(1,31) = 3.55, P < 0.07$]; ambiguous sentences containing THE induced longer reaction times than those containing ONLY, although the effect only approached significance in the analysis by items. There was also a main effect of ADJ/NOADJ [$F_1(1,31) = 3.54, P < 0.07$; $F_2(1,31) = 8.14, P < 0.01$]; longer reaction times were found in ambiguous sentences with an intervening adjective than in those without one, although in this case the effect by subjects only approached significance. There was no interaction between these factors, however. The lack of interaction is presumably due to the high variance associated with the sentences that evoked longer reaction times. Of the four versions of test sentences, three ("The-amb", "The-adj-amb", "Only-adj-amb") were expected to produce garden path effects, but "Only-amb" sentences should not have induced a garden path effect according to the referential theory. This interpretation is supported by a comparison of standard deviations ("Only-amb" = 123.23, "The-amb" = 335.61, "The-adj-amb" = 356.78, "Only-adj-amb" = 336.91). Indeed, a planned contrast between the reaction times of the "Only-amb" sentences and an average of the reaction times of the other three versions revealed that the former (mean = 502.11 msec) was significantly shorter than the latter (mean = 695.76 msec) [$t(1) = 18.94, P < 0.01$].

While there were no significant main effects or interactions among the four versions of the unambiguous control sentences at Region 5, a planned comparison between ambiguous and unambiguous sentences (hereafter AMB/UNAMB) revealed a significant main effect [$F(1,31) = 45.49, P < 0.01$], with longer reaction times associated with the ambiguous sentences. Sentences with THE were also found to induce longer reaction times than those with ONLY [$F(1,31) = 5.90, P < 0.03$]. The difference between sentences with or without an adjective was non-significant [$F(1,31) = 3.22, P < 0.09$]. There was a significant interaction between the factors THE/ONLY and AMB/UNAMB [$F(1,31) = 8.70, P < 0.01$]; whereas sentences with THE induced longer reaction times in the ambiguous cases,

⁸To simplify the exposition, we will use the following conventions for expressing *P*-values in ordinary language: $P < 0.01$ = highly significant; $P < 0.05$ = significant; $P < 0.1$ but > 0.05 = approaching significance; $P > 0.1$ = not significant. Also, because Experiments 1, 2 and 3 used a mixed design, analysis by subjects (F_1) and analysis by items (F_2) were not both carried out in all cases. The symbol *F* will be used if only one analysis was possible, either by subjects or by items.

this was not true of the unambiguous cases. The interaction between THE/ONLY and ADJ/NOADJ approached significance [$F(1,31) = 3.84$, $P < 0.06$]; an intervening adjective induced longer reaction times in the ambiguous sentences but not in the unambiguous ones. There was no significant three-way interaction (i.e. AMB/UNAMB \times THE/ONLY \times ADJ/NOADJ).

Turning to the analyses of the error data, we conducted separate two-way (THE/ONLY \times ADJ/NOADJ) analyses of variance for the ambiguous test sentences and for the unambiguous controls. The profiles by region are depicted in Fig. 2. As Fig. 2 indicates, there were no significant differences between versions in the first three regions. At Region 4, the word before the disambiguating region, more errors were found for ambiguous sentences with an adjective than ones without [$F(1,31) = 4.29$, $P < 0.05$]. For the unambiguous sentences, the error rate was higher for sentences containing THE than for sentences containing ONLY [$F(1,31) = 8.75$, $P < 0.01$].

At Region 5, for ambiguous sentences, there was a main effect of THE/ONLY [$F_1(1,31) = 36.78$, $P < 0.01$; $F_2(1,31) = 34.55$, $P < 0.01$] and a main effect of ADJ/NOADJ [$F_1(1,31) = 27.35$, $P < 0.01$; $F_2(1,31) = 23.37$, $P < 0.01$]. In addition, there was a significant interaction of THE/ONLY and ADJ/NOADJ [$F_1(1,31) = 5.20$, $P < 0.03$; $F_2(1,31) = 8.49$, $P < 0.01$]. More errors were made on ambiguous sentences with THE than on ones with ONLY, and for those with an adjective than those without one. For the unambiguous controls, no such effects existed. There was a significant main effect of AMB/UNAMB at Region 5 [$F(1,31) = 253.47$, $P < 0.01$], with ambiguous sentences inducing more errors than unambiguous sentences. There was no significant effect of interaction by three-way ANOVA (AMB/UNAMB \times THE/ONLY \times ADJ/NOADJ).

Finally, we carried out two planned comparisons on the reaction time and error data between the "Only-amb" test sentences and an average based on the four versions of the unambiguous control sentences. With mean reaction time as the dependent measure, the difference between the "Only-amb" sentences and the unambiguous controls approached significance [$t(1) = 4.04$, $P < 0.06$]. However, there was a robust difference between them when error rate was the dependent measure [$t(1) = 30.88$, $P < 0.01$]. This result points to the fact that while subjects responded with greater accuracy and read the "Only-amb" sentences faster than other versions of the ambiguous sentences, they nevertheless responded less accurately and read these sentences somewhat slower than the unambiguous control sentences. This pattern is predicted by the referential theory, as discussed below.

To summarise, the results of Experiment 1 lend support to the referential theory. Most importantly, there was a significant decrease in reaction times and errors on the "Only-amb" sentences at the disambiguating Region 5, as compared to the "The-amb" sentences. The requirement posed by ONLY

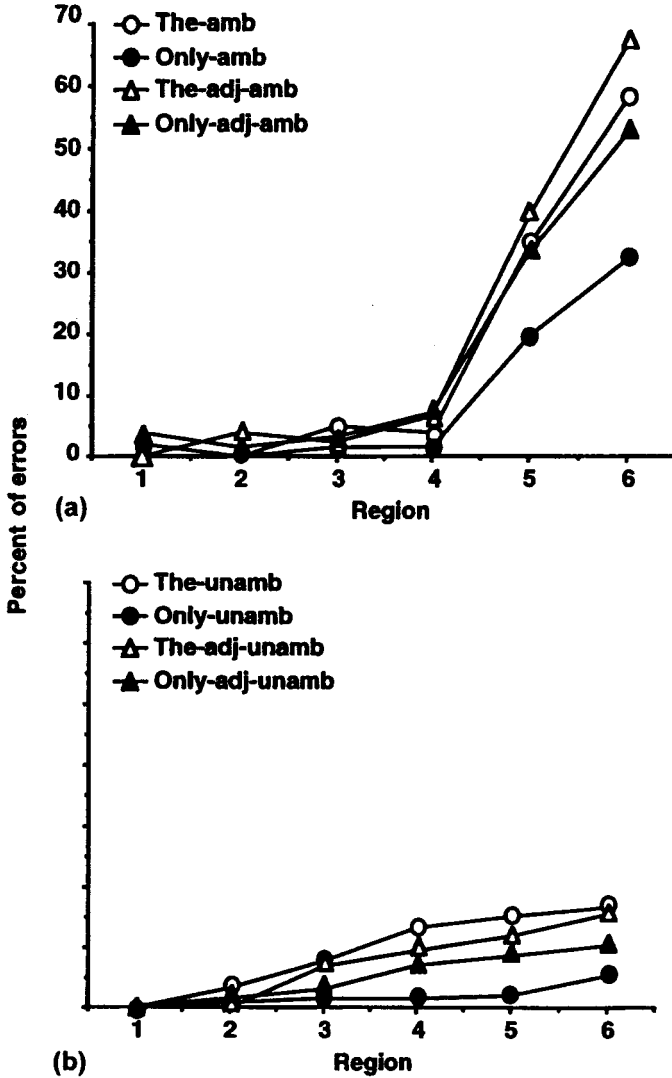


FIG. 2. Experiment 1: Percent of errors at each region for (a) ambiguous test sentences and (b) unambiguous control sentences.

for a contrast set apparently directed the parser to opt for the reduced relative clause alternative within the ambiguous region and hence the main verb came as no surprise—a potential garden path effect was averted. By contrast, in “The-amb” sentences, the main verb analysis was generally chosen in the ambiguous region, resulting in a severe garden path effect at the disambiguating region. There was also a marked difference between

“Only-amb” sentences and “Only-adj-amb” sentences, with the latter, but not the former, showing a garden path effect at the disambiguating region. This suggests that the appearance of an adjective satisfied the requirement for a contrast set and, accordingly, the parser preferred to analyse the first verb phrase as the main verb, leading to the garden path effect.

It remains to comment on the comparison between the “Only-amb” sentences and the unambiguous control sentences. As we saw, compared with unambiguous controls, there was a slight elevation in reaction times for the “Only-amb” sentences at the disambiguating region; these sentences also produced significantly more errors. This, we suggest, is chiefly a consequence of parallel processing. When an input string is ambiguous, albeit briefly, more than one analysis is entertained. Although the principles of the referential theory direct the parser in its decisions, even a dispreferred interpretation will be accepted by some proportion of the subjects. The roughly 20% increase in errors for the “Only-amb” sentences, as compared to the unambiguous control sentences, suggests that while the presence of ONLY sufficiently promoted the reduced relative clause analysis in the majority of cases, the main clause reading was not ruled out on every occasion. This resulted in some elevation in reaction times and errors, as compared to the unambiguous controls.

The findings of the present experiment would not be expected on a serial processing model. If a parsing strategy such as Minimal Attachment was always applied first, then a lexical substitution in the initial noun phrase would not affect the structural analysis at the onset of an ambiguity, because at that point each of the ambiguous sentences has the same structure. As noted earlier, however, on the model proposed by Mitchell et al. (1992), relevant semantic information could be used very rapidly to override a brief initial misanalysis within the ambiguous region. This could explain the decrease in reaction times on the “Only-amb” sentences at Region 5 (the so-called disambiguating region). Note that if reanalysis is performed, it would occur following a garden path effect, however brief it might be. Although the findings of the present experiment do not indicate the existence of a garden path effect in the ambiguous region of the “Only-amb” sentences, in contrast to the “The-amb” sentences, it is conceivable that non-structural information was used within the ambiguous region to override a brief and mild garden path effect, but that the word-by-word reading measure used in this experiment was insufficiently sensitive to detect such an effect.

By its nature, the self-paced, word-by-word reading technique has an inherent limitation. Because a decision is called for at every word, reading speed is slowed to levels far below normal reading rates. Therefore, measures of word-by-word reading may include the cumulative influences of a number of factors, and some of these factors may be used by the parser

earlier than others. It has been suggested repeatedly, on grounds such as these, that measures of word-by-word reading may be insensitive to the exact timing of the availability and application of different sources of information at the potentially most informative points in sentence processing, and that measures of eye movements may afford greater precision (e.g. Rayner, 1993; Rayner et al., 1989; Rayner & Morris, 1991; but see Ferreira and Clifton, 1986, for evidence that eye movement results often confirm findings from word-by-word experiments).⁹ Based on these considerations, Experiment 2 repeated Experiment 1, using the technique of recording subjects' eye movements during reading.

EXPERIMENT 2

As noted in the Introduction, the influence of non-syntactic information in sentence interpretation is not in dispute. At issue, however, is the exact time-course of the application of these sources of information in on-line sentence processing. A central question is whether non-syntactic information is available and used immediately to resolve ambiguities, or whether structurally based strategies suffice to explain the parser's early decision making. To answer this question, a more time-sensitive measure of parsing is required than that afforded by the word-by-word reading method. Accordingly, in Experiment 2, we adopted the technique of eye movement recording. Tracking subjects' eye movements while they are reading arguably affords the required precision. First, this technique permits normal, uninterrupted reading. Second, it permits the experimenter to identify specific fixation locations in the line of print. When reading materials are analysed by region, we can ascertain not only how long a subject's eyes remain in a region when that region is read for the first time, but also how often a regressive eye movement is initiated from that region and where it lands.

The capability to examine both first-pass reading and the incidence of regressive eye movements is important in addressing the question of *when* different sources of information are used by the parser. It is our working assumption that first-pass fixations are most indicative of on-line processes; that is, they indicate the influence of information that is immediately accessed by the reader. We will assume that regressions are not indicative of on-line processing because they occur only sporadically in normal reading. The disparate patterns of regression on different types of sentences are therefore informative. Frequent regressions may signal difficulties that lead the parser to reprocess earlier material. Exploiting the reciprocity between

⁹Eye movement patterns may reflect pre-cognitive processes, however, as pointed out by Altmann and Steedman (1988, p. 217).

reading times and regressions, we can use eye movement tracking as an aid to infer which sources of linguistic information are used immediately, and which sources are used somewhat later in processing.

The object of Experiment 2 was to find out whether the information contributed by the focus operator ONLY is used in first-pass reading. In this experiment, subjects' eye movements were monitored while they read test materials similar to those used in Experiment 1. The predictions of the referential theory were straightforward: (1) the parser should follow semantic/referential principles in constructing discourse representations, and (2) garden path effects should be modulated by the presence or absence of the focus operator ONLY.

Method

Subjects. Twenty-two undergraduate students participated in the experiment, all of whom were native speakers of English. All subjects reported normal vision or normal vision with soft contact lenses.

Materials. Twenty-four test sentences and 16 control sentences were randomly selected from the materials used in Experiment 1. Each sentence had two versions, the sole difference between them being the alternation between THE and ONLY in the initial NP. Some minor revisions were made to the test materials so that none of the sentences exceeded 76 characters in length. This enabled us to present each sentence on a single line beginning at the left margin of centre screen. A sample set of test sentences and their corresponding controls is as follows:

- | | |
|------------|------------------------------------------------------------------------------------|
| The-amb | <i>The businessmen loaned money at low interest were told to record expenses.</i> |
| Only-amb | <i>Only businessmen loaned money at low interest were told to record expenses.</i> |
| The-unamb | <i>The vans stolen from the parking lot were found in a back alley.</i> |
| Only-unamb | <i>Only vans stolen from the parking lot were found in a back alley.</i> |

A counterbalanced design was used, with test and control sentences evenly distributed in each of the two experimental conditions, which contained either a pre-nominal THE or ONLY. No subject read the same sentence with THE and with ONLY, and all test and control sentences occurred in either condition over the two stimulus lists. Eleven subjects were tested on each list in which test and control sentences were intermixed

among 60 fillers in a pseudo-random fashion, such that each test or control sentence was followed by at least one filler. Each stimulus list was divided into two halves containing an equal number of test and control sentences, which were presented in separate sessions, preceded by 10 warm-up trials.

Equipment. The subjects' eye movements were recorded using the IRIS infrared-light eye movement system (SKALAR model 6500). The IRIS system uses a differential reflection method of eye movement recording. In this technique, infrared-emitting diodes and infrared-sensitive detectors are positioned in front of the eye so that their receptive fields match the iris-sclera boundary, both on the nasal side and on the temporal side. Upon horizontal rotation of the eye, the nasally positioned detector measures an increase in scleral infrared reflection, while the detector on the temporal side measures a decrease in infrared reflection. Subtraction of the signals from the nasal and temporal detectors gives eye position relative to head position. Eye position is sampled every millisecond by a computer equipped with an analogue-to-digital conversion board. Each eye fixation is represented by an x and y screen coordinate, a starting time and an ending time. Eye movements are recorded from the right eye only, but viewing is binocular. The stimuli are displayed on a 13 inch High Resolution RGB monitor set 64 cm from the subject's eyes. In our test materials (mixed case in Courier 14 point font), the visual angle of each character was slightly greater than 12 min of arc, permitting a resolution of less than one character width.

Procedure. The subjects were given written instructions that contained a brief description of the eye movement monitoring technique. Since eye positions are determined relative to the position of the head, head movements were kept to a minimum. Head stabilisation was achieved by a bite-bar and a forehead rest. After the eye tracker was calibrated, the subject was told to begin reading sentences from the starting point at the centre left of the screen, where a fixation cross was presented before each sentence. The experimenter emphasised that the sentence should be read at the subject's normal rate. When the subject finished reading the sentence, he or she pressed on the mouse and the sentence was erased. On one-third of the filler trials, a comprehension question appeared on the screen after a sentence disappeared. The subject answered the question by moving a mouse-controlled arrow to a "YES" box or a "NO" box on the screen and clicked on it. Feedback was given by the computer, informing the subject whether or not the answer was correct. After each trial, there was a calibration-checking routine, at which time the subject fixated consecutively on five coordinates on the screen. Adjustments of signal strength were occasionally performed, but recalibration was rarely needed. A pre-test containing six sentences was conducted.

Results and Discussion

We present analyses based on the recorded eye fixation durations and percent of regressions. An eye fixation was recorded if a subject's eye dwelled upon a character for 8 msec or longer. Total fixation duration in milliseconds was computed for each scoring region of the test and control sentences. The incidence of regressions from these regions was also recorded. For the purpose of data analysis, the test and control sentences were divided into the same six regions as those described in Experiment 1: Region 1 contained the subject NP; Region 2 included the first verb, which was morphologically ambiguous for test sentences; Region 3 contained the remainder of the first verb phrase except the last word, which was the sole content of Region 4; Region 5 contained either an auxiliary verb and the main verb, or the main verb and the following word. The final region consisted of the remainder of the sentence.

Planned factorial comparisons at each region were carried out by means of two-way ANOVA, testing the two types of pre-nominal word (THE/ONLY) and the two types of initial verb phrase (AMB/UNAMB). The measure of first-pass reading time included fixation durations by subjects who read a certain region of the sentence for the first time, provided that they had not read beyond that region. The dependent variable used in the analyses was residual reading time (RRT).¹⁰ The RRT was calculated by conducting a regression analysis for each subject, using the length of each region (the number of letters and spaces) as the independent variable and the total duration of all eye fixations at each region as the dependent variable. This measure statistically removes the length of a region as a factor. The incidence of regressive eye movements was based on the percent of subjects' first-pass readings of a region that ended in a leftward regression to a portion of the sentence that had either been visited before or skipped.

First-pass Reading Times. Figure 3 presents a profile of mean first-pass RRTs at each region for the test sentences and controls. A significant main effect of THE/ONLY was found at Region 1, where reading times on sentences beginning with ONLY were significantly greater than those beginning with THE [$F(1,21) = 6.72, P < 0.02$]. No significant main effect of AMB/UNAMB was found, nor was the interaction significant between

¹⁰Trueswell et al. (1994) maintain that per-character reading time, a widely used measure, may provide a misleading basis for comparison across regions of different length. Per-character reading time progressively distorts the accuracy of the analyses as the length of a region decreases. If the regions being compared are identical in length and/or in material, then the measure of uncorrected total reading time suffices. If the portions being compared are different (especially in length), then residual reading time provides a more accurate measure.

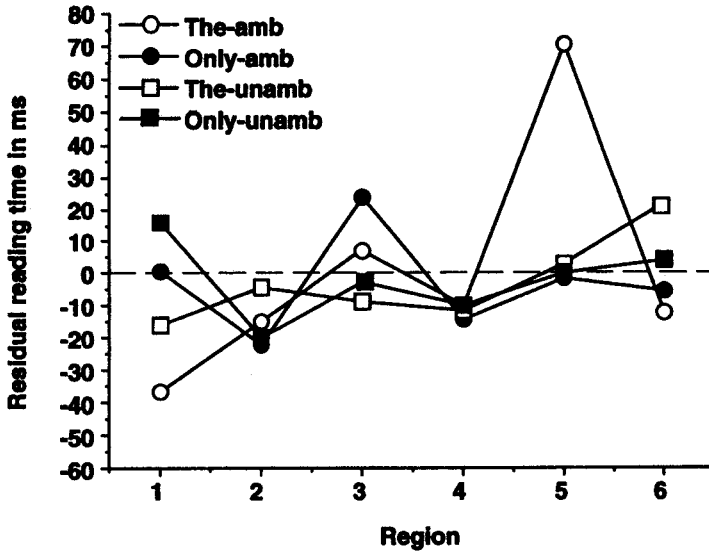


FIG. 3. Experiment 2: Mean first-pass residual reading time (RRT) at each region.

THE/ONLY and AMB/UNAMB. It is clear that the focus operator ONLY had an impact on the parser such that it induced longer reading times than did the definite determiner THE.

The only significant effect in the ambiguous region (Regions 2–4) was a main effect of AMB/UNAMB at Region 3, where reading times on the ambiguous test sentences were longer than those on the unambiguous controls [$F(1,21) = 6.17$, $P < 0.02$]. Ambiguous sentences apparently presented more difficulties than unambiguous sentences within the ambiguous region. The convergence of reading times on all versions of sentences at Region 4 suggests that resolution of the ambiguity was reached by the end of the ambiguous region.

At Region 5, where the main verb appeared, there was a main effect of THE/ONLY: reading times for sentences beginning with THE were significantly longer than for those beginning with ONLY [$F(1,21) = 23.14$, $P < 0.01$]. There was also a significant main effect of AMB/UNAMB: ambiguous test sentences took significantly longer to read than unambiguous controls [$F(1,21) = 9.91$, $P < 0.01$]. Moreover, there was a significant interaction of THE/ONLY \times AMB/UNAMB [$F(1,21) = 7.05$, $P < 0.02$]. A pairwise comparison showed that reading times for test sentences with THE (“The-amb”) were significantly greater than those for the “Only-amb” sentences [$F_1(1,21) = 21.83$, $P < 0.01$; $F_2(1,23) = 13.65$, $P < 0.01$]. On the other hand, no significant difference was found between the two versions of the unambiguous controls ($P > 0.1$). Reading times did

not differ between the "Only-amb" version and either of the two versions of the controls ($P > 0.1$).

Reading times at Region 5 suggest that a garden path effect occurred for test sentences beginning with THE; long fixation durations in that region indicated that subjects were apparently surprised to see that the material (the main verb) could not be incorporated into the analysis they had adopted and, as a result, they were forced to pause. For sentences that began with ONLY, on the other hand, no significant rise in fixation durations was recorded, suggesting that the main verb was expected, in keeping with the cases of the unambiguous controls. This pattern leads us to suppose that the reduced relative clause analysis was adopted in the ambiguous test sentences beginning with ONLY. Considering that these are first-pass fixation times, the information carried by ONLY must have been used extremely rapidly. However, in order to conclude that the processing of "Only-amb" sentences is genuinely different from that of "The-amb" sentences, subjects' eye-regression patterns must also be considered.¹¹

Incidence of Regression. Two-way ANOVAs were performed on the percent of first-pass readings that resulted in regressive eye movements. Inspection of the percent of regressions at each region revealed a significant effect of AMB/UNAMB at Region 5 [$F(1,21) = 10.05, P < 0.01$] and at Region 6 [$F(1,21) = 8.73, P < 0.01$]. In each case, ambiguous test sentences induced more regressions than unambiguous controls. There was no main effect of THE/ONLY, nor was there an interaction at any region. Figure 4 displays the pattern of regressions for each of the four versions of sentences at each region.

The ambiguous test sentences were found to induce more regressions than their unambiguous controls, despite the fact that first-pass reading times were significantly different between the two types of ambiguous test sentences at Region 5. This is reminiscent of the results of Experiment 1, where reaction times showed a significant difference between "The-amb"

¹¹Closely related to regression patterns are second-pass reading times, which are calculated from the fixation durations in the portions of sentences that have either been read earlier or skipped entirely. Our data showed that, overall, ambiguous test sentences required longer reading times and were re-read more often than their unambiguous controls. There was no indication that the two types of test sentences were re-read differently, at least in cases where re-reading did occur. However, in this paper, we report regression patterns rather than second-pass reading times. As pointed out by the editor and an anonymous reviewer, it is often difficult to interpret second-pass reading time data, because what counts as a second-pass fixation depends on how a region is defined. For instance, a particular fixation Y that follows a fixation X but lands to the left of X may be counted as a first-pass fixation if the landing site is within the boundary of a particular region. On the other hand, if the region boundary is redefined so that it falls between fixations X and Y, then Y is counted as a second-pass fixation.

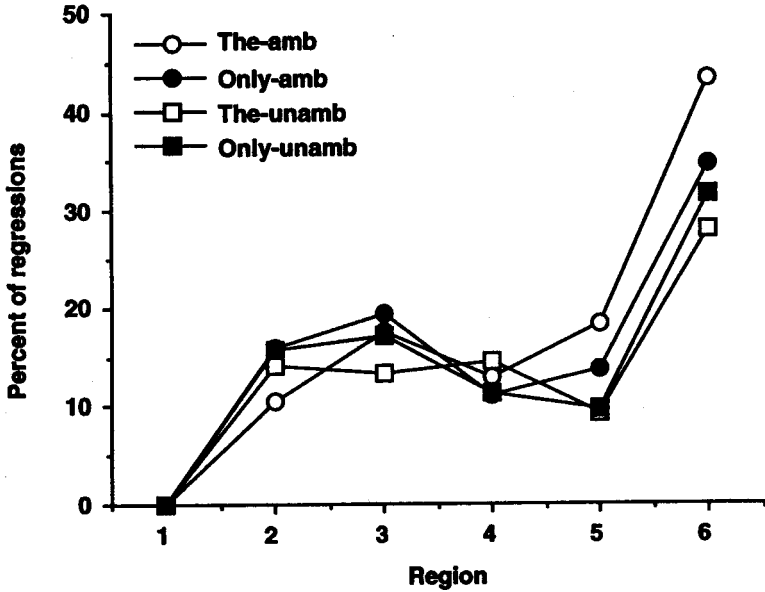


FIG. 4. Experiment 2: Percent of regressions from each region.

and “Only-amb” sentences at the point of disambiguation, but both versions showed a higher error rate than the unambiguous controls.

In light of the fact that ambiguous sentences induced more regressions than unambiguous ones, we also analysed the data from first-pass reading times by creating a new data set that excluded any trial on which there was a regression; that is, we conducted a regression-contingent analysis of first-pass reading times at Region 5 (see Altmann et al., 1992, for a detailed discussion of this method). The purpose of this analysis was to determine whether a difference in reading times persisted at Region 5 between “The-amb” sentences and “Only-amb” sentences using an uncontaminated measure of first-pass reading. Shorter reading times in a region sometimes result when there are many regressions from that region. In the present analysis, we wanted to ensure that this was not the source of the relatively faster reading times at Region 5 on the “Only-amb” sentences. We expect, on the basis of the referential theory, that the difference in processing ambiguous sentences beginning with THE versus those beginning with ONLY should be present *in the absence of a regression*.

The results of the regression-contingent analysis closely resembled the reading time patterns reported earlier—there was a main effect of THE/ONLY [$F(1,21) = 15.39, P < 0.01$] as well as AMB/UNAMB [$F(1,21) = 7.76, P < 0.02$]. The effect of THE/ONLY \times AMB/UNAMB interaction approached significance [$F(1,21) = 3.55, P < 0.08$]. A comparison between

sentence versions showed that reading times on "The-amb" sentences were significantly longer than those on "Only-amb" sentences [$F_1(1,21) = 12.15$, $P < 0.01$; $F_2(1,23) = 7.54$, $P < 0.02$]. On the other hand, there was no difference between the two kinds of controls ($P > 0.1$). No difference existed between test sentences beginning with ONLY and either of the two kinds of unambiguous controls ($P > 0.1$).

In sum, the results of Experiment 2 confirm the main findings of Experiment 1; that is, they support the contention of the referential theory that referential effects occur on-line in ambiguity resolution. The referential information carried by the pre-nominal focus operator ONLY strongly influenced the parser's initial analysis of the ambiguity, as attested by the eye movement patterns, especially the significant interaction between THE/ONLY and AMB/UNAMB at the disambiguating region on the reading time measure. First-pass reading times in the absence of regressions showed a similar result. These findings are as predicted by the referential theory. The theory also allows that, because more than one analysis is entertained for ambiguous sentences, these sentences prove to be more difficult than their unambiguous counterparts; this is shown by a significantly higher rate of regressions in the former than in the latter.

In discussing the findings of Experiment 1, we noted the difficulty in distinguishing predictions of the referential theory from predictions of the recently revised, structurally based garden path model (Mitchell et al., 1992). According to this account, reanalyses can be implemented early, using semantic and/or discourse information: If non-syntactic sources of information are delayed only briefly, then a Minimal Attachment analysis can be rapidly overridden within the ambiguous region in cases like the "Only-amb" sentences of the present experiment. As a result, no garden path effect would be expected to occur when the structural disambiguating material is encountered later. Because Experiment 1 used a word-by-word measure of reading, it may have lacked sufficient sensitivity to evaluate this possibility. With the measurement of eye movements, however, we are in a better position to reconstruct the pattern of processing within the ambiguous region. In fact, there was an apparent elevation of reading times for ambiguous sentences in this region.¹² The referential theory contends that

¹²A finer-grained, word-by-word analysis of first-pass eye fixation durations within the ambiguous region was conducted to look for signs of any effect of reanalysis for the "Only-amb" sentences. The most we could find was a non-significant elevation that occurred in the middle of the ambiguous region for "Only-amb" sentences relative to "The-amb" ones. This was probably because ambiguous sentences with ONLY required the construction and evaluation of a contrast set in addition to the difficulties imposed by parallel processing. There was also an effect of AMB/UNAMB which approached significance [$F(1,21) = 4.20$, $P < 0.06$] in the middle of the ambiguous region, with both ambiguous sentences taking longer to read than their unambiguous controls. By the end of the ambiguous region, however, reading times for all of the versions had converged.

this elevation is a consequence of parallel processing (see Discussion of Experiment 1). The effect within the ambiguous region could be reconciled with the model proposed by Mitchell et al. (1992), but only with the addition of two assumptions. The first is that reanalysis can be triggered and completed within the ambiguous region. The second assumption is that there are two distinct kinds of garden path effects: one kind that is sensitive to structural information and is costly of processing resources, and another kind that is responsive to non-syntactic sources of information and is cost-free for reanalysis. Although such a distinction is possible in principle, it would have to be motivated on independent grounds. In addition, some empirical way of distinguishing a costly garden path effect from a cost-free reanalysis would be required (cf. Fodor & Inoue, 1994; Frazier, 1994).

Experiments 1 and 2 converge on the same conclusion: garden path effects can be modulated by referential factors within the test sentences. The results from eye movement recording closely parallel those found with word-by-word, self-paced reading. As noted earlier, we employed within-sentence manipulations in the present study to enable us to examine referential effects in resolving sentences containing a main-verb/reduced-relative-clause ambiguity, a structure that has proven to be relatively impervious to extra-sentential context (Ferreira & Clifton, 1986; Murray & Liversedge, 1994; but also see Spivey-Knowlton & Tanenhaus, 1994, for evidence of the influence of referential context for this construction). We focused on the referential contributions of pre-nominal modifiers (ONLY versus THE) within a test sentence while holding constant factors such as the frequency of a verb (with an “-ed” ending) being used as a past-participle. Therefore, although the findings are entirely consistent with the constraint satisfaction model, it is important to note that the results were predicted by the principles of the referential theory.

In Experiments 3 and 4, we pursued a related issue: the use of general world knowledge (plausibility) in on-line sentence processing. The focus of these experiments was on the locus of attachment of prepositional phrases in structurally ambiguous sentences. In Experiment 3, we asked whether plausibility is used to resolve structural ambiguities. Experiment 4 investigated the time-course of the availability and application of this source of information, using the eye movement recording technique.

EXPERIMENT 3

The parser's preference in resolving local ambiguities involving the site of attachment of prepositional phrases has received much attention. As the examples below illustrate, one option in the sentences under consideration is to attach a prepositional phrase to the preceding verb phrase. This will be

referred to as the VP-attachment analysis. A second option is to attach the prepositional phrase to the immediately preceding noun phrase. This is referred to as the NP-attachment analysis. The following examples show that while the PP *with new brushes* in (A) can only be attached to the verb to make the sentence felicitous, the PP *with large cracks* in (B) can only be used to modify the immediately preceding noun:

- A. *The man [painted the doors [with new brushes]]
before the festival. VP-attachment*
- B. *The man [painted [the doors with large cracks]]
before the festival. NP-attachment*

It has been demonstrated repeatedly that subjects find sentences like (B) more difficult than those like (A) (e.g. Altmann & Steedman, 1988; Britt et al., 1992; Ferreira & Clifton, 1986; Perfetti, 1991; Rayner et al., 1983, 1992).¹³ Various explanations have been advanced to explain this difference. According to the referential theory, both VP-attachment and NP-attachment analyses are computed by the syntax at the onset of the preposition *with*, and both are evaluated by the semantic processor in a word-by-word fashion. The parser pursues the analysis that best fits the context (e.g. Altmann & Steedman, 1988). In the absence of prior linguistic context, however, a mental model of the discourse is set up which includes, for examples (A) and (B) above, a set of doors as required by the definite noun phrase, *the doors*. Having augmented the mental model with a set of doors, the VP-attachment analysis of the following PP is pursued. The alternative NP-attachment analysis requires the parser to further modify the mental model by distinguishing a subset of doors with specifications from other doors. The principle of parsimony therefore predicts that this analysis should be dispreferred because it requires more extensions to the mental model than the VP-attachment analysis. The pursuance of the VP-attachment analysis results in a temporary anomaly in sentences like (B), since *cracks* are not things that one can use to *paint* with, and reanalysis is instigated, leading to increased reading difficulty.

However, if the noun *doors* is preceded by ONLY, such as in (C) and (D), then the referential theory predicts a reversal in parsing preferences;

¹³A distinction has been made in recent studies (e.g. Britt, 1991; Britt et al., 1992; Frazier & Clifton, 1996) between attachment preferences due to an argument and ones due to an adjunct of a PP: A PP that follows a verb which requires a Goal argument (e.g. *put*) prefers to attach to the verb, while one that follows a verb like *throw*, which does not require an argument, may be more responsive to contextual manipulations. Other factors such as definiteness of NPs also affect parsing decisions (Crain & Steedman, 1985; Sedivy & Spivey-Knowlton, 1994). The present research did not consider these factors, which were held constant in the experimental manipulations.

namely, the NP-attachment analysis will be preferred, not the VP-attachment one:

- C. *The man [painted only doors [with new brushes]]
before the festival. VP-attachment*
- D. *The man [painted [only doors with large cracks]]
before the festival. NP-attachment*

According to the referential theory, the presence of the pre-nominal focus operator ONLY invites the parser to assume the existence of a set of entities that contrasts with those referred to by the noun. The most parsimonious way to construct a contrast set is to divide an existing set into subsets. Pursuing this option, an NP-attachment analysis provides the needed information for a contrast set, namely a specific set of doors (with cracks). As a consequence, sentences like (D) will be easy to process, but those like (C) will induce a temporary anomaly because it is infelicitous to modify *doors* with *brushes*, and reanalysis is required. Experiment 3 was designed to test these predictions, using a word-by-word reading paradigm.

Method

Subjects. Forty-four undergraduate students participated in the experiment, all of whom were native speakers of English and were naive as to the purposes of the experiment. The subjects were randomly assigned to one of four groups.

Materials. The experiment included 20 sets of test sentences in each of four versions: VP-attachment sentences with THE ("The-VP") or with ONLY ("Only-VP") and NP-attachment sentences with THE ("The-NP") or with ONLY ("Only-NP"), as shown in the examples below. A full list of test sentences can be found in Appendix 2. Forty filler sentences were interspersed among the test sentences. Four lists of stimuli were constructed. VP-attachment sentences were rotated through two lists, and NP-attachment sentences were rotated through the other two lists.¹⁴ Each list was tested on a different group of 11 subjects.

¹⁴This experiment used a mixed design in which half of the subjects read only VP-attachment sentences (with THE and ONLY), and the other half read only NP-attachment sentences (with THE and ONLY). As a result, THE/ONLY was a within-subject variable but VP/NP was a between-subject variable. Most of the data analyses, therefore, were by-item analyses (F_2). The mixed design was used because a pilot test showed that subjects were confused when both VP- and NP-attachment sentences were present, and there was a spillover effect. This was probably caused by the task that asked subjects to judge whether or not the sentence continued to make sense at every word. In Experiment 4, which used the same material in an eye movement monitoring study, a fully crossed design was used.

- The-VP *The man painted the doors with new brushes before the festival.*
 The-NP *The man painted the doors with large cracks before the festival.*
 Only-VP *The man painted only doors with new brushes before the festival.*
 Only-NP *The man painted only doors with large cracks before the festival.*

Procedure. The subjects read sentences displayed on the computer screen one word at a time, and the words remained on the screen until the sentence ended. The instructions given to the subjects were similar to those given in Experiment 1, except for one change. Since both VP-attachment and NP-attachment were grammatical, the subjects were asked to decide whether or not the sentence continued "to make sense" as each consecutive word appeared (see Boland et al., 1990, for a more detailed discussion of this "stop making sense" task). The computer recorded the duration (in milliseconds) between the onset of each new word and the subject's key press. Subjects' responses ("YES" or "NO") were also recorded for each word. A pre-test was conducted that contained eight example sentences.

Results and Discussion

Two-way ANOVAs were carried out, testing the effects of the two types of pre-nominal modifiers (THE/ONLY) and the two sites of attachment (VP/NP). The dependent variables were mean reaction time (RT) per word and percent of errors. Mean RT included the time subjects took to correctly accept each newly presented word to be a sensible continuation of the ongoing sentence fragment. A "NO" response to any word in a test sentence was interpreted as indicating that the subject erroneously deemed the sentence nonsensical. Reaction times on any sentence in which an error occurred were excluded from the analyses. Error analyses included all the responses from the subjects.

For the purpose of conducting statistical analyses, the test sentences were divided into five regions.¹⁵ Region 1 contained the subject NP [*The man*] and Region 2 contained the main verb [*painted*]. Region 3 contained the object noun phrase that was preceded by either the definite determiner THE or the focus operator ONLY, and followed by the preposition [*the/only doors with*]. Region 4 contained the object NP of the prepositional phrase, the content of which either confirmed or disconfirmed a particular attachment [*new brushes/large cracks*]. Region 5 contained the remainder of the sentence minus the last word [*before the*]. We report results from all five regions, focusing on Region 4; at this region, the referential theory predicts an interaction of THE/ONLY \times VP/NP.

¹⁵The division of regions was based on those used by Rayner et al. (1983).

Analyses of reaction times revealed no significant effect or interaction at either Region 1 or Region 2, as expected, since all the test sentences were identical in these regions. At Region 3, where the sentences diverged as to whether the definite determiner THE or the focus operator ONLY preceded the object noun, there was a significant main effect of THE/ONLY [$F_1(1,43) = 10.76, P < 0.01$; $F_2(1,19) = 5.64, P < 0.03$]: reaction times on sentences with ONLY were longer than those with THE. At Region 4, there was a significant interaction between THE/ONLY and VP/NP [$F(1,19) = 13.47, P < 0.01$]. A similar pattern existed at Region 5, where the THE/ONLY \times VP/NP interaction was significant [$F(1,19) = 18.99, P < 0.01$]. Because the design of this experiment used VP/NP as a between-subjects variable, the effects of VP/NP (and THE/ONLY \times VP/NP interaction) were calculated by items only. Figure 5 depicts the mean reaction times at each region.

Region 4 merits further analysis, because it contains the pragmatic information that either confirmed or disconfirmed the subjects' earlier parsing decisions. A pairwise comparison between sentence versions revealed that, as expected, reaction times on "The-NP" sentences were longer than those on "The-VP" sentences, although the effect only approached significance [$F(1,19) = 3.71, P < 0.07$]. A reversal occurred between the two versions with ONLY: reaction times to the "Only-NP" sentences were much shorter than those to the "Only-VP" sentences [$F(1,19) = 9.56, P < 0.01$]. Reaction times to the "Only-VP" version were longer than those to the "The-VP" version [$F_1(1,21) = 8.27, P < 0.01$; $F_2(1,19) = 8.85, P < 0.01$]; reaction times to the "Only-NP" version, on the

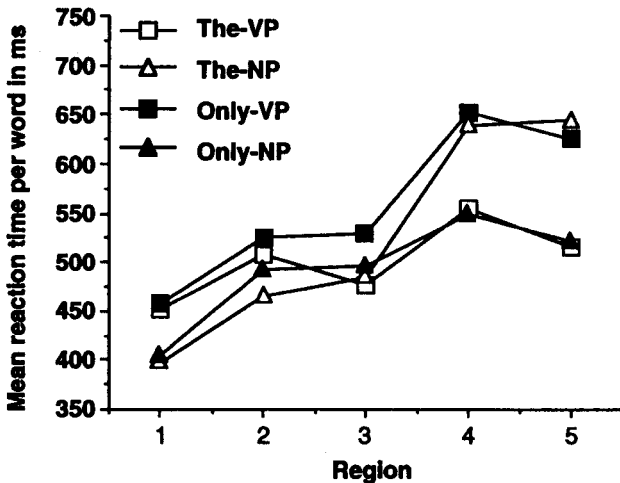


FIG. 5. Experiment 3: Mean reaction time per word at each region.

other hand, were shorter than those to the "The-NP" version [$F_1(1,21) = 6.82, P < 0.02; F_2(1,19) = 4.36, P < 0.05$].

Figure 6 depicts the mean error rate at each region. The error count across all regions revealed a highly significant interaction of THE/ONLY \times VP/NP [$F(1,19) = 126.47, P < 0.01$]. The effect occurred mainly at Regions 4 and 5. ANOVAs carried out on combined scores at these regions revealed a significant main effect of THE/ONLY [$F_1(1,43) = 39.56, P < 0.01; F_2(1,19) = 38.32, P < 0.01$]. The effect of VP/NP approached significance [$F(1,19) = 3.42, P < 0.08$]. The interaction between the two factors was significant [$F(1,19) = 39.87, P < 0.01$]. A pairwise comparison showed that "Only-VP" sentences induced significantly more errors than "Only-NP" sentences [$F(1,19) = 18.33, P < 0.01$]. The difference between "The-VP" and "The-NP" sentences approached significance [$F(1,19) = 4.02, P < 0.06$], with "The-NP" sentences inducing more errors. "Only-VP" sentences induced significantly more errors than "The-VP" sentences [$F_1(1,21) = 27.69, P < 0.01; F_2(1,19) = 69.79, P < 0.01$]. There was no difference between the "The-NP" and the "Only-NP" versions.

The reaction times and error rates confirmed the predictions of the referential theory. According to the referential theory, both VP-attachment and NP-attachment analyses are computed when the prepositional phrase is encountered. The focus operator ONLY makes the parser anticipate a modification of the noun phrase, opting for the NP-attachment analysis. The anticipation is met when the NP *large cracks* is encountered, because the prepositional phrase is a plausible modifier of the object NP *doors*. On the

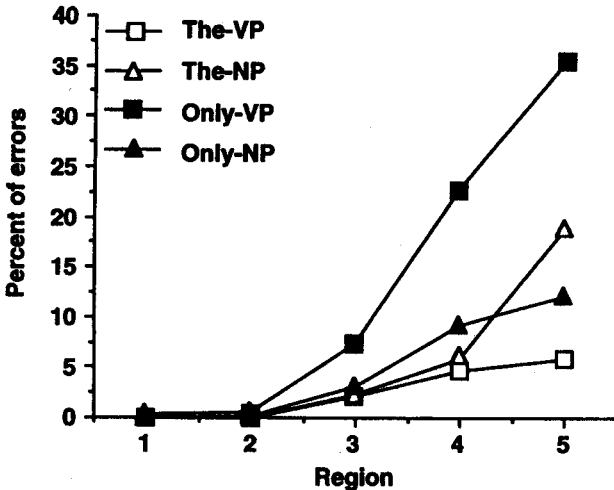


FIG. 6. Experiment 3: Percent of errors at each region.

other hand, the anticipation of an NP modifier in "Only-VP" sentences leads the parser to encounter the noun phrase *new brushes*, which makes the NP-attachment analysis of the PP anomalous. Reanalysis is therefore required, resulting in a significant elevation of reaction times, as well as increased erroneous rejections of the sentence.

The findings of Experiment 3 indicate that a simple substitution between THE and ONLY in a noun phrase followed by a prepositional phrase can change the parser's immediate decision as to how a prepositional phrase should be attached. It was found in Experiments 1 and 2 that the semantic information carried by the focus operator ONLY is used on-line, establishing an immediate preference for an analysis that sets up a contrast set. The effects obtained in this experiment, however, hinge on the added assumption that subjects use *a priori* plausibility elicited by the noun phrase of the prepositional phrase as their criterion for PP-attachment. The disambiguating factor here is real-world knowledge. It is clear from Experiment 3 that information governing the *a priori* plausibility of the alternative representations of an ambiguous phrase is used quickly in reaching decisions about PP-attachment. The manner in which such information is used by the parser was assessed with greater precision in Experiment 4, which used the eye movement recording technique.

EXPERIMENT 4

The eye movement recording methodology offers an improvement in measurement precision. With this technique, we are in a better position to detect subtle garden path effects, or to detect the speedy recovery from a misanalysis involving a semantic anomaly. This methodology also has the advantage that it can reveal the time-course of the availability and use of different sources of information during on-line sentence processing. The issue of timing is at the crux of current research on sentence processing. If some sources of information are used earlier than other sources, this circumstance could be used to decide between competing models of ambiguity resolution (cf. Fodor et al., 1996). As we saw earlier, it is a basic tenet of the referential theory that information about specific conversational context takes precedence over the use of real-world knowledge. In this experiment, we asked whether the query of the store of world knowledge can keep pace with the assimilation of semantic information contributed by the focus operator ONLY.

Another issue we addressed in this experiment concerned the origin of individual differences in profiles of sentence processing. Previous research by MacDonald et al. (1992) has established that individual differences in

working memory capacity constrain the ability of a subject to process ambiguous sentences. The authors interpret this finding as evidence that subjects with high memory spans can maintain multiple syntactic representations for ambiguous sentences, whereas subjects whose memory capacities are more limited can maintain only the representation that is most frequently used. Pearlmutter and MacDonald (1995) also found that high span subjects were more sensitive to the relative plausibility of alternative representations of an ambiguous phrase. On the basis of these findings, we are led to infer that in ambiguity resolution, subjects with a higher memory capacity may be more efficient in using diverse sources of information than those with more limited memory capacity. Arguably, rapid decision making is facilitated by maintaining alternative representations in memory, based on whatever information the parser has at its disposal. Therefore, individual differences in working memory may be correlated with differences in the time at which various sources of information in a sentence are used. A subject who finds it difficult to maintain alternative representations of an ambiguous sentence will have difficulties in resolving the ambiguity, especially if late-arriving information is critical to the decision. This subject will be expected to "look back" more frequently in reading. As noted earlier, it can be inferred from the principles of the referential theory that persons with limited memory capacity will have greater difficulty than those with a higher memory capacity in resolving ambiguities that require information about the relative plausibility of the alternative meanings of a sentence.

The test materials used in this experiment were the same as those used in Experiment 3. An example set is repeated here:

- The-VP *The man painted the doors with new brushes before the festival.*
- The-NP *The man painted the doors with large cracks before the festival.*
- Only-VP *The man painted only doors with new brushes before the festival.*
- Only-NP *The man painted only doors with large cracks before the festival.*

The predictions of the referential theory remain the same as those for Experiment 3; namely, that while "The-VP" sentences will not cause processing difficulty and "The-NP" sentences will, the reverse effect should be seen for sentences with ONLY. "Only-VP" sentences will prove difficult to read, because the focus operator ONLY will lead the parser to favour an NP-attachment analysis, resulting in an anomaly at the noun phrase of the prepositional phrase. On the other hand, the NP-attachment analysis of the PP is anticipated for "Only-NP" sentences, and no effect of anomaly should occur.

Method

Subjects. Thirty-two undergraduate students participated in the experiment, all of whom were native speakers of English. They were not informed of the purpose of the experiment. Individuals were selected who had uncorrected vision or wore soft contact lenses.

Materials. The same 20 sets of test sentences and 60 fillers used in Experiment 3 were used in this experiment. Four stimulus lists were generated. The four versions of each set of test sentences (The-VP, The-NP, Only-VP, Only-NP) were rotated through the four lists. Each list was tested on a different group of eight subjects. This experiment used a fully factorial repeated-measure design, with both THE/ONLY and VP/NP as within-subjects and within-items factors. Eight warm-up sentences preceded each stimulus list.

Equipment. The equipment and data analysis programs used in this experiment were the same as those employed in Experiment 2.

Procedure. The procedure was as in Experiment 2. In addition, however, a memory span test was administered following the reading test.¹⁶ In the memory span test, the subjects listened to recorded materials over headphones. Their task was to report in order the last word of each sentence in a set of spoken sentences when prompted by a non-verbal signal. They were also required to judge, after each sentence within a set, whether it was "True" or "False". The number of sentences in each set was gradually increased from two to five. The subjects were encouraged to say as many terminal words as they could, even if they were unsure about their order of occurrence. A total of 42 words was solicited and the subjects were ranked according to the number of words they correctly reported. Based on span length, the subjects were divided into two groups of 16 by a median split. "High span" versus "low span" was treated as a separate factor in the statistical analysis.

Results and Discussion

The analyses were based on recorded eye fixations at five regions of the test sentences, as defined in Experiment 3: Region 1 was the subject noun phrase; Region 2 contained the main verb; Region 3 contained the object noun phrase (preceded either by THE or ONLY) and the following preposition; Region 4 contained the object noun phrase of the prepositional phrase; and

¹⁶Daneman and Carpenter (1980) report a correlation between subjects' performance on measures of language comprehension and memory span.

Region 5 contained the rest of the sentence. Region 4 was the focus of interest, because attachment preferences are contingent upon the processing of the semantic content of this region.

Two-way ANOVAs were performed, testing the two types of pre-nominal word (THE/ONLY) and the two types of attachment (VP/NP), with first-pass residual reading time (RRT) and percent of regressions as the dependent variables. As was discussed in Experiment 2, RRT provides a more adequate measure than total fixation durations, since reading time comparisons are made between sentences that contain different words of variable length (especially at Region 4).

First-pass Reading Times. Collapsing across regions, ANOVAs performed on first-pass residual reading times revealed a significant main effect of THE/ONLY: sentences with ONLY yielded significantly longer reading times than those with THE [$F_1(1,31) = 5.51, P < 0.03$; $F_2(1,19) = 6.33, P < 0.01$]. There was no main effect of VP/NP ($P > 0.1$), nor was there a significant THE/ONLY \times VP/NP interaction ($P > 0.1$). Separate ANOVAs were performed on reading times at each region. A profile by region is shown in Fig. 7.

No significant effects were found at either Region 1 or Region 2. At Region 3, there was a significant main effect of THE/ONLY [$F_1(1,31) = 10.72, P < 0.01$; $F_2(1,19) = 10.97, P < 0.01$]. Reading times were significantly longer in sentences with ONLY. At Region 4, the effect of VP/NP approached significance [$F_1(1,31) = 4.09, P < 0.06$; $F_2(1,19) = 3.59, P < 0.08$], but there was no effect of THE/ONLY. The interaction of THE/ONLY \times VP/NP was significant in the analysis by subjects

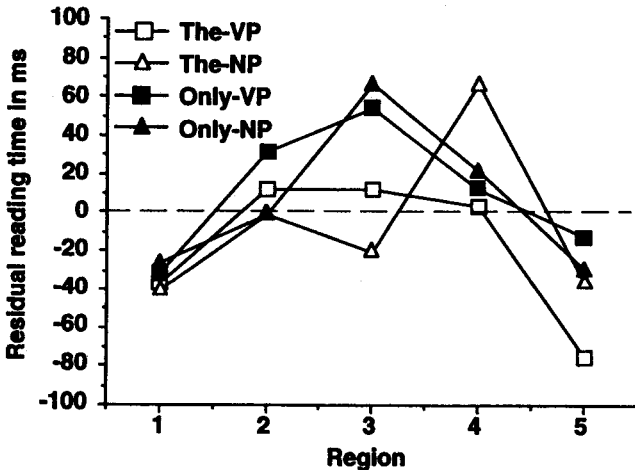


FIG. 7. Experiment 4: Mean first-pass residual reading time (RRT) at each region.

$[F_1(1,31) = 5.09, P < 0.04]$, but not in the analysis by items ($P > 0.1$). A planned comparison between sentence versions at Region 4 revealed that reading times for "The-NP" sentences were significantly longer than those for "The-VP" sentences [$F_1(1,31) = 10.01, P < 0.01; F_2(1,19) = 7.39, P < 0.02$]. There was no significant difference between the two versions with ONLY ($P > 0.1$). Region 5 revealed no significant effects.

In sum, first-pass reading times at Region 4, which is the region of interest, did not seem to meet the specific predictions of the referential theory on sentences with ONLY. Reading times on "Only-NP" sentences were *not* shorter than those on "Only-VP" sentences, as predicted by the referential theory. But neither were reading times for "Only-NP" sentences longer than those for "Only-VP" sentences, which would go against the referential theory.

Since the critical information needed to recover from a misanalysis of PP-attachment turns on *a priori* plausibility, its use is expected to be sensitive to individual differences in memory capacity. We therefore turned to the analyses which partitioned subjects according to their scores on the memory span test. As illustrated in Fig. 8, the reading-time profiles at Region 4 revealed an apparent discrepancy between the two subject groups. Compared with the pattern with all subjects combined (Fig. 8a), high-span subjects (Fig. 8b) acted in accordance with the predictions of the referential theory, whereas low-span subjects (Fig. 8c) did not.

For high-span subjects, there was a significant interaction of THE/ONLY \times VP/NP in the analysis by subjects [$F_1(1,15) = 11.78, P < 0.01$]; the analysis by items approached significance [$F_2(1,19) = 3.78, P < 0.07$]. As expected, a significant delay at Region 4 existed in reading "The-NP" sentences compared with reading the "The-VP" sentences [$F_1(1,15) = 6.32, P < 0.02; F_2(1,19) = 4.32, P < 0.05$]. However, reading times on "Only-VP" sentences were longer than those on "Only-NP" sentences, producing an effect that approached significance in the analysis by subjects [$F_1(1,15) = 3.65, P < 0.08$], though not in the analysis by items ($P > 0.1$).

For low-span subjects, there was a main effect of VP/NP [$F_1(1,15) = 4.61, P < 0.05; F_2(1,19) = 7.01, P < 0.02$]: NP-attachment sentences were read more slowly than VP-attachment ones at Region 4. The main effect of THE/ONLY was significant in the analysis by items [$F_2(1,19) = 5.31, P < 0.04$] and approached significance in the analysis by subjects [$F_1(1,15) = 3.92, P < 0.07$]. The interaction between THE/ONLY and VP/NP was not significant ($P > 0.1$). Note, however, that the reading times on the "Only-NP" sentences were about the same as those on the "The-VP" sentences, and both were markedly different from reading times on the "The-NP" sentences.

The discrepancy between high-span and low-span subjects on first-pass reading time profiles may be attributable to the way these two groups use

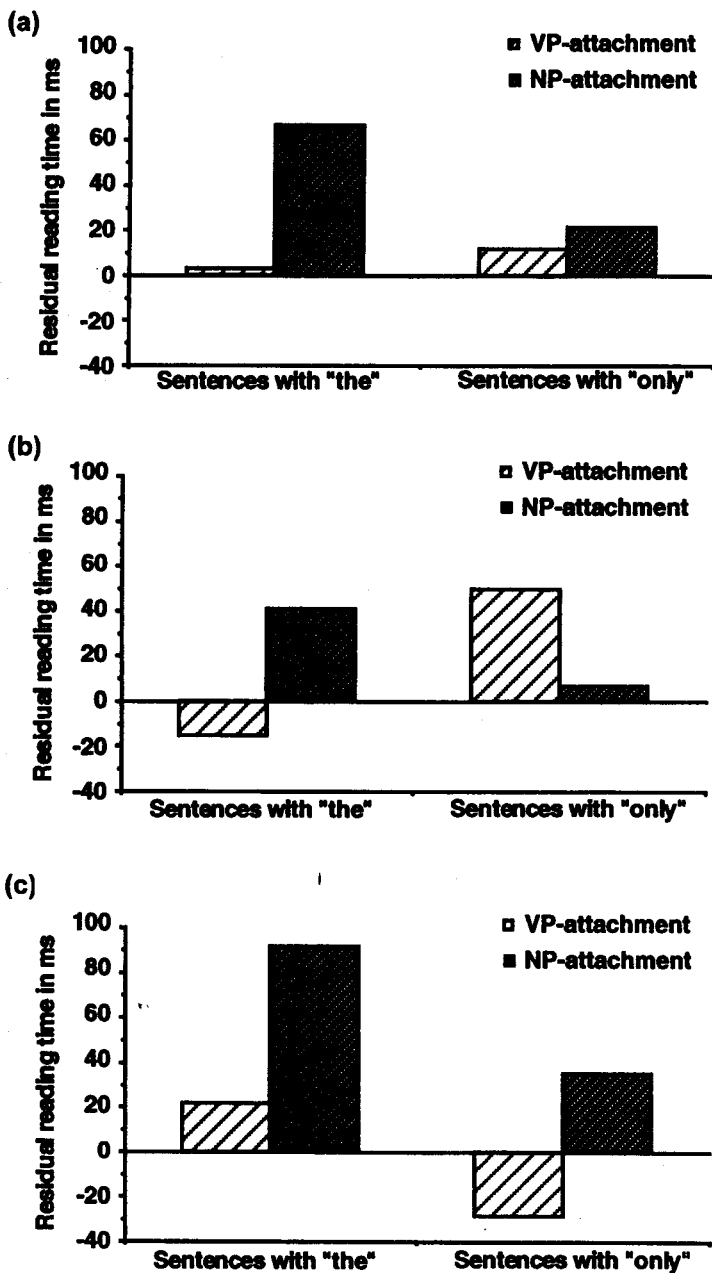


FIG. 8. Experiment 4: Mean first-pass residual reading time (RRT) at Region 4 by all subjects (a), high-span subjects (b) and low-span subjects (c).

general world knowledge in reading Region 4. As we indicated, appeal to *a priori* plausibility is required in order to make and revise attachment decisions. It is useful, therefore, to look at the incidence of regressive eye movements from Region 4, which may indicate reading difficulties.

Incidence of Regression. As with the profiles of reading times, the pattern of regressions did not yield any systematic effect in the undifferentiated data set, as shown in the region-by-region profile (Fig. 9). However, distinctive patterns of regression did occur for high- and low-span groups at Region 4, as depicted in Fig. 10. For the high-span subjects (Fig. 10b), there was no systematic effect ($P > 0.1$ in all analyses). These subjects seemed to have recovered from misanalysis in the course of their first-pass reading and, presumably because of this, they showed little difference in regression patterns among the different versions of the test sentences. By contrast, the low-span subjects (Fig. 10c) displayed dissociated patterns. There was a main effect of VP/NP [$F_1(1,15) = 9.28$, $P < 0.01$; $F_2(1,19) = 8.41$, $P < 0.01$]; that is, VP-attachment sentences induced significantly more regressions than NP-attachment ones. There was no main effect of THE/ONLY ($P > 0.1$), but the interaction between VP/NP and THE/ONLY was significant [$F_1(1,15) = 7.38$, $P < 0.02$; $F_2(1,19) = 5.43$, $P < 0.03$]. A pairwise comparison revealed that low-span subjects made significantly more regressions on the "Only-VP" sentences than on the "Only-NP" sentences [$F_1(1,15) = 26.42$, $P < 0.01$; $F_2(1,19) = 11.34$, $P < 0.01$].

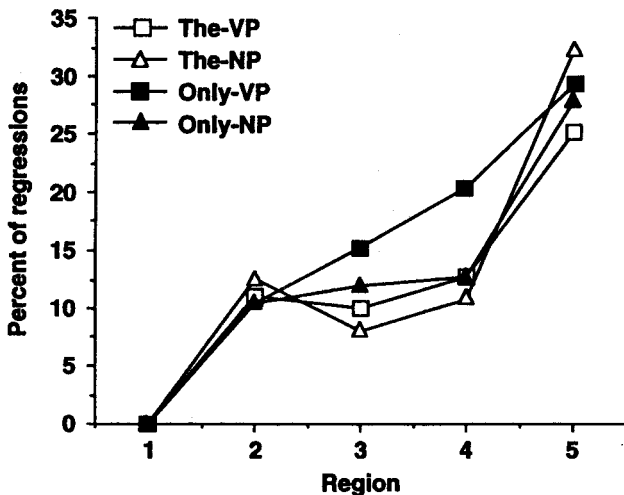


FIG. 9. Experiment 4: Percent of regressions at each region.

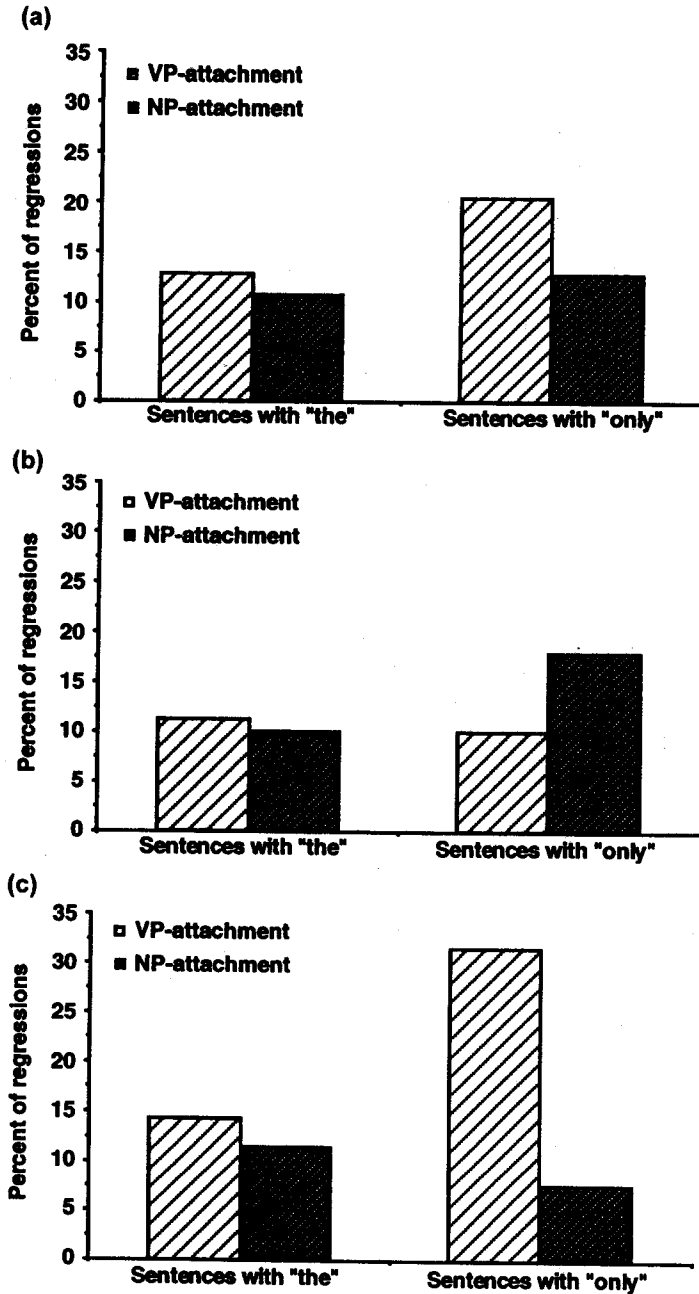


FIG. 10. Experiment 4: Percent of regressions at Region 4 by all subjects (a), high-span subjects (b) and low-span subjects (c).

The following is our explanation of the combined results of first-pass reading times and the incidence of regressions. We maintain that the divergent processing patterns by the subject groups resulted from discrepancies in the time-course of application of plausibility information associated with individual differences in memory capacity. Let us look first at the sentences with the definite determiner, THE. As expected on the referential theory, it turned out to be relatively easy for both high- and low-span subjects to construct a mental representation for these sentences. Having pursued the referentially simple VP-attachment analysis of both "The-VP" and "The-NP" sentences, both groups of subjects detected the implausibility of the noun phrase *large cracks* as a modifier of the verb in "The-NP" sentences. Reanalysis was therefore initiated, triggering the long first-pass reading times. That the relatively simple recovery process was generally accomplished on-line is also attested by the relatively low incidence of regressions by either high-span or low-span subjects.

Differences in the pattern of responses of the high- and low-span subjects were observed, however, in processing sentences with ONLY. The high-span subjects were apparently able to recover on-line from the anomaly in "Only-VP" sentences, where the PP *with new brushes* was not a plausible modifier for the NP *doors*. The long first-pass reading times at Region 4 on "Only-VP" sentences by these subjects, coupled with the absence of appreciable regressions from that region, indicated that this group successfully reanalysed the sentence without looking back to earlier regions. The low-span subjects had relatively fast reading times at Region 4, but made a greater number of regressive eye movements than the high-span group. This pattern of results suggests that the low-span subjects had difficulties recovering on-line from an initial misanalysis; hence, they were compelled to review material that they had read earlier.

These findings related to memory span can be readily accommodated within the referential theory. Recall that in order to interpret the sentences with ONLY, a contrast set must be constructed. Failing to locate the contrast set within the current mental model of the conversational context, where it is expected to be found, the perceiver's next option is to examine the incoming string of words to see if a contrast set can be motivated on this basis. (Failing that, the alternative is to conjure up a contrast set. However, as we saw earlier, this would require unmotivated additions to the mental model, and therefore this option is dispreferred, according to the referential theory.) Attempts to locate the contrast set in the incoming string require the perceiver to attach the ensuing PP as a modifier of the preceding NP containing ONLY. Having made the decision to attach the PP to the preceding NP, the parser continues its search for the contrast set within the PP. It turns out, however, that when the noun phrase of the PP is encountered in the "Only-VP" version of the test sentences in Experiment 4,

the contrast set that presents itself is semantically anomalous. To achieve a semantically coherent interpretation, not only must the parser revise the structure of the earlier portion of the sentence, but it must also pursue its last-resort option for constructing a contrast set by making one up from scratch. Not surprisingly, the combined effort in making these computations proves highly demanding of memory resources and, therefore, pushes apart groups of subjects who differ in memory span.

The rapid first-pass reading by the low-span subjects at Region 4 on "Only-VP" sentences suggests that, although they detected the pragmatic incompatibility of the noun *brushes* with their initial analysis, their memory resources had already been exhausted, thus triggering repeated regressive eye movements. Frequent resort to looking back would seem to imply that low-span subjects are unable to use real-world knowledge effectively in on-line recovery from a misanalysis.¹⁷ The referential theory gives a parsimonious explanation for this complementary pattern of eye fixations and regressions—the use of semantic information (carried by the focus operator ONLY) makes relatively light demands on memory resources. This explains why there were no significant differences between subject groups on the "Only-NP" sentences and no differences in any region preceding the anomaly in the "Only-VP" sentences. It is apparent that in the presence of ONLY, all subjects pursued the NP-attachment analysis, following the principle of parsimony. It was the specific requirement that plausibility information be invoked to disconfirm the initial parsing decision in the "Only-VP" sentences that distinguished subjects with different memory capacities. [See Crain, Shankweiler, Macaruso and Bar-Shalom (1990) for discussion of other related effects of working memory differences for sentence processing.]

In keeping with the preceding experiments, Experiment 4 also provides support for the referential theory. What is new is that the findings indicate that while the semantic information carried by ONLY is used on-line in resolving ambiguities, the use of plausibility information may not be. If and when plausibility is used depends upon the memory capacity of the individual. For high-span subjects, plausibility information seems to be used rapidly to recover from a misanalysis, whereas its use appears to be delayed by low-span subjects. How quickly it is used depends on two factors that can be identified, the first of which is the memory capacity of the reader. An

¹⁷The alternative is to suppose that the low-span subjects followed the Minimal Attachment strategy and did not initially pursue the NP-attachment analysis of prepositional phrases in sentences with ONLY. On this account of the findings, the significant number of regressions by these subjects to "Only-VP" sentences is unexplained. Mitchell (pers. comm.) suggests that the garden path model can account for this kind of result if it stipulates that early corrections occur when resources are available. However, there is no evidence yet that a revision was attempted after a brief garden path effect within the prepositional phrase.

additional factor is the point within the sentence at which plausibility information is available for use. If information pertaining to plausibility is encountered before the point of ambiguity, it can be effective in resolving local ambiguities that are encountered subsequently (Trueswell et al., 1994). According to the referential theory, however, plausibility is used only to adjudicate among competing alternative partial structural analyses. Plausibility does not compete with more specific information about the conversational context if the latter is operative. The present findings conform to this expectation. Apparently, if information about plausibility is encountered *after* the point of ambiguity, then even though this information may be available to the parser, its implementation in decision making may be delayed. In contrast, semantic (focus) information is rapidly used to adjudicate among competing structural analyses. The distinction between availability and use of plausibility information is consistent with the Modularity Hypothesis (see Fodor et al., 1996).

GENERAL DISCUSSION

We have presented findings from four experiments that were designed to test predictions of the referential theory, and to assess its explanatory scope in the context of the current debate on the relative timing of the use of the varieties of non-syntactic information in on-line sentence processing. The first experiment confirmed a major prediction of the referential theory, namely that referential principles are applied in resolving ambiguities involving main-verb/reduced-relative-clause analyses. A word-by-word reading test revealed that a simple substitution of ONLY for THE in the subject noun phrase substantially reduced garden path effects. The findings provide circumstantial evidence for the contention that ambiguity resolution is influenced by properties of discourse representations that are assigned to the alternative analyses of an ambiguous phrase.

Experiment 2 confirmed the results of Experiment 1, using records of subjects' eye movements. First-pass reading times on garden path sentences containing the focus operator ONLY did not differ to a significant degree from those of their unambiguous control sentences in the disambiguating region. There was a significant difference at this region, however, when garden path sentences beginning with the definite determiner THE were compared with appropriate controls. The results were interpreted as evidence for the rapid on-line use of the semantic contribution carried by the focus operator ONLY.

It should be noted that in Experiment 1 all ambiguous sentences, with THE or with ONLY, produced significantly more errors than their respective unambiguous controls. These sentences also produced more

regressions in Experiment 2. In our view, this effect is the result of parallel processing. Because more than one representation is computed within the ambiguous region, subjects occasionally select the one that is inconsistent with the discourse context. However, the presence of ONLY was sufficient to promote the reduced relative clause analysis on the majority of trials. Compared with the unambiguous control sentences, there was about a 20% increase in overall errors (in Experiment 1) and a 4% increase in regressions (at Region 5 in Experiment 2) for ambiguous sentences with ONLY. In contrast, for ambiguous sentences with THE, there was a 46% increase in overall errors in Experiment 1 and a 9% increase in regressions in Experiment 2.

Taken together, the findings of Experiments 1 and 2 provide evidence for the influence of referential content of noun phrases in the resolution of garden path sentences such as the infamous *The horse raced past the barn fell*. The findings of Experiments 1 and 2 support the joint predictions of the referential theory that semantic (focus) information is used to decide among competing syntactic analyses (Experiments 1 and 2) and that this information is used on-line (Experiment 2). These findings would not be anticipated and are difficult to explain on any account of ambiguity resolution that ignores the referential properties of sentences in the initial decisions made by the parser. Although some researchers have shown themselves willing to acknowledge the involvement of referential factors in resolving ambiguities involving prepositional phrase attachments, it is probably generally believed that the effects of structurally based strategies such as Minimal Attachment predominate in sentences with a main-verb/reduced-relative-clause ambiguity (see, e.g. Tanenhaus & Trueswell, 1995). To the contrary, the findings of the present research show referential effects to be as strong, if not stronger, in the resolution of the main-verb/reduced-relative-clause ambiguity, than in resolution of ambiguities in the attachment of prepositional phrases, which was investigated in Experiments 3 and 4.

Returning to the method of word-by-word reading in Experiment 3, we found that subjects used information about the *a priori* plausibility of alternative representations of an ambiguous phrase in arriving at their preferred attachment of a prepositional phrase. A simple substitution of ONLY for THE in a noun phrase followed by a prepositional phrase changed the parser's decisions as to where to attach the prepositional phrase within an ambiguous sentence. The subsequent effects were seen to depend on subjects' use of plausibility information contributed by the head noun of the prepositional phrase. The results were interpreted as further confirmation for the referential theory, which, as we noted, maintains the view that the parser bases its initial decisions on semantic/referential principles.

Experiment 4 was designed to estimate the time at which plausibility information is used by the sentence processing system. Adopting the standpoint of the referential theory, we anticipated that the use of this source of information would be pre-empted by referential factors contributed by the focus operator ONLY. This expectation is based on the principle of *a priori* plausibility, which maintains that plausibility information is invoked only if the semantic/referential content of a sentence does not offer a sufficient basis for selecting among the competing interpretations. The technique of eye movement recording was used to test this prediction using the same sentences as in Experiment 3. In analysing the results of Experiment 4, subjects were grouped according to working memory span, to ask under what conditions the resolution of ambiguities would show individual differences in memory. The results suggested that the time-course of the use of plausibility information did in fact co-vary with memory span. Such information was used rapidly by individuals who had relatively high memory spans, but its use was delayed and was probably less effectively used by low-span individuals. These results therefore support the view that the ability to take advantage of *a priori* plausibility is highly resource-dependent. Moreover, if the relevant information is encountered after the onset of ambiguity, as is the case with ambiguities of prepositional phrase attachment, its value for ambiguity resolution is more limited than if it is encountered prior to the onset of ambiguity. We take this result, too, as support for the referential theory, according to which semantic (focus) information is used on-line in the construction of discourse representations, with plausibility information exerting its influence only with ambiguities that remain unresolved after semantic principles have been applied.

Let us now take stock. There is a degree of consensus among researchers concerned with sentence processing that some kinds of non-syntactic information are very rapidly assimilated by the parser, as when it goes about the business of ambiguity resolution. There are simply too many empirical facts in the literature that point in this direction to make a denial plausible. But, as we have emphasised, non-syntactic influences are not all of the same kind, and researchers who adhere to the referential theory characteristically differ from adherents to the garden path model and the constraint satisfaction model, in how and when various kinds of non-syntactic information are incorporated by the parser. The referential theory makes a case for the primacy of discourse considerations. An integrated set of experiments was designed to find out how well the referential theory could predict and explain findings from experimental paradigms that permit comparison of the times at which discourse principles and factors governing real-world knowledge become operative. Our findings consistently confirmed that discourse principles are operative on-line in resolution of two kinds of structural ambiguities, and that they take precedence over

plausibility. We further clarified the costs associated with the use of these two kinds of information, showing that the parser's appeals to real-world knowledge, but probably not its application of discourse principles, are heavily resource-consuming and are dependent on processing resources that vary greatly among individuals. We therefore maintain that the findings form a coherent package that can adequately be explained by the referential theory. Can the alternative accounts of sentence processing deal as well with findings such as these? We leave it for future research to decide.

Manuscript received July 1994

Revised manuscript received January 1996

REFERENCES

- Altmann, G.T.M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.
- Altmann, G.T.M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, 31, 685-712.
- Altmann, G.T.M., Garnham, A., & Henstra, J.A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 209-216.
- Beach, C. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, 30, 644-663.
- Bever, T.G. (1970). The cognitive basis for linguistic structure. In J.R. Hayes (Ed.), *Cognition and the development of language*. New York: John Wiley.
- Boland, J.E., Tanenhaus, M.K., & Garnsey, S.M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, 29, 413-432.
- Britt, M.A. (1991). *The role of referential uniqueness and argument structure in parsing prepositional phrases*. PhD dissertation, University of Pittsburgh, Pittsburgh, PA.
- Britt, M.A., Perfetti, C.A., Garrod, S., & Rayner, K. (1992). Parsing and discourse: Context effects and their limits. *Journal of Memory and Language*, 31, 293-314.
- Clifton, C., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, 4, 77-103.
- Crain, S., & Ni, W. (1991). Parsermony. In *Proceedings of Texas Conference on Cognitive Science*, Vol. X, University of Texas, Austin, TX, April.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D.R. Dowty, L. Karttunen, & A.M. Zwicky (Eds), *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.
- Crain, S., Shankweiler, D., Macaruso, P., & Bar-Shalom, E. (1990). Working memory and comprehension of spoken sentences. Investigations of children with reading disorders. In G. Vallar & T. Shallice (Eds), *Neuropsychological impairments of short-term memory*. Cambridge: Cambridge University Press.
- Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.

- Fodor, J.D., & Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research*, 23, 407-434.
- Fodor, J.D., Ni, W., Crain, S., & Shankweiler, D. (1996). Tasks and timing in the perception of linguistic anomaly. *Journal of Psycholinguistic Research*, 25, 25-57.
- Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 203-218.
- Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. PhD dissertation, University of Connecticut. Distributed by the Indiana University Linguistics Club, Bloomington, IN.
- Frazier, L. (1994). Sentences (re)analysis. Paper presented at the 7th CUNY Conference on Human Sentence Processing, CUNY, New York, March.
- Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- Gibson, E. (in press). *Memory limitations and sentence processing breakdown*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1972). *Semantic interpretation of generative grammar*. Cambridge, MA: MIT Press.
- Juliano, C., & Tanenhaus, M.K. (1993). Contingent frequency effects in syntactic ambiguity resolution. In *Proceedings of the 15th Annual Conference of the Cognitive Society*, pp. 593-598. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Juliano, C., & Tanenhaus, M.K. (1994). A constraint-based lexicalist account of the subject/object attachment preference. *Journal of Psycholinguistic Research*, 23, 459-471.
- Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1, 181-194.
- Kennedy, A., & Murray, W.S. (1984). Inspection times for words in syntactically ambiguous sentences under three presentation conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 833-849.
- Krifka, M. (1991). A compositional semantics for multiple focus constructions. *Linguistische Berichte, Sonderheft*, 4. Also in *Proceedings of Semantics and Linguistic Theory (SALT) I*, Cornell Working Papers II.
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339-359.
- MacDonald, M.C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 32, 692-715.
- MacDonald, M.C., Just, M.A., & Carpenter, P.A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56-98.
- MacDonald, M.C., Pearlmutter, N., & Seidenberg, M.S. (1994). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*, pp. 123-153. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Marslen-Wilson, W.D., Tyler, L.K., Warren, P., Grenier, P., & Lee, C.S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology*, 45A, 73-87.
- Merlo, P. (1994). A corpus-based analysis of verb continuation frequencies for syntactic processing. *Journal of Psycholinguistic Research*, 23, 435-457.
- Mitchell, D.C., Corley, M.M.B., & Garnham, A. (1992). Effects of context in human sentence

- parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 69–88.
- Murray, W., & Liversedge, S. (1994). Referential context and syntactic processing. In C. Clifton, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*, pp. 359–388. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Nagel, H.N., & Shapiro, L.P. (1994). Prosody and the processing of filler–gap sentences. *Journal of Psycholinguistic Research*, 23, 473–485.
- Ni, W., & Crain, S. (1990). How to resolve structural ambiguities. In *Proceedings of the North Eastern Linguistic Society*, 20/2, pp. 414–427. Amherst, MA: University of Massachusetts.
- Pearlmutter, N., & MacDonald, M.C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, 34, 521–542.
- Perfetti, C.A. (1991). The cooperative language processors: Semantic influence in autonomous syntax. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds), *Comprehension processes in reading*, pp. 205–230. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, 90, 2956–2970.
- Rayner, K. (1993). Eye movements in reading: Recent developments. *Current Directions in Psychological Science*, 2, 81–85.
- Rayner, K., & Morris, R. (1991). Comprehension processes in reading ambiguous sentences: Reflections from eye movements. In G.B. Simpson (Ed.), *Understanding word and sentence*. Amsterdam: North-Holland.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction between syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358–374.
- Rayner, K., Sereno, S.C., Morris, R.K., Schmauder, A.R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 3/4, 21–49.
- Rayner, K., Garrod, S., & Perfetti, C.A. (1992). Discourse influence during parsing is delayed. *Cognition*, 45, 109–139.
- Rooth, M. (1985). *Association with focus*. Unpublished PhD dissertation, University of Massachusetts, Amherst, MA.
- Sedivy, J.C., & Spivey-Knowlton, M.J. (1994). The use of structural, lexical, and pragmatic information in parsing attachment ambiguities. In C. Clifton, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*, pp. 389–413. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Soames, S. (1982). How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry*, 13, 483–545.
- Speer, S.R., Crowder, R.G., & Thomas, L.M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language*, 32, 336–358.
- Spivey-Knowlton, M.J., & Sedivy, J.C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227–267.
- Spivey-Knowlton, M.J., & Tanenhaus, M.K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*, pp. 415–439. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Spivey-Knowlton, M.J., Trueswell, J.C., & Tanenhaus, M.K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, 47, 276–309.
- Stalnaker, R. (1974). Pragmatic presuppositions. In M.K. Munitz & P.K. Unger (Eds), *Semantics and philosophy*. New York: New York University Press.
- Tabossi, P., Spivey-Knowlton, M.J., McRae, K., & Tanenhaus, M.K. (1994). Semantic effects

- on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umiltà & M. Moscovitch (Eds), *Attention and performance XV*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Tanenhaus, M.K., & Trueswell, J.C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds), *Handbook of perception and cognition, Vol. 11: Speech, language and communication*. New York: Academic Press.
- Tanenhaus, M.K., Garnsey, S.M., & Boland, J.E. (1991). Combinatory lexical information and language comprehension. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.
- Taraban, R., & McClelland, J.L. (1990). Parsing and comprehension: A multiple-constraint view. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds), *Comprehension processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Trueswell, J.C., & Tanenhaus, M.K. (1994). Toward a constraint-based lexicalist approach to syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds), *Perspectives on sentence processing*, pp. 155–179. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285–318.

APPENDIX 1

*Test materials for Experiments 1 and 2
(Experiment 2 used a subset of these sentences with some modifications)*

Test sentences

(Each sentence has four types: The-amb, Only-amb, The-adj-amb, Only-adj-amb)

1. The/Only (smart) people taught new math will pass the test.
2. The/Only (wealthy) businessmen loaned money at low interest were told to record their expenses.
3. The/Only (brave) soldiers killed in the line of duty were mourned.
4. The/Only (frequent) visitors issued passes used them to leave on weekends.
5. The/Only (protesting) generals presented copies of the report blamed the government for cutting defense spending.
6. The/Only (greedy) wives left during the first month of marriage demanded alimony at the hearing.
7. The/Only (bold) students pushed into the flow of traffic got badly hurt.
8. The/Only (inexperienced) boxers punched hard in the early rounds were unable to finish the bout.
9. The/Only (poor) shopkeepers charged for repairs thought they were being cheated.
10. The/Only (old) union workers warned about possible layoffs picketed the company's main office.
11. The/Only (trained) paratroopers dropped into the dense jungle were captured by guerrilla forces.
12. The/Only (gourmet) chefs asked to have food ready refused to do so.
13. The/Only (cranky) children rocked to sleep soon woke up.
14. The/Only (new) homeowners hurt because of the increase in taxes decided to join forces against the administration.
15. The/Only (junior) pilots delivered the warning notice went out on strike.
16. The/Only (senior) doctors stopped while driving to work were not fined by the police.
17. The/Only (skinny) clowns tripped during the skit remained on the ground until the end of the performance.

18. The/Only (new) owners offered tempting food gulped it down.
19. The/Only (dishonest) students furnished answers before the exam received high marks.
20. The/Only (adventurous) swimmers drowned in the icy lake were not found until the spring.
21. The/Only (social) organizations donated emergency supplies helped to provide shelters to the earthquake victims.
22. The/Only (irate) listeners called during prime time programs didn't answer their phones.
23. The/Only (fishing) ships salvaged during the hurricane returned to the dock.
24. The/Only (fresh) turkeys roasted for under three hours were ready in time for the banquet.
25. The/Only (trained) social workers lectured about the dangers of smoking tried to help their own friends quit.
26. The/Only (chocolate) cookies baked in the brick ovens were sold at the carnival.
27. The/Only (big) boulders rolled down the mountain stopped the approaching trucks.
28. The/Only (crooked) dealers sold forgeries went straight to the police.
29. The/Only (heavy) boats floated down many rivers failed to get over the rapids.
30. The/Only (senior) senators elected to hold dinners for fundraisers were allowed to miss the vote.
31. The/Only (famous) actors paid for the entertainment performed in an outdoor theater.
32. The/Only (retired) men delivered junk mail threw it in the trash.

Control sentences

(Each sentence has four types: The-unamb, Only-unamb, The-adj-unamb, Only-adj-unamb)

1. The/Only (young) hunters bitten by ticks worried about getting lime disease.
2. The/Only (dangerous) criminals taken into custody at the riot were not released the next day.
3. The/Only (homeless) people shaken by the earthquake feared an aftershock.
4. The/Only (miserly) jewelers given huge diamonds cut them into small stones.
5. The/Only (pretty) models drawn by the illustrator were used for a magazine cover.
6. The/Only (candidate) managers chosen by the company answered every question at the interview.
7. The/Only (white) strangers seen at the time of the robbery had scars.
8. The/Only (indecent) scientists proven to be incorrect faked their data.
9. The/Only (short) hymns sung with great emotion were worth listening to.
10. The/Only (long) speeches written by the candidate were hard to understand.
11. The/Only (blue) vans stolen from the parking lot were found in a back alley.
12. The/Only (small) crops grown by farmers were damaged by the frost.
13. The/Only (strong) horses ridden past the finish line were given the prizes.
14. The/Only (sports) cars driven at high speeds were found to be defective.
15. The/Only (dance) shoes worn by the famous actress were put on display.
16. The/Only (fried) poultry eaten at the fair gave people an upset stomach.

APPENDIX 2

Test materials for Experiments 3 and 4

(Each sentence has four types: The-VP, The-NP, Only-VP, Only-NP)

1. The burglar blew open (the/only) safes with (high quality dynamite/high quality diamond) and fled.
2. The historian studied (the/only the) maps with (a magnifying glass/large print) carefully.
3. The cleaners wiped (the/only the) windows with (a heavy durable cloth/a heavy covering of dirt) every day.

4. The workman opened (the/only the) valves without (much effort/much rust) during the flood.
5. The monkey tried to eat (the/only the) bananas with (silverware/bruises) during the show.
6. The dressmaker cut (the/only) material with (unusual scissors/unusual patterns) for the wedding.
7. The little girl cut (the/only the) oranges with (pairing knives/thin skins) before dinner.
8. The drunk smashed (the/only the) windows with (an empty bottle/stained glass) last night.
9. The vet tranquilized (the/only the) tigers with (a dart gun/bad tempers) before treating them.
10. The tribesmen killed (the/only) lions with (poison arrows/sharp teeth) as part of the ritual.
11. The company demolished (the/only) buildings with (huge bulldozers/cement foundations) last weekend.
12. The craftsman stripped (the/only) cabinets with (paint stripper/brass hinges) for the building.
13. The doctor examined (the/only) women with (a cold stethoscope/high fever) before the surgery.
14. The secretary typed up (the/only) reports with (an IBM typewriter/few diagrams) after lunch.
15. The thief opened (the/only the) doors with (a credit card/faulty locks) during the robbery.
16. The woman repaired (the/only the) socks with (some thread/large holes) during the TV show.
17. The fireman broke (the/only) windows with (the ax/rusty hinges) during the fire.
18. The gardener cut down (the/only the) trees with (the chainsaw/the disease) last fall.
19. The detective was watching (the/only) women with (binoculars/straw hats) at the station.
20. The man decided to paint (the/only) doors with (new brushes/large cracks) for the festival.