

Probabilistic Constraints and Syntactic Ambiguity Resolution

Maryellen C. MacDonald

Departments of Linguistics and Psychology, University of Southern California, Los Angeles, California, USA

Natural languages contain probabilistic constraints that influence the resolution of ambiguities. Current models of sentence processing agree that probabilistic constraints affect syntactic ambiguity resolution, but there has been little investigation of the constraints themselves—what they are, how they differ in their effects on processing, and how they interact with one another. Three different types of probabilistic constraints were investigated: “pre-ambiguity” plausibility information, information about verb argument structure frequencies, and “post-ambiguity” constraints that arrive after the introduction of the ambiguity but prior to its disambiguation. Reading times for syntactically ambiguous sentences were compared to reading times for unambiguous controls in three self-paced reading experiments. All three kinds of constraints were found to be helpful, and when several constraints converged, ambiguity resolution was facilitated compared to when constraints conflicted. The importance of these constraint interactions for ambiguity resolution models is discussed.

INTRODUCTION

In a review of a documentary in the *New York Times*, film critic Janet Maslin wrote the following sentences: “The homeless people interviewed in the film are exceptionally calm, articulate and intelligent by any standard. Even the former psychiatric patients heard here sound unusually

Requests for reprints should be addressed to Maryellen C. MacDonald, Hedco Neuroscience Building, University of Southern California, Los Angeles, California 90089-2520, USA. E-mail: mcm@gizmo.usc.edu.

This research was supported by Grant No. DBS-9120415 from the National Science Foundation, an award from the Zumberge Research Innovation Fund at the University of Southern California, and by BRSF No. 1711 awarded to the Massachusetts Institute of Technology. I thank Jennifer Freed, Paula O'Day, Karen Katzman, Greta Buck and Alex Rosen for help in developing stimuli and running subjects, and Mark Seidenberg, Neal Pearlmuter, Shari Speer, Gerry Altmann, Don Mitchell, Tom Bever and an anonymous reviewer for helpful discussions and comments on earlier drafts of the manuscript.

sane" (Maslin, 1991, p. C8). Maslin's two sentences are interesting from a psycholinguistic perspective because they each contain a temporary "main verb/reduced relative" (MV/RR) ambiguity that is resolved in favour of a syntactically complex and relatively infrequent "reduced relative clause" structure. In this structure, the initial verbs in the two sentences, *interviewed* and *heard*, must not be interpreted as the past-tense main verbs of their sentences but rather as past-participles that modify the subject nouns, yielding the meanings *people who were interviewed* and *patients who were heard*. Bever's (1970) well-known "garden-path" sentence *The horse raced past the barn fell* has exactly this structure—*raced* is a past-participle modifying *horse*. It is quite striking that Bever's example is much more difficult to comprehend than are Maslin's sentences, despite the fact that all three sentences contain the identical syntactic ambiguity with the identical resolution. This paper investigates the factors that cause comprehension to fail or succeed for ambiguous sentences such as these. The identification of factors that affect comprehension of syntactic ambiguities furthers the development of a general theory of ambiguity resolution in human language comprehension.

Bever (1970) noted the exceptional difficulty of his *horse raced* sentence in comparison to other examples of this ambiguity, and he suggested that differences in processing difficulty for different sentences were the result of comprehenders' application of different strategies. More recently, other theories of syntactic ambiguity resolution have offered two alternative accounts of why the Maslin sentences are more easily comprehended than the Bever sentence. The first alternative is provided by the garden-path model developed by Frazier and her colleagues (Frazier, 1987; Rayner, Carlson, & Frazier, 1983). In this model, both the Maslin and Bever sentences would initially be misanalysed by the syntactic processor or *parser*, which has access only to major lexical category information (e.g. Determiner, Noun) and constructs only one syntactic structure at a time, guided by phrase structure rules and simplicity metrics. The difference between the easy and difficult sentences rests in the *thematic processor*, which operates at a second stage of analysis. This processor is hypothesised to have access to more than syntactic information, so that it "presumably uses thematic role preferences, discourse context, and world knowledge to choose the preferred analysis of the sentence. It considers, in parallel, all possible assignments of arguments and adjuncts to syntactic positions within syntactic domains defined by the syntactic processor, and proposes attractive alternative assignments to the syntactic processor" (Clifton & Ferreira, 1989, p. 87). This model suggests that there is something about the Maslin sentences that allows the thematic processor to find the correct interpretation and redirect the parser, and there is something about the Bever sentence that causes the thematic processor to continue with the

misanalysis that the syntactic processor initiated. It is not clear what these differences between the sentences are, how the thematic processor finds and uses them, nor at what point during the course of processing the feedback to the parser is provided, because, as Clifton and Ferreira (1989) note, there has been little investigation of the proposed thematic processor.

The second account is provided by a number of interactive or "constraint-based" models of sentence processing (Altmann & Steedman, 1988; MacDonald, Pearlmutter, & Seidenberg, submitted; McClelland, St. John, & Taraban, 1989; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993). Although the specific models differ in detail, all of them hypothesise that ambiguity resolution is a continuous process and is not divided into two temporally distinct stages as in the garden-path model. On this view, as Steedman and Altmann (1989) have noted, the whole of the ambiguity resolution process could look much like Clifton and Ferreira's (1989) description of the thematic processor, in that probabilistic lexical and discourse information constrain which alternative interpretation is preferred. Syntactic information provides an additional source of constraint and is not assigned to a special first stage, unlike in the garden-path model.

These alternative hypotheses have been assessed in a number of studies in which non-syntactic information was manipulated in ambiguous sentences in order to determine whether it could indeed allow the comprehender to avoid the initial misanalysis that is predicted by the garden-path model. The results to date have been mixed, with some studies reporting evidence for a first-stage parser that is oblivious to non-syntactic information (Britt, Perfetti, Garrod, & Rayner, 1992; Ferreira & Clifton, 1986; Mitchell, Corley, & Garnham, 1992; Rayner et al., 1983; Rayner, Garrod, & Perfetti, 1992), and others reporting evidence against the existence of the first-stage parser (Altmann, Garnham, & Dennis, 1992; Altmann, Garnham, & Henstra, in press; Altmann & Steedman, 1988; Pearlmutter & MacDonald, 1992; Spivey-Knowlton et al., 1993; Taraban & McClelland, 1988; Trueswell, Tanenhaus, & Garnsey, in press). One reason for these mixed results is the underspecification of the kinds of information that may constrain ambiguity resolution and how the constraint-satisfaction mechanisms work, within both the constraint-based models and the garden-path model. As a consequence, there are few experimental results that could not be accommodated within either type of model. For example, a finding in which non-syntactic information does not appear to influence the initial stages of ambiguity resolution, generally taken as evidence for the garden-path model, could be handled by a purely constraint-based model with an appeal to constraint strength—the non-syntactic constraints that were manipulated in a particular experiment were simply too weak or confounded with other factors to have an early effect

(for arguments of this sort, see Altmann et al., in press; Taraban & McClelland, 1988; Trueswell et al., in press; Tyler, 1989). Similarly, a result that appears to support constraint-based models, that helpful non-syntactic information removes all effects of ambiguity, could be handled by the garden-path model by claiming that the thematic processor corrects the parser's misanalysis so quickly that it is impossible to detect the initial misanalysis in reading times. This claim has been offered for both self-paced reading measures (e.g. Ferreira & Clifton, 1986; Rayner et al., 1992) and for eyetracking (Clifton & Ferreira, 1989).

One way to move beyond this stalemate might be to further specify how probabilistic constraints affect ambiguity resolution. Indeed, any progress in this domain would make either type of model more complete and more testable, because research to date has focused on the time course issue—whether or not a first-stage parser operates ahead of probabilistic constraints—at the expense of in-depth investigation of the constraints themselves. There is little evidence available about the range of probabilistic constraints that affect ambiguity resolution, the relative strength of these constraints, or how they interact with one another. The primary goal of this paper is to investigate how several probabilistic constraints might affect the interpretation of syntactic ambiguities. This research is crucial to an understanding of ambiguity resolution, independent of whether that process is viewed from the garden-path model or a constraint-based perspective. Additional knowledge about constraint strength and interaction should also illuminate the debate concerning the existence of a first-stage parser, although other issues that will not be addressed here, such as specification of the lag between the garden-path model's parser and thematic processor, will also be essential for the eventual resolution of that debate.

Because little discussion has been devoted to the mechanisms whereby the constraints are brought to bear on the ambiguity resolution process in either type of model, a framework for investigating these constraints here will be taken from MacDonald et al. (submitted). This framework will provide some helpful terminology, although these concepts are also incorporated into other constraint-based models (e.g. Trueswell et al., in press) or even in the garden-path model's thematic processor.¹ The remainder of the Introduction is devoted to a brief sketch of an account of how constraints might apply and interact, and the introduction of several

¹For example, knowledge of the relative frequencies of alternative interpretations of an ambiguity are described in terms of partial activation below, but Don Mitchell has suggested to me that the comprehender might instead have direct access to probabilities of occurrence in the language of the different interpretations and could build a single structure and evaluate its plausibility on the basis of this knowledge of probabilities.

constraints that appear to affect ambiguity resolution in the MV/RR ambiguity. The specific predictions for the effects of these constraints concern the availability of *competitor interpretations* to the (correct) reduced relative interpretation of the MV/RR ambiguity: Ambiguity resolution in favour of the RR interpretation is easy when competitors are rapidly inhibited or when available competitors are weak. Ambiguity resolution becomes more difficult when strong competitors are not inhibited or when this inhibition is delayed. The operation and interaction of these constraints are then explored in three self-paced reading experiments.

Activation of Multiple Alternatives

Following much work in syntactic theory arguing for a rich lexical representation to constrain syntactic well-formedness (e.g. see discussion and references cited in Altmann & Steedman, 1988; Bresnan, 1982; Levin & Pinker, 1991; Pollard & Sag, 1988), and work in psycholinguistics stressing the importance of a rich lexical representation for language comprehension (Carlson & Tanenhaus, 1988; Ford, Bresnan, & Kaplan, 1982; Tanenhaus & Carlson, 1989; Tyler, 1989), MacDonald and co-workers' account of constraint use places much of the burden of syntactic ambiguity resolution on lexical representations.² In this model, lexical representations contain syntactically relevant information, such as verb argument structure information, lexical category information and morphological information such as tense and number. Whenever a word is encountered in the input, the various aspects of its lexical representation become activated to differing degrees. The notion of partial activation for lexical representation has long been invoked in models of lexical processing (e.g. Morton's, 1969, Logogen Model, in which words with different frequencies had different thresholds); the extension here is merely that lexical-syntactic knowledge, such as argument structures, are hypothesised to be among the lexical representations that are frequency-sensitive and subject to partial activation. If a word is ambiguous in some dimension, then alternative interpretations are partially activated as a function of their frequencies in the language. The alternative interpretations compete for activation, and strongly activated interpretations may strongly inhibit competitors, whereas weaker interpretations can exert only weak inhibition on competitors, as in interactive activation models (e.g. Rumelhart & McClelland,

²This is not to say that no phrase structure representation exists in MacDonald and co-workers' model, but rather that comparatively more work is done in lexical representations in this model than in other proposals. It may be that complete phrase markers need not always be built, or that constraints of X-bar syntax allow phrasal skeletons (e.g. an NP) to be activated in the lexicon, rather than constructed anew each time (for additional discussion, see Berg, 1991; MacDonald et al., submitted).

TABLE 1
Four Argument Structures for Ambiguous Verbs, with Example Sentences

Active transitive	⟨ <u>Agent</u> , Theme⟩	The patient heard the music.
Intransitive	⟨ <u>Agent</u> ⟩	The patient heard with the help of a hearing aid.
Sentential complement	⟨ <u>Agent</u> , Proposition⟩	The patient heard (that) the nurses were leaving.
Reduced relative	⟨e], <u>Theme</u> ⟩	The patient heard in the cafeteria was complaining.

Note: Thematic grids for the argument structures are given in angled brackets, with the external argument underlined. The reduced relative argument structure is a passive structure in which the Theme is the external argument and there is a possibly null internal argument corresponding to the Agent of the passive.

1982). Also, as in interactive activation models, one interpretation eventually "wins" the competition when its activation exceeds some threshold level of activation, and all alternatives are completely inhibited (Rumelhart & McClelland, 1982).

Because lexical representations are hypothesised to contain syntactically relevant information, many syntactic ambiguities are associated with lexical ambiguities. The MV/RR ambiguity is an example of a syntactic ambiguity that is triggered by a lexical ambiguity, in this case by a verb with several possible argument structures. Argument structures encode information about what kinds of phrases occur with certain words. An argument structure for a verb contains a list of the *thematic* roles that must appear with the verb, including roles such as Agent (the doer of the action specified by the verb), Theme (what is acted upon), and others. The four alternative argument structures that are of interest for the MV/RR ambiguity are illustrated in Table 1 for the verb *heard*, with example sentences. The argument structure notation in Table 1 is similar to one that is commonly used in descriptions of verb argument structure representations (e.g. Levin & Rappaport, 1986). In each case, the arguments that are assigned by the verb are contained in angled brackets, with the *external* argument (i.e. external to the verb phrase) underlined, followed by all (verb phrase) internal arguments.³

Verbs vary in the number of alternative argument structures they permit. Some verbs, like *heard*, will pass activation to all four of the

³The reduced relative argument structure will also accommodate an unreduced passive, as in *The patient was heard* or *The patient who was heard*... In this case, additional information in the sentence (e.g. the presence or absence of *who/that* and *was*) indicates which surface form of the argument structure is correct. The term "reduced relative" for the argument structure, rather than "passive", is used only because reduced relatives, and not passives in general, are of interest here.

argument structures in Table 1, but other verbs will have fewer options. The verbs *interviewed* and *raced*, for example, do not take sentential complements, so this argument structure will not be activated by these verbs. Individual verbs can differ both in the number of alternative argument structures they will activate, and in relative frequency (and therefore activation) of the alternatives.

This application of partial activation to ambiguity resolution has previously been invoked for lexical semantic ambiguities (Kawamoto, 1988), past vs. past-participle verb tenses in MV/RR ambiguity (Trueswell et al., in press; Burgess & Hollbach, 1988), and lexical category ambiguities (words that are ambiguous in their grammatical category, e.g. *promise* can be a noun or verb: MacDonald, 1993). A related account has been suggested for quantifier scope ambiguities (Kurtzman & MacDonald, 1993). The proposed model thus embodies the claim that ambiguity resolution at many linguistic levels invokes the same basic processing mechanisms, in which alternative lexical representations (e.g. alternative lexical semantic representations, alternative lexical categories, alternative argument structures, etc.) are partially activated as a function of both the frequency of the alternatives and the compatibility of the alternatives with other information in the input. The proposed model is also obviously related to several activation-type models of word recognition, including the Logogen Model (Morton, 1969), the Cohort Model (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978), a number of interactive activation models of cognitive functions (e.g. Rumelhart & McClelland, 1982), and other proposals for the use of probabilistic constraints (e.g. MacWhinney & Bates, 1989).

Additional Constraints on Activation Levels. In addition to information concerning the frequencies of argument structures associated with verbs, other information in the input can also affect ambiguity resolution. For the MV/RR ambiguity, most investigations of contextual constraints have focused on semantic and pragmatic influences from context prior to the ambiguity—what will be termed *pre-ambiguity constraints* here. For example, Trueswell et al. (in press) investigated the influence of animacy of the subject noun phrase on resolving the MV/RR ambiguity. Comprehenders who were presented with temporarily ambiguous reduced relative sentences with inanimate subjectnouns, such as *The evidence examined by the lawyer turned out to be unimportant*, showed no comprehension difficulty (measured by reading time) compared to unambiguous controls. The ambiguous sentence became difficult, however, if the animate noun *defendant* replaced the inanimate *evidence*. The explanation, phrased in terms of the alternative argument structures in Table 1, rests on the argument structure frequency information from the verb *examined* in

conjunction with information about the subject noun phrase. Because *examined* cannot be intransitive or take a sentential complement, these two argument structures are not activated, leaving only the active transitive and reduced relative argument structures. When the subject noun is *evidence*, the inanimate status of this noun makes the active transitive argument structure implausible (this interpretation is the one in which *evidence* would be the external argument and receive the Agent thematic role, with the meaning that the evidence was examining something). Activation of the active transitive interpretation is therefore inhibited, the only remaining alternative is the reduced relative, and ambiguity resolution is easy. With *defendant*, however, the active transitive argument structure therefore accrues activation and competes with the (correct) reduced relative argument structure, so that ambiguity resolution in favour of the reduced relative is difficult.

Evidence that such plausibility-based constraints are not the only ones that affect resolution of this ambiguity is provided by the fact that these constraints do not explain the difference in difficulty between the Maslin and Bever sentences. First, the subject noun phrases of all three sentences are animate and make plausible agents for their verbs: Bever's horse can race, Maslin's homeless people can interview, and mental patients can hear. Second, pragmatic context (Altmann & Steedman, 1988) might have helped comprehenders who read Maslin's original article, but her sentences are still perfectly comprehensible when quoted out of context above. Thus other constraints may be operating to aid comprehenders in the Maslin sentences. One of these constraints arrives after the ambiguity is introduced, and will therefore be termed a *post-ambiguity constraint*.

A Post-ambiguity Constraint. The post-ambiguity constraint that is investigated here is predicted to help parsing of the reduced relative construction by inhibiting the active transitive argument structure. Comprehension difficulty then depends on the extent to which pre-ambiguity and verb constraints inhibit the incorrect alternatives, so that ambiguity resolution in favour of the correct reduced relative argument structure can be achieved.

The post-ambiguity constraint affects the active transitive interpretation via some restrictions on the formation of English transitive sentences. In English, verbs are usually adjacent to their direct objects, as illustrated in sentences (1)–(3). In the (a) versions, the verb and the noun phrase (NP) direct object are adjacent, but in the (b) versions, a phrase intervenes between the verb and its NP direct object. Most native speakers of English judge the (b) versions to be awkward or ungrammatical, indicated here by asterisks. These intervening phrases in the (b) versions have a number of different structures; for ease of exposition, they will be called "not-direct

object" (not-DO) phrases, because the one thing they share is that they are *not* the direct object of the preceding verb:

- 1a. The horse raced the donkey past the barn.
- b. *The horse raced past the barn the donkey.
- 2a. The patients heard the doctor here.
- b. *The patients heard here the doctor.
- 3a. The homeless people interviewed a doctor in the film.
- b. *The homeless people interviewed in the film a doctor.

The knowledge that a verb and its direct object are typically adjacent in English could be important for ambiguity resolution. If an ambiguous verb is followed immediately by a not-DO phrase [as in (1b)–(3b)], then a NP direct object is unlikely to appear later in the verb phrase. In lexical activation terms, the appearance of a not-DO phrase inhibits the active transitive argument structure.

This information is probabilistic, because a not-DO phrase can sometimes appear between the verb and the direct object. Some examples can be seen in the active transitive sentences (4) and (5):

4. I gave to Jim my last copy of the conference proceedings.
5. I packed away in a box all those stupid figurines that Grandma used to keep on the mantle.

In these examples, the not-DO phrases (*to Jim*, *in a box*) do not prevent subsequent direct objects. Virtually all speakers of English agree that (4) and (5) sound much better than the unacceptable (1b)–(3b). Ross (1967) coined the term "heavy NP-shift" for sentences such as those in (4) and (5), reflecting the observation that when the NP direct object is particularly long or "heavy", it may appear at the end of the sentence rather than adjacent to its verb. Given the relative infrequency of heavy NP-shifted sentences (Hawkins, 1990), the presence of a not-DO phrase makes the active transitive interpretation less likely, but not impossible.

The other three alternative argument structures shown in Table 1 are not affected by the not-DO phrase. In (6), for example, the active intransitive interpretation of *raced* is fine with the not-DO phrase *past the barn*, and in (7), the not-DO phrase *here* does not affect the sentential complement interpretation. And in Maslin's first sentence, repeated in (8), the not-DO phrase *in the film* does not disrupt the reduced relative interpretation:

6. *Active intransitive*: The horse raced past the barn.
7. *Sentential complement*: The patients heard here that they would have to file insurance forms.
8. *Reduced relative*: The homeless people interviewed in the film are exceptionally calm . . .

Given the fact that the not-DO phrase affects only the active transitive argument structure, and not the other two competitors to the reduced relative shown in (6) and (7), it is clear that comprehension of reduced relatives could still be very difficult unless some other constraints inhibit these other two competitors. The argument structure options of the ambiguous verb become crucial here, because some verbs do not permit the intransitive and sentential complement structures, while others do permit one or both of these structures. The combined effects of verb argument structure information and of the not-DO phrase can be illustrated with the Maslin and Bever sentences. In the case of Maslin's first sentence, repeated in (8), the verb *interviewed* cannot take a sentential complement, and it is only rarely intransitive in English (virtually only with the preposition *for*, e.g. *for a job*). The not-DO phrase *in the film*, then, inhibits the active transitive interpretation. Because the properties of *interviewed* result in little or no activation for the active transitive and the sentential complement argument structures, and with the active transitive interpretation constrained by the not-DO phrase, the *only* remaining argument structure is the reduced relative, and so this interpretation can accrue substantial activation. By contrast, Bever's *raced* sentence permits two alternatives to remain active. Here, the not-DO phrase again inhibits the active transitive interpretation [as in (1b)], and *raced* cannot take a sentential complement, so this interpretation is eliminated. However, *raced* may be intransitive, and so the active intransitive argument structure [as in (6)] remains a strong competitor to the reduced relative, which as a result cannot acquire much activation. Thus when verb argument structure frequencies and post-ambiguity constraints combine to inhibit competitor interpretations, ambiguity resolution is predicted to be easy, but weaker constraints result in more competitors and difficult ambiguity resolution. This hypothesised combined action of the verb argument structure frequencies and the post-ambiguity not-DO constraint is the focus of Experiment 1.

EXPERIMENT 1

This experiment has two parts. In Experiment 1A, reading times are measured in ambiguous and unambiguous sentences in the presence of verb and post-ambiguity constraints. Experiment 1B is an off-line norming study that assesses the strength of pre-ambiguity constraints in the materials. Regression analyses are used to evaluate the combined effects of pre- and post-ambiguity constraints on reading times.

The stimuli for Experiment 1A contrast two kinds of ambiguous verbs with an unambiguous condition. One ambiguous condition employed *transitive-only* verbs such as *captured* and *admired*, which do not permit

sentential complements. When the not-DO constraint makes the active transitive interpretation less probable for such verbs, the *only* grammatical alternative is the reduced relative interpretation. This condition is similar to Maslin's *people interviewed* sentence. In a second ambiguous condition, the first verb in the sentence can be optionally transitive or intransitive, as in *fought* and *raced*. In this *optional verb* condition, the presence of the not-DO constraint still inhibits the active transitive interpretation, but two alternatives remain: the active intransitive interpretation and the reduced relative interpretation. Bever's *horse raced* sentence is an example. The finding that ambiguity resolution in favour of the reduced relative interpretation is more difficult in the optional verb condition than in the transitive-only condition would be evidence for the claim that difficulty in ambiguity resolution varies with the number of alternative interpretations that remain available to compete with the reduced relative.

A second variable that is manipulated in this experiment is the time between the introduction of the ambiguity and the availability of constraint from the not-DO phrase. All sentences contain three-word not-DO phrases that begin immediately after the ambiguous verb, but by carefully choosing the words in the phrase, it is possible to manipulate when constraining information arrives. In one condition, the not-DO phrase contains words that begin to inhibit the active transitive interpretation immediately. In the other condition, the not-DO phrase is constructed so that the inhibition does not begin until the second or third word of the phrase, so that the active transitive interpretation remains viable for a longer time after the ambiguity is introduced at the verb. The claim that alternative interpretations compete for activation and inhibit one another yields the prediction that ambiguity resolution in favour of the reduced relative will benefit from early inhibition of alternative interpretations relative to late inhibition. The early inhibition condition is therefore termed a *good post-ambiguity constraint*, whereas the late inhibition condition is a *poor post-ambiguity constraint*.

The effect of the constraints on ambiguity resolution will be assessed with comparisons of reading times in ambiguous and matched unambiguous sentences in two regions: the *ambiguous* region, containing the ambiguous verb and post-ambiguity constraint, and the *disambiguation* region, containing the syntactic disambiguation of the sentence. Most previous research has focused on the disambiguation region. The goal has been to determine whether the presence of helpful constraints would produce ambiguous reading times similar to unambiguous times or whether an "ambiguity effect" (longer reading times in ambiguous than unambiguous conditions) persists even in the face of helpful constraints. Some previous research suggests that the ambiguous region will also be informative, however. For example, in a different ambiguous construction contain-

ing noun/verb lexical category ambiguities such as *fires*, MacDonald (1993, Experiment 2) found that reading times in both ambiguous and disambiguation regions were sensitive to which interpretation was promoted by probabilistic constraints. The ambiguous conditions had the form *warehouse/corporation fires*, in which the pre-ambiguity word biased the ambiguous word either to join with the previous word in a simple noun phrase structure (as in [*the warehouse fires*]_{NP} . . .), or it biased the ambiguous word to be a verb, yielding a more complex NP + VP analysis for these two words (as in [*the corporation*]_{NP} [*fires* . . .]_{VP}). The unambiguous condition always contained a different number marking that forced the more complex NP + VP structure for these words—*the warehouses/corporations fire*. When this same interpretation was promoted in the ambiguous condition, as in *corporation fires*, there were no differences in reading times as a function of ambiguity either at the ambiguous region (the word *fires*) or at the later disambiguation region. When the ambiguous sequence was more plausible as a simple NP (*warehouse fires*), however, what will be termed a “reverse ambiguity effect” appeared in the ambiguous region: Reading times on *fires* were shorter in the ambiguous condition than in the unambiguous control. Then in the disambiguation region, the opposite pattern appeared, with longer reading times in the ambiguous condition than in the unambiguous control—an ambiguity effect. This pattern was attributed to the relative complexity of the two interpretations: *warehouse* promoted the simple NP interpretation, with the result that reading times were shorter in the ambiguous region in comparison to the more complex unambiguous condition, and then at the disambiguation, the misparse was discovered, resulting in longer reading times for ambiguous than unambiguous conditions.

Exactly the same pattern was found by Altmann et al. (1992) in another syntactic ambiguity. These researchers compared reading times in a syntactically complex unambiguous condition and an ambiguous condition for which pragmatic constraints favoured either this complex interpretation or a simpler interpretation. Two eyetracking experiments demonstrated effects of the pragmatic context. In the absence of helpful context, first-pass reading times in the ambiguous region showed the reverse ambiguity effect—ambiguous condition reading times were reliably shorter in the ambiguous region than in the more complex unambiguous condition. The subsequent disambiguation region then produced robust ambiguity effects—longer reading times in ambiguous than unambiguous conditions. When the helpful pragmatic context was available, however, there was little or no reverse ambiguity effect in the ambiguous region and little or no ambiguity effect at the disambiguation.

These results indicate the importance of examining processing in the ambiguous region and suggest that the pattern of reading times across the

two regions can be diagnostic of the strength of probabilistic constraints. A pattern of reverse ambiguity and ambiguity effects that is similar to that found in MacDonald (1993) and Altmann et al. (1992) is predicted here, because the unambiguous condition is syntactically complex (a relative clause), and constraints in the input can support either this complex interpretation or a simple main verb interpretation of the ambiguity. When the complex reduced relative interpretation is promoted, reading times should resemble those in the unambiguous condition at both regions, but when simpler interpretations are promoted, reading times in the ambiguous region should be shorter than in the unambiguous condition (the reverse ambiguity effect), followed at the disambiguation by longer reading times in the ambiguous condition than in the unambiguous condition (the ambiguity effect). The extent to which this pattern holds is predicted to vary with the strength of the constraints—the more alternative simple interpretations that are permitted (that is, the weaker the constraints), the larger the reverse ambiguity and ambiguity effects are predicted to be.

Method: Experiment 1A

Subjects. Forty-eight MIT undergraduates were paid for their participation. All the subjects in this and all subsequent experiments were native speakers of English. Five additional subjects were tested but rejected for missing more than 20% of the comprehension questions in the experiment.

Materials. Eighteen experimental items were created. Six verb triples were chosen to manipulate argument structure constraints in the verb, so that each triple contained one unambiguous past-participle verb (e.g. *over-thrown*), one transitive-only ambiguous verb (e.g. *captured*)⁴ and one optional ambiguous verb that can be used either transitively or intransitively (e.g. *fought*). Three unrelated sentences were written for each verb triple so that any member of the triple could sensibly appear as the fourth word in the sentence, introducing a reduced relative clause. An example item is contained in Table 2; all sentences are contained in Appendix 1.

As can be seen in Table 2, the first verb was followed by a three-word not-DO phrase (e.g. *in the coup, just after dawn*). Good and poor constraint versions of this phrase differed in the point at which constraining

⁴The verb *found*, which permits a sentential complement structure in addition to the transitive structure, was accidentally included in the transitive-only condition. Eliminating this item from the analyses in Experiment 1 had no effect on the results. This result is not surprising, as *found* is used transitively more frequently than it is used with sentential complements (Holmes et al., 1989)—Experiment 3 explores verb bias effects.

TABLE 2
Example Sentences and Question from Experiment 1

Unambiguous verb, good constraint

The ruthless dictator/overthrown in the coup/was hated/throughout the country.

Transitive-only verb, good constraint

The ruthless dictator/captured in the coup/was hated/throughout the country.

Optional verb, good constraint

The ruthless dictator/fought in the coup/was hated/throughout the country.

Unambiguous verb, poor constraint

The ruthless dictator/overthrown just after dawn/was hated/throughout the country.

Transitive-only verb, poor constraint

The ruthless dictator/captured just after dawn/was hated/throughout the country.

Optional verb, poor constraint

The ruthless dictator/fought just after dawn/was hated/throughout the country.

Comprehension question for all versions

Did the people in the country love the dictator? Answer: NO

Note: Slashes indicate analysis regions for the reading time data; subjects did not see the slashes.

information became available. In the good constraint condition, the first word of the phrase was a preposition such as *in* or *during*, which could not introduce a direct object for the preceding verb. Here, the word immediately following the ambiguous verb indicates that a direct object is not adjacent to the verb, and so the constraining information from the not-DO phrase begins immediately. This constraint arrives later in the poor constraint condition because the first one or two words of the not-DO phrase in this condition did not immediately rule out the possibility that a direct object was adjacent to the verb. In the poor constraint example in Table 2, *just after dawn*, the first word of the phrase may introduce a direct object (e.g. *the dictator captured just one soldier*). The not-DO constraint is available only at the next words *after*, since it is unlikely that the sequence *just after* could introduce a direct object. The active transitive interpretation is therefore inhibited at the first word of the not-DO phrase in the good constraint condition, but the inhibition is delayed for one or two words in the poor constraint condition.

"Yes/no" comprehension questions were prepared for all experimental items and for all 10 practice and 72 filler items, some of which were

experimental items for an unrelated study.⁵ The questions for the experimental items did not focus on the reduced relative clause or the not-DO phrase, so that the same question could be used in all six versions of an experimental sentence. The correct answer was "Yes" for half of the experimental and filler trials.

Six lists were prepared so that there were three experimental items of each type in a list. Each sentence appeared with only one of its triple of verbs, and each member of the verb triple appeared only once in a list. Each subject saw a different randomisation of experimental and filler items.

Procedure. Each subject was tested individually. The subjects read sentences in a moving window display in which all non-space characters of the sentence initially appeared as dashes on the computer screen (Just, Carpenter, & Woolley, 1982). They were instructed to press a key to see each new word of the sentence. The first keypress revealed the first word, and with the second keypress, the first word reverted to dashes and the second word was revealed, and so on. Following the keypress terminating the last word of the sentence, the comprehension question appeared. The subjects pressed one of two keys to indicate a "yes" or "no" answer and did not receive feedback. The subjects completed the experiment without a break in a 30 min session.

Predictions and Analyses. As discussed above, the extent to which reverse ambiguity and ambiguity effects appear is hypothesised to be a function of constraint strength, with the strongest constraints yielding no such effects and the weakest constraints yielding the largest effects. Table 3 makes this hypothesis more explicit by indicating which competitor argument structures remain in each ambiguous condition after the constraints have had their predicted effects, as well as the predicted reverse ambiguity and ambiguity effects. The four combinations of verbs and constraints are ordered in Table 3 according to the amount of constraining information they are predicted to contain and the extent to which they eliminate competitor interpretations. The combination of transitive-only verbs and the good post-ambiguity constraint is predicted to inhibit all of the reduced

⁵In all three experiments, less than half of the filler items were experimental items in a wholly unrelated experiment. Experiment 1 contained 16 fillers with embedded wh-questions, as in *John didn't know which book . . .*, and 32 fillers were modifications of experimental items from Frazier and Rayner's (1987) lexical category ambiguities and are reported in Experiment 1 of MacDonald (1993). Experiment 2 contained 24 items investigating the noun phrase/sentential complement ambiguity, and Experiment 3 contained 20 items with wh-questions.

TABLE 3

Predictions for which Argument Structure Competitors are Eliminated and for Reading Time Differences Between Ambiguous and Unambiguous Conditions in Experiment 1

Verb Type and Post-ambiguity Constraint	Competitors Remaining after Constraints Apply	Differences from Unambiguous Condition in Two Regions	
		Ambiguous	Disambiguation
Transitive-only verb + good constraint	None	≈ 0	≈ 0
Transitive-only verb + poor constraint	Active transitive (temporarily)	< 0	> 0
Optional verb + good constraint	Active intransitive	< 0	> 0
Optional verb + poor constraint	Active intransitive, active transitive (temporarily)	< 0 (largest)	> 0 (largest)

Note: Verb and post-ambiguity constraint conditions are given in the order of increasing predicted difficulty. Poor not-DO constraints temporarily permit the active transitive interpretation, but this interpretation becomes unlikely by the end of the ambiguous region.

≈ 0 indicates little or no difference predicted between ambiguous and unambiguous conditions.

< 0 indicates a reverse ambiguity effect (negative ambiguous-unambiguous difference score).

> 0 indicates an ambiguity effect (positive ambiguous-unambiguous difference score).

relative interpretation's competitors, so that reading times should mimic those of the unambiguous reduced relative condition. The three other ambiguous conditions are predicted to be less constraining to varying degrees, so that reading times in these conditions should differ to varying degrees from those in the unambiguous controls, as a function of the number and strength of the competitor interpretations. When the constraints do not eliminate all competitors, it is predicted that reading times will reflect the dominance of these simpler alternative structures over the rarer and more complex reduced relative structure that is developed for the unambiguous sentence, yielding a reverse ambiguity effect in the ambiguous region. Then having pursued a simple, incorrect structure in the face of misleading constraints, comprehenders pay a price at the disambiguation. Such effects should be small for the transitive-only verb and poor constraint, which only temporarily admits the active transitive competitor, and largest for the optional verb + poor constraint, which admits the most competitors. Given this progression of difficulty over four stages, it is unlikely that each condition will always turn out to be reliably different from the adjacent conditions in Table 1, but it is reasonable to expect that

difficulty should progress in the direction indicated, so that at least some conditions at the bottom of Table 1 will be reliably more difficult than those at the top.

To better compare across the various conditions, reading times were adjusted for length using a method common in self-paced reading studies (Ferreira & Clifton, 1986). For each subject, a linear regression equation was calculated to predict reading times in each word from word length in all experimental sentences. Reading time at each word can then be expressed as a difference score from the predicted reading time for that subject. This procedure does not affect the regions at and after the disambiguation, which contain identical words in all conditions, and it serves to reduce the effects of varying the words in the ambiguous region.

Results

Only sentences for which the comprehension question was answered correctly were included in reading time analyses. The length-adjusted reading times were trimmed at each word by removing all data points over 5 standard deviations above the condition mean for that word. This procedure removed 1.2% of the data. Figure 1 shows difference scores between ambiguous and unambiguous conditions, so that positive and negative differences in the predictions in Table 3 can be checked with positive and negative values in Fig. 1. Appendix 2 contains the length-adjusted reading times that were used to calculate the difference scores in

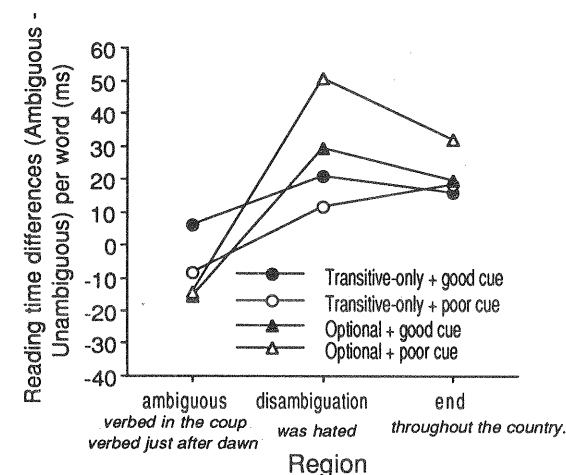


FIG. 1 Length-adjusted reading time differences between ambiguous and unambiguous sentences in Experiment 1.

Fig. 1, as well as the raw reading times. The sentences are divided into three regions: the ambiguous region, consisting of the verb and three-word not-DO phrase, a two-word disambiguation, and the rest of the sentence.

Planned comparisons were conducted to assess the hypotheses given in Table 3 concerning specific reading patterns at each region. Turning first to the ambiguous region, it was predicted that the reading times for the more constraining transitive-only verbs would more closely resemble those of the unambiguous condition than would times for the optional verbs, which were predicted to produce a reverse ambiguity effect. Figure 1 shows that the predictions for the effects of verb were supported, in that there was no difference in reading times between unambiguous and transitive-only conditions in this region (t 's < 1), but the two optional verb conditions yielded a mean 15 msec per word reverse ambiguity effect [$t_1(47) = 2.5$, $P < 0.01$; $t_2(17) = 2.4$, $P > 0.05$]. An alternative way to examine these data is to compare the size of the ambiguous-unambiguous difference scores for the two types of ambiguous verbs. This analysis showed that the mean difference score for the transitive-only verbs was reliably different from that of the optional verbs [$t_1(47) = 2.39$, $P < 0.05$; $t_2(17) = 1.83$, $P > 0.05$]. In contrast to these verb effects, the effects of the post-ambiguity not-DO constraint were not reliable in this region (all t 's < 1).

In the disambiguation region, reading times reflected the joint effects of the verb and post-ambiguity constraint factors. The combination of the optional verb + poor constraint, which was predicted to be the most difficult condition, gave a robust ambiguity effect at the disambiguation [$t_1(47) = 3.91$, $P < 0.001$; $t_2(17) = 2.60$, $P < 0.01$]. The optional verb + good constraint was predicted to be next hardest, and it produced the next largest effect; the ambiguity effect here was reliable in the subjects analysis [$t_1(47) = 2.00$, $P < 0.05$], but not in the items analysis [$t_2(17) = 1.57$, $P < 0.10$]. A comparison of the size of the difference scores for these two conditions reveals that the difference between the ambiguity effect with the optional verb + poor constraint (50 msec per word) and for the optional verb + good constraint (29 msec) approached significance [$t_1(47) = 1.41$, $P < 0.08$]. The two transitive-only verbs, shown with circles in Fig. 1, produced smaller ambiguity effects, as predicted. The combination of these verbs + good constraints turned out to be slightly more difficult than the transitive-only verbs with poor constraints, so that the ambiguity effect at the disambiguation was reliable in the subjects analysis when the constraint was good [$t_1(47) = 2.10$, $P < 0.05$; $t_2(17) = 1.08$, $P > 0.20$], but no effect when it was poor (t 's < 1). Comparing the difference scores, the size of the ambiguity effects in these two conditions did not differ reliably from one another (t 's < 1), though they were both reliably smaller than the ambiguity effect produced by the most difficult optional verb + poor constraint condition (compared with transitive-only verb + poor con-

straint: $t_1(47) = 2.76$, $P < 0.001$; compared with transitive-only verb + good constraint: $t_1(47) = 1.98$, $P < 0.05$].

Almost all effects of ambiguity disappeared by the end region of the sentence, with only the most difficult condition, the optional verb + poor constraint, showing an ambiguity effect that approached significance [$t_1(47) = 2.64$, $P < 0.01$; $t_2(17) = 1.22$, $P < 0.15$].

The subjects were uniformly very accurate in answering the comprehension questions (a mean of 96.9% correct), and there were no reliable differences across conditions.

Discussion

The purpose of Experiment 1A was to test the hypothesis that processing difficulty during ambiguity resolution is a function of the number of alternative interpretations that can compete with the correct interpretation, as modulated by the strength of verb and not-DO constraints. This hypothesis was clearly supported for the verb manipulation: The conditions with transitive-only verbs, predicted to be quite constraining, yielded no reverse ambiguity effect in the ambiguous region, at most a marginal ambiguity effect at the disambiguation, and no effect in the end region. By contrast, the optional verbs produced reliable reverse ambiguity effects in the ambiguous region, followed by ambiguity effects at the disambiguation and, when combined with a poor post-ambiguity cue, in the end region. The effects of the post-ambiguity constraint, which was predicted to modulate how rapidly the active transitive argument structure was inhibited, had little effect on the transitive-only verbs, but did have some effect on the optional verbs in the disambiguation and end regions, in that poor constraint conditions were more difficult than good constraint conditions. These results were in the predicted direction but were rather small, perhaps because the manipulation itself was subtle, with as little as one word difference in the arrival of constraining information in the two conditions. A stronger manipulation of this constraint is employed in Experiment 2.

The results of Experiment 1A are important in several respects. First, they demonstrate that the number of alternative argument structures permitted by a verb is important for ambiguity resolution for this construction. This result shows that constraints inherent within the lexical item that triggers the syntactic ambiguity must be incorporated in any model of ambiguity resolution. Second, Experiment 1A confirmed the relationship between reverse ambiguity effects in the ambiguous region and ambiguity effects at the disambiguation—if a reverse ambiguity effect appears in the first region, then a sizeable ambiguity effect appears in the second region, but if reading times in the first region show little difference as a function of

ambiguity, then reading times in the second region also show little difference. This linkage of effects in the two regions, which replicate MacDonald's (1993) and Altmann and co-workers' (1992) data for two other ambiguous constructions, suggests that reading times are sensitive to the activation of alternative argument structures: When simple, high-frequency alternative structures are available to compete with the reduced relative structure, reading times are short in the ambiguous region, then long at the disambiguation when it becomes clear that the simple structure is incorrect. When the constraints eliminate most or all competitors to the reduced relative, however, reading times for ambiguous conditions show little difference from unambiguous controls. These patterns of reading times thus appear to be diagnostic of the strength of probabilistic constraints in the input.

Though the general pattern of results supports the hypotheses concerning the importance of the number of alternative interpretations, the pattern is not ideal. The combination of the transitive-only verb + good constraint, which was predicted to be the most constraining ambiguous condition, did not completely remove the ambiguity effect at the disambiguation. One possible explanation is that pre-ambiguity information may have conflicted with the good post-ambiguity constraints. All of the subject noun phrases were animate in Experiment 1, which has been shown to promote an incorrect active interpretation rather than the correct reduced relative interpretation (Trueswell et al., in press). In Experiment 1B, ratings were collected using a procedure like the one used by Pearlmutter and MacDonald (1992) to examine pre-ambiguity constraint strength, and regression analyses were performed to assess the extent to which both pre- and post-ambiguity constraints were contributing to reading times in Experiment 1A.

Method: Experiment 1B

Subjects. Forty-four MIT undergraduates, who had not participated in Experiment 1A, were paid for their participation.

Materials. The ambiguous stimuli from Experiment 1A were modified to enable the subjects to rate how natural the stimuli sounded as relative clauses. So that ratings were based on only the pre-ambiguity information, only the first few words of the sentences were presented to the subjects. Additional words were added to the sentence fragments to create unambiguous relative clause stimuli for the subjects to rate. A relative pronoun (*who/that*) and *was/were* were added before the verb, and *by the* was added

after the verb to create unambiguous relative clause structures, as in *The ruthless dictator who was captured by the . . .*

The combination of 18 sentence fragments \times 2 ambiguous verbs (transitive-only and optional) yielded 36 items for rating. Two rating booklets were designed, so that a fragment was paired with only one of its two ambiguous verbs in a booklet (e.g. no subject rated both *the dictator who was captured* and *the dictator who was fought*). Each booklet contained half of the 36 fragments, intermixed with 30 other filler items, including items from an unrelated experiment that were specifically designed to have strong pre-ambiguity constraints. Because there were 18 experimental items in each booklet but only 12 different verbs (because the unambiguous verbs were not rated), 6 verbs appeared twice on each list (with different noun phrases), separated from one another by at least 7 items.

Procedure. Each subject received a six-page booklet of 48 items; the order of pages in the booklet was randomised for each subject. The subjects were shown examples and were asked to rate each fragment on a scale of 1–7 (1 = the best rating). The subjects were instructed to use their first impressions, based on “whether the situation that is described makes sense”. Each fragment was rated by 22 subjects.

Results and Discussion

Because of an error in materials construction, one fragment was never rated with its transitive-only verb. The transitive-only means given below therefore reflect only 17 items, whereas the means for the optional items include all 18 sentences. Removing this item's optional verb rating did not change any of the analyses reported below.

There was a small effect of verb type on the ratings, with fragments containing transitive-only verbs rated as better relative clauses (1.96) than were optional verb fragments (2.37) [$t(33) = 1.97$, $P < 0.10$]. This result presumably reflects the transitive–intransitive biases of the verbs; the relative clause requires a transitive usage, and this usage is by definition higher in relative frequency for the transitive-only verbs than for the optional verbs, which permit an intransitive interpretation. This trend complicates the ratings for the present analyses: The effects of the pre-ambiguity information must be assessed *within* levels of the verb factor, because verb type and relative clause rating were correlated. The question is thus whether, within each level of verb, the pre-ambiguity information and the good–poor constraint manipulation affected ambiguity resolution.

The ratings were entered into regression analyses predicting ambiguous–unambiguous reading time difference scores for each stimulus item from

TABLE 4
Correlations Between Ratings and Ambiguous-Unambiguous Reading Time Differences in Experiment 1

Verb and Constraint	Region		
	Ambiguous	Disambiguation	Rest of Sentence
Transitive-only			
good constraint	0.30 ^c	0.09	0.09
poor constraint	-0.45 ^{a,c}	-0.50 ^b	-0.32
Optional			
good constraint	0.05	-0.03	-0.02
poor constraint	-0.09	-0.01	0.19

Note: A positive correlation indicates that favourable relative clause ratings are associated with small (or negative) ambiguous-unambiguous difference scores.

^a0.05 < *P* < 0.10; ^b*P* < 0.05.

^cThese conditions are shown in Fig. 2.

Experiment 1A. Difference scores were used because the question of interest is whether the pre-ambiguity information specifically affected the size of the reverse ambiguity and ambiguity effects observed in Experiment 1A. Table 4 presents the correlations between the relative clause rating and ambiguous-unambiguous difference scores for each condition. Smaller values on the rating scale indicated more favourable ratings for relative clauses, so positive correlations indicate that as the rating for relative clauses improves, the ambiguous-unambiguous difference score decreases. However, the interpretation of a positive or negative correlation depends on which region is being considered—whether the mean difference score in a region is positive (an ambiguity effect) or negative (a reverse ambiguity effect).

Consider first the correlations for the ambiguous region in the left-hand column of Table 4. When the verb was transitive-only, the correlation between the relative clause rating and the difference score was (non-significantly) positive when the constraint was good ($r = 0.30$) and negative when the constraint was poor ($r = -0.45$); these two correlations differ from one another ($P < 0.07$) using the Z_1^* statistic for dependent correlations (Steiger, 1980). The fact that the same set of ratings correlated positively with reading time difference scores in one post-ambiguity constraint condition and negatively in another suggests that there is an interaction between pre- and post-ambiguity constraints. A closer examination of these two conditions, shown in Fig. 2, illuminates these effects. In the good constraint condition, the differences between the unambiguous

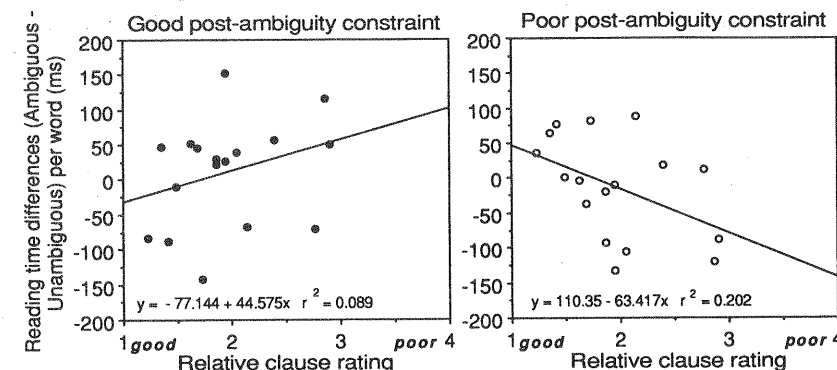


FIG. 2 Effects of pre- and post-ambiguity information in Experiment 1. Ambiguous-unambiguous difference scores as a function of pre-ambiguity constraints at two levels of post-ambiguity constraint.

and ambiguous conditions tended to be near zero when the fragment was rated as a good relative clause, but when the fragment was rated as poor (so that the pre-ambiguity information conflicted with the good post-ambiguity constraint), reading times tended to increase in the ambiguous condition over the unambiguous control. In the poor constraint condition, however, a different effect appeared. The stimuli that were rated as better relative clauses again produced similar reading times in ambiguous and unambiguous conditions, but for the items with poorer relative clause ratings, the result was *shorter* reading times for ambiguous conditions than unambiguous conditions. This, of course, is the reverse ambiguity effect, and Fig. 2 illustrates that this pattern emerged most strongly when a poor pre-ambiguity constraint and a poor post-ambiguity constraint together conspired to promote a simple active structure over the correct relative clause interpretation.

The negative correlation continues in the poor constraint condition to the disambiguation region and less strongly to the end of the sentence for the transitive-only verbs, as shown in Table 4. The positive correlations in the good constraint condition drop off sharply in these regions. Interestingly, the optional verb condition did not produce any reliable correlations between relative clause ratings and reading time. Though caution is necessary when interpreting non-significant effects, these results may indicate that the active intransitive interpretation, which is allowed by the optional verbs and is never ruled out by a not-DO phrase, is so strong that the plausibility of a relative clause structure has very little influence.

In sum, the important result in Experiment 1B was the fact that in the transitive-only conditions, the same set of pre-ambiguity ratings correlated positively or negatively with difference scores in the ambiguous region

depending on whether that region contained a good or poor post-ambiguity constraint. The correlational data are tentative, but they suggest that the reading patterns in the ambiguous region are sensitive to the extent to which both pre- and post-ambiguity constraints are consistent in promoting one interpretation over others. Thus reading times appear to be sensitive to whether or not constraints converge or conflict. This convergence and conflict was measured only *post hoc* in Experiment 1B; Experiment 2 examines these effects with explicit manipulations of converging and conflicting constraints.

EXPERIMENT 2

This experiment investigates ambiguity resolution in sentences in which both pre- and post-ambiguity constraints are manipulated at two levels, good and poor. The results of Experiment 1B suggest that there are far more than two levels of these constraints, and the exact tracing of their joint effects will eventually require more than a simple factorial manipulation. At this stage, however, it is important to replicate the major results of Experiment 1 and examine the effects of converging and conflicting constraints. Accordingly, the present experiment attempted to create very strong constraint manipulations—the good constraints were designed to be as good as possible, and the poor constraints were designed to be as poor as possible while still permitting the reduced relative interpretation. As in the previous experiment, a progression of difficulty is predicted: Ambiguity resolution should be easiest when both constraints converge in promoting the reduced relative interpretation, the two conflict conditions (one good and one poor constraint) are predicted to be more difficult, and the most difficult condition should be the one where two poor constraints converge in promoting the incorrect main verb interpretation of the ambiguity.

Method

Subjects. Forty MIT undergraduates were paid for their participation. An additional three subjects were tested but rejected for high comprehension question error rates.

Materials and Procedure. Stimuli were manipulated in a 2 (pre-ambiguity constraint) \times 2 (post-ambiguity constraint) \times 2 (ambiguity) design. Thirty-two sentences were constructed, each with a different ambiguous verb. These verbs were chosen primarily for their properties in contributing to the pre-ambiguity constraint manipulation; both transitive-only and optional verbs were used, and verb type was not manipulated.

TABLE 5
Example Sentences and Questions from Experiment 2

Both constraints good

The management team believed that the shipment (that was)/transported to the polluted beaches/would help/clean up the oil spill.

Good pre-ambiguity constraint, poor post-ambiguity constraint

The management team believed that the shipment (that was)/transported almost two thousand miles/would help/clean up the oil spill.

Poor pre-ambiguity constraint, good post-ambiguity constraint

The management team believed that the workers (that were)/transported to the polluted beaches/would help/clean up the oil spill.

Both constraints poor

The management team believed that the workers (that were)/transported almost two thousand miles/would help/clean up the oil spill.

Comprehension questions

Good pre-ambiguity constraint: Was the shipment used to transport something? Answer: NO
Poor pre-ambiguity constraint: Were the workers used to transport something? Answer: NO

Note: The unambiguous conditions were created with the addition of the parenthetical material. Slashes indicate divisions for reading time analyses.

Because there are not sufficient numbers of unambiguous past-participle verbs in English to produce sensible unambiguous conditions for all sentences in the present experiment, as had been done in Experiment 1, the unambiguous conditions were created here by adding *that was* (or *were*) before the ambiguous verb. Introductory and ending phrases were longer than those in Experiment 1, in order to provide reasonable contexts for the combined manipulations of pre- and post-ambiguity constraints. Examples can be seen in Table 5, and all stimuli are contained in Appendix 3.

Pre-ambiguity constraints were manipulated through the animacy of the subject noun (Trueswell et al., in press). A good pre-ambiguity constraint contained an inanimate noun that made a good Theme but a poor Agent for the following verb, promoting the reduced relative interpretation. In the inanimate example in Table 5, shipments are more likely to be transported (Theme) than to transport (Agent). The poor pre-ambiguity constraint contained an animate noun that was a plausible agent for the following verb, promoting an active structure (e.g. *workers transported*).

The post-ambiguity constraint was again manipulated, but the not-DO phrases were lengthened to four words (compared to three in Experiment 1). This change was designed to better differentiate good and poor post-ambiguity constraints. The longer not-DO phrases allowed additional

delay before helpful information arrived in the poor constraint condition. The third or fourth word of these phrases introduced the constraining information here, compared to the second or third word in Experiment 1. Good post-ambiguity constraints again contained constraining information in the first word of the not-DO phrase.

Comprehension questions for each item varied slightly with the pre-ambiguity manipulation. The questions focused on the subject noun phrase and verb more closely than in Experiment 1, so that the question could better assess how the ambiguities were interpreted. Eight stimulus lists were created to fully counterbalance all factors, with four examples of each condition in each list. Each list contained 5 practice items and 60 filler items, some of which were experimental items from an unrelated experiment.

Predictions. Difficulty of ambiguity resolution is again predicted to be a function of the strength of the constraints and the number of alternative interpretations that are available to compete with the reduced relative interpretation. The progression of difficulty is predicted to appear at the disambiguation, with no differences between ambiguous and unambiguous conditions when both constraints are good, then small ambiguity effects when one constraint is good and one is poor, and the largest ambiguity effects when both constraints are poor.

Unlike in Experiment 1, the ambiguous region may not be informative in this experiment, where unambiguous conditions were created by adding two extra words (*that was/were*) to the ambiguous condition. One disadvantage of this manipulation is that the addition of these words has been shown to produce large effects in reading times on the next several words (Spivey-Knowlton et al., 1993; Trueswell et al., in press), which may prevent the detection of any reverse ambiguity effects. Previous studies using this manipulation have found that reading times at the verb and perhaps for several words thereafter tend to be shorter in the condition that contains *that was* compared to a condition without these words. The effect of these additional words is independent of ambiguity, as it appears even in otherwise unambiguous sequences—reading times on *shown near the* are shorter in *the horse that was shown near the . . .* compared with *the horse shown near the . . .* (Spivey-Knowlton et al., 1993; Trueswell et al., in press). These researchers argue that the source of this difference rests in the early notification of a relative clause structure that is provided by *that was* in the “extra-words” condition, so that by the time *shown near the* are read, much of the syntactic processing required for the relative clause structure has already been done. By contrast, in the absence of *that was*, the work of establishing a relative clause representation must be done while the words *shown near the* are read, and the longer reading times here

reflect this extra effort. On this view, the extra-words disambiguation is not ideal, because the unambiguous *that was* condition is different from the ambiguous condition not only in ambiguity, but also because the addition of two words displaces a locus of processing difficulty to an earlier point in the sentence, so that comparisons of reading times in the ambiguous region are confounded with this additional factor. Previous comparisons between ambiguous and extra-words unambiguous conditions have not revealed any reverse ambiguity effects (Ferreira & Clifton, 1986; Rayner et al., 1983; Spivey-Knowlton et al., 1993; Trueswell et al., in press), and it is unlikely that any reverse ambiguity effects will be observed in the ambiguous region in this experiment either.

Results

The length adjustment was performed on reading times as in Experiment 1, except that word length was recorded for both experimental and filler sentences during this experiment, so that a subject's regression equation was based on all of these sentences, instead of just on the experimental sentences, as in Experiment 1. The reading times were trimmed as in Experiment 1, removing less than 3% of the data. Only those reading times for which the comprehension question was answered correctly were entered into the analyses.

The differences between ambiguous and unambiguous conditions can be seen in Fig. 3. The disambiguation region shows the predicted progression of difficulty as a function of the constraint goodness. When both con-

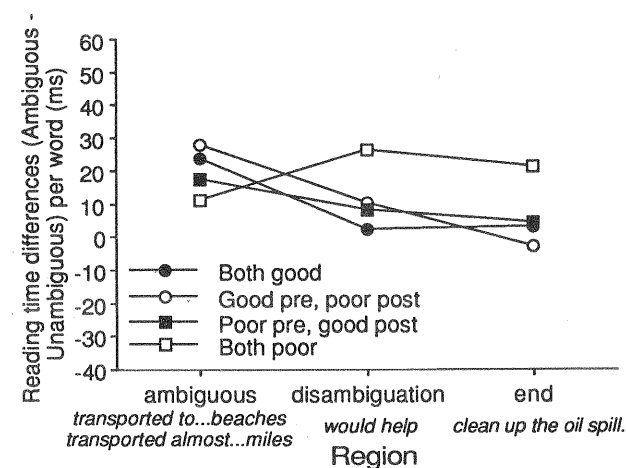


FIG. 3 Length-adjusted reading time differences between ambiguous and unambiguous sentences in Experiment 2.

straints were good, there was no difference between ambiguous and unambiguous reading times at the disambiguation (t 's < 1). Next, the two conflict conditions, with one good and one poor constraint, produced ambiguity effects of 8–10 msec per word at the disambiguation that were reliable in only one of the subjects or items analyses [good pre-, poor post-ambiguity constraint: $t_1(39) = 2.16$, $P < 0.05$, $t_2 = 1.0$, $P > 0.15$; poor pre-, good post-ambiguity constraint: $t_1 < 1$, $t_2(31) = 1.69$, $P < 0.05$]. Finally, when both constraints were poor, there was a reliable 26 msec per word ambiguity effect at the disambiguation [$t_1(39) = 2.69$, $P < 0.005$; $t_2(31) = 2.78$, $P < 0.005$]. The end region showed a pattern that is similar to that in Experiment 1, in that only the most difficult condition, with both poor constraints, produced a reliable ambiguity effect [$t_1(39) = 3.20$, $P < 0.01$; $t_2(31) = 3.12$, $P < 0.01$].

Reading times in the ambiguous region showed none of the reverse ambiguity effects that were present in Experiment 1. Indeed, reading times were longer in all the ambiguous conditions than in the unambiguous conditions (all P 's < 0.01). This result seems directly attributable to the nature of disambiguation in the present experiment, which added two words to the sentence (in contrast to the disambiguation in Experiment 1).

Comprehension Questions. The subjects' comprehension accuracy, shown in Table 6, was slightly lower than in Experiment 1. This result is probably due to the somewhat harder questions that were used in the present experiment; accuracy declined for both unambiguous and ambiguous items compared with Experiment 1. An omnibus ANOVA revealed that accuracy on unambiguous items (85.3%) was not reliably better than on ambiguous items (82.7%) [$F_1(1,39) = 2.39$, $P > 0.10$; $F_2(1,31) = 2.86$, $P = 0.10$]. The subjects were slightly more accurate answering questions with a good pre-ambiguity constraint (87.3% correct) than a poor pre-ambiguity constraint (80.6% correct) [$F_1(1,39) = 7.19$, $P < 0.05$; $F_2(1,31) = 3.91$, $P < 0.10$]; this effect did not interact with ambiguity (P 's > 0.10). No other effects were reliable.

TABLE 6
Percent Correct Responses to Comprehension Questions in Experiment 2

Ambiguity	Pre- and Post-Ambiguity Constraints			
	Both Good	Good Pre-, Poor Post-	Poor Pre-, Good Post-	Both Poor
Ambiguous	87.50	86.88	80.60	75.60
Unambiguous	88.12	86.88	85.00	81.25

Discussion

These results confirm the results from Experiment 1 that ambiguity resolution grows progressively more difficult as the constraints become weaker. When the constraints were at their strongest, there was no ambiguity effect. The weaker "conflict" conditions, with one good constraint and one poor constraint, yielded small ambiguity effects, and the combination of two poor constraints yielded large ambiguity effects. This progression of effects indicates that multiple probabilistic constraints can modulate the ambiguity resolution process, and the extent to which constraints converge or conflict is an important factor in ambiguity resolution difficulty.

The disambiguation region in this experiment was much more informative than the ambiguous region, in contrast to Experiment 1, in which both regions were informative. In this respect, Experiment 2 more closely resembles other results in the literature. Most previous studies have added extra words to create the unambiguous control condition, and most have found the ambiguous region uninformative. The failure to find a reverse ambiguity effect in Experiment 2 is therefore not surprising, but it is still important to replicate the effect from Experiment 1, in order to substantiate the claim that reading times in the ambiguous region reflect the constraints available at that point. One of the goals of Experiment 3 is therefore to replicate Experiment 1's pattern of a reverse ambiguity effect followed by an ambiguity effect in the presence of unhelpful constraints. An additional goal is to extend the investigation of constraints to include situations in which strong or weak competitors with the reduced relative interpretation remain after post-ambiguity constraints have arrived in the input.

EXPERIMENT 3

The results of Experiment 1 supported the hypothesis concerning the effect of the number of competitor interpretations on ambiguity resolution, via manipulations of transitive-only or optional verbs, but a more general hypothesis has not yet been addressed—that it is not simply the number of competitor interpretations that modulates difficulty of ambiguity resolution, but also the strength of those alternatives relative to the intended interpretation. Comprehension of the reduced relative is predicted to be easier when there is one weak competitor argument structure than when there is one strong competitor. Experiment 3 investigates this prediction by holding constant the number of competitors and varying their strength relative to the reduced relative interpretation.

The ambiguous verbs in this experiment are all optional verbs, permitting both an active transitive and an active intransitive competitor to the

correct reduced relative interpretation. The verbs differ in the relative strength of these alternatives, however. Half of the ambiguous verbs are more frequently used transitively than intransitively in English (e.g. *pushed*), so that the active transitive interpretation is strong, and the active intransitive interpretation is weak. The other half have a more frequent intransitive usage in English (e.g. *moved*), so that for these verbs, the active intransitive interpretation is strong and the active transitive interpretation is weak.

It is predicted that when these verbs are combined with a not-DO phrase, ambiguity resolution will be easier for the transitive-biased verbs than for the intransitive-biased verbs. In the case of transitive-biased verbs like *pushed*, the not-DO phrase inhibits this highly preferred interpretation, so that the parser is left with a choice between the reduced relative argument structure and the weak active intransitive argument structure. With only a weak competitor, the (correct) reduced relative argument structure will be able to accrue activation fairly easily. With the intransitive-biased verbs like *moved*, however, the not-DO phrase again inhibits the active transitive argument structure, but this interpretation was already weak, so that the parser is faced with a choice between the very frequent intransitive usage and the reduced relative interpretation. In this case, the active intransitive interpretation should be strongly preferred, producing a reverse ambiguity effect in the ambiguous region as comprehenders pursue this simple structure, followed by an ambiguity effect at the disambiguation when they discover their error.

Method

Subjects. Forty-eight MIT undergraduates were paid for participating. Two additional subjects were rejected for high error rates on the comprehension questions in the experiment.

Materials and Procedure. Stimulus construction procedures were similar to those in Experiment 1, in that verb triples were selected, and three unrelated sentences were written for each triple. The three levels of the Verb factor were unambiguous, biased-transitive ambiguous and biased-intransitive ambiguous. The transitive- and intransitive-biased verbs were drawn from norms collected by Connine et al. (1984). Eight verb triples were selected, yielding 24 experimental sentences, many of which were variants of the items in Experiment 1. Table 7 contains an example sentence for the verb triple *driven/pushed/moved*, and all stimuli are contained in Appendix 4. The good and poor not-DO phrases were four words long, as in Experiment 2. Pre-ambiguity constraints were not

TABLE 7
Example Sentences and Question from Experiment 3

Unambiguous, good constraint

The rancher could see that the nervous cattle/driven into the crowded pen/were afraid/of the cowboys.

Biased-transitive, good constraint

The rancher could see that the nervous cattle/pushed into the crowded pen/were afraid/of the cowboys.

Biased-intransitive, good constraint

The rancher could see that the nervous cattle/moved into the crowded pen/were afraid/of the cowboys.

Unambiguous, poor constraint

The rancher could see that the nervous cattle/driven a short distance away/were afraid/of the cowboys.

Biased-transitive, poor constraint

The rancher could see that the nervous cattle/pushed a short distance away/were afraid/of the cowboys.

Biased-intransitive, poor constraint

The rancher could see that the nervous cattle/moved a short distance away/were afraid/of the cowboys.

Comprehension question varied with the verb manipulation

Did someone drive/push/move the cattle? Answer: YES

manipulated; all subject nouns were animate. This choice is conservative, as animate subjects have been shown to promote an incorrect active structure (Trueswell et al., in press; Experiment 2, this study).

Comprehension questions for each sentence focused on the verb and subject noun phrase, as in Experiment 2. Six stimulus lists were prepared, with four sentences in each of the six conditions on each list.

The procedure was identical to that in the previous experiments. The subjects read 5 practice items and 54 filler items (some of which were experimental items for an unrelated experiment), in addition to the 24 experimental items.

Results

Before turning to the reading time data, it is necessary to examine the comprehension question data, which produced rather different results than in the previous experiments. Table 8 shows that although comprehension

TABLE 8
Percent Correct Responses to Comprehension Questions in Experiment 3

Post-ambiguity constraint	Verb		
	Unambiguous	Biased-transitive	Biased-intransitive
Good	91.15	88.54	74.48
Poor	91.15	83.85	64.58

of unambiguous sentences was comparable to that in the previous experiments, the ambiguous sentences that contained poor post-ambiguity constraints were comprehended very infrequently.

Analyses of the comprehension data revealed a robust main effect of verb [$F_1(2,94) = 43.80, P < 0.001$; $F_2(2,46) = 10.02, P < 0.001$]. Accuracy in the biased-intransitive condition (69.5%) was lower than in both the unambiguous and biased-transitive conditions (91.2 and 86.2%, respectively; all P 's < 0.001); these latter two conditions were only marginally different from one another [$F_1(1,47) = 4.56, P < 0.05$; $F_2(1,23) = 2.70, P > 0.10$]. Comprehension accuracy was also higher for good-constraint sentences (84.7%) than for poor-constraint sentences (79.9%), producing a reliable effect of post-ambiguity constraint [$F_1(1,47) = 8.02, P < 0.01$; $F_2(1,23) = 5.47, P < 0.05$]. This effect of constraint was strongest for the biased-intransitive verbs and non-existent for the unambiguous items, but the verb \times constraint interaction was not reliable [$F_1(2,94) = 2.30, P > 0.10$; $F_2(2,46) = 1.67, P > 0.10$].

Although the accuracy data reveal the predicted effects of the verb and post-ambiguity constraint manipulations, the low accuracy in the more difficult conditions poses a problem for the reading time analyses. It is not clear how to interpret the reading time data for those conditions in which accuracy was scarcely better than chance, particularly when some cells were empty or had only one observation when error trials were removed. Reading times were therefore examined only for the good constraint conditions, as that would still allow hypotheses about the effects of the verb bias to be tested. Only those items with correct comprehension question responses were included in the analyses. Length adjustment regressions and trimming were conducted as in Experiment 2, removing less than 3% of the data.

Reading times in the good-constraint condition are shown in Fig. 4, which indicates that as predicted, the biased-intransitive sentences proved more difficult than the unambiguous controls, but the biased-transitive sentences were not different from the unambiguous condition. The effect of verb bias can be seen most clearly at the disambiguation, where the

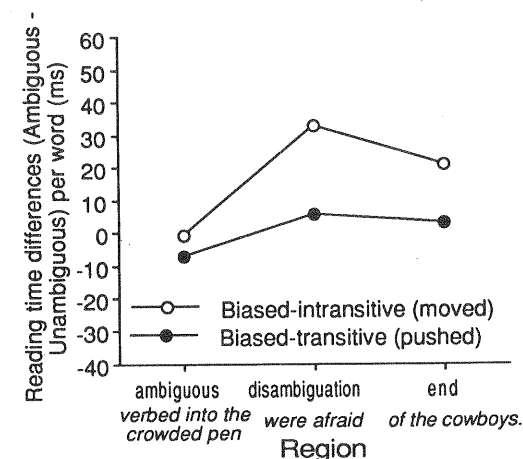


FIG. 4 Length-adjusted reading time differences between ambiguous and unambiguous sentences in Experiment 3.

biased-intransitive condition produced a significant ambiguity effect [$t_1(47) = 2.83, P < 0.01$; $t_2(23) = 4.00, P < 0.001$], but the biased-transitive condition did not differ from the unambiguous condition (t 's < 1). At the end region, there is again no difference between the biased-transitive condition and the unambiguous condition (t 's < 1), and the effect in the biased-intransitive condition is not reliable [$t_1(47) = 1.46, P < 0.10$; $t_2(23) = 1.25, P > 0.10$].

The effects in the ambiguous region do not show the longer reading times with ambiguity that were found with the extra-words disambiguation in Experiment 2, but neither do they appear to show any reverse ambiguity effects. Because these data were collected in the good-constraint condition, in which the constraining information arrives at the second word of this five-word region, any reverse ambiguity effect would be expected to be short-lived. A closer examination of this region supports this hypothesis. Figure 5 shows the ambiguous-unambiguous difference scores at each of the five words in the ambiguous region, revealing a reverse ambiguity effect in the first two words of the region that is gone by the third word. This pattern produced a robust verb \times word interaction [$F_1(8,376) = 4.50, P < 0.001$; $F_2(8,184) = 3.42, P < 0.001$].

The effects at each word in the ambiguous region were analysed. The reverse ambiguity effect was robust at the first word (the ambiguous verb) when it was a biased-transitive verb [$t_1(47) = 2.13, P < 0.05$; $t_2(23) = 2.63, P < 0.01$], but not when the verb was biased-intransitive [$t_1(47) = 1.07, P > 0.10$; $t_2(23) = 1.09, P > 0.01$]. Though reading times are adjusted for

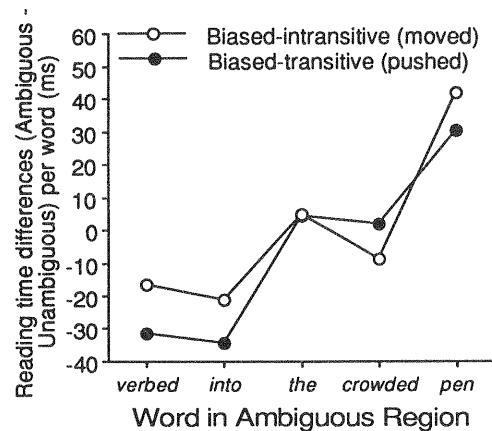


FIG. 5 Length-adjusted reading time differences at each word in the ambiguous region in Experiment 3.

length, these comparisons are across three different verbs, and so any analysis of this position must remain tentative. The rest of the region contains identical words in all three conditions, however. At the next word, which starts the good-constraint not-DO phrase, both ambiguous conditions produced reliable reverse ambiguity effects [biased-transitive: $t_1(47) = 3.61$, $P < 0.001$; $t_2(23) = 2.99$, $P < 0.01$; biased-intransitive: $t_1(47) = 1.72$, $P < 0.05$; $t_2(23) = 1.99$, $P < 0.05$]. These effects disappear in words 3–4 (*the crowded* in Fig. 5; all t 's < 1). At the last word of the region, however, both ambiguous conditions have begun to show an ambiguity effect, producing longer reading times in the ambiguous condition compared to the unambiguous condition [biased-transitive: $t_1(47) = 3.50$, $P < 0.001$; $t_2(23) = 2.48$, $P < 0.05$; biased-intransitive: $t_1(47) = 3.25$, $P < 0.001$; $t_2(23) = 2.40$, $P < 0.05$]. The presence of early reverse ambiguity effects confirms the prediction that when a simple active structure is permissible, reading times will be faster in the ambiguous condition compared to the unambiguous reduced relative clause. As more of the good constraint was encountered, the reverse ambiguity effect disappeared and was replaced by an ambiguity effect. The exact sizes and locations of these effects are probably also influenced by pre-ambiguity constraints, as Experiment 1B demonstrated, but these were not assessed here.

Discussion

This experiment demonstrated that ambiguity resolution becomes more difficult as the competitor interpretations become stronger: Ambiguous verbs with biased-intransitive interpretations produced an ambiguity effect

at the disambiguation, but ambiguous verbs with biased-transitive interpretations did not. Both ambiguous conditions allowed both active transitive and active intransitive alternative interpretations, but the strength of these competitors differed. When the not-DO phrase eliminated a strong competitor, leaving only weak competitors, parsing was relatively easy, but when a strong competitor remained, ambiguity resolution was difficult.

These results confirm the hypothesis that it is not the absolute number of competitors that makes an ambiguous sentence difficult to comprehend, but rather the strength of these alternatives relative to the correct interpretation. This is the outcome that is predicted if alternatives are competing for activation: The correct reduced relative argument structure can accrue activation when its competitors are weak, but not when a competitor is strong.

Returning to the contrast between the Maslin and Bever sentences, it is now clearer why Maslin's second sentence, *Even the former mental patients heard here sound . . .*, is not very difficult. This sentence is interesting because its ambiguous verb *heard* permits three competitors to the reduced relative, shown in Table 1. Thus *heard* actually permits one more competitor than *raced*, yet Bever's sentence is the more difficult. However, as Experiment 3 has demonstrated, it is not the number of competitors that matters in ambiguity resolution—it is the strength of the alternatives. The most frequent interpretation of *heard* is transitive (Connine et al., 1984; Holmes, Stowe, & Cupples, 1989), and the other two competitors, intransitive and sentential complement, are rarer. Therefore, when the not-DO phrase *here* in Maslin's sentence inhibits the strong active transitive interpretation, the two weak competitors do not prevent the correct reduced relative structure from accruing activation. Thus the several weak competitors in Maslin's sentence interfere less with ambiguity resolution than does the one strong competitor in Bever's sentence. Indeed, both *heard* and *noticed*, which also takes a sentential complement, were included in the biased-transitive condition in Experiment 3, yet this condition did not differ from the unambiguous condition.

A second important finding in this experiment was the appearance of reverse ambiguity effects early in the ambiguous region. This result supports the argument that these effects were absent in Experiment 2 only because of the type of unambiguous condition used in that experiment. It thus appears that when ambiguous and unambiguous conditions contain the same number of words, the pattern of reading times in both ambiguous and disambiguation regions can reflect the effect of constraints on ambiguity resolution.

GENERAL DISCUSSION

In these experiments, the difficulty in resolving syntactic ambiguities was shown to be modulated by three types of probabilistic constraints in the input. These constraints were: (1) the frequencies of the alternative argument structures of ambiguous verbs; (2) the post-ambiguity not-DO phrases that make the active transitive interpretation less likely; and (3) pre-ambiguity information concerning the plausibility of alternative interpretations of the ambiguity. Different combinations of constraints were examined across the three experiments. First, Experiments 1 and 3 demonstrated the effects of the verb constraints. In both of these experiments, the more often a verb was used intransitively in English, the more difficult it was to interpret this verb in a reduced relative construction. The intransitive interpretation is a competitor to the reduced relative and is not affected by the not-DO constraint, and the higher the frequency of a verb's intransitive argument structure, the stronger this competitor becomes. This result confirms the hypothesis that difficulty in ambiguity resolution varies with the strength of the alternative interpretations available to compete with the reduced relative interpretation.

Second, Experiments 1 and 2 examined effects of the post-ambiguity and not-DO constraints. These effects were subtle and appeared primarily in the disambiguation and end regions: Ambiguity effects were generally smaller with good post-ambiguity constraints than with poor constraints. Comprehension accuracy in Experiment 3 was consistent with this pattern, so that accuracy for ambiguous sentences with poor post-ambiguity constraints was too low to permit analysis of the reading time data. Third, the effects of the pre-ambiguity plausibility constraints, already demonstrated in previous research (Trueswell et al., in press; Pearlmutter & MacDonald, 1992), were investigated with regression analyses in Experiment 1B and with a factorial manipulation in Experiment 2; good pre-ambiguity constraints were more helpful for ambiguity resolution than were poor pre-ambiguity constraints.

In addition to demonstrating effects of the three constraints independently, these experiments demonstrated how multiple probabilistic constraints could act together to promote or discourage the reduced relative interpretation. Experiments 1A and 2 showed that combinations of two poor constraints (verb and post-ambiguity in Experiment 1A and pre- and post-ambiguity in Experiment 2) produced the largest ambiguity effects at the disambiguation, and these combinations of converging poor constraints were the only conditions to produce ambiguity effects at the end region in both experiments. The comprehension data in Experiments 2 and 3 are consistent with the reading time effects (the comprehension questions in Experiment 1 did not specifically assess interpretation of the ambiguity). The correlational data in Experiment 1B also demonstrated the interaction

of multiple constraints, in that ambiguous region reading times were similar to unambiguous times when the pre- and post-ambiguity constraints supported the correct reduced relative interpretation, but a reverse ambiguity effect emerged when the two constraints promoted the simpler main verb interpretation.

Finally, Experiments 1 and 3 showed a relationship between reading times in the ambiguous region and the disambiguation. Reading times in both regions were modulated by the strength of the probabilistic constraints, so that when constraints strongly promoted a simpler interpretation than the relative clause interpretation that is forced by the unambiguous condition, a reverse ambiguity effect emerged in the ambiguous region, followed by an ambiguity effect at the disambiguation. By contrast, helpful constraints produced reading times in both regions that closely mimicked the unambiguous condition.

Taken together, these results provide important information about how several probabilistic constraints affect the ambiguity resolution process. The results revealed that a single type of constraint can vary markedly in strength, that multiple constraints do operate together, and that reading times vary as a function of whether constraints converge or conflict (e.g. Fig. 2). This investigation does not exhaust the list of probabilistic constraints for ambiguity resolution in the MV/RR construction. Other candidates include pragmatic constraints on the felicity of noun modification (Altmann & Steedman, 1988; Altmann et al., 1992; in press; Crain & Steedman, 1985; Spivey-Knowlton et al., 1993), tense shifts in the discourse (Trueswell & Tanenhaus, 1991; 1992), and the frequency of past tense versus past-participle uses of the ambiguous verb (Burgess & Hollbach, 1988). The limitations imposed by factorial designs make it extremely difficult to investigate more than one or two constraints at a time, yet it may turn out that the typical situation in natural speech or writing, as with Maslin's sentences, contains far more than two constraints to guide the ambiguity resolution process. Researchers who manipulate only one probabilistic constraint at a time (in an attempt to evaluate the garden-path and interactive models, or for other purposes) must therefore be sensitive to the possibility that a multitude of other probabilistic constraints may be varying or even working against the manipulated constraint.

The tradition in syntactic ambiguity research has been to relate work on probabilistic constraints to the question of whether or not an autonomous syntactic parser operates ahead of the point at which these constraints have their influence. The research here departs from that tradition, in that it was argued that the investigation of multiple probabilistic constraints is an important topic in its own right, one which generates testable predictions concerning the difficulty of ambiguity resolution. These investigations of processing difficulty, constraint strength, and the number of and strength of alternative interpretations available during ambiguity resolution, have

revealed some of the types of information that comprehenders can use during on-line language comprehension, including verb argument structure frequency information and not-DO constraints, which had not been studied previously for this ambiguity. Demonstrations of constraint use and constraint interaction have yielded important insights concerning the nature of the ambiguity resolution process, independent of the issue of whether syntactic information is a similar constraint or the province of an autonomous parser.

Though this research concerning the kinds of probabilistic constraints that are brought to bear on ambiguity resolution is interesting apart from the debate between the garden-path and more interactive constraint-based models concerning the time-course of constraint use, the opposite is not true: The time-course debate is itself dependent on research investigating the nature of the constraints. The continued exploration of probabilistic constraints, both within a language and cross-linguistically (e.g. Cuetos & Mitchell, 1988; MacWhinney & Bates, 1989), should eventually form an important part of the resolution of the time-course debate, because when researchers are better able to assess the range and strength of probabilistic constraints that affect ambiguity resolution, they will be better able to determine whether a first-stage parser ignores these constraints. This question is only one of several crucial questions in the study of syntactic ambiguity resolution, however, and further investigations of probabilistic constraints, their interactions and their limitations appear to offer much promise in illuminating the ambiguity resolution process.

Manuscript received December 1992
Revised manuscript received July 1993

REFERENCES

- Altmann, G.T.M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-328.
- 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. (in press). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory and Cognition*.
- Berg, G. (1991). *Learning recursive phrase structure: Combining the strengths of PDP and x-bar syntax*. (Tech. Report). Albany: SUNY. Department of Linguistics and Cognitive Science.
- Bever, T.G. (1970). The cognitive basis for linguistic structure. In J.R. Hayes (Ed.), *Cognitive development of language*. New York: John Wiley.
- Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Britt, M.A., Perfetti, C.A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: Context effects and their limits. *Journal of Memory and Language*, 31, 293-314.
- Burgess, C., & Hollbach, S.C. (1988). A computational model of syntactic ambiguity as a lexical process, pp. 263-269. *Proceedings of the 10th Annual Conference of the Cognitive Science Society*.
- Carlson, G.N., & Tanenhaus, M.K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Syntax and semantics, Vol. 21: Thematic relations*. London: Academic Press.
- Clifton, C., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, 4, 77-104.
- Connine, C.M., Ferreira, F., Jones, C., Clifton, C., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, 13, 307-319.
- Crain, S., & Steedman, M. (1984). On not being led up the garden path: The use of context by the psychological syntax processor. In D.R. Dowty, L. Karttunen, & A.M. Zwicky (Eds), *Natural language processing*. Cambridge: Cambridge University Press.
- Cuetos, F., & Mitchell, D.C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30, 83-106.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.
- Ford, M., Bresnan, J., & Kaplan, R.M. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*, pp. 559-586. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, 26, 505-526.
- Hawkins, J.A. (1990). A parsing theory of word order universals. *Linguistic Inquiry*, 21, 223-261.
- Holmes, V.M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language*, 28, 668-689.
- Just, M.A., Carpenter, P.A., & Woolley, J.D. (1982). Paradigm and processes in reading comprehension. *Journal of Experimental Psychology: General*, 3, 228-238.
- Kawamoto, A. (1988). Interactive processes in the resolution of lexical ambiguity. In S. Small, G. Cottrell, & M.K. Tanenhaus (Eds), *Lexical ambiguity resolution: Computational, linguistic, and psychological perspectives*. New York: Morgan Kaufman.
- Kurtzman, H.A., & MacDonald, M.C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, 48, 243-279.
- Levin, B., & Pinker, S. (1991). Introduction to special issue of *Cognition* on lexical and conceptual semantics. *Cognition*, 41, 1-7.
- Levin, B., & Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry*, 17, 623-661.
- MacDonald, M.C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692-715.
- MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (submitted). The lexical based nature of syntactic ambiguity resolution.
- MacWhinney, B., & Bates, E. (1989). *The cross-linguistic study of sentence processing*. Cambridge: Cambridge University Press.
- Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71-102.
- Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions during word-recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.

- Maslin, J. (1991). Problems of homeless: The individual stories. *New York Times*, 22 March, p. C8.
- McClelland, J.L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, 4, 287–336.
- 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.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165–178.
- Pearlmutter, N.J., & MacDonald, M.C. (1992). Plausibility and syntactic ambiguity resolution. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Society*, pp. 498–503. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Pollard, C., & Sag, I. (1988). *Information-based syntax and semantics, Vol. 1: Fundamentals*. Chicago, IL: University of Chicago Press.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior*, 22, 358–374.
- Rayner, K., Garrod, S., & Perfetti, C.A. (1992). Discourse influences during parsing are delayed. *Cognition*, 45, 109–139.
- Ross, J.R. (1967). *Constraints on variables in syntax*. Unpublished PhD dissertation, Department of Linguistics, Massachusetts Institute of Technology, Cambridge, MA.
- Rumelhart, D., & McClelland, J.L. (1982). An interactive activation model of context effects in letter perception, Part II: The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89, 60–94.
- Spivey-Knowlton, M., Trueswell, J., & Tanenhaus, M.K. (1993). Context and syntactic ambiguity resolution. *Canadian Journal of Experimental Psychology*, 47, 276–309.
- Steedman, M.J., & Altmann, G.T.M. (1989). Ambiguity in context: A reply. *Language and Cognitive Processes*, 4, S105–122.
- Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- Tanenhaus, M.K., & Carlson, G. (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.
- Trueswell, J.C., & Tanenhaus, M.K. (1991). Tense temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes*, 6, 339–350.
- Trueswell, J.C., & Tanenhaus, M.K. (1992). Consulting temporal context during sentence comprehension: Evidence from the monitoring of eye movements in reading. In *14th Annual Conference of the Cognitive Science Society*, pp. 492–497. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Taraban, R., & McClelland, J.L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, 27, 597–632.
- Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (in press). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*.
- Tyler, L.K. (1989). The role of lexical representations in language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

APPENDIX 1: ITEMS IN EXPERIMENT 1

Verbs are in the order unambiguous/transitive-only/optional, and the not-DO phrases are in the order (good constraint/poor constraint).

1. The ruthless dictator overthrown/captured/fought (in the coup/just before dawn) was hated throughout the country.
2. The thoroughbred horses shown/admired/raced (at the exhibition/all day long) wore black and red ribbons.
3. The airforce cadet drawn/presented/painted (in full uniform/every other weekend) was a symbol of patriotism.
4. The little girl chosen/selected/applauded (for the role/just after auditioning) was a very natural actress.
5. The suspected muggers seen/found/watched (in their car/around three o'clock) gave no evidence of guilt.
6. The tiny insects eaten/devoured/attacked (in the night/every single night) are the bats' major meal.
7. The Indian leaders overthrown/captured/fought (in their villages/just after regrouping) weren't prepared for European colonists.
8. The sleek greyhound shown/admired/raced (at the track/nearly every evening) won four enormous gold trophies.
9. The nude model drawn/presented/painted (in art class/every art class) stood completely still for hours.
10. The guest speaker chosen/selected/applauded (in the committee/almost without hesitation) spoke with humour and intelligence.
11. The rebellious children seen/found/watched (during recess period/every recess period) had already skipped six classes.
12. The Protestant missionaries eaten/devoured/attacked (in their camp/just after arriving) hadn't heard the cannibals approaching.
13. The ineffective emperor overthrown/captured/fought (in the revolution/just after midnight) feared he would be killed.
14. The giant bullfrogs shown/admired/raced (at the fair/every July fourth) were a favourite of children.
15. The handsome actor drawn/presented/painted (for publicity brochures/every Tuesday morning) attracted the movie producer's attention.
16. The young educators chosen/selected/applauded (in regional committees/every single year) received many awards for teaching.
17. The tropical birds seen/found/watched (in the jungle/all summer long) captured the young boy's imagination.
18. The numerous caribou eaten/devoured/attacked (during the winter/every winter night) could not escape the wolves.

APPENDIX 2: RAW AND LENGTH-ADJUSTED READING TIMES PER WORD IN THREE REGIONS IN EXPERIMENT 1

Length-adjusted reading times (given as deviations from predicted): Verb and post-ambiguity constraint

Region	Good Post-ambiguity Constraint			Poor Post-ambiguity Constraint		
	Unambiguous	Transitive-only	Optional	Unambiguous	Transitive-only	Optional
Ambiguous	-18.78	-13.00	-34.57	-7.54	-15.87	-22.00
Disambiguous	-35.47	-14.46	-6.21	-10.46	1.17	40.24
End	-10.73	5.06	8.85	-4.38	13.74	27.57

Raw reading times: Verb and post-ambiguity constraint

Region	Good Post-ambiguity Constraint			Poor Post-ambiguity Constraint		
	Unambiguous	Transitive-only	Optional	Unambiguous	Transitive-only	Optional
Ambiguous	378.37	390.80	366.67	402.46	399.70	390.78
Disambiguous	356.90	378.11	387.49	382.74	395.54	431.93
End	409.19	425.53	429.77	416.69	432.92	445.28

APPENDIX 3: ITEMS IN EXPERIMENT 2

Subject noun phrases are given in the order animate (poor pre-ambiguity constraint)/inanimate (good pre-ambiguity constraint). The four-word post-ambiguity not-DO phrases are given in the order good constraint/poor constraint. The unambiguous condition was created by adding that was/were between the subject noun and the ambiguous verb.

1. The news reports stated that the spy/microfilm concealed inside the secret passageway/most of the night was discovered when the maid began cleaning.
2. The head coach heard that the athlete/game observed from the busy sidelines/almost all evening long did not impress the professional recruiter.
3. The woman quickly realised that the boy/lamp carried up the long staircase/a very long distance seemed to get heavier with each step.
4. The contest results showed that the inspector/cheese praised by the international experts/just about every day was considered the best in the land.
5. The merchant's records revealed that a slave/necklace sold in the crowded bazaar/about four weeks earlier had once belonged to the beautiful queen.
6. The researchers carefully investigated why the emperor/carving worshipped on the tiny island/many long years ago was supposed to have had magical powers.
7. Although things were quiet now, the child/crate moved out of the way/several dozen yards away had been a nuisance all morning long.
8. Though the facts weren't clear, the detective/murder investigated on the evening news/some time before April appeared to be connected with the Mafia.
9. The management team believed that the workers/shipment transported to the polluted beaches/almost two thousand miles would help clean the oil spill.
10. The ship's captain believed that the stowaway/cargo surrendered to the port authorities/a short time ago was taken to the police station.
11. It annoyed the workers that the foreman/production supervised during the efficiency campaign/some of the time was ignored by management on other occasions.
12. The secret documents indicated that the scientists/chemicals guarded in the genetics lab/practically all day long were the key to the project's success.
13. The community was pleased that an intruder/crime witnessed from a nearby apartment/a few days earlier was not being described in the newspapers.
14. The dance teacher noticed that the ballerina/ballet studied during the lengthy rehearsals/many hours a day was an inspiration to the young dancers.
15. The woman could see that the child/shirt washed in the deep sink/quite a long time was still smudged with patches of soot.
16. After some very intense negotiations, the soldiers/weapons captured during the fierce battle/one night last month were exchanged for hostages the next day.
17. It was a shame that the poet/poetry admired in the English class/most of the time was not discussed in the new textbooks.
18. Amidst a dozen video cameras, the defendant/evidence examined in the crowded courtroom/almost all day long was damaging to the district attorney's case.
19. The concerned nurse noticed that the patient/wheelchair pushed down the long hallway/a short distance away seemed about to collapse.
20. Grandma and Grandpa remembered that the relatives/beaches visited during the spring break/just about every summer were a favourite with the teenagers.
21. Although the club was small, the freshman/award accepted after the campus review/a few weeks earlier would enhance the group's reputation for excellence.
22. The art teacher remarked that the children/flowers painted by the famous artist/many mornings in June were perfect subjects for an Impressionist work.

23. With no attempt at anonymity, the pupil/project graded by the physics professor/about ten minutes earlier was used as an example of plagiarism.
24. The historian now understood why the farmers/farms attacked during the moonless night/only twelve miles away were important to the tiny pioneer settlement.
25. Although the crew worked smoothly, the experts/equipment replaced during the press conference/almost every single week was/were necessary for news broadcasts from Kuwait.
26. The townspeople didn't know why the criminal/ransom paid for Timmy's safe release/a long time ago was never traced by the police.
27. The town was lucky, because the terrorist/bomb identified in the dark alleyway/one dark winter evening could have done terrible damage.
28. The school board learned that the lecturer/textbook approved in last week's meeting/only two weeks ago has now been criticised by several educators.
29. The manager finally realised why the clients/messages overlooked in the waiting room/a number of times would become extremely important to the company.
30. The videotape coverage showed that the hostages/supplies unloaded from the blue van/about an hour ago appeared to be in pretty good condition.
31. After some last minute arrangements, the student/essay selected in the writing contest/one day last week would be presented to the college president.
32. The magazine article implied that the actor/exhibit reviewed in the local papers/several nights on TV was praised by most of the critics.

APPENDIX 4: ITEMS IN EXPERIMENT 3

Verbs are in the order unambiguous/biased-transitive/biased-intransitive, and the not-DO phrases are in the order (good constraint/poor constraint).

1. The secret CIA report indicated that the ruthless dictator overthrown/chased/fought (in the violent coup/several hours before dawn) was hated throughout the country.
2. The stable hands knew that the thoroughbred horses shown/noticed/raced (at the important exhibition/almost all day long) wore black and red ribbons.
3. According to the press bulletin, the airforce cadet drawn/painted/studied (in his dress uniform/about two weeks ago) was a symbol of patriotism.
4. The music director complained that the clarinet player chosen/criticised/lectured (during the band practice/one day last week) was a very poor musician.
5. The police report revealed that the suspected muggers seen/heard/watched (in the stolen car/a few hours ago) gave no evidence of guilt.
6. The rancher could see that the nervous cattle driven/pushed/moved (into the crowded pen/a short distance away) were afraid of the cowboys.
7. The news reports indicated that the Kuwaiti hostages overtaken/passed/marched (on the desert highway/a short time ago) were very hungry and thirsty.
8. The young lieutenant suspected that the desperate stowaway taken/attacked/surrendered (on the lower deck/one night last week) was dragged before the captain.
9. According to the ancient legends, the Indian leaders overthrown/chased/fought (in their small camps/many long years ago) weren't prepared for European colonists.
10. The owners were pleased that the huge greyhound shown/noticed/raced (at the crowded track/many times a day) won four enormous gold trophies.
11. The students were amazed that the nude model drawn/painted/studied (in the art class/almost all day long) stood completely still for hours.
12. Some people weren't surprised that the guest speaker chosen/criticised/lectured (in the committee meeting/only two minutes earlier) replied in a haughty voice.

13. Everyone noted with satisfaction that the rebellious children seen/heard/watched (during the morning recess period/most of the morning) were well-behaved in the afternoon.
14. The warden didn't know that the dangerous prisoner driven/pushed/moved (into the dark compound/several dozen yards away) was hated by the guards.
15. The den mother realised that the cub scouts overtaken/passed/marched (on the winding trail/a few hours ago) would not finish the trek.
16. The deputy sheriff reported that the three criminals taken/attacked/surrendered (at the prison gate/only a few hours ago) were taken to the infirmary.
17. The TV documentary said that the ageing emperor overthrown/chased/fought (in the palace courtyard/a short time ago) feared he would be killed.
18. The local newspaper noted that the giant bullfrogs shown/noticed/raced (at the country fair/one warm Saturday afternoon) were a favourite of children.
19. The agent was pleased that the handsome actor drawn/painted/studied (for Hollywood publicity brochures/one day last week) attracted the movie producer's attention.
20. The experienced administrator saw that the young educators chosen/criticised/lectured (in the regional committees/almost every single year) had implemented very bold policies.
21. In a new children's story, the tropical birds seen/heard/watched (in the Brazilian jungle/many mornings in Brazil) captured a young boy's imagination.
22. The nurse was pleased that the hospital patient driven/pushed/moved (to the busy cafeteria/one morning last week) was very eager for breakfast.
23. The principal didn't know that the school children overtaken/passed/marched (in the wide hallway/a few minutes earlier) were being punished for tardiness.
24. No-one could confirm whether the fighter pilot taken/attacked/surrendered (at the crash site/only six miles away) was later given his freedom.