JOURNAL OF MEMORY AND LANGUAGE 33, 285–318 (1994)

Semantic Influences on Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution

JOHN C. TRUESWELL AND MICHAEL K. TANENHAUS

University of Rochester

AND

SUSAN M. GARNSEY

University of Illinois at Urbana

Ferreira and Clifton (1986, Experiment 1) found that readers experienced equal difficulty with temporarily ambiguous reduced relatives clauses when the first noun was animate (e.g., "The defendant examined by the lawyer was . . .") and when it was inanimate and thus an unlikely Agent (e.g., "The evidence examined . . ."). This data pattern suggested that a verb's semantic constraints do not affect initial syntactic ambiguity resolution. We repeated the experiment using: (1) inanimate noun/verb combinations that did not easily permit a main clause continuation, (2) a baseline condition with morphologically unambiguous verbs (e.g., "stolen"), (3) a homogeneous set of disambiguating prepositional phrases, and (4) a display in which all of the critical regions were presented on the same line of text. In two eye-movement experiments, animacy had immediate effects on ambiguity resolution: only animate nouns showed clear signs of difficulty. Post-hoc regression analyses revealed that what little processing difficulty readers had with the inanimate nouns varied with the semantic fit of individual noun/verb combinations: items with strong semantic fit showed no processing difficulty compared to unambiguous controls, whereas items with weak semantic fit showed a pattern of processing difficulty which was similar to Ferreira and Clifton (1986). The results are interpreted within the framework of an evidential (constraint-based) approach to ambiguity resolution. Analyses of reading times also suggested that the millisecond per character correction for region length is problematic, especially for small scoring regions. An alternative transformation is suggested.

Language comprehension takes place rapidly and, to a first approximation, incrementally. Because the interpretation of a sentence is strongly constrained by its syntactic structure, readers must make at least partial syntactic commitments as a sentence unfolds, even when faced with local syntactic indeterminacies. However, a variety of syntactically relevant constraints are likely to be available which could be used to inform these commitments. How and when these constraints are actually used remains an important but currently unresolved question.

A great deal of research on syntactic ambiguity resolution has been guided by comparisons among two classes of models: (1) *autonomous* (modular) models in which an initial commitment is made to a single syntactic structure using only a restricted domain of syntactic information, and (2) *inter-*

This work was partially supported by NIH Grants HD-22271 and HD27206, NSF Grant BNS-8617738, and the NIH/NEI Grant EY-01319 to the Center for Visual Science. We are grateful to Maryellen MacDonald, Ken McRae, Neal Pearlmutter, and Michael Spivey-Knowlton for allowing us to include our materials in their rating study. We are also grateful to Curt Burgess for his sentence completion norms. Address correspondence and reprint requests to John C. Trueswell, Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104-6196. Email: trueswell@cattell.psych.upenn.edu.

Accepted by Dr. Shoben.

0749-596X/94 $6.00

active models in which more than one alternative analysis can be evaluated using multiple sources of evidence.

In autonomous models, initial syntactic commitments are made using structure-building procedures which make use of only a restricted domain of syntactic information. For example, in the "garden-path" theory originally proposed by Frazier and Rayner (1982), the syntactic processor or "parser" builds the simplest syntactic structure consistent with phrase structure rules of the language. These structures are initially built without reference to other potentially relevant constraints, including discourse context, the local sentence context, and, in some versions of the model, information specific to particular lexical items (e.g., Frazier, 1987, 1989; Ferreira & Henderson, 1990; Clifton, 1993). Constraints ignored in the first stage of parsing are then used to evaluate, and if necessary, revise mistaken commitments.

In this class of models, the indeterminacy that arises during syntactic processing is initially resolved within the syntactic module. This explanation has the advantage of allowing initial syntactic decisions to be restricted to information internal to that module. However, in doing so, other relevant constraints are temporarily ignored. Thus, a characteristic of two-stage models is that they predict that parsing involves frequent temporary syntactic misanalysis or "garden paths" (Frazier & Rayner, 1982).

Interactive models have adopted a "constraint-based" or "evidential" approach to syntactic processing (e.g., Tarraban & McClelland, 1988; McClelland, Tarraban, & St. John, 1989). In recent variants of evidential models, ambiguity resolution is viewed as a continuous process in which the most likely syntactic alternatives are evaluated with respect to evidence provided by syntactic input as well as salient semantic and discourse-based constraints (e.g., MacDonald, 1992; Pearlmutter & MacDonald, 1992; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, in press; Trueswell, Tanenhaus, & Kello, 1993). Unlike autonomous models, evidential approaches allow for multiple constraints to be used in the initial resolution of local syntactic indeterminacy. Readers and listeners may experience difficulty when the evidence is inconsistent, or when a structure is of low probability. Garden paths arise when the correct analysis of a local ambiguity is temporarily unavailable.

Distinguishing between restricted domain and evidential approaches has proved to be difficult, in part because few, if any, models are explicit enough to make clear quantitative predictions. As a consequence, many results that would seem to provide evidence in favor of one approach can be accounted for by the other approach using different assumptions. Nonetheless, there are some results in the literature that would seem to provide clear support for one approach or the other (for recent reviews see Mitchell, 1994; Tanenhaus & Trueswell, 1994).

Perhaps the most compelling evidence in favor of restricted domain models comes from Experiment 1 in Ferreira and Clifton (1986). Ferreira and Clifton monitored eye movements while subjects read sentences beginning with reduced or unreduced relative clauses which began with noun phrases with animate or inanimate nouns. A sample set of materials is illustrated in:

1. a. The defendant examined by the lawyer turned out to be unreliable.
   b. The defendant that was examined by the lawyer turned out to be unreliable.
   c. The evidence examined by the lawyer turned out to be unreliable.
   d. The evidence that was examined by the lawyer turned out to be unreliable.

Sentences with reduced relative clauses such as (1a) are temporarily ambiguous in English when the same form (usually verb + "ed") can be used for the past tense and the passive participial forms of the verb. Thus a noun phrase followed by a morphologically ambiguous verb (e.g., "The defendant examined . . .") will be temporarily ambiguous between a main clause with a past tense verb and a reduced relative clause with a participial verb. In a main clause, the noun phrase will be the subject of the verb and it will typically play the semantic/thematic role of Agent in the event described by the verb. In a reduced relative clause, the noun phrase is the grammatical object of the verb and it plays the role of Theme or Patient. In a full or unreduced relative such as (1b) and (1d), the relative clause is unambiguously marked by the relative pronoun "who" or "that."

Sentence (1c) is also a reduced relative clause. However, a main clause with "evidence" as the subject of "examined" would be semantically incongruous because "evidence" does not have the appropriate semantic properties to be the Agent of an examining event. In contrast, "evidence" is a plausible "Theme." Thus sentence (1c) provides strong semantic constraints against the main clause structure and in favor of a reduced relative clause. In general, animate nouns are good Agents for most events, whereas inanimate nouns are often poor Agents and typically good Themes.

Restricted domain and constraint-based approaches each would predict that readers will experience some difficulty in reading reduced relative clauses with animate noun phrases such as (1a) compared to full relative clauses such as (1b), although they would do so for different reasons. In the garden-path model, the verb "examined" will be parsed as the matrix verb in a main clause because that attachment is syntactically simpler than treating the verb as part of a relative clause. Thus, the processing system will be garden-pathed when it encounters the following prepositional phrase ("by the lawyer") which is inconsistent with a main clause. Other types of explanations for the main clause bias for NP V sequences have been proposed by Ford, Bresnan, and Kaplan (1982), Pritchett (1988, 1992), Gibson (1991), and Weinberg (1993).

Constraint-based approaches predict a main clause bias on the basis of frequency and supporting contextual constraints. An NP followed by an "ed" verb at the beginning of a sentence is more likely to be the start of a main clause containing a past-tense verb than it is to be a reduced relative clause containing a participial verb (Tabossi et al., in press). This was the basis for the NVN strategy proposed by Bever (1970). Thus, at the verb, the main clause would be more active than the reduced relative. The main clause would receive additional contextual support because the animate noun would have the right semantic properties to be the Agent of the verb. In addition, a main clause would be more consistent with the presuppositions present in the null context than would a reduced relative clause (Crain & Steedman, 1985; Ni & Crain, 1991; Trueswell & Tanenhaus, 1991, 1992). When the prepositional phrase "by the lawyer" was encountered, the reader would experience difficulty because the phrase provides strong evidence against the more active main clause structure and strong evidence in support of the less active reduced relative clause structure.

The predictions for reduced relative clauses with inanimate nouns are more complicated. Within a restricted domain framework, there is a choice about whether or not to include thematic information such as animacy as part of the input to the first-stage parser. The argument for including animacy is that it is morphologically marked in many languages, i.e., it is a semantic feature that can have clear morphological/syntactic reflexes. The arguments against including animacy are more compelling. Inanimate noun phrases can often be subjects in main clauses (e.g., "The car towed the trailer."). In addition, attempts to include syntactically relevant semantic features as part of the syntax run into a slippery slope problem. It is difficult to draw a clear boundary between semantic features that have special syntactic characteristics and those that do not. This is the primary reason why the notion of semantic features which provide constraints on the selection of verb arguments within the syntax (i.e., "selectional restrictions") was largely abandoned within linguistic theory. As McCawley (1971) pointed out, "a person who utters 'My toothbrush is alive and is trying to kill me' should be referred to a psychiatric clinic not to a remedial English course."

If thematic constraints are used in evaluating and revising initial analyses, as Rayner, Carlson, and Frazier (1983) proposed, then readers should still experience a garden path when reading a sentence beginning with a reduced relative clause with an inanimate subject NP. However, the animacy information might speed recovery from the misanalysis.

While this is the most natural prediction for restricted domain models, exactly *when* semantic effects will be seen will depend on specific assumptions about how quickly a syntactic commitment can be evaluated and revised. If semantic effects are delayed, there should be a temporal window of several words in which animacy would not interact with ambiguity resolution. This pattern was obtained by Rayner et al. (1983) using pragmatically biasing contexts. The limiting case, however, is that revisions can occur almost immediately after a commitment is made. Under these conditions, an interaction between animacy and ambiguity could still be taken as evidence for a temporary garden path as long as the ambiguous sentences with inanimate nouns were still more difficult than the unambiguous controls.

Constraint-based models predict that reduced relative clauses with inanimate nouns will, in general, be easier to process than reduced relatives with animate nouns. In order to see why, it will be helpful to sketch out a few details about the type of model that we have been developing. We are assuming that a verb makes available in parallel all of its possible thematic roles, or thematic grids (Carlson & Tanenhaus, 1988; Tanenhaus & Carlson, 1989; Tanenhaus, Carlson, & Trueswell, 1989; Pearlmutter & MacDonald, 1992). For ambiguous verbs, the alternative sets of thematic roles for each possible reading will be activated. For the past-tense reading of most -ed verbs, the preceding NP will typically play the role of Agent or Instrument, whereas for the participial reading, the preceding NP will be the Patient or Theme. The semantic fit between the NP and its possible thematic roles provides evidence that is used to evaluate the alternative argument structures for the past tense and participial readings (cf. Pearlmutter, Daugherty, MacDonald, & Seidenberg, 1993). An inanimate noun that is both a poor Agent for a particular verb and a good Patient or Theme will provide evidence in support of the argument structures linked to the participial form and thus evidence in support of the reduced relative structure. At the same time it would provide evidence against the past-tense/main clause. In the limiting case, the semantic constraint might be strong enough to completely eliminate any difficulty with reduced relative clauses. However, even a strong semantic constraint may not be able to completely override a frequency asymmetry as strong as that between the main clause and the reduced relative clause. Thus, the most natural prediction for constraint-based models is that "pre-ambiguity" cues such as animacy and "post-ambiguity" cues such as the disambiguating prepositional phrase will interact (MacDonald, 1992).

Thus far we have argued that both restricted domain and constraint-based models predict that readers will experience difficulty with reduced relative clauses with animate agents. In addition both classes of models could account for a pattern of results in which animacy interacted with ambiguity. However, constraint-based models would have difficulty explaining the complete absence of an interaction, whereas restricted domain models would have trouble explaining an interaction with no residual effect of ambiguity for reduced relatives with inanimate noun phrases.

Ferreira and Clifton (1986) found the pattern of data that is most problematic for constraint-based models. They analyzed reading times for three scoring regions in their experimental sentences: (1) the first noun phrase ("the defendant"); (2) the ambiguous verb ("examined"); and (3) the disambiguating prepositional phrase ("by the lawyer"). Animacy had immediate effects: reading times were longer for the reduced relative clauses at the ambiguous verb when the preceding noun was inanimate. Nonetheless, animacy did not interact with ambiguity resolution at the prepositional phrase. First and second pass reading times were longer for the reduced relatives sentences compared to the unreduced relatives, regardless of animacy. Moreover, the size of the reduction effect for sentences with inanimate nouns was virtually the same as the reduction effect for sentences with animate nouns.

These results would seem to offer compelling evidence in favor of restricted domain models. The syntactic module is apparently making decisions without taking into account a local constraint that is both relevant, and clearly available—as indicated by the animacy effects at the verb. This finding is difficult to reconcile with any version of a constraint-based model of ambiguity resolution.

However, there are aspects of the materials used by Ferreira and Clifton (1986) and the presentation mode that was used to display the sentences that, taken together, might account for the results. First, the manipulation of thematic constraint was weak for many of the test sentences. Inspection of the materials indicates that about half of the sentences with inanimate nouns could have been plausibly continued as a main clause up through the verb, either because the inanimate noun could play the role of Instrument, e.g., "The car towed . . . ," or because the verb had an "ergative" reading, e.g., "The trash smelled . . .". In a constraint-based model like the one we sketched earlier, an inanimate noun provides strong evidence against a main clause analysis *only* insofar as it is incompatible with the argument structures that can occur with the past-tense form of the verb. Thus, the relevant dimension is actually degree of semantic fit between the noun phrase and its possible thematic roles.

Second, the display mode used by Ferreira and Clifton (1986) might have introduced differences between the reduced and unreduced sentences. Ferreira and Clifton used a 42-character per line display. As a consequence, the test sentences, which averaged about 70 characters in length, had to be displayed across more than one line. Sentences with unreduced relative clauses typically had more than 40 characters up through the disambiguating region. Thus line breaks had to be placed in the middle of the disambiguating region (the prepositional phrase) for these sentences. Given this constraint, Ferreira and Clifton made the reasonable decision to equate the position of the line break for reduced and unreduced relative clauses with the same content. However, this meant that the first line of a sentence with a reduced relative clauses, which was usually nine characters shorter than its unreduced counterpart, ended well before the end of the screen, resulting in an unnatural looking display. Pilot work we have conducted shows that reading times are longer to sentences with early line breaks. A similar effect in the Ferreira and Clifton (1986) experiment would have elevated reading times in the disambiguating region of the sentences with reduced relative clauses.

Finally, unreduced relative clauses may not provide an appropriate baseline for reduced relative clauses. Readers often do not fixate on short function words. Thus the landing sites for saccades to the unreduced relative clauses might have been different than those to the reduced relative clauses. Trueswell et al. (1993) have recently demonstrated that landing site differences can result in differences in fixation duration which superficially look like syntactic misanalysis effects. In addition, unreduced relative clauses might be somewhat easier to process than reduced relatives for reasons that are unrelated to ambiguity. Information about the structure, and the resulting processes, are distributed across several words for unreduced relatives ("that was examined") but concentrated at the verb for reduced relatives (Trueswell & Tanenhaus, 1991).

Because of the theoretical importance of Ferreira and Clifton's results (1986) results, we repeated the experiment using materials in which the inanimate nouns were inconsistent with main clause continuations. We used an 80-character display in order to present all of the scoring regions of the test sentences on a single line. In addition, we included conditions with morphologically unambiguous verbs to allow us to factor out any differences in difficulty between reduced and unreduced relative clauses that are unrelated to ambiguity.

Experiment 1

Method

Subjects

Twenty-four undergraduates from the University of Rochester participated in the experiment for course credit. All subjects were native English speakers with uncorrected vision.

Equipment

The eye movements of each subject were recorded using a Stanford Research Institute Dual Purkinje Eyetracker (Fifth Generation). The eyetracker transmitted information about horizontal and the vertical eye position angle to a PDP-11 computer equipped with an analog to digital conversion board. Eye position was determined by sampling these signals every other millisecond. At the end of each trial, fixation positions and durations were computed and stored. Each fixation was represented by an x and y screen coordinate, a starting time, and an ending time. Two eyetrackers of the same type, one for the left eye, one for the right eye, were in the lab. Occasionally, one or the other was unavailable for use. Eye movements were recorded from either the left or right eye and viewing was binocular. Stimuli were presented on a Hewlett-Packard CRT.

Materials

Two groups of target sentences were generated (see Appendix A). The first group used 16 morphologically ambiguous verbs such as "examined" which use the same form for both the participle and the past tense. The second group used 12 morphologically unambiguous verbs such as "drawn" which have unique participle forms. Example sentences from each group are presented in Table 1. All of the verbs were embedded in subject–position relative clauses containing a noun phrase (e.g., "The defendant"), a verb ("examined"), a "by" phrase introducing an explicit Agent ("by the lawyer"), and a main clause verb phrase ("turned out . . ."). For the ambiguous verbs, four types of sentences were generated by combining two factors. The first factor was whether the first noun phrase was animate ("The defendant") or inanimate ("The evidence"). The criterion for selecting inanimate nouns was that none of the experimenters could think of a plausible main clause continuation of the sentence fragment containing the noun phrase and the verb (e.g., "The evidence examined . . . ?"). The second factor was whether the relative clause was a reduced relative or an unreduced relative.

For the unambiguous verbs, reduced and unreduced sentences with inanimate nouns were generated. We initially tried to develop a set of materials in which the same sentences were used with both ambiguous and unambiguous verbs, but this proved unfeasible. We chose not to manipulate an-

TABLE 1
EXAMPLE STIMULI FROM EXPERIMENT 1

Verb type	Noun type	Clause type	Example
Ambig.	Animate	Reduced	The defendant examined by the lawyer turned out to be unreliable.
		Unreduced	The defendant that was examined by the lawyer turned out to be unreliable.
	Inanimate	Reduced	The evidence examined by the lawyer turned out to be unreliable.
		Unreduced	The evidence that was examined by the lawyer turned out to be unreliable.
Unambig.	Inanimate	Reduced	The poster drawn by the illustrator was used for a magazine cover.
		Unreduced	The poster that was drawn by the illustrator was used for a magazine cover.

imacy for the unambiguous verbs because there are relatively few commonly known verbs with distinct past tense and past participle forms, and many of these verbs cannot be used naturally with both animate and inanimate objects. We settled upon the inanimate conditions because these provide the critical baseline needed to factor out any processing difficulty with reduced relatives with inanimate nouns that is unrelated to ambiguity resolution.

Four presentation lists were constructed by combining the 28 target sentences with 62 distractor sentences. The word length and frequency of the head noun phrase of the "by" phrase was controlled across the four stimulus lists. Each test sentence was followed by at least one distractor sentence. The distractor sentences contained a variety of sentence types including sentences using main clauses with past tense verbs. The four sentence types for each of the ambiguous verbs and the two sentence types for the unambiguous verbs were rotated through the four lists. Each subject was presented with 10 practice sentences and one of the four lists.

Procedure

Because small head movements decrease the accuracy of the eyetracker, a bite bar was made for each subject at the beginning of the testing session. Subjects were given instructions and seated in front of the computer screen, with the subject's eyes approximately 75 cm from the screen. All sentences appeared in upper case. The visual angle of each character was approximately 18 min of arc. At the beginning of the experiment, the brightness of the CRT was adjusted to the subject's comfort. The eyetracker was aligned and the signal from the eyetracker was calibrated with the screen coordinates. During the calibration procedure, the subject fixated on a series of screen positions, with the computer sampling the eyetracker at each position. These samples were then used by the computer to derive a set of linear equations that converted the horizontal eye position signal into horizontal screen coordinates and the vertical signal into vertical screen coordinates.

Each trial consisted of the presentation of a single sentence. All of the sentences fit on a single line of the screen. Before the sentence was presented, a fixation cross was displayed at the starting position of the sentence. The subject fixated on the cross and pressed a button on a button box to display the sentences. The subject read the sentence silently and then pressed the button to signal that he or she was finished. After each sentence, the calibration was checked by displaying a moving cross-hair controlled by the subject's eye movements along with the stationary fixation cross of the next trial. If the experimenter judged that the eye position did not line up adequately, the computer was recalibrated. (Recalibrations were typically not necessary.) During each trial, the experimenter monitored the SRI control panel and recorded the number of any trial in which there was a track loss. After about a fifth of the sentences, the experimenter asked a yes/no question about the sentence just read. Subjects answered by tapping on the table in front of them. They were given feedback as to whether their answer was correct. Each reading session lasted approximately 30 min. Subjects were allowed to release the bite bar between sentences at any time during the experiment. Subjects usually took one or two breaks.

Results

Before presenting the reading time results, we will briefly address two methodological issues. First, we will discuss the ms/character transformation which is often used in reading studies, and was used by Ferreira and Clifton (1986). Second, we will present an analysis of the landing site probabilities for initial fixations across the different conditions in order to determine whether the probability of initially fixating on the crucial regions of the relative clause was similar for the reduced and unreduced relatives.

Milliseconds per Character (ms/char)

Many eye-tracking studies, including Ferreira and Clifton (1986), report reading time in ms/char in order to adjust reading times for string length within and across scoring regions. In Appendix B, we show that this transformation is problematic, especially for small scoring regions. The ms/char transformation is computed by dividing the reading time for a region by the number of characters in the region, including spaces and punctuation. This transformation assumes that reading time is a linearly increasing function of the number of characters with a value of 0 for a region length of 0 characters. Ferreira and Clifton (1986) argued that the zero-intercept assumption is incorrect for self-paced reading, since each button-press adds a constant amount to each reading time. They recommended an alternative procedure, namely analyzing deviations from expected reading times as determined by the best linear fit for reading times as a function of character length. Appendix B shows that the zero-intercept assumption is also incorrect for eye-movement data. The ms/char transformation when applied to our data: (a) leaves a significant amount of variance due to string length, (b) reverses the relationship between string length and reading time, and (c) introduces a non-linearity. For this reason, we report analyses using unadjusted reading times, because string length was controlled for within regions.

First Fixation Probabilities

The design of this study calls for comparisons between reading times to the same words in reduced relative clauses (e.g., "The evidence *examined by the lawyer* . . .") and full relative clauses (e.g., "The evidence that was *examined by the lawyer* . . ."). A potential problem with this comparison is that readers often skip short function words (e.g., "that" or "was"). This can affect the landing sites and fixation durations on subsequent words (Trueswell et al., 1993). To determine whether the landing sites were similar for the reduced and unreduced relatives, first fixation probabilities were computed for the four words following the insertion point (verb, "by," "the," and noun). The results appear in Table 2. As can be seen in the table, the fixation probabilities were similar for the ambiguous and unambiguous sentences. To conserve space, a complete analysis of variance will not be reported here. However, the analysis did not reveal any reliable effects or interactions at any of the positions for the sentences with ambiguous verbs. For the unambiguous verbs, the effect of relative clause type was significant in the item analysis but not in the sub-

TABLE 2
THE PROBABILITY OF A FIRST FIXATION FOR THE VERB, "BY," "THE," AND THE NOUN (EXPERIMENT 1)

			Relative clause type			
Word	Verb type	NP type	Reduced	Unreduced	Δ	Mean
Verb	Ambig.	Animate	.99	.96	+.03	.98
		Inanimate	.98	.92	+.06	.95
		Mean	.99	.94		
	Unambig.		.93	.86	+.07	.90
"*by*"	Ambig.	Animate	.27	.21	+.06	.24
		Inanimate	.21	.20	+.01	.21
		Mean	.24	.21		
	Unambig.		.24	.20	+.04	.22
"*the*"	Ambig.	Animate	.58	.59	−.01	.59
		Inanimate	.56	.58	−.02	.57
		Mean	.57	.59		
	Unambig.		.60	.58	+.02	.59
Noun	Ambig.	Animate	.89	.87	−.02	.88
		Inanimate	.85	.81	−.04	.83
		Mean	.87	.84		
	Unambig.		.87	.83	+.04	.85

ject analysis at the verb ($F1(1,10) = 1.79$; $F2(1,10) = 8.44$, $p < .05$), whereas the same effect was significant in the subject analysis but not the item analysis at the "by" ($F1(1,10) = 5.92, p < .05; F2(1,10) = .60$). The similarities in landing sites across conditions make it unlikely that fixation patterns could contribute to differences between the reduced and unreduced reading times. Nevertheless, for the most crucial comparisons, we will also report statistics comparing only the sentences with reduced relatives.

Initial Processing (First Pass Reading Times)

Following Ferreira and Clifton (1986), the test sentences were divided into four scoring regions: (1) the *noun phrase*, e.g., "The defendant"; (2) the *verb*, e.g., "examined"; (3) the "*by*"*-phrase*, e.g., "by the lawyer"; and (4) the *first two words of the main verb phrase*, e.g., "turned out."

When the reader's eye position entered a scoring region, fixation durations were considered to be part of a first pass reading if the subject had not read the region before, and if the subject had not already read any of the words beyond that region. A first pass reading time was obtained by summing the durations of all left-to-right fixations in a region plus any regressions made to other points *within* that region. First pass reading was considered complete when the reader made either a regressive eye movement to a prior region or a forward movement to a following region (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

First pass reading times. Figure 1 presents first pass reading times for each of the four scoring regions. As shown in Fig. 1a, the semantic properties of the noun clearly influenced reading times for sentences with ambiguous verbs. For sentences with animate noun phrases, reduced relatives took longer to read than unreduced relatives in the disambiguating "*by*"*-phrase* region. For sentences with inanimate noun phrases, the verb prior to the "*by*" *phrase* was the only position that showed any hint of longer reading times to reduced relatives as compared to unreduced relatives. However, sentences with unambiguous verbs (Fig. 1b) showed the same pattern.

Analysis of sentences with ambiguous verbs. Subject and item first pass means for sentences with ambiguous verbs were entered into separate ANOVAs with four fac-

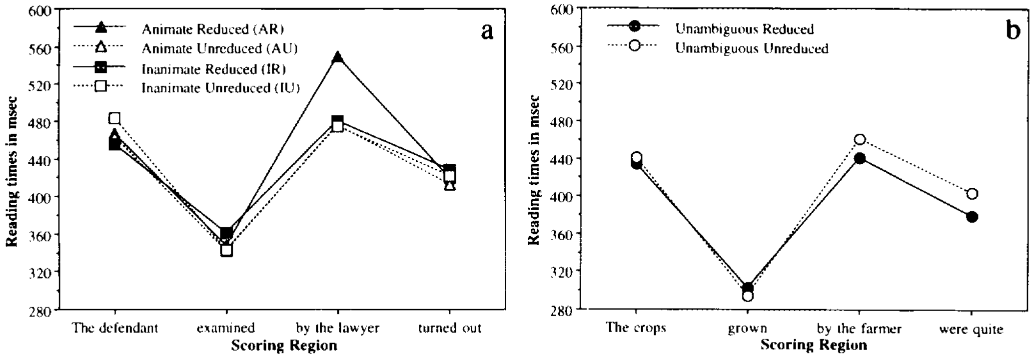

FIG. 1. Mean first pass reading times in ms for (a) sentences with Ambiguous verbs and (b) sentences with Unambiguous verbs. (Experiment 1.)

tors: list (four lists) or item group (four groups), region (the verb, the "by" phrase, and the first two words of the main verb phrase), animacy (animate and inanimate), and relative clause type (reduced and unreduced). A separate analysis of the prior noun phrase region showed no reliable effects or interactions and thus was not included.

When collapsing across all conditions, there was a reliable effect of scoring region ($F1(2,19) = 39.83$, $p < .01$; $F2(2,11) = 15.13$, $p < .01$). (We report Hyunh-Feldt adjusted probability levels for all effects involving region, since it has more than two levels which are not independent of one another.)

ANALYSIS BY REGION. Similar ANOVAs were conducted at each scoring region. At the *verb*, reading times for all four conditions were similar, resulting in no significant effects or interactions (all $Fs < 1$).

At the "*by*" *phrase*, there was a reliable interaction between animacy and the relative clause type in the subject analysis but not in the item analysis ($F1(1,20) = 6.17$, $p < .05$; $F2(1,12) = 2.80$, $p = .11$). Animate (reduced and unreduced) relative clauses were on average 36 ms longer than the inanimate conditions, resulting in a marginal effect of animacy ($F1(1,20) = 3.31$, $p < .1$; $F2(1,12) = 2.68$). Also, reduced relatives were on average 41 ms longer than unreduced relatives, resulting in a reliable effect of relative clause type ($F1(1,20) = 3.39$, $p < .1$; $F2(1,12) = 4.78$, $p < .05$). The difference between reduced and unreduced relatives was +75 ms for the animates and +6 ms for inanimates, resulting in a significant effect of relative clause type for the animates ($F1(1,20) = 6.59$, $p < .05$; $F2(1,12) = 7.23$; $p < .05$), but not for the inanimates ($Fs < 1$). We also looked at the effect of animacy for the reduced and unreduced relative clauses separately. Animate relative clauses were longer than inanimate relative clauses when the clause was reduced (70 ms), but not when it was unreduced (1 ms), resulting in a reliable effect of animacy for reduced relatives ($F1(1,20) = 8.47$, $p < .05$; $F2(1,12) = 7.96$, $p < .05$), but not for unreduced relatives (both $Fs < 1$).

At the *first two words of the verb phrase*, reading times for all four conditions were similar and there were no reliable effects or interactions (all $Fs < 1$).

Analysis of sentences with unambiguous verbs. Subject and item first pass means for sentences with unambiguous verbs were entered into separate ANOVAs with three factors: list (two lists) or item group (two groups), region (the verb, the "by" phrase, and the first two words of the main verb phrase), and relative clause type (reduced and unreduced). The ambiguous sentence types had four sentence conditions, whereas the unambiguous sentence types had only two conditions. Thus, the same unambiguous sentence items within a condition were repeated in one of the other lists; resulting in two lists for these conditions. A separate analysis of the prior noun phrase region showed no reliable effects or interactions, and thus was not included. (One subject was not included in the ANOVA because he or she consistently skipped over the verb during first pass reading in the unambiguous conditions.) When collapsing across both conditions, there was a reliable effect of scoring region ($F1(2,20) = 55.67$, $p < .01$; $F2(2,9) = 21.34$, $p < .01$). ANOVAs were also performed at each scoring region. There were no reliable effects or interactions at any of these positions.

Reprocessing Effects (Second Pass Reading Times)

Figure 2 presents second pass reading times in milliseconds for each of the four scoring regions. Second pass reading times reflect any re-reading of these regions. As shown in Fig. 2a, the semantic properties of the noun clearly influenced reading times for sentences with ambiguous verbs. For sentences with animate noun phrases, reduced relatives took longer to read than unreduced in all four regions, indicating that subjects often reread these regions in the reduced condition. For sentences with inanimate noun phrases, the *verb* prior to the "*by*"-*phrase* was the only position that showed longer reading times to reduced relatives as compared to unreduced. Again, sentences with unambiguous verbs (Fig. 2b) showed a pattern similar to the ambiguous inanimate conditions. Reduced relatives were slightly longer than unreduced relatives at the verb region.

Analysis of sentences with ambiguous verbs. Subject and item second pass means for sentences with ambiguous verbs were entered into separate ANOVAs with four factors: list (four lists) or item group (four groups), region (the noun phrase, the verb, the "by"-phrase, and the first two words of the main verb phrase), animacy (animate and inanimate), and relative clause type (reduced and unreduced).

When collapsing across all four regions, there was a marginal interaction between animacy and type of relative clause ($F1(1,20) = 3.82$, $p < .1$; $F2(1,12) = 3.03$, $p < .1$). In total, second pass reading times were 178 ms longer for animate conditions as compared to the inanimate conditions, resulting in a significant effect of animacy ($F1(1,20) = 7.91$, $p < .05$; $F2(1,12) = 14.94$, $p < .05$). Overall, reduced relatives were 201 ms longer than unreduced relatives, resulting in a reliable effect of relative clause type ($F1(1,20) = 13.29$, $p < .05$; $F2(1,12) = 7.37$, $p < .05$). The total difference between reduced and unreduced relatives was +326 ms for the animates and +75 ms for the inanimates, resulting in a reliable effect of relative clause type for the animates ($F1(1,20) = 8.74$, $p < .01$; $F2(1,12) = 9.19$, $p < .05$), but not for the inanimates ($F1(1,20) = 2.64$; $F2(1,12) =$

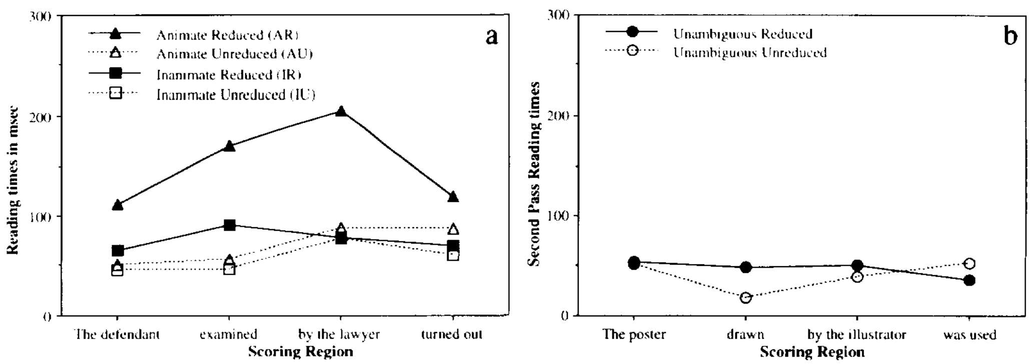

FIG. 2. Mean second pass reading times in ms for (a) sentences with Ambiguous verbs and (b) sentences with Unambiguous verbs. (Experiment 1.)

.07). We also looked at the effect of animacy separately for the reduced and unreduced relatives. The total difference between animate and inanimate conditions was +303 ms for reduced relatives compared to +52 ms for unreduced relatives, resulting in a significant effect of animacy for the reduced relatives ($F(1,20) = 6.78$, $p < .05$; $F2(1,12) = 9.40$, $p < .01$), but not for the unreduced relatives (both $Fs < 1$).

ANALYSIS BY REGION. Similar ANOVAs were conducted at each scoring region. To conserve space, we will only report analyses from the *verb* and the "*by*"-*phrase*. Effects similar to those reported for the "by" phrase were found in the first and last regions, although some effects only approached significance.

At the *verb,* the interaction between animacy and type of relative clause was not significant ($F1(1,20) = 2.36$; $F2(1,12) = 2.94$). Animate relative clauses were on average 45 ms longer than the inanimate conditions, resulting in a significant effect of animacy in the subject analysis but not the item analysis ($F1(1,20) = 6.39$, $p < .05$; $F2(1,12) = 2.50$). Also, reduced relatives were 80 ms longer than unreduced relatives, resulting in a reliable effect of relative clause type ($F1(1,20) = 12.74$, $p < .01$; $F2(1,12) = 14.10$, $p < .01$). Simple effects revealed a reliable effect of relative clause type when the preceding noun was animate ($F1(1,20) = 7.83$, $p < .05$; $F2(1,12) = 10.37$, $p < .01$). This same effect was only reliable in the subject analysis when the preceding noun was inanimate ($F1(1,20) = 5.53$, $p < .01$; $F2(1,12) = .85$).

At the "*by*"-*phrase,* there was a reliable interaction between animacy and the type of relative clause ($F1(1,20) = 4.26$, $p = .05$; $F2(1,12) = 4.86$, $p < .05$). Animate relative clauses were on average 69 ms longer than the inanimate conditions, resulting in a significant effect of animacy ($F1(1,20) = 7.20$, $p < .05$; $F2(1,12) = 9.52$, $p < .05$). Also, reduced relatives were on average 60 ms longer than unreduced relatives. This effect of relative clause type was marginal in the item analysis, and not significant in the subject analysis ($F1(1,20) = 2.64$; $F2(1,12) = 3.81$, $p < .1$). The difference between the reduced and unreduced relatives was +118 ms for the animates and +1 ms for the inanimates, resulting in a significant effect of relative clause type when the preceding noun was animate ($F1(1,20) = 5.15$, $p < .05$; $F2(1,12) = 8.01$, $p < .05$), but not when it was inanimate ($Fs < 1$). The difference between the animate and inanimate conditions was +127 ms for the reduced relatives and +10 ms for the unreduced relatives, resulting in a significant effect of animacy for reduced relative clauses ($F1(1,20) = 6.96$, $p < .05$; $F2(1,12) = 12.35$, $p < .01$), but not for unreduced relative clauses ($Fs < 1$).

Analysis of unambiguous verbs. Subject and item second pass means for sentences with unambiguous verbs were entered into separate ANOVAs with three factors: list (two lists) or item group (two groups), region (the noun phrase, the verb, the "by"-phrase, and the first two words of the main verb phrase), and relative clause type (reduced and unreduced). When collapsing across all four regions, there were no reliable effects or interactions (all $Fs < 1$). The only reliable effects at any position occurred at the *verb*. At this position, there was a reliable effect of relative clause type, but only in the subject analysis ($F1(1,22) = 9.57$, $p < .01$; $F2(1,10) = .40$).

Discussion

The results provide clear evidence that semantic constraints had immediate effects on ambiguity resolution. Reduced relative clauses with animate nouns had longer first and second pass reading times compared to their unreduced counterparts and compared to reduced relatives with inanimate nouns. Reduced relatives with inanimate nouns had similar first pass reading times to their unreduced counterparts. There was some suggestion that second pass reading times were slightly longer at the verb for reduced relatives with inanimate nouns compared to unreduced relatives, however, the same pattern occurred for the stimuli with unambiguous verbs.

The hint of some difficulty with the inanimates might be taken as evidence that the garden path was not completely eliminated. Recall, however, that for a constraint-based system, it is the semantic fit of the first noun phrase as the Agent or the Patient/Theme of the verb rather than animacy per se that is crucial. We will return to this issue in the discussion of Experiment 2. To anticipate, the difficulty with inanimates varies across items and reflects differences in the degree to which particular inanimate nouns were poor Agents and good Patients or Themes for the verbs they were paired with.

EXPERIMENT 2

The purpose of this experiment was to replicate the results of Experiment 1. In addition, we made three potentially important changes. First of all, we used mostly new target sentences in which semantic constraint was established empirically rather than by the experimenter intuition. We modified sentences that Burgess (1991) had constructed using completion norms. Burgess had subjects complete noun/verb sequences, such as "The evidence examined. . . ." He then constructed sentences with animate nouns that typically resulted in main clause completions and sentences with inanimate nouns that typically resulted in relative clause completions. Second, we used fewer relative clauses in the experiment in order to reduce any possible set effects. We eliminated the conditions with morphologically unambiguous verbs. Although these conditions provided important information in Experiment 1, they increased the total number of sentences with relative clauses, and perhaps more importantly, increased the number of sentences with reduced relatives that began with inanimate nouns. We also presented two sentences on each trial, with the test sentence always occurring first. Ferreira and Clifton (1986) had used this procedure. Finally, we presented the sentences in mixed case rather than in upper case. It is possible that upper case presentation enhanced context effects because word recognition is slower than it is with mixed case presentation.

Method

Subjects

Twenty students from the University of Rochester participated in the experiment. Subjects were paid $7 for 1 h of their time. All subjects were native English speakers with uncorrected vision.

Equipment

The equipment and laboratory set up were the same as Experiment 1, with the following exceptions. For all subjects, the right eye position was monitored. A Macintosh II computer equipped with an analog to digital converter board was used to collect data. The track loss signal from the eye tracker was collected on a third channel. Sampling of all three channels occurred every millisecond. The start time and end time of any track loss were stored for each trial. Stimuli were displayed on a 13-in. AppleColor High Resolution RGB monitor.

Materials

Sixteen sentences with morphologically ambiguous verbs were generated that were similar (and in some cases identical) to those used in the previous experiment. Verbs and nouns were selected from the Burgess (1991) sentence completion norms. In the Burgess study, subjects generated sentence completions to noun phrase–verb pairs (e.g., "The defendant examined . . ."). All animate noun phrases in this study had 100% main clause sentence completions when paired with a particular verb (e.g., "The defendant examined . . ." was completed as a main clause 100% of the time). All inanimate noun phrases had no more than 30% main clause completions when paired with the same verb (e.g., "The evidence examined . . ." had a 20% main clause completion rate). Appendix A presents a list of the stimuli and the completion percentages. Each trial consisted of reading two sentences. The second sentence was a natural continuation of the story established by the first sentence. Trials containing a target sentence always began with the target sentence.

Animate unreduced sentences in this study used "who was" instead of "that was" (e.g., "The defendant who was examined by . . ."). We made this change because of concerns about the naturalness of "that" as a relative pronoun for the animate noun phrases. Inanimate unreduced conditions used "that was" as in the previous experiment.

Four presentation lists were constructed by combining the 16 target sentence pairs with 44 distractor sentence pairs. Each test sentence was followed by at least one distractor sentence. The distractor sentences contained a variety of sentence types including sentences with verbs used in the past tense. The four sentence types were rotated through the four lists. Each subject was presented with five practice sentences and one of the four lists.

Procedure

The procedure was the same as Experiment 1, with the following exceptions. The subject's eyes were approximately 64 cm from the screen. All sentences appeared in mixed case in Courier fourteen point font. The visual angle of each character was approximately 12 minutes of arc, still allowing for one character resolution from the eyetracker position signals. Each trial consisted of the presentation of the sentence pair. Each line contained no more than 65 characters, and all the scoring regions appeared on the first line of text. The Macintosh mouse button was used instead of a button box. Comprehension questions appeared after about a third of the trials. Subjects responded by moving a mouse arrow into a YES or NO box and clicking the mouse button. Subjects were given feedback via the computer. The experimenter did not record track losses since the computer did this automatically. Finally, a line trace was drawn out, instead of a moving cross hair, to determine the accuracy of the eyetracker between trials.

Results

The reading time results are divided into three sections: first fixation probabilities, initial processing effects (first pass reading times), and reprocessing effects (second pass reading times).

First Fixation Probabilities

First fixation probabilities were determined for the verb and each word after the verb (Table 3). As can be seen in the table, fixation probabilities showed little or no differences. To conserve space, a complete analysis of variance will not be presented. However, at each word position, none of the possible effects or interactions were significant or approached significance. As in the previous study, the lack of reliable differences in fixation probabilities makes it unlikely that first pass reading times were contaminated by changes in landing sites. Nevertheless, for the most crucial compar-

TABLE 3
THE PROBABILITY OF A FIRST FIXATION FOR THE VERB, "BY," "THE," AND THE NOUN (EXPERIMENT 2)

		Relative clause type			
Word	NP type	Reduced	Unreduced	Δ	Mean
verb	Animate	.98	.92	+.06	.95
	Inanimate	.89	.92	−.03	.91
	Mean	.94	.92		
"*by*"	Animate	.26	.26	.00	.26
	Inanimate	.37	.24	+.13	.31
	Mean	.32	.25		
"*the*"	Animate	.60	.69	−.09	.65
	Inanimate	.60	.58	+.02	.69
	Mean	.60	.64		
noun	Animate	.86	.86	.00	.86
	Inanimate	.89	.91	−.02	.90
	Mean	.88	.89		

isons, we will also report statistics comparing only the sentences with reduced relative clauses.

Initial Processing Effects (First Pass Reading Times)

The test sentences were divided into four scoring regions: (1) the *noun phrase*, e.g., "The defendant"; (2) the *verb*, e.g., "examined"; (3) the "*by*"-*phrase*, e.g., "by the lawyer"; and (4) the *first two words of the main verb phrase*, e.g., "turned out."

Figure 3 presents first pass reading times in milliseconds for each of the four scoring regions. The same data pattern obtained as in the first experiment. Reduced relatives took longer to read than unreduced relatives in the disambiguating "*by*"-*phrase* region when the preceding noun phrase was animate, but not when it was inanimate.

Subject and item first pass means were entered into separate ANOVAs with four factors: list (four lists) or item group (four groups), region (the verb, the "by"-phrase, and the first two words of the main verb phrase), animacy (animate and inanimate), and relative clause type (reduced and unreduced).[1] When collapsing across all three regions, reduced relatives were in total 87 ms longer than unreduced relatives, resulting in a reliable effect of relative clause type ($F1(1,16) = 6.72$, $p < .05$; $F2(1,12) = 12.82$, $p < .01$). Animate conditions were in total 90 ms longer than inanimate conditions, resulting in an effect of animacy that was significant in the subject analysis and marginal in the item analysis ($F1(1,16) = 9.58$, $p < .01$; $F2(1,12) = 4.23$, $p < .1$). The interaction between these two effects was not significant (Fs < 1). There was also a

[1] First pass reading times to the initial noun phrase were excluded from the analysis. However, ANOVAs including this region revealed a reliable effect of Animacy in the subject analysis but not the item analysis ($F1(1,16) = 5.90$, $p < .05$; $F2(1,12) = 1.52$). This effect is likely to be a string length effect, since inanimate nouns were slightly longer than animate nouns in this experiment.

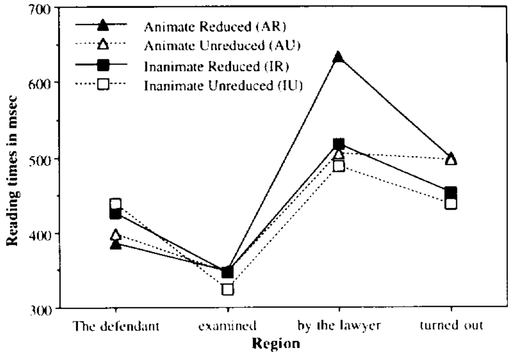

FIG. 3. Mean first pass reading times in ms (Experiment 2).

reliable effect of scoring region ($F1(2,15) = 29.02$, $p < .01$; $F2(2,11) = 22.01$, $p < .01$).

Analysis by region. Similar ANOVAs were conducted at each scoring region. At the *verb*, reading times were similar for all four conditions and there were no significant effects or interactions.

At the "*by*"-*phrase*, the interaction between animacy and relative clause type approached significance in the subject analysis, but not the item analysis ($F1(1,16) = 3.13$, $p < .1$; $F2(1,12) = 1.99$). Animate relative clauses were on average 67 ms longer than the inanimate conditions, resulting in a significant effect of animacy ($F1(1,16) = 7.97$, $p < .05$; $F2(1,12) = 5.27$, $p < .05$). Also, reduced relatives were 79 ms longer than unreduced relatives, resulting in a reliable effect of relative clause type ($F1(1,16) = 13.60$, $p < .01$; $F2(1,12) = 16.10$, $p < .01$). The difference between reduced and unreduced relatives was +128 ms for the animates and +29 ms for the inanimates, resulting in a significant effect of relative clause type when the preceding noun phrase was animate ($F1(1,16) = 9.01$, $p < .01$; $F2(1,12) = 6.73$, $p < .05$), but not when it was inanimate ($F1(1,16) = 1.85$; $F2(1,12) = 1.18$). The difference between animate and inanimate conditions was +116 ms for the reduced relatives and +17 ms for the unreduced relatives, resulting in a significant effect of animacy when the relative clause was reduced ($F1(1,16) = 6.84$, $p <$

.05; $F2(1,12) = 4.49$, $p = .05$), but not when it was unreduced (Fs < 1).

At *the first two words of the verb phrase,* there was an effect of animacy that was significant only in the subject analysis ($F1(1,16) = 7.22, p < .05$; $F2(1,12) = 2.86$). No other effects or interactions were significant (all Fs < 1).

Reprocessing Effects (Second Pass Reading Times)

Figure 4 presents second pass reading times in milliseconds for each of the four scoring regions. As shown in the figure, a similar data pattern obtained as in the first experiment. For sentences with animate noun phrases, reduced relatives took longer to read than unreduced in all four regions, indicating that subjects often re-read these regions in the reduced condition. For sentences with inanimate noun phrases, the noun phrase and the verb prior to the "by"-phrase were the only positions that showed longer reading times to reduced relatives compared to unreduced. The only result in the present study which was clearly different from the previous study was that the unreduced sentences showed longer second pass reading times in the "by"-phrase region for the animate condition as compared to the inanimate condition.

Subject and item second pass means for sentences with ambiguous verbs were entered into separate ANOVAs with four factors: list (four lists) or item group (four groups), region (the noun phrase, the verb, the "by"-phrase, and the first two words of the main verb phrase), animacy (animate and inanimate), and relative clause type (reduced and unreduced). When collapsing across all of four regions, the ANOVA revealed a marginal interaction between animacy and the type of relative clause ($F1(1,16) = 3.89, p < .1$; $F2(1,12) = 4.25$, $p < .1$). In total, animate conditions were 344 ms longer than inanimate conditions, resulting in a significant effect of animacy ($F1(1,16) = 28.80$, $p < .01$; $F2(1,12) = 7.72$, $p < .05$). Reduced relatives were in total 181 ms longer than unreduced relatives, resulting a significant effect of relative clause type ($F1(1,16) = 5.08, p < .05$; $F2(1,12) = 8.29, p < .05$). There was also a significant effect of scoring region ($F1(3,14) = 7.35, p < .01$; $F2(3,10) = 8.87, p < .01$). The total difference between reduced and unreduced relatives was +350 ms for animates and +11 ms for inanimates, resulting in a significant effect of relative clause type when the preceding noun was animate ($F1(1,16) = 6.02$, $p < 0.05$; $F2(1,12) = 6.92$, $p < .05$), but not when it was inanimate (Fs < 1). The total difference between animate and inanimate conditions was +513 ms for reduced relatives and +174 ms for unreduced relatives, resulting in a significant effect of animacy when the relative clause was reduced ($F1(1,16) = 29.31$, $p < .01$; $F2(1,12) = 9.42, p < .01$), but not when it was unreduced ($F1(1,16) = 2.11$; $F2(1,12) = 1.00$).

Analysis by region. Similar ANOVAs were conducted at each scoring region. To conserve space, we will only report analyses from the *verb* and the *"by"-phrase*.

At the *verb,* the interaction between animacy and type of relative clause was not significant (Fs < 1). Animate conditions were 62 ms longer than inanimate conditions, resulting in a significant effect of animacy in the subject analysis but not in the item analysis ($F1(1,16) = 5.97$, $p < .05$; $F2(1,12) = 2.32$). Reduced relatives were

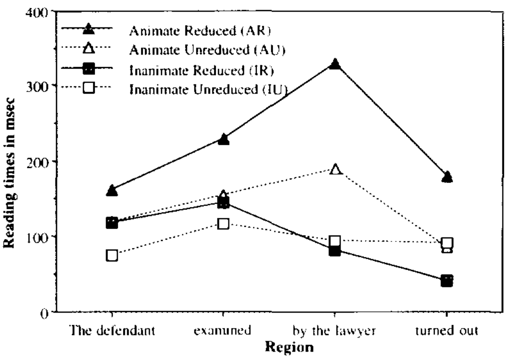

FIG. 4. Mean second pass reading times in ms (Experiment 2).

51 ms longer than unreduced relatives, resulting in a significant effect of relative clause type in the item analysis but not in the subject analysis ($F1(1,16) = 2.45$; $F2(1,12) = 5.01$, $p < .05$). Simple effects revealed the effect of relative clause type to be unreliable when the preceding noun was animate ($F1(1,16) = 2.66$; $F2(1,12) = 3.16$), and when it was inanimate (Fs < 1).

At the *"by"-phrase*, there was a marginal interaction between animacy and type of relative clause ($F1(1,16) = 3.82$, $p < .1$; $F2(1,12) = 3.96$, $p < .1$). Animate conditions were 172 ms longer than inanimate conditions, resulting in a reliable effect of animacy ($F1(1,16) = 33.64$, $p < .01$; $F2(1,12) = 11.06$, $p < .01$). Reduced relatives were 64 ms longer than unreduced relatives, resulting in a marginal effect of relative clause type ($F1(1,16) = 4.00$, $p < .1$; $F2(1,12) = 3.01$, $p = .1$). The difference between reduced and unreduced relatives was +140 ms for the animates and −12 ms for the inanimates, resulting in an effect of relative clause type which was reliable in the subject analysis and marginal in the item analysis when the preceding noun was animate ($F1(1,16) = 4.63$, $p < .05$; $F2(1,12) = 4.28$, $p < .1$), and not significant when the noun was inanimate (Fs < 1). The difference between animate and inanimate conditions was +248 ms for reduced relatives and +96 ms for unreduced relatives, resulting in a significant effect of animacy for the reduced relative clauses ($F1(1,16) = 23.81$, $p < .01$; $F2(1,12) = 13.99$, $p < .01$), and a marginal effect of animacy in the subject analysis for the unreduced relatives ($F1(1,16) = 4.08$, $p < .1$; $F2(1,12) = 1.63$).

Discussion

The most important results from Experiment 1 were replicated. Readers again had more difficulty reading sentences with reduced relative clauses when they began with animate nouns than when they began with inanimate nouns. In addition, reading times to reduced relatives with inanimate nouns were not reliably longer than reading times to control sentences with full relatives. However, the interaction between animacy and the type of relative clause was not reliable at the "by"-phrase, although the size of the interaction was similar to Experiment 1. The most likely explanation is that the data from this experiment were a bit noisier because we tested fewer subjects than in Experiment 1. This would suggest that an analysis that combined first pass reading times from the two experiments would produce a more robust interaction. In fact this was the case. At the "by"-phrase, there was a reliable interaction between animacy and type of relative clause in both the subject and item analysis ($F1(1,42) = 7.68$, $p < .01$; $F2(1,25) = 4.42$, $p < .05$). (In the item analysis, we averaged scores for the six items used in both experiments, which is why there are 25 rather than 31 degrees of freedom.)

Even though the pattern of results across both experiments clearly indicates that animacy had immediate effects on ambiguity resolution, there were suggestions that readers experienced slight difficulties for at least some of the reduced relative with inanimate nouns. First pass reading times for these sentences were a little longer at the verb and at the prepositional phrase compared to the full relatives, though these differences never approached significance. In addition, the interaction between animacy and relative clause type was somewhat weaker in the item analysis for each experiment than in the subject analysis. The same pattern held for the combined analysis. This suggests that the items might have varied along a dimension that was related to ambiguity resolution.

Correlations with Semantic Fit

Earlier, we pointed out that from a constraint-based perspective, the strongest semantic constraints in support of a reduced relative clause would come from nouns which were both poor Agents *and* good Patients or Themes. In a system in which both the main clause and reduced relative structures are partially activated, a noun with both of these properties would simultaneously provide strong evidence *against* the main clause structure and in *favor* of the reduced relative structure. In both experiments, we selected inanimate nouns which were poor Agents for their respective ambiguous verbs. However, we did not attempt to equate the inanimate nouns for goodness of fit as a Patient or Theme. In fact, the materials did vary along this dimension. They were subsequently rated as part of a large norming project, with the norms collected at the University of Southern California by Maryellen MacDonald and Neal Pearlmutter in collaboration with the second author, Ken McRae and Michael Spivey-Knowlton. A total of 107 subjects rated the typicality of the Patient/Theme relation for each individual verb/noun pair by rating a question such as "How typical is it for evidence to be examined by someone?" on a 7-point scale with 1 as not typical at all and 7 as very typical. Subjects made similar ratings for typicality of the inanimate noun as the Agent of the verb. The mean ratings for the inanimate nouns are presented in Appendix A. The ratings confirmed that the inanimate nouns were nearly always poor Agents. The mean Agent rating was 1.4. Only a few items had Agent ratings higher than 2.0. However, there was quite a bit of variability in the Patient Typicality ratings as shown in Table 4.

In order to see whether the typicality of the noun as the semantic object (Patient or Theme) of the verb influenced ambiguity resolution, we conducted a series of stepwise linear regressions. We conducted separate analyses at each scoring region using the first and second pass reading time data for the sentences with inanimate nouns. For each position, reading times for the unreduced sentences were first entered into the regression to predict the reading times for the corresponding sentences with reduced relatives. Then, Patient/Theme typicality ratings were entered into the regression to determine the extent to which they accounted for the remaining variance.

Correlations for Experiment 1. As shown in Table 5, the regressions revealed significant *negative* correlations with typicality ratings for first pass reading times at the "by"-phrase ($R = -.51$, $F(1,13) = 4.62$, $p < .05$) and second pass reading times at both the initial noun phrase ($R = -.52$, $F(1,13) = 4.82$, $p < .05$) and the ambiguous verb ($R = -.56$, $F(1,13) = 5.89$, $p < .05$). In addition, suggestive but non-significant negative correlations were found for first pass reading times at the ambiguous verb and second pass reading times at the final two positions. These negative correlations are consistent with immediate but graded use of semantic constraints; i.e., when semantic constraints were most consistent with a relative clause interpretation (high Patient typicality ratings) processing difficulty was reduced.

The conditions with morphologically *unambiguous* verbs in Experiment 1 provide an important control. If the correlations with typicality are really reflecting ambiguity resolution, then typicality should *not* predict reading times for the reduced relative clauses with morphologically unambiguous verbs. In fact, there were no significant correlations in either the first pass or

TABLE 4

MEAN PATIENT TYPICALITY RATINGS FOR INANIMATE NOUNS, ON A SCALE OF 1 (NOT TYPICAL) TO 7 (VERY TYPICAL) (RANGE IN PARENTHESES)

Experiment	Verb type	Agent typicality	Patient/theme typicality
Experiment 1	Ambiguous	1.4 (1 to 2.2)	4.7 (1.8 to 6.5)
	Unambiguous	1.7 (1.1 to 3.8)	5.8 (1.8 to 6.5)
Experiment 2	Ambiguous	1.6 (1 to 5.2)	5.7 (4.1 to 6.6)

TABLE 5

PARTIAL CORRELATIONS BETWEEN "PATIENT TYPICALITY" RATING AND VARIANCE LEFT AFTER PREDICTING INANIMATE REDUCED READING TIMES FROM INANIMATE UNREDUCED READING TIMES AT EACH SCORING REGION

		Region			
Exp.	RT	"The evidence"	"examined"	"by the lawyer"	"turned out"
Exp. 1	Fpass	—	$R = -.42, p = .12$	$R = -.51, p < .05$	—
	Spass	$R = -.52, p < .05$	$R = -.56, p < .04$	$R = -.40, p = .15$	$R = -.38, p = .17$
Exp. 2	Fpass	—	$R = -.45, p < .1$	$R = -.57, p < .03$	$R = -.51, p = .05$
	Spass	—	—	—	—

second pass reading times for these conditions (all Fs < 1, except for first pass reading times at the "by"-phrase and the final region, where there were weakly positive correlations ($R = .53$, $F(1,9) = 3.43$ and $R = .45$, $F(1,9) = 2.29$, respectively).

Correlations for Experiment 2. As is shown in Table 5, regressions on the reading times for Experiment 2 revealed significant negative correlations with typicality ratings, but with a slightly different pattern. A reliable negative correlation was found for first pass times at the "by"-phrase ($R = -.57$, $F(1,13) = 6.25$, $p < .05$). Also, marginally significant negative correlations were found for first pass times at both the ambiguous verb ($R = -.45$, $F(1,13) = 3.31$, $p < .1$) and the final regions ($R = -.51$, $F(1,13) = 4.50$, $p = .05$). No other positions showed significant correlations. Typicality ratings for these stimuli had a smaller range than the ratings for Experiment 1, and therefore may have reduced the likelihood of finding reliable correlations.

Comparisons of strong and weak semantic fit. The results presented here demonstrate that the semantic constraints provided by the noun have immediate effects on ambiguity resolution. In order to illustrate this more clearly, we used the typicality ratings to select those items that have the strongest semantic constraints, that is, those inanimate nouns that are both poor Agents and good Patients/Themes. Fourteen of the 26 items met the criterion of having an Agent rating less than 2.0 and a Patient rating greater than 5.0. Figure 5a plots the reduction effect (reading times to the reduced relative minus the reading times to the unreduced relative) for first pass reading times at the *verb* and the

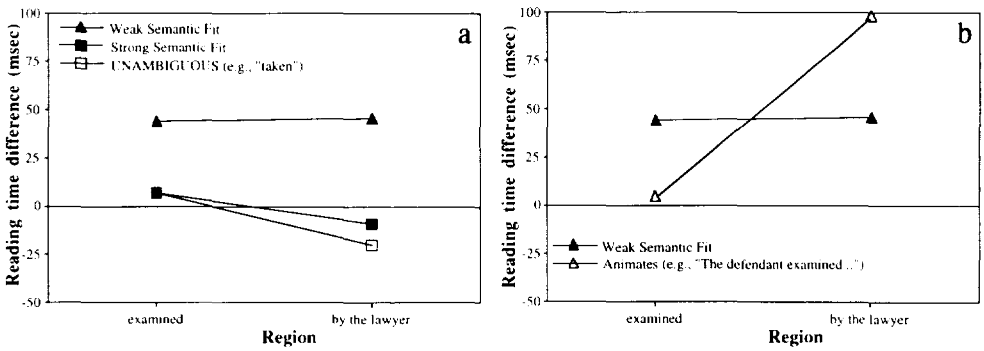

FIG. 5. Mean first pass "reduction" effects (reduced relative relatives minus unreduced relatives) for the *verb* and the "*by*"-*phrase*. Positive numbers indicate increases in processing difficulty for the reduced relative. (a) Inanimate nouns with weak semantic fit and strong semantic fit as compared to unambiguous controls. (b) Inanimate nouns with weak semantic fit as compared to the animate nouns.

"*by*"-*phrase* for both the more constraining inanimate items (filled squares) and the remaining twelve inanimate items, which were less constraining (filled triangles). As can be seen in the figure, the less constraining items show increases in processing difficulty at both positions (+44 and +46 ms, respectively), whereas the more constraining items show little or no increases at either position (+7 and −9 ms). In fact, the more constraining items show a pattern very similar to the morphologically unambiguous verbs (plotted as open squares), which are +7 ms at the *verb* and −20 ms at the "*by*"-*phrase*. When semantic constraints are strongest, reduced relatives with ambiguous and unambiguous verbs behave similarly.

Finally, Fig. 5b replots the less constraining inanimates (filled triangles) and compares them with the animate noun condition (open triangles). At the verb, there is processing difficulty for the inanimates but not the animates. Then, at the "*by*"-*phrase*, both show some processing difficulty. This is similar to the first pass data pattern reported in Ferreira and Clifton (1986). Ferreira and Clifton (1986) interpreted this pattern as support for the delayed use of semantic information. They argued that readers were initially aware that the inanimates were poor Agents of the verb, but could still not use this information to avoid a syntactic misanalysis. However, the results for the more constraining items (Fig. 5a) show that these items fall at one end of a continuum of semantic constraint. The fact that the Ferreira and Clifton's (1986) items behave similarly to our less constraining items is not surprising, given that their semantic manipulation was relatively weak.

General Discussion

The experiments reported here demonstrate that a strong local semantic constraint, the semantic fit of a noun to potential argument positions, had immediate effects on syntactic ambiguity resolution for reduced relative clauses. In addition, correlations with typicality ratings showed that the effects of the semantic constraint were related to the strength of the constraint. In order to provide a clearer framework for discussing the results, we will outline an evidential model of syntactic ambiguity resolution, focusing on the relative clause ambiguity. The model is similar in spirit to recent proposals developed independently by MacDonald and colleagues (e.g., MacDonald, 1992, 1993; Pearlmutter & MacDonald, 1992).

The principles that underlie the approach are simple. Structures are partially activated with the strength of activation dependent upon their likelihood given the input. The effects of a contextual constraint will depend upon its strength and the availability of the alternative structures. To a first approximation, these are the same factors that are important for lexical ambiguity resolution (MacDonald, 1993; Tabossi et al., in press; Tanenhaus, Dell, & Carlson, 1987). Syntactic and lexical ambiguity resolution are viewed as similar and interrelated processes because many syntactic ambiguities depend upon lexical ambiguities. This becomes particularly clear when one takes into account the alternatives provided by aspects of combinatory lexical information such as argument structure (see Tanenhaus & Carlson, 1989; Tanenhaus et al., 1989; Tanenhaus, Garnsey, & Boland, 1991).

The main clause/relative clause ambiguity hinges upon a morphological ambiguity between a past-tense form and a passive participial form. Consider the evidence that the processing system would receive when it encounters the ambiguous verb in the context of a noun phrase. We will assume that the ambiguous verb will provide partial evidence/activation for both the past tense and participial forms, with the strength of the evidence determined by relative frequency (Burgess & Hollbach, 1988; Tabossi et al., in press). The question of how to calibrate frequency for lexical/structural ambiguity is just beginning to be explored (Hindle & Rooth, 1990; MacDonald, 1993; Mitchell & Cuetos, 1991; Tanenhaus & Juliano, 1992; Juliano, Trueswell, & Tanenhaus, 1992; Trueswell, Tanenhaus, & Kello, 1993). However, to a first approximation it appears that frequency is determined by the specific lexical item (i.e., how frequent the past-tense and participial forms are for the particular verb) conditionalized on the frequency of the syntactic category in the syntactic environment (Juliano & Tanenhaus, 1992). Preliminary corpus analyses indicate that at the beginning of a sentence, a morphologically ambiguous verb that follows a noun phrase is far more likely to be a past tense verb in a main clause than a passive participial in a reduced relative clause (Tabossi et al., in press). Thus at the verb, there will be a clear frequency-based bias in favor of the past tense/main clause structure, though the participial/relative clause structure will also be partially activated. In addition, we will assume that verb forms activate the set of thematic/conceptual roles associated with the verb and their corresponding syntactic-mappings (Carlson & Tanenhaus, 1988; Cottrell, 1988; McClelland & Kawamoto, 1986; Pearlmutter & MacDonald, 1992; Tanenhaus, Carlson, & Trueswell, 1989). Thematic assignment is immediate, with the thematic fit of a potential argument evaluated with respect to the active alternatives.

When there is a good thematic fit between the noun and the Agent role ("The defendant examined . . ."), the main clause structure will be further supported and the reduced relative structure will be only weakly activated. Under these conditions, readers will experience difficulty when they encounter a "by"-phrase that is inconsistent with a main clause. In contrast, when the noun is a poor Agent and a good Patient ("The evidence examined . . ."), thematic fit will provide positive evidence for the relative clause and negative evidence for the main clause. In this case, one would expect to find less processing difficulty when readers encounter the "by"-phrase. Whether there is any residual difficulty will depend upon the strength of the semantic constraint and the relative availability of the past tense and participial forms. This is exactly what we found at the "by" phrase. Animate nouns which strongly support a main clause reading showed increased processing difficulty. Inanimate nouns which strongly supported a relative clause interpretation showed no processing difficulty compared to unambiguous controls, whereas there was still some difficulty with the less constraining inanimates.

The data pattern at the preceding ambiguous verb is also explained by this account. Eye movement research has demonstrated that reading times to semantically ambiguous words which have two equiprobable meanings are longer when prior context is neutral or only weakly biasing as compared to when context strongly biases one meaning. Reading times for an ambiguous word are also longer when context supports the subordinate rather than the dominant meaning (Duffy, Morris, & Rayner, 1986; Rayner, Pacht, Sereno, & Duffy, 1991; see also Rayner & Polletsek, 1987).

As we pointed out earlier, in an NP V context, there is a large frequency asymmetry in favor of the past tense/main clause, making the past tense the dominant alternative and the participle the subordinate alternative. Thus, one might expect to see difficulty at the verb when a noun is a poor Agent, even when it is also a good Theme. This would be an example of a context biasing a subordinate alternative. However, this does not take into account the effects of parafoveal information on availability. The first fixation data that we presented earlier demonstrated that readers rarely fixated on the "by," indicating that it was read parafoveally during fixations on the verb. It is likely that this would have increased the availability of the participial/relative clause structure, making the situation more like that of an ambiguous word with two equally frequent senses. Note, however, that the word "by" is itself ambiguous and it is not necessarily inconsistent with a past tense/main clause. Tabossi et al. (in press) show that "by" following an NP V is frequently taken to be a manner or a locative preposition in a main clause when the noun is a good Agent, but it is nearly always taken to be an agentive preposition in a relative clause when the noun is a poor Agent. Thus, elevations at the verb for weakly biasing inanimates are simply an ambiguity effect related to the nature of the context.

Recent work by Burgess (1991, Burgess & Tanenhaus, in preparation) provides important empirical support for some of our conjectures about the importance of parafoveal support. Burgess conducted two self-paced reading studies using the materials that we modified slightly for Experiment 2. In one study, the sentences were presented one word at a time. In the other study they were presented two words at a time, with the verb and "by" presented together (e.g., /The evidence/ examined by/ . . .). Burgess found immediate effects of animacy with two-word presentation, but not with one-word presentation. Thus strongly biased nouns had immediate effects only when there was parafoveal support for the less frequent participial form.

At this point, we have identified two factors that should affect ambiguity resolution for reduced relative clauses, the strength of the contextual bias as determined by the semantic fit of the noun and the availability of the past tense and participial forms. Under conditions where there is a strong semantic constraint in favor of a strongly activated alternative, there will be no difficulty, i.e., no ambiguity effect. However, ambiguity effects will be observed when the constraint is weak or when the biased alternative is initially less active. Subsequent input that is inconsistent with the more active alternative structure will lead to increased difficulty. Conscious garden paths represent the extreme end of a continuum in which a conspiracy of evidence makes the correct alternative so inactive that it is inaccessible at the point of disambiguation.

When viewed from the perspective of a two-stage model, effects of either weak constraints or constraints that bias a low frequency syntactic alternative will look like revisions after an initial syntactic misanalysis. However, as the work presented here demonstrates, the effects of semantic constraint are actually continuous and they interact with availability. Moreover, they are strikingly similar to results obtained with lexical ambiguity.

The model we have sketched accounts for the patterns of data that we reported for first pass reading times. The model also provides a framework for understanding the Ferreira and Clifton (1986) first pass reading data. Recall that they found a large reduction effect at the verb and at the prepositional phrase for inanimate nouns. This, in fact, is an exaggerated version of the pattern that we found for our less constrained inanimate items. Many of the inanimate nouns used by Ferreira and Clifton were only weakly constraining. In addition, the prepositional phrases contained a variety of prepositions, including some long prepositions (e.g., "from," "about," etc.). These prepositions would normally receive separate fixations. Thus, they would presumably not have been recognized parafoveally during fixations on the verb. Therefore Ferreira and Clifton's materials combined a relatively weak constraint with low availability of the relative clause. It is just these conditions that would lead to ambiguity effects at the verb and the prepositional phrase. It is still somewhat puzzling that the first pass reading times at the prepositional phrase for the reduced relatives with inanimate nouns were not reliably faster than the reading times for the reduced relatives with animate nouns. However, it is possible that two factors might have masked a small animacy effect. First, early line breaks for the reduced relative clauses might have inflated reading times to the prepositional phrase. Second, because these line breaks occurred in the middle of the prepositional phrase, reading times in this scoring region would have been quite variable.

Even though the framework that we are proposing is clearly constraint-based, it is consistent with many results in the literature that have been used to argue for restricted-domain approaches to parsing. Contextual constraints will have weak and/or delayed effects when they are not particularly strong or when they support a structure that is not highly available because of local frequency factors. Likewise, pragmatic knowledge will not override clear syntactic constraints. And, contextual dimensions that are not correlated with syntactic alternatives will rarely provide much constraint. We would suggest that many of inconsistencies in the literature on local context effects in parsing are due to failure to take into account both the strength and local relevance of the contextual constraint and the local availability of the alternative structures. Similar arguments apply to the literature on discourse context effects in parsing (Spivey-Knowlton et al., 1993).

It is important to note that strength of constraint and frequency-based availability run the risk of becoming theoretical wild cards if they are not made explicit. However, current corpus-based work suggests that it is possible to quantify availability and frequency in ways that clearly predict psycholinguistic performance (MacDonald, 1992, 1993; Mitchell, 1992; Tanenhaus & Juliano, 1992; Juliano et al., 1992; Trueswell et al., 1993). Similarly, factors like semantic fit can be quantified using metrics like typicality ratings.

The questions addressed in this article have typically been discussed within the ongoing debate about modularity in language processing (Fodor, J. A., 1983; Fodor, J. D., 1991; Clifton & Ferreira, 1987; Ferreira & Clifton, 1986; Frazier, 1991; Marslen-Wilson & Tyler, 1987; McClelland, 1987; Tanenhaus, Carlson, & Seidenberg, 1985; Tanenhaus, Dell, & Carlson, 1987; Thompson & Altmann, 1991). Research on modularity in language processing has often conflated two important, but distinct types of questions (Garnham, 1985; Tanenhaus & Lucas, 1987). The first concerns representational modularity, e.g., are there distinct types of linguistic representations and can one distinguish representations that are inherently grammatical from those that are not. The second concerns processing modularity, e.g., are there informationally encapsulated subsystems within the language processing system. The model we have proposed presupposes representational modularity, and it assumes that language is processed along at least partially independent dimensions. However, the model is non-modular in that local indeterminacies within subsystems are resolved using correlated information from different domains.

One of the attractive aspects of completely modular systems is that they allow a complex task to be divided into subcomponents which can be performed by local experts. This works best when there is clear (unambiguous) input to each expert and the tasks performed by the experts are autonomous. In natural systems complex stimuli are often organized along various dimensions of the input. However, this segregation results in pervasive ambiguity within each subsystem. Thus, consistent solutions require the coordination of multiple constraints. For example, in visual processing it is well established that the segregation of motion information from other stimulus dimensions (color, form, etc.) by the visual system results in indeterminacy. The motion of a line is initially represented in parallel by a set of equally possible speeds and directions (Movshon, Adelson, Gizzi, & Newsome, 1985; Poggio, Torre, & Koch, 1985). The system finds a single solution by grouping together the possible motions from other contours. However, it has become clear that this grouping is not done by the motion system alone based on domain-specific principles or constraints. Rather, information about color, orientation, spatial frequency, and depth is recruited to determine grouping (e.g., Adelson & Movshon, 1982, 1984; Movshon et al., 1985; Krauskopf & Farrell, 1990; Kooi, De Valois, & Switkes, 1991).

We are arguing for similar state of affairs in sentence processing. Linguistic input is encoded along partially independent dimensions (e.g., syntactic information, semantic/thematic information). However, online processing results in local indeterminacy within each subsystem. Within the syntactic domain, one influential approach has been to incorporate domain-specific decision principles that eliminate or restrict initial ambiguity (Kimball, 1973; Frazier, 1978; Pritchett, 1988, 1992). These decisions then have to be coordinated and reconciled with potentially conflicting information. For example, in some proposals independent decisions about highly correlated dimensions (e.g., the semantic and syntactic aspects of argument structure) are initially made by different subsystems (Rayner et al., 1983; Frazier, 1987, 1991). However, the experiments reported here and elsewhere (MacDonald, 1992, 1993; Spivey-Knowlton et al., 1992; Tanenhaus, Garnsey, & Boland, 1991; Trueswell & Tanenhaus, 1991, 1992) suggest an alternative approach in which the process of constraining ambiguity in one domain is accomplished by recruiting information from other relevant domains. As in visual processing, this selection-based account maintains many of the characteristics of autonomous models. However, it also allows the system to use correlated information from different domains, thereby reducing the problem of inconsistent solutions.

Appendix A: Target Sentences from Experiments 1 and 2

Experiment 1

Ambiguous Verbs

Each item is listed with both the animate/inanimate noun pair (e.g., "defendant/evidence") and the optional unreduced form (e.g., "that was"). The mean Agent typicality rating and mean Patient/Theme typicality rating for the inanimate nouns are provided in parenthesis.

1. The defendant/evidence (that was) examined by the lawyer turned out to be unreliable. (1.4, 6.3)
2. The prisoner/gold (that was) transported by the guards was closely watched. (1.1, 5.5)
3. The teacher/textbook (that was) loved by the class was very easy to understand. (1.0, 1.9)
4. The workers/bricks (that were) lifted by the crane were deposited on the roof. (1.3, 5.4)
5. The student/paper (that was) graded by the professor received a low mark. (1.1, 6.5)
6. The contestant/recipe (that was) selected by the judges did not deserve to win. (2.2, 4.4)
7. The specialist/equipment (that was) requested by the hospital finally arrived. (1.1, 5.1)
8. The thief/jewelry (that was) identified by the victim was held for questioning/as evidence. (1.7, 4.4)
9. The soldier/valley (that was) captured by the enemy was closely guarded. (1.4, 1.8)
10. The troops/power plant (that were/that was) attacked by the terrorists suffered heavy losses. (1.7, 3.2)
11. The artist/painting (that was) studied by the historian was a complete unknown. (1.3, 5.4)
12. The boy/necklace (that was) described by the lady was quite handsome/beautiful. (1.2, 5.9)
13. The mailman/package (that was) expected by the secretary arrived too late. (1.3, 5.5)
14. The woman/sofa (that was) scratched by the cat was badly injured/damaged. (1.3, 4.8)
15. The man/van (that was) recognized by the spy took off down the street. (1.0, 4.1)

16. The client/account (that was) wanted by the advertiser was worth a lot of money. (2.1, 4.6)

Unambiguous Verbs

Each item is listed with the inanimate noun followed by the optional unreduced form (e.g., "that was"). The mean Agent typicality rating and mean Patient/Theme typicality rating are provided in parenthesis.

1. The money (that was) taken by the student was finally returned. (1.2, 5.6)
2. The poster (that was) drawn by the illustrator was used for a magazine cover. (1.1, 5.9)
3. The work (that was) done by the carpenter was quite good. (1.5, 6.4)
4. The ball (that was) thrown by the boy broke a window. (1.4, 6.2)
5. The crops (that were) grown by the farmer were damaged by the frost. (3.8, 6.0)
6. The letter (that was) written by the teacher was hard to understand. (1.4, 6.1)
7. The vase (that was) broken by the child was worth a fortune. (1.1, 5.9)
8. The computer (that was) chosen by the company was a good investment. (3.3, 4.3)
9. The wallpaper (that was) shown by the salesman was a perfect choice. (2.7, 5.3)
10. The truck (that was) seen by the policeman did not have a license plate. (1.1, 6.1)
11. The van (that was) stolen by the thief was later found in a back alley. (1.1, 5.2)
12. The poultry (that was) eaten by the guest gave him an upset stomach. (1.2, 6.6)

Experiment 2

Each item is listed with both the animate/inanimate noun pair (e.g., "defendant/evidence") and the optional unreduced form (e.g., "that was"). The percentage of main clause completions from Burgess (1991) for inanimate nouns and the mean Agent typicality rating and mean Patient/Theme typicality rating for the inanimate nouns are provided in parenthesis.

1. The speaker/solution (who/that was) proposed by the group would work perfectly for the program. (10%, 1.5, 6.6)
2. The man/ransom (who/that was) paid by the parents was unreasonable. (0%, 2.4, 4.5)
3. The lawyer/package (who/that was) sent by the governor arrived late. (10%, 1.6, 6.2)
4. The student/award (who/that was) accepted by the school was very pleased. (0%, 1.1, 6.6)
5. The woman/portrait (who/that was) sketched by the artist was very beautiful. (10%, 1.4, 4.9)
6. The defendant/evidence (who/that was) examined by the lawyer turned out to be unreliable. (20%, 1.4, 6.3)
7. The specialist/equipment (who/that was) requested by the hospital had finally arrived. (20%, 1.1, 5.1)
8. The artist/painting (who/that was) studied by the historian was relatively unknown. (20%, 1.3, 5.4)
9. The man/van (who/that was) recognized by the spy took off down the street. (20%, 1.0, 4.1)
10. The man/message (who/that was) recorded by the secretary could not be understood. (0%, 1.8, 5.6)
11. The author/book (who/that was) read by the student was very difficult to understand. (10%, 1.3, 6.6)
12. The director/building (who/that was) watched by the cop was in a bad part of town. (30%, 1.1, 4.9)
13. The scientists/alternatives (who/that were) considered by the committee each had limitations. (28%, 1.3, 6.6)
14. The student/paper (who/that was) graded by the professor was very interesting. (30%, 1.1, 6.5)
15. The mailman/package (who/that was) expected by the secretary arrived too late. (30%, 1.3, 5.5)

16. The man/car (who/that was) towed by the garage was parked illegally. (30%, 5.2, 6.1)

Appendix B: The Relationship Between Region String Length and First Pass Reading Times

Reading times for a scoring region are generally longer when the length of the region is longer. When an experiment calls for comparisons across scoring regions with different linguistic content, an experimenter will typically control for string length effects either by making sure all conditions have the same average string length or by adjusting reading times based on string length. However, even when average region lengths are controlled, it may be important to adjust reading times for string length if there is considerable variation in length across items.

The standard string length adjustment of reading times in eye movement studies has been milliseconds per character (ms/char). This metric was first introduced into the literature by Frazier and Rayner (1982) and has been used in many subsequent studies (e.g., Altmann, Garnham, & Dennis, 1992; Ferreira & Clifton, 1986; Rayner, Carlson, & Frazier, 1983; Rayner, Garrod, & Perfetti, 1992). Milliseconds per character is computed by dividing the reading time (in ms) of a region by the number of characters that make up that region including spaces and punctuation. This transformation makes two assumptions about the relationship between string length and unadjusted reading times. First, the relationship is linear, i.e., the addition of one character to a region results in the same increase in reading times regardless of whether the region is small or large. Second, the relationship has a y intercept of 0 ms, i.e., a hypothetical region of zero characters should have a 0 ms reading time. This latter assumption is of critical importance for this discussion and is highlighted by the *ideal* reading time and string length relation plotted in Fig. 6A.

In Fig. 6A, string length in characters is on the x-axis and reading time in milliseconds is on the y axis. The function plotted on the figure corresponds to the hypothetical linear relation between string length and ms reading time, with a y intercept of 0 ms. This function can be represented by the equation

$$y = mx \qquad [1]$$

where m is the slope of the line, x is string length in characters, and y is reading time in ms. Milliseconds per character can be obtained by dividing ms reading times (y) by string length (x):

$$y/x = mx/x \text{ or} \qquad [2]$$

$$y/x = m. \qquad [3]$$

Figure 6B plots this relationship between string length and ms/char. Since m is a constant, there is no relationship between string length and ms/char. This is the desired adjustment of reading times; the variance associated with string length has been removed.

However, this outcome is not obtained if the relationship between string length and millisecond reading times does *not* contain a y intercept of 0 ms (Fig. 6C). This can be represented by

$$y = mx + b \qquad [4]$$

where b is a positive constant. Dividing by string length (x) to get ms/char results in:

$$y/x = m + b/x. \qquad [5]$$

This relationship between string length and ms/char is plotted in Fig. 6D. Clearly, if the millisecond y intercept is not zero (Fig. 6C), ms/char does not remove all of the variance due to string length (Fig. 6D). Moreover, as shown in 6D, the relationship between reading time and string length is reversed (shorter regions tend to have longer reading times), and the relationship is no longer linear. The non-linearity arises from the inverse relationship ($1/x$, or x^{-1}) in the equation.

Thus, if there is a positive y-intercept between string length and *unadjusted* reading times in milliseconds, there will be an in-

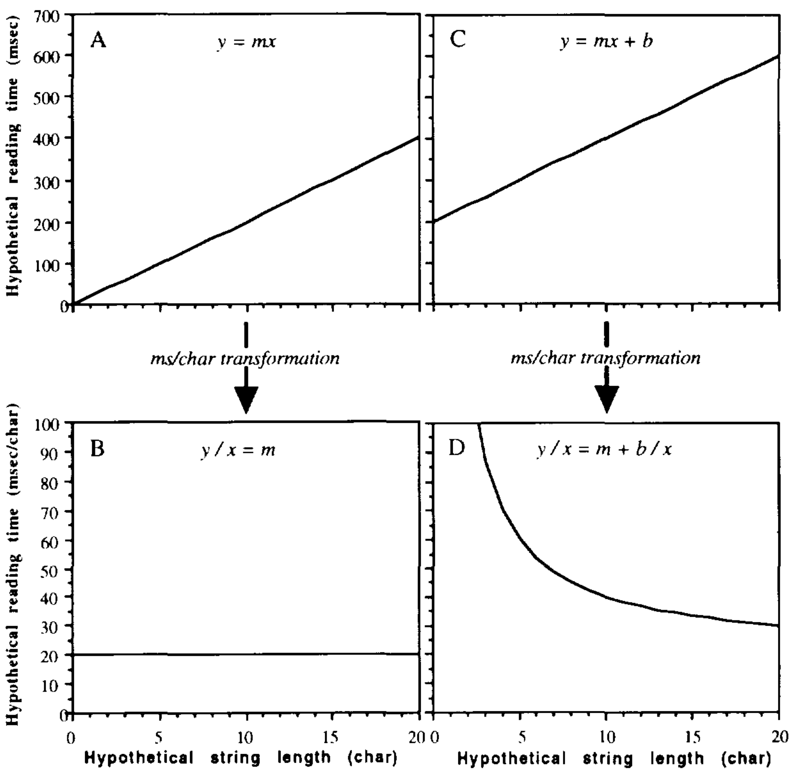

FIG. 6. (A) Hypothetical reading times in ms as a function of region string length, with a y-intercept of zero ms. (B) Hypothetical reading times in ms/char as a function of region string length, transformed from (A). (C) Hypothetical reading times in ms as a function of region string length, with a positive y-intercept. (D) Hypothetical reading times in ms/char as a function of region string length, transformed from (C).

verse function distorting the relationship between string length and ms/char reading times. As shown in Fig. 6D (and Eq. 5 above), reading times will be *less* distorted for *large* scoring regions (i.e., as length, x, approaches infinity, reading time, y, approaches the constant m), but reading times will be *more* distorted for *small* scoring regions (i.e., as length, x, approaches 0, reading times, y, approaches infinity). Thus, the extent to which *actual* ms/char reading times are distorted will depend upon: (1) the size of the y intercept of unadjusted reading times and (b) the scoring region sizes used in the study.

First Pass Reading Times and Ms/Char

In this section we will examine actual first pass reading times from Experiment 1 as a function of region string length. Figure 7a plots first pass reading times (in ms) as function of scoring region length (in characters) for all 24 subjects in Experiment 1. To determine the best linear fit to this data, a linear regression was computed with string length as the independent variable and reading time as the dependent variable. The regression revealed a reliable positive linear relation between string length and reading time, represented by the dotted in Fig. 7a ($y = 21.9x + 190$, Pearson $R = .360$, $F(1,2521) = 375.84$, $p < .001$). This equation clearly has a non-zero intercept (190 ms), suggesting that the ms/char transformation would produce a negative correlation with string length. In fact, this was the case. Figure 7b presents the same first pass reading times in *ms/char* as a function of scoring region length (in char). The best linear fit, represented by the dotted line in

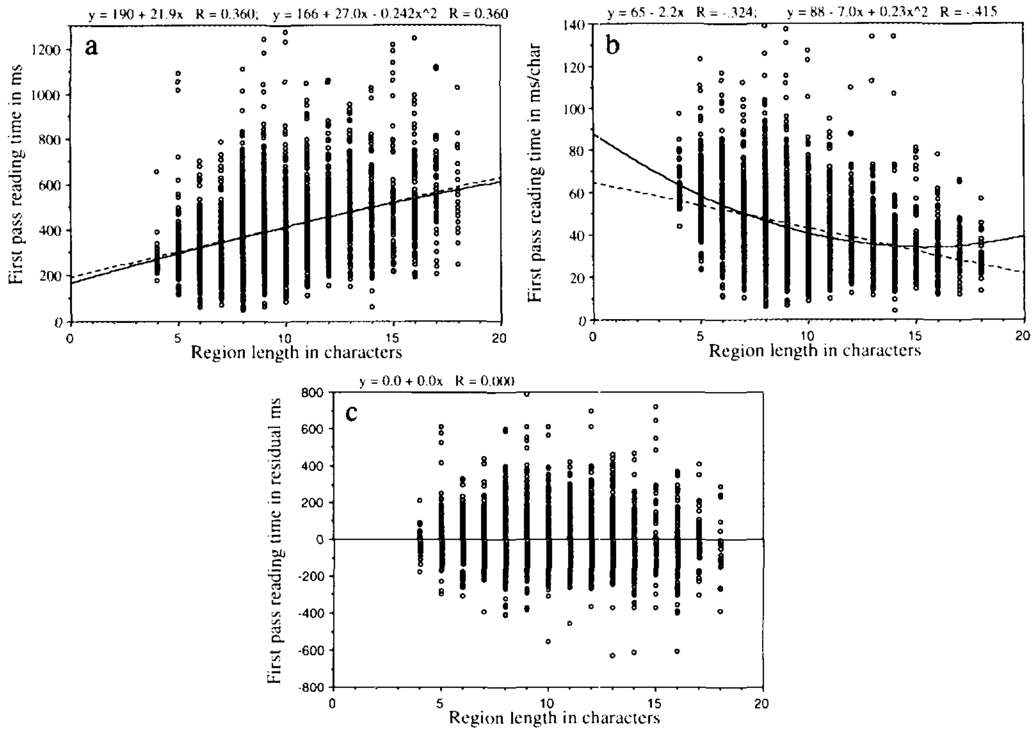

FIG. 7. (a) First pass reading times in ms from Experiment 1 as a function of string length. (b) First pass reading times in ms/char from Experiment 1 as a function of string length. (c) First pass reading times in residual ms from Experiment 1 as a function of string length.

Fig. 7b, had a reliable *negative* correlation ($y = -2.2x + 65, R = -.324, F(1{,}2521) = 295.11, p < .001$).

As Lorch and Myers (1990) point out, however, regressions performed across all subjects, like those presented here, can increase the probability of a Type I error in assessing the reliability of the linear coefficients. Following Lorch and Myers (1990), separate millisecond and ms/char regression equations were computed for each subject (see Table 6). Averaging the coefficients and the Rs for each equation provide similar results (ms: $y = 22.1x + 193, R = .421$; ms/char: $y = -2.2x + 69, R = -.378$). A single-group t test was performed on each coefficient to test whether each was reliably different from zero and is reported in the table (see Lorch & Myers, 1990). All coefficients, in both equations, were reliably different from zero. Crucially, the y intercept (B_0) of the millisecond relation is reliably positive, and, the ms/char transformation has a reliable negative slope (B_1) with string length.

We also examined the extent to which there is a non-linear component to the relationships shown in Fig. 7. The solid line in Fig. 7a shows the best-fit second-order equation for the relationship between region length in char and reading time in ms ($R^2 = .130$; $p < .001$). Although this is a non-linear equation, the fit is surprisingly linear. Following Lorch and Myers (1990), we tested the validity of this polynomial fit by finding the best fit for each subject's data. Again, a single-group t test was performed on each coefficient to test whether each was reliably different from zero and is reported in Table 7. As can be seen in the table, there was a reliable linear component, and only a marginal non-linear component. This indicates that first pass reading times have, to a very good first approx-

TABLE 6
COEFFICIENTS FOR FIRST-ORDER LINEAR EQUATION $y = B_0 + B_1x$

Subject	List	B_0 (ms)	B_1 (ms)	B_0 (ms/char)	B_1 (ms/char)
1	1	167.3	19.80	58.3	−1.99
2	1	153.1	26.78	60.2	−1.66
3	1	189.5	22.94	69.4	−2.51
4	1	220.9	8.11	51.1	−1.94
5	1	79.3	31.01	49.0	−.92
6	1	434.8	11.22	106.4	−4.72
1	2	152.3	20.97	153.1	−1.79
2	2	153.1	21.22	56.2	−1.81
3	2	133.7	26.74	56.6	−1.5
4	2	230.1	18.11	70.0	−2.63
5	2	183.6	18.16	59.4	−2.09
6	2	146.1	32.91	65.6	−1.649
1	3	110.1	23.29	51.2	−1.54
2	3	147.1	20.70	53.9	−1.68
3	3	309.7	26.17	91.6	−3.09
4	3	291.4	37.07	107.8	−3.80
5	3	255.2	12.59	68.7	−2.78
6	3	184.2	21.90	60.5	−1.82
1	4	218.3	31.51	80.1	−2.42
2	4	192.6	19.89	62.5	−2.13
3	4	202.6	23.94	67.4	−2.12
4	4	185.1	10.84	53.0	−2.16
5	4	60.9	34.02	50.3	−.94
6	4	237.4	10.69	61.5	−2.46
		$m = 193.3$	$m = 22.11$	$m = 69.3$	$m = -2.17$
		$SD = 78.4$	$SD = 7.87$	$SD = 24.0$	$SD = .84$
		$t = 12.08$	$t = 13.77$	$t = 14.14$	$t = -12.74$
		$p < .001$	$p < .001$	$p < .001$	$p < .001$

Note. m, mean; SD, standard deviation.

imation, a linear relationship with region string length. However, the *ms/char* transformation, adds a non-linear component to this relation. The solid line in Fig. 7b shows the best-fit second-order equation for the relationship between region length in char and reading time in ms/char ($R^2 = .173; p < .01$). As shown in Table 7, there were reliable linear and non-linear components to this equation. The non-linearity should be expected, since the ms relation with string length has a reliable positive intercept, violating the ms/char assumptions.

As expected, this distortion is greater for smaller scoring regions than for larger scoring regions. This can be demonstrated by performing two separate linear regressions on the ms/char data: one for large scoring regions (10 or more characters) and another for small scoring regions (fewer than 10 characters). When the scoring regions were *large*, the best linear fit between string length and reading times had a very shallow slope (−1.66 ms/char) and accounted for little of the variance ($y = 53 - 1.66x$; $R^2 = .029$; $F(1,1102) = 33.32$, $p < .001$). When the scoring regions were *small*, the best linear fit between string length and reading times was quite steep (slope = −4.59) and accounted for more of the variance ($y = 83 - 4.59x$; $R^2 = .075$; $F(1,1417) = 114.79$, $p < .001$).

Overall, these results indicate that the ms/char correction for length is problematic, especially for small scoring regions. The transformation, when applied to our data: (a) leaves a significant amount of variance due to string length, (b) reverses the

TABLE 7

COEFFICIENTS FOR SECOND-ORDER POLYNOMIAL $y = B_0 + B_1x + B_2x^2$

Subject	List	B_1 (ms)	B_2 (ms)	B_1 (ms/char)	B_2 (ms/char)
1	1	12.918	.328	−7.079	.241
2	1	31.626	−.225	−5.350	.172
3	1	−8.474	1.481	−12.868	.488
4	1	51.521	−1.972	−1.181	−.034
5	1	32.359	−.064	−3.177	.107
6	1	86.021	−3.625	−8.000	.159
1	2	5.869	.704	−7.306	.257
2	2	14.329	.324	−6.969	.243
3	2	41.883	−.718	−4.173	.127
4	2	30.170	−.569	−8.284	.267
5	2	15.933	.104	−7.268	.243
6	2	31.091	.085	−5.621	.186
1	3	−3.745	1.281	−8.839	.346
2	3	19.539	.054	−6.210	.211
3	3	67.923	−1.994	−2.209	−.042
4	3	40.480	−.160	−19.665	.745
5	3	4.418	.374	−10.860	.369
6	3	69.961	−2.283	−1.293	−.025
1	4	47.531	−.766	−5.502	.148
2	4	58.566	−1.894	−3.477	.066
3	4	49.937	−1.229	−4.686	.121
4	4	17.191	−.307	−6.708	.220
5	4	9.784	1.148	−6.913	.283
6	4	36.328	−1.194	−5.567	.150
		m = 31.798	m = −.463	m = −6.634	m = .210
		SD = 24.364	SD = 1.237	SD = 3.913	SD = .169
		t = 6.39	t = −1.84	t = −8.31	t = 6.09
		$p < .001$	$p = .08$	$p < .001$	$p < .001$

Note. m, mean; SD, standard deviation.

relationship between string length and reading time, and (c) introduces a nonlinearity that is most pronounced at shorter scoring regions (less than 10 characters).

The practical importance of this can be highlighted by comparing first pass reading times to the inanimate ambiguous and inanimate unambiguous sentence types in Experiment 1. These two conditions have a clear difference in string lengths at the verb region (ambiguous verbs averaged 9.0 characters in length, whereas unambiguous averaged 6.3 characters). Table 8 presents reading times to these two conditions at this position. As would be expected, reading times in ms were longer to the longer (ambiguous) verbs, as compared to the shorter (unambiguous) verbs (a difference of +53 ms; $F(1,19) = 8.64$, $p < .01$; $F2(1,26) = 11.82$, $p < .01$). However, the ms/char reading times show the opposite relation (a difference of −10.1 ms/char; $F1(1,19) = 10.38$, $p < .01$, $F2(1,26) = 7.35$, $p < .05$). This reversal should be expected, since, as we have already pointed out, the slope relating string length to reading times in ms/char is highly negative for smaller scoring regions (see Fig. 7b).

TABLE 8

READING TIMES TO AMBIGUOUS AND UNAMBIGUOUS VERBS (EXPERIMENT 1)

Verb type	Milliseconds	Milliseconds per character	Residual milliseconds
Ambiguous	352	46.3	−18
Unambiguous	299	56.4	−16
Difference	+53*	−10.1*	−2

* $p < .01$.

In fact, a more appropriate string length correction (residual milliseconds in Table 8) revealed no reliable difference between these scoring regions (Fs < 1). This suggests that the difference in unadjusted reading times in ms was indeed due to string length. The residual correction was first introduced by Ferreira and Clifton (1986) for self-paced reading time data. In order to compute residuals, the best linear fit between region length and first pass reading times was determined for each subject (Table 6). Then, for each individual reading time on a region, the predicted reading time from the linear fit was subtracted from the actual reading time. This correction, by definition, removes all linear variance that has been related to string length, as shown in Fig. 7c. It is important to note that this adjustment should be made *for each subject,* since as Table 6 shows, there is subject-to-subject variation in the best linear fit between string length and reading time.

We should note that use of the ms/char transformation is not likely to have distorted the results of any prior studies which have either compared reading times across conditions with the same string length (e.g., the same words) or used relatively large scoring regions. However, the increasing interest in initial processing within the field is leading many researchers to use smaller scoring regions. Using ms/char to compare across small scoring regions with even slight differences in string length (e.g., a one character difference) is likely to introduce artifacts. These artifacts can exaggerate, mask, and even reverse real effects.

Finally, it is important to note that the ms/char transformation does not provide an accurate index for comparing reading rates across experiments, unless one is comparing regions of the same length. An unbiased measure of reading rate is provided by the y intercept reading time and the slope in ms/char.

In sum, we strongly recommend that future eye movement studies follow one of two procedures when reporting reading times. The first option is to use uncorrected reading times, with string length tightly controlled. The second is to report residual reading times along with the average reading rate, where reading rate is the average slope and intercept of each subject's best linear fit between string length and reading time.

References

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature,* **300,** 523–525.

Adelson, E. H., & Movshon, J. A. (1984). Binocular disparity and the computation of two-dimensional motion. *Optical Society of America, Annual Technical Meeting Program, 56.*

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.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language.* New York: Wiley.

Britt, M. A., Perfetti, C. A., Garrod, S., & Rayner, K. (1992). Parsing and discourse: Context effects and their limits. *Journal of Memory and Language,* **31,** 293–314.

Burgess, C. (1991). *Interaction of semantic, syntactic and visual factors in syntactic ambiguity resolution.* Ph.D. dissertation, Department of Psychology, University of Rochester, Rochester, NY.

Burgess, C., & Hollbach, S. C. (1988). A computational model of syntactic ambiguity as a lexical access process. *Proceedings of the Tenth Annual Cognitive Science Society Meetings,* 263–269.

Burgess, C., & Tanenhaus, M. K. (in preparation). The interaction of semantic and parafoveal information in syntactic ambiguity resolution.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkens (Eds.), *Syntax and semantics, 21.* New York, NY: Academic Press.

Clifton, C., Jr. (1993). The role of thematic roles in sentence processing. *Canadian Journal of Psychology,* **47,** 222–246.

Clifton, C., Jr., & Ferreira, F. (1987). Modularity in sentence comprehension. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding.* Cambridge, MA: MIT Press.

Clifton, C., Jr., Speer, S., & Abney, S. P. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language,* **30,** 251–271.

Cottrell, G. (1985). *A connectionist approach to* *word sense disambiguation*. Unpublished Ph.D. dissertation. University of Rochester, Rochester, NY.

CRAIN, S., & STEEDMAN, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*. Cambridge, England: Cambridge University Press.

DUFFY, S. A., MORRIS, R. K., & RAYNER, K. (1983). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, **27**, 429–446.

GIBSON, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Unpublished Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA.

FERREIRA, F., & CLIFTON, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, **25**, 348–368.

FERREIRA, F., & HENDERSON, J. M. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **16**, 555–568.

FERREIRA, F., & HENDERSON, J. M. (1991). How is verb information used during syntactic processing? In G. B. Simpson (Ed.), *Understanding Word and Sentence*. North Holland: Elsevier Science Publishers.

FODOR, J. A. (1983). *The modularity of the mind*. Cambridge, MA: MIT Press.

FODOR, J. D. (1991). Thematic roles and modularity: Comments on chapters by Frazier and Tanenhaus et al. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing*. Cambridge, MA: MIT Press.

FORD, M., BRESNAN, J., & KAPLAN, R. (1982). A competence based theory of syntactic closure. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

FRAZIER, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. Thesis, University of Connecticut. Indiana University Linguistics Club.

FRAZIER, L. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence Processing*. Hillsdale, NJ: Erlbaum.

FRAZIER, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural Language Parsing*. Cambridge University Press, Cambridge.

FRAZIER, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.

FRAZIER, L. (1989). Against lexical generation of syntax. In W. D. Marslen-Wilson (Ed.), *Lexical Representation and Process*. Cambridge, MA: MIT Press.

FRAZIER, L. (1991) Exploring the architecture of the language-processing system. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing*. Cambridge, MA: MIT Press.

FRAZIER, L., & RAYNER, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14**, 178–210.

GARNHAM, A. (1985). *Psycholinguistics: Central Topics*. New York: Methuen.

HINDLE, M., & ROOTH, M. (1990). Structural ambiguity and lexical relations. *Proceedings of the 28th Annual Meeting of the Association of Computational Linguistics*.

JULIANO, C., TRUESWELL, J. C., & TANENHAUS, M. K. (1992). What can we learn from "that"? Paper presented at the *33rd Annual Meeting of the Psychonomic Society*. St. Louis, MO.

KAPLAN, R., & BRESNAN, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

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

KOOI, F., DE VALOIS, K. K., & SWITKES, E. (1991). Higher order factors influencing the perception of sliding and coherence of a plaid. In *From pigments to perceptions: Advances in understanding visual processes*. New York: Plenum.

KRAUSKOPF, J., & FARRELL, B. (1990). Influence of color on the perception of coherent motion. *Nature*, **348**, 328–331.

LORCH, R. F., JR., & MYERS, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **16**, 149–157.

MARSLEN-WILSON, W. D., & TYLER, L. K. (1987). Against modularity. In J. Garfield (Ed.), *Modularity in knowledge representations and natural language understanding*. Cambridge, MA: MIT Press.

MACDONALD, M. (1992). Probabilistic constraints and syntactic ambiguity resolution. Manuscript submitted for publication.

MACDONALD, M. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32(5), 692–715.

McCawley, J. D. (1971). Where do noun phrases come from? In D. D. Steinberg and L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics & psychology*. Great Britain: Cambridge University Press.

McClelland, J. L. (1987). The case of interactionism in language processing. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Erlbaum.

McClelland, J. L., & Kawamoto, A. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In J. McClelland and D. Rumelhart (Eds.), *Parallel distributed processing, part 2: Psychological and biological models*. Cambridge, MA: MIT Press.

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. (1994). Syntactic Processing. In M. Gernsbacher (Ed.), *The Handbook of Psycholinguistics*. In press. New York, NY: Academic Press.

Mitchell, D. C., & Cuetos, F. (1991). The origins of parsing strategies. Unpublished manuscript.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattas, & C. Gross (Eds.), *Pattern recognition mechanisms*. New York, NY: Springer.

Ni, W., & Crain, S. (1990). How to resolve structural ambiguities. *Proceedings to NELS, 20, Vol 2*. Pp. 414–427.

Pearlmutter, N. J., Daugherty, K., MacDonald, M. C., & Seidenberg, M. S. (1993). Constraint satisfaction in main verb/reduced relative ambiguities. *Poster presented at the Sixth Annual CUNY Sentence Processing Conference*. University of Massachusetts, Amherst, MA.

Pearlmutter, N. J., & MacDonald, M. C. (1992). Plausibility effects in syntactic ambiguity resolution. *Proceedings of 14th Annual Conference of the Cognitive Science Society*.

Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, 64, 539–576.

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago, IL: The University of Chicago Press.

Poggio, T., Torre, V., & Koch, C. (1985). Visual perception and regularization theory. *Nature*, 317, 314–319.

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.

Rayner, K., Pacht, J. M., Sereno, S. C., & Duffy, S. A. (1991). Lexical ambiguity and eye fixations: Frequency and familiarity effects. *Paper presented at the 32nd Annual Meeting of the Psychonomic Society*. San Francisco, CA.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Erlbaum.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processing, Special Edition, Parsing and Interpretation*, 4, 21–49.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Parsing reduced relative clauses. *Canadian Journal of Psychology*, 47, 276–309.

Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1993). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. To appear in *Attention & Performance XV*.

Tanenhaus, M. K., Boland, J. E., Garnsey, S. M., & Carlson, G. N. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research*, 18, 37–50.

Tanenhaus, M. K., Boland, J. E., Mauner, G., & Carlson, G. N. (in press). More on combinatory lexical information: Thematic structure in parsing and interpretation. In G. T. M. Altmann & Shillock (Eds.), *Cognitive models of speech and processing: The second Sperlonga workshop*. East Sussex, England: Erlbaum.

Tanenhaus, M. K., & Carlson, G. N. (1989). Lexical structure and language comprehension. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. (1985). Do listeners compute linguistic representations? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing*. Cambridge: Cambridge University Press.

Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processing, Special Edition, Parsing and Interpretation*, 4, 211–234.

Tanenhaus, M. K., Dell, G. S., & Carlson, G. (1987). Context effects in lexical processing: A connectionist perspective on modularity. In J. Garfield (Ed.), *Modularity in Knowledge Repre-*

sentation and Natural Language Understanding. Cambridge, MA: MIT Press.

TANENHAUS, M. K., GARNSEY, S. M., & BOLAND, J. (1991). Combinatory Lexical Information and Language Comprehension. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing.* Cambridge, MA: MIT Press.

TANENHAUS, M. K., & JULIANO, C. (1992). What to do about "that": Use of co-occurrence information in parsing. *Paper presented at the 5th Annual CUNY Conference on Sentence Processing.* City University of New York, New York, NY.

TANENHAUS, M. K., & LUCAS, M. M. (1987). Context effects in lexical processing. *Cognition,* **25**, 213–234.

TANENHAUS, M. K., & TRUESWELL, J. C. (1994). Sentence comprehension. To appear in *Handbook of Perception and Cognition: Volume 11 Speech, Language and Communication.*

TARABAN, R., & MCCLELLAND, J. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *The Journal of Memory and Language,* **27**, 1–36.

THOMPSON, H., & ALTMANN, G. T. M. (1991). Modularity Compromised: Selecting Partial Hypotheses. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing.* Cambridge, MA: MIT Press.

TRUESWELL, J. C., & TANENHAUS, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes,* **6**, 303–338.

TRUESWELL, J. C., & TANENHAUS, M. K. (1992). Consulting temporal context in sentence comprehension: Evidence from the monitoring of eye movements in reading. *Proceedings of the 14th Annual Meeting of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

TRUESWELL, J. C., TANENHAUS, M. K., & KELLO, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **19**(3), 528–553.

WEINBERG, A. (in press). Minimal commitment: A parsing theory for the nineties. *Language and Cognitive Processes.*

(Received June 7, 1989)
(Revision received May 19, 1993)