Serial versus Parallel Sentence Comprehension

Neal J. Pearlmutter and Aurora Alma Mendelsohn
Northeastern University

Send correspondence to either author:
Psychology Dept., 125 NI
Northeastern University
Boston, MA 02115

E-mail: pearlmutter@neu.edu, amendels@lynx.dac.neu.edu
Phone: (617) 373-3040, (617) 373-5551
Fax: (617) 373-8714

Running head: SERIAL VERSUS PARALLEL SENTENCE COMPREHENSION

Manuscript submitted August 31, 1999; comments welcome.

Abstract

Two experiments examined whether the human sentence processing system handles ambiguities by constructing exactly one possible interpretation and switching to another only if necessary (a serial system), or if, instead, it maintains multiple ranked interpretations in parallel. In the main experiment, comprehenders read sentences containing temporary ambiguities, in which the preferred interpretation (established in the second experiment) was plausible throughout and turned out to be correct, and in which the unpreferred interpretation either remained plausible or became implausible prior to disambiguation. Readers displayed difficulty when the unpreferred interpretation became implausible, contradicting the predictions of deterministic serial systems, in which the unpreferred alternative should never have been considered, and supporting the predictions of ranked-parallel systems. Correlational analyses showed that the two alternative interpretations were not in competition with each other, contradicting the predictions of both probabilistic serial systems, in which a single alternative is selected, but not always the same one across trials; and competitive ranked-parallel systems, in which increased support for one alternative necessarily weakens support for the other alternative(s). The full pattern of results supports a ranked-parallel system in which interpretations compete only to the extent that they are incompatible with each other, where compatibility depends on the amount of overlap between the representations required for the ambiguity's alternatives. We propose such a system, in which subcomponents of representations can be separately activated, can decay independently, and can be shared across multiple alternatives for an ambiguity.

Keywords: sentence comprehension, ambiguity resolution, parsing, competition, compatibility, structure sharing, argument structure frequency

Theories of human sentence comprehension attempt to explain how people combine the meanings of individual words to create the more complex meaning of a complete sentence. The intuitively smooth and seamless nature of this process, however, belies substantial complexity, which can be seen in one of the central problems in sentence processing, the resolution of temporary ambiguity. Temporary ambiguity is ubiquitous in natural language, but for the most part comprehenders are able to interpret sentences that resolve with either possible meaning, typically without ever becoming aware of alternative interpretations.

The reason that temporary ambiguities create a problem is that they require the comprehension system to make decisions under uncertainty which are potentially costly in terms of computational resources. Two basic kinds of models have been proposed to handle ambiguity: In a serial model, the comprehension system selects a single interpretation on some basis, and only considers other alternatives if its first choice turns out to be wrong. When the system's first choice is correct, this model is optimally efficient, expending no resources on possibilities that turn out to be irrelevant. However, if its first choice turns out to be wrong, reanalyzing the ambiguity may be time-consuming and disruptive. In a parallel model, the comprehension system attempts to maintain all possibilities (or at least the a priori most likely ones) to guarantee that the right one will be available when required. Parallel models avoid the need for costly reanalysis, but they can require substantially more resources during the processing of the ambiguity.

The choice between a serial and a parallel system has always been considered critical in defining sentence comprehension theories (e.g., Crain & Steedman, 1985; Frazier, 1979, 1987; Gibson, 1991; Gorrell, 1987; Hickok, 1991; Just & Carpenter, 1992; Kurtzman, 1985; MacDonald, Just, & Carpenter, 1992; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994), but the issue has rarely been directly investigated. Frazier and Rayner (1982) examined the direct object versus sentence complement ambiguity exemplified in (1), where the noun phrase *the mayor's position* can initially be interpreted as either the direct object of the matrix verb *argued*, as eventually required in (1a), or instead as the subject of an embedded sentence complement, as eventually required in (1b).

(1) a. The city council argued the mayor's position forcefully.

b. The city council argued the mayor's position was incorrect.

Readers spent more time (adjusted for text length) reading the text following the ambiguous noun phrase when the sentence complement interpretation was required than when the direct object interpretation was. Numerous other experiments as well as intuitions have revealed similar preferences for one alternative in an ambiguity over others (e.g., Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Gibson, 1991; Trueswell, Tanenhaus, & Garnsey, 1994; Trueswell, Tanenhaus, & Kello, 1993). These kinds of results rule out a model in which multiple alternatives are all equally available, as would be the case in the simplest version of a parallel model.

However, while these results are consistent with serial models, they are also compatible with ranked-parallel models, in which multiple alternatives are computed and then ranked according to various criteria, possibly including grammatical category frequency, argument structure frequency, syntactic structure frequency, syntactic complexity, plausibility, and discourse complexity (e.g., Boland, 1997; Gibson, 1991, 1998; Jurafsky, 1996; Just & Carpenter, 1992; MacDonald et al., 1994; Spivey & Tanenhaus, 1998; Stevenson, 1994; Tabor, Juliano, & Tanenhaus, 1997). In general, the existence of a preference at disambiguation for one alternative over another does not indicate whether the non-preferred alternative was never constructed and thus must be computed at some cost when required (as in a serial model), or whether instead the non-preferred alternative is available, but the reordering process imposes a cost (as in a ranked-parallel model).

Evidence supporting a ranked-parallel model, however, is also quite limited. Gorrell (1987), using a syntactic-context priming paradigm (Wright & Garrett, 1984), and Hickok (1991), using a gap-priming paradigm (e.g., Swinney, Ford, Frauenfelder, & Bresnan, 1988), both showed that lexical decisions to words which were compatible with the non-preferred interpretation of an ambiguity were fast relative to various controls, suggesting that a non-preferred interpretation was computed. However, lexical decision tasks in general (e.g., West & Stanovich, 1982, 1986), and the gap-priming paradigm in particular (McKoon, Ratcliff, & Ward, 1994), have been criticized. It is also possible that relatively rapid reanalysis within a serial model, possibly caused by the lexical decision probe itself, might be able to account for such results.

MacDonald et al. (1992) found individual differences in the processing of ambiguities and argued that readers with higher working memory capacities (as measured by Daneman & Carpenter's, 1980, reading span task) were more likely to be able to maintain multiple interpretations simultaneously, although at a substantial cost in reading time. The effects of span appeared most clearly at disambiguation, however, and Pearlmutter and MacDonald (1995) suggested that the results might be due to higher-span readers' apparent ability to make use of more complex information sources on-line, rather than to an increased capacity for considering multiple interpretations simultaneously. A variety of other studies have also shown sensitivity at disambiguation to various information sources, and specifically to properties associated with either the preferred or non-preferred interpretation (e.g., Garnsey et al., 1997; Schmauder & Egan, 1998; Trueswell et al., 1993, 1994), suggesting that both interpretations were being considered. But because all of these effects occur at disambiguation, they can also be explained in a serial model as the result of rapid correction of an earlier incorrect selection, with the difficulty of the correction process depending on properties of the alternatives (Frazier, 1995).

Thus to provide direct evidence about ranked-parallel versus serial models, we need to test for the presence of the non-preferred interpretation of an ambiguity prior to its disambiguation (see also Frazier, 1995). We must therefore be able to manipulate some property of that interpretation during the ambiguity while holding constant properties of the preferred interpretation. If the manipulation then has an influence, this would provide evidence that the non-preferred interpretation was being maintained, which would be evidence against a serial model. The ambiguity we used is the noun-triggered sentence complement versus relative clause ambiguity, an example of which is shown in (2).

(2) a. The claim that the cop ignored the informant might have affected the jury. (SC)

b. The claim that the cop ignored might have affected the jury. (RC)

In (2), at the word *ignored*, the sentence so far is ambiguous between the sentence complement alternative eventually required in (2a) and the relative clause alternative eventually required in (2b). In the sentence complement version (2a), an embedded sentence (*that the cop ignored the informant*) follows *claim* and specifies the content of the claim. In the relative clause version (2b), *report* is followed by a modifying relative clause (*that the cop ignored*). The most critical difference for current purposes is that the entity being ignored by the cop in the two interpretations is different: In (2a) it is the informant, whereas in (2b) it is the claim.

Although this ambiguity has received very little empirical attention (cf. Gibson, 1991), most theories which make an a priori prediction predict that the relative clause alternative will be non-preferred (e.g., Abney, 1989; Altmann, Garnham, & Dennis, 1992; Altmann & Steedman, 1988; Frazier, 1979, 1987; Frazier & Clifton, 1996; Pritchett, 1988). Experiment 2 provides on-line evidence supporting these predictions, and thus our strategy is to manipulate the relative clause interpretation prior to disambiguation, while holding constant relevant properties of the sentence complement interpretation.

Because the critical difference between the two interpretations is in which noun phrase is interpreted as the direct object of the embedded verb, we can vary the embedded verb and differentially affect the two interpretations, as shown by a comparison of (2) above to (3) below. In (2), with the embedded verb *ignored*, both the sentence complement and relative clause interpretations are plausible (a cop might reasonably ignore either an informant or a claim). In (3), however, the embedded verb *shot* selectively renders the relative clause interpretation (3b) implausible: A cop could possibly shoot an informant (the sentence complement; 3a), but a cop shooting a claim (3b) is semantically odd. Thus in (2), both interpretations are plausible and remain so at least until disambiguation. In (3), on the other hand, both interpretations are initially plausible, but the relative clause interpretation (only) becomes odd at the embedded verb.

(3) a. The claim that the cop shot the informant might have affected the jury. (SC)

b. The claim that the cop shot might have affected the jury. (RC)

The stimuli used in Experiment 1 were always resolved with the sentence complement alternative (2a or 3a), and that interpretation was always plausible; but the potential relative clause interpretation, though never ultimately correct, either remained plausible (2) or became implausible (3) at the embedded verb (e.g., *ignored/shot*). Thus the plausibility of the unpreferred relative clause interpretation is manipulated while the plausibility of the sentence complement interpretation is held constant.

In order to examine the predictions of serial and ranked-parallel models, we need to divide each class of model into two subtypes. Serial models can be differentiated in terms of whether they are deterministic or probabilistic. In deterministic serial models (e.g., Frazier, 1979, 1987; Frazier & Rayner, 1982; Frazier & Clifton, 1996, which is deterministic with respect to this ambiguity, although it is probabilistic for some others), the comprehension system always selects the same interpretation for a given ambiguity. If the preferred interpretation turns out to be correct, these models always consider only that interpretation, and thus manipulations of the non-preferred interpretation should have no influence on comprehension. In a sentence like (3a), where the sentence complement interpretation turns out to be correct, a deterministic serial model will select that interpretation initially and will never consider any others. Manipulations of the relative clause alternative should not be noticed. Deterministic serial models are addressed in Experiment 1A.

In probabilistic serial models (e.g., Ferreira & Henderson, 1990; Mitchell, 1987, 1989; van Gompel, Pickering, & Traxler, 1999), a single alternative is chosen, but while the preferred interpretation is the more likely choice, the non-preferred interpretation will occasionally be selected instead. Thus in (3a), the sentence complement will usually be the only interpretation considered, and on these trials, the relative clause implausibility will be irrelevant. However, on some small percentage of trials, the relative clause interpretation (only) will be considered instead, and thus its implausibility will be noticed in these cases. Because the sentence complement interpretation is eventually correct, some reanalysis will take place as well, either at disambiguation or when the implausibility of the relative clause is detected. (Specific models may or may not permit implausibility to trigger reanalysis.) Probabilistic serial models are addressed in Experiment 1B.

Ranked-parallel models can also be divided into two types: competitive and non-competitive. In either case, these models should be sensitive to manipulations of either interpretation, because they will compute them both. Thus at the embedded verb in (3a), readers should notice that one of the alternatives has become implausible, and this should create some processing difficulty. This prediction is tested in Experiment 1A.

The difference between competitive and non-competitive ranked-parallel models is in whether manipulations of one alternative necessarily influence the other alternative as well. For example, in a competitive model, when some source of information (e.g., argument structure frequency) renders the preferred alternative relatively more likely, this also has the effect of rendering the non-preferred alternative less likely. The two alternatives compete with each other for computational resources (e.g., working memory capacity or activation), or because they are mutually incompatible, and so any factor that affects the ranking of one also affects the ranking of the other. In contrast, in a non-competitive model, factors which influence the support for one alternative may or may not have corresponding effects on other alternatives. In such a model, as one alternative becomes more strongly preferred, other alternatives might show increased support, no change, or a weakening in strength. Most ranked-parallel models at least permit competition (e.g., Boland, 1997; Gibson, 1991; Jurafsky, 1996; Just & Carpenter, 1992; MacDonald et al., 1994; Tabor et al., 1997), and some rely on it for ambiguity resolution (e.g., Spivey & Tanenhaus, 1998; Stevenson, 1994). In some cases (e.g., Gibson, 1991; Just & Carpenter, 1992), the degree to which competition exists might vary with working memory capacity and from ambiguity to ambiguity, as some ambiguities will be more taxing than others. Experiment 1B will examine the distinction between competitive and non-competitive ranked-parallel models.

Experiment 1A

This experiment made use of the noun-triggered sentence complement versus relative clause ambiguity (as in (2) and (3)) to examine whether readers were sensitive to manipulations of the non-preferred (relative clause) interpretation. We manipulated the plausibility of the non-preferred (relative clause) interpretation while holding constant the plausibility of the preferred (sentence complement) interpretation. The stimuli were always resolved with the preferred sentence complement interpretation. In a deterministic serial model, only the sentence complement alternative should ever be considered, and thus manipulations of the relative clause should be irrelevant. In a ranked-parallel model, however, although the sentence complement interpretation will be preferred, the relative clause possibility will also be constructed, and readers should experience some difficulty if it then becomes implausible. (See Frazier, 1995, for related discussion.)

Stimuli like those in (2a) and (3a) were used. Each of these was compared to a corresponding unambiguous control, which forced the sentence complement interpretation of the embedded clause in advance. This was to insure that uncontrolled differences between the critical verbs (*ignored* in (2a) and *shot* in (3a)) were not confounded with effects predicted by the different models. Thus any potential difficulty resulting from the implausibility of the relative clause interpretation can be measured as the difference between the ambiguous version in which the relative clause becomes implausible (3a) and its unambiguous control. If the implausibility of the relative clause is noticed, the size of this difference should be larger than the corresponding difference for the versions in which both interpretations are plausible at the critical verb (as in (2a)).

Method

Participants. One hundred twenty-five Northeastern University undergraduates participated; one was excluded for having comprehension performance below 60%. All participants in this and the following experiments were native English speakers and received either class credit or $8 for their participation. No participant provided data for more than one study.

Materials and Design. Thirty-six stimulus sets like that shown in (4) were selected from a larger pool of candidate items. The stimuli always began with *the* and a noun (e.g., *claim* in (4)) which could potentially take a sentence complement or a relative clause beginning with *that*. Following *that* was a two-word noun phrase subject for the embedded clause (*the cop* in (4)) and then one of two possible embedded verbs (*ignored*, *shot*). After the embedded verb was a two-word direct object noun phrase (*the informant*) and then a matrix verb phrase. The matrix verb phrase always began with an unambiguous past tense or modal form (*might* in (4)) to rule out the possibility of treating it as part of a passive reduced relative clause. All 36 sentence sets are shown in Appendix A.

(4) a. The claim that the cop ignored the informant might have affected the jury.
b. The claim alleging that the cop ignored the informant might have affected the jury.
c. The claim that the cop shot the informant might have affected the jury.
d. The claim alleging that the cop shot the informant might have affected the jury.

The four versions were created by varying two factors. The first was ambiguity. In the ambiguous versions (4a and 4c), either interpretation was possible prior to encountering the embedded direct object. In the unambiguous versions (4b and 4d), the sentence complement interpretation was forced in advance by inserting a progressive verb immediately after the initial noun (e.g., *alleging*; cf. **the claim alleging that the cop ignored might have...*).

The second factor was the plausibility of the relative clause interpretation, which was manipulated by varying the embedded verb. In (4a), prior to encountering *the informant*, either the sentence complement or relative clause interpretation was possible, and both were plausible. In (4c), however, changing *ignored* to *shot* selectively renders the relative clause interpretation implausible. The plausibility of the different interpretations was measured in a separate norming

Table 1: Embedded Verb Properties and Rated Stimulus Plausibility for Experiments 1 and 2

	Embedded Verb			Plausibility			
Condition	Length	Log Freq	Transitivity	SC	RC	SV	Unamb
Plausible	7.1 (2.3)	3.38 (1.84)	92.9 (17.2)	5.7 (.76)	5.4 (.75)	5.4 (.56)	4.9 (.89)
Implausible	6.5 (1.7)	2.77 (1.48)	93.4 (12.5)	5.8 (.86)	2.6 (.74)	5.4 (.71)	4.8 (.87)

Note. Values in parentheses are standard deviations. Transitivity is the percentage of transitive uses out of all others (see text). Plausibility was rated on a scale of 1 (least plausible) to 7 (most plausible). SC = sentence complement interpretation. RC = relative clause interpretation. SV = subject-verb fragment. Unamb = unambiguous control.

study (see below) and is shown in Table 1. The length and log frequency (Francis & Kučera, 1982) of the plausible and implausible embedded verbs, which did not differ reliably (ps $> .10$), are also shown in Table 1, along with the verbs' transitivity bias (see below).

The 36 experimental stimuli were combined with 64 fillers to form four 100-item lists such that each experimental stimulus appeared exactly once in each list, and each list contained the same number of experimental items (nine) in each condition. The fillers had a variety of different syntactic structures, including 20 items which began with sentence-complement-taking nouns like those in the experimental items but did not continue with a similar embedded clause. Each experimental stimulus and filler also had a Yes/No comprehension question. For the experimental items, the question asked about some part of the item other than the ambiguity.

Plausibility Norming. To insure that changing the embedded verb selectively altered the relative clause interpretation's plausibility, a plausibility norming study was conducted. Four simple subject-verb-object (transitive) norming sentences were created from each stimulus set like (4), and corresponding subject-verb-*the* sentence fragments were also created. The norming stimuli for the stimulus set in (4) are shown in (5).

(5) a. The cop ignored the informant.

b. The cop ignored the claim.

c. The cop shot the informant.

d. The cop shot the claim.

e. The cop ignored the...

f. The cop shot the...

The subject of each norming stimulus was always the subject of the embedded clause (e.g., *the cop*). The different versions for norming were created by varying the verb and the direct object (when it was present). The verb was either the plausible (*ignored*) or implausible (*shot*) verb from the original stimulus set. The direct object noun, when present, was either the direct object which actually appeared in the stimulus set (*the informant*), or else the initial noun phrase of the stimulus set (*the claim*), which was interpreted as the direct object of the embedded verb in the relative clause interpretation of the ambiguity. Thus ratings of (5a) reflect the plausibility of the sentence complement interpretation with the plausible verb, and ratings of (5b) reflect the plausibility of the relative clause interpretation with the same plausible verb. Similarly, ratings of (5c) and (5d) reflect the plausibility of the sentence complement and relative clause interpretations, respectively, with the implausible verb. For the plausibility manipulation to be successful, the plausibility of the two sentence complement interpretations (5a and 5c) should be matched, as should the ratings of the two subject-verb fragments (5e and 5f), but the implausible relative clause (5d) should be less plausible than the plausible relative clause (5b).

The norming stimuli were placed into six lists such that each list contained exactly one version created from each original stimulus set, and full-sentence items to be rated (5a–5d) were never mixed with fragment items. Each list also contained an additional 79 items with the same structure (full transitive sentence or subject-verb fragment). Ninety-two participants each rated the stimuli in a single list for plausibility, using a 1–7 scale printed to the right of each item (1 = least plausible). Each list was 4 pages long, and the order of the pages was separately randomized for each participant.

The ratings of the two interpretations and the subject-verb fragment for both the plausible and implausible conditions are shown in Table 1; ratings for individual stimulus sets are shown in Appendix A. Paired comparisons confirmed the expected patterns: The two sentence complement interpretations (5a and 5c) were equally plausible ($t(35) = 1.14$, $p > .25$), as were the two fragments (5e and 5f) ($t(35) = .37$), but the implausible relative clause interpretation (5d) was reliably less plausible than the plausible relative clause interpretation (5b) ($t(35) = -17.86$, $p < .001$).

In addition to the plausibility ratings for the interpretations of the ambiguous stimuli, we also collected plausibility ratings of the unambiguous controls, using sentence-initial fragments like those in (6). The fragments were identical to the stimuli in the main experiment, except that they ended with the word following the critical verb, plus ellipsis.

(6) a. The claim alleging that the cop ignored the...

b. The claim alleging that the cop shot the...

The 36 norming items, each with a plausible-verb and implausible-verb version, were placed into two lists such that each list contained exactly one version from each original stimulus set. The order of the items in the lists was separately randomized, and 50 participants each rated the stimuli in a single list for plausibility, using a 1–7 scale identical to that described above (1 = least plausible). The mean ratings of the two versions are shown in Table 1; they did not differ ($t(35) = .86$). Individual stimulus ratings are shown in Appendix A.

Transitivity Norming. In addition to norming the stimuli for the desired plausibility differences, we also normed the embedded verbs for transitivity. For the relative clause plausibility manipulation to be effective, the initial noun must be interpreted as the direct object of the embedded verb. If the embedded verb is intransitively biased (e.g., *lecture*, Connine, Ferreira, Jones, Clifton, & Frazier, 1984; cf. *The claim that the cop lectured about...*), then the implausibility of the relevant relative clause interpretation might never be considered, regardless of whether processing is serial or ranked-parallel.

The 69 embedded verbs (3 verbs were used in more than one stimulus set) were placed into a single 6-page list in a random order. Each verb was in the past tense and was presented by itself. Twenty-five participants completed the survey. They were instructed to write a complete sentence using each verb, and the order of the pages was randomized separately for each participant.

The sentence tokens were coded for the presence of a direct object (transitivity). Standard direct object noun phrases, passives, and datives were counted as transitive. Constructions in which no direct object was present, including sentence complements (finite and infinitival) and prepositional phrases, were counted as intransitive. A sentence token was not counted if the verb was used in a different grammatical category (e.g., adjective), or if the intended meaning of the verb in the token was clearly different from the one in the experimental stimuli. Transitivity bias was computed as the percentage of transitive tokens out of all tokens counted, and each verb's bias is shown in Appendix A. The mean bias for the plausible and implausible verbs is shown in Table 1; the two means did not differ ($t(35) = -.14$).

Apparatus and Procedure. An IBM-compatible personal computer presented the stimuli and collected all data in the main experiment. Participants read 10 initial practice items followed by one of the 100-item lists in a random order. The experimental stimuli were displayed on 2 lines with the linebreak always just after the third word of the main verb phrase (e.g., after *affected* in (4)), and the fillers were displayed similarly. The stimuli were presented using a non-cumulative self-paced moving-window procedure (Just, Carpenter, & Woolley, 1982). At the beginning of a trial, an item was displayed on the screen with all non-space characters replaced by dashes. When the participant pressed the space bar, the first word of the item was displayed, replacing the corresponding dashes. When the participant pressed the space bar a second time, the first word reverted to dashes, and the second word was displayed in place of the appropriate dashes. Each subsequent press of the space bar revealed the next word and removed the previous word. Pressing the space bar on the last word of the item caused the item to be replaced by its Yes/No comprehension question, which the participant answered by pressing one of two keys above the space bar on the keyboard. The computer recorded the time between each button-press as well as the comprehension question response and presented feedback about the participant's answer to the question. Most participants completed the experiment in approximately 30 minutes.

Results

All data were analyzed in 2 (plausibility) $\times$ 2 (ambiguity) repeated-measures ANOVAs, with either participants (F_1) or items (F_2; Clark, 1973) as the random factor. During analysis, the comprehension question for one item (see Appendix A) was discovered to be ambiguous. Comprehension performance on this item was below 50% in each condition, and it was excluded from all the analyses reported below.

Comprehension Question Performance. Comprehension performance was 91% correct for the implausible ambiguous condition, 90% correct for the implausible unambiguous condition, and 89% correct for the two plausible conditions. These values did not differ reliably, although the main effect of plausibility was marginal by participants only ($F_1(1,123) = 3.05$, $MS_e = 96.29$, $p < .10$; $F_2(1,34) = 1.43$, $MS_e = 67.10$, $p > .20$; all other Fs < 1).

Reading Times. To adjust for differences in word length across conditions as well as overall differences in participants' reading rates, a regression equation predicting reading time from word length was constructed for each participant, using all filler and experimental items (Ferreira & Clifton, 1986; see Trueswell et al., 1994, for discussion). At each word position, the reading time predicted by the participant's regression equation was subtracted from the actual measured reading time to obtain a residual reading time. Thus each participant's mean reading time per word across the entire experiment was transformed to 0 ms residual reading time, and negative residual times indicate faster than average times. Trials on which the comprehension question was answered incorrectly were excluded; and residual reading times beyond 2.5 *SD* from the corresponding condition $\times$ position cell mean were also excluded, affecting less than 2.3% of the data. All analyses were conducted on the resulting data set. Appendix B reports the raw reading times trimmed at 2.5 *SD*.

Figure 1 shows residual reading time per word by condition for three analysis regions: the critical embedded verb, where effects of plausibility could first appear; the following direct object, which syntactically disambiguated the embedded clause to the sentence complement interpretation, and the following two words, which formed the beginning of the main verb phrase.

At the critical embedded verb (e.g., *ignored*, *shot*), while the two plausible conditions were identical (Fs < 1), the implausible ambiguous condition was reliably slower than its unambiguous control ($F_1(1,123) = 8.78$, $p < .01$; $F_2(1,34) = 8.91$, $p < .01$). This resulted in a plausibility $\times$ ambiguity interaction ($F_1(1,123) = 8.27$, $MS_e = 4352$, $p < .01$; $F_2(1,34) = 6.07$, $MS_e = 1702$, $p < .05$) and no main effects (plausibility: Fs < 1; ambiguity: Fs < 3, ps $> .10$).

At the direct object (e.g., *the informant*), ambiguous conditions were slower than unambiguous ones ($F_1(1,123) = 5.76$, $MS_e = 1986$, $p < .05$; $F_2(1,34) = 6.00$, $MS_e = 602$, $p < .05$), although neither of the paired comparisons within the main effect was reliable separately (plausible: $F_1(1,123) = 3.17$, $p = .08$; $F_2(1,34) = 2.26$, $p = .14$; implausible: $F_1(1,123) = 2.65$, $p = .11$;

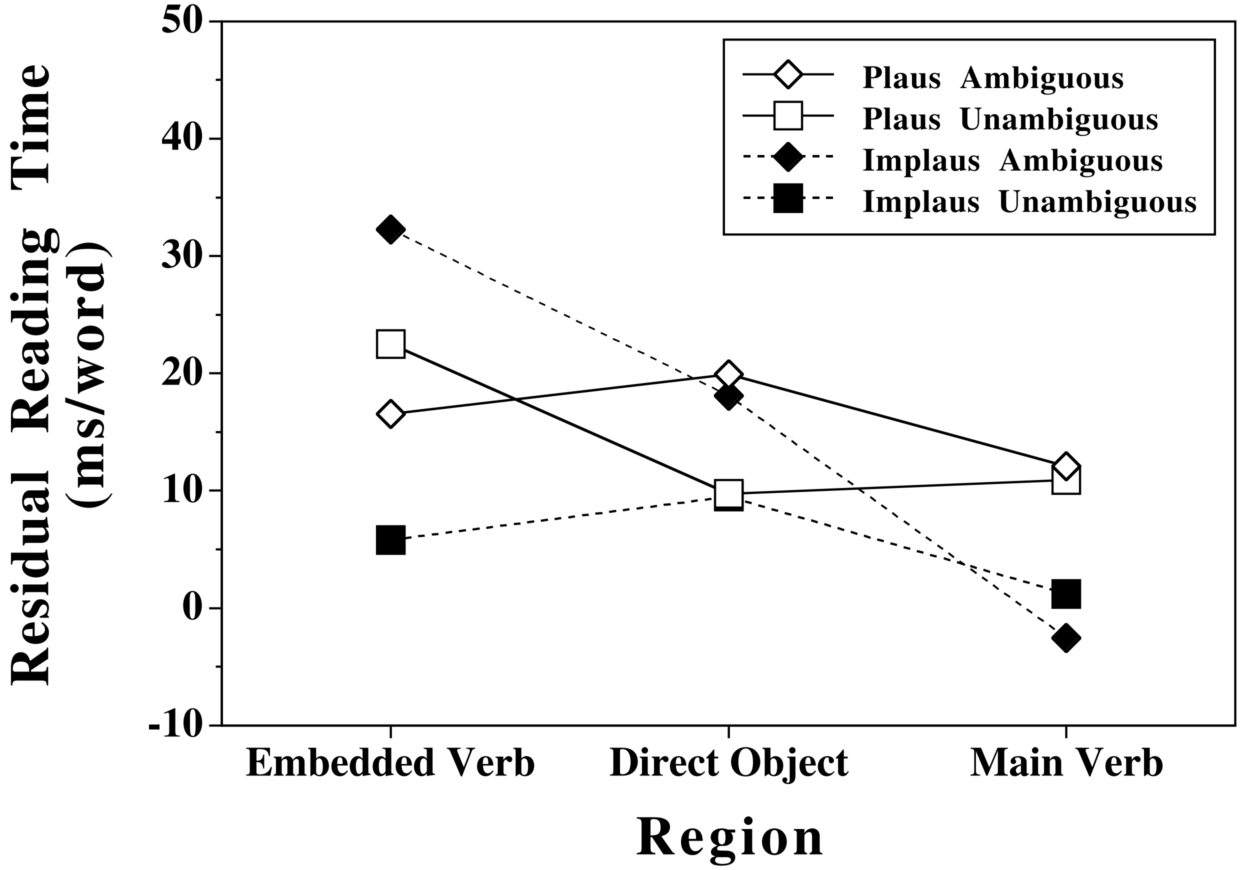

Figure 1: Experiment 1A residual reading time per word by condition and region.

$F_2(1, 34) = 3.49$, $p = .07$). There was no interaction or effect of plausibility (Fs < 1).

At the first two words of the main verb phrase (e.g., *might have*), a main effect of plausibility was present ($F_1(1, 123) = 13.20$, $MS_e = 1410$, $p < .01$; $F_2(1, 34) = 8.70$, $MS_e = 541$, $p < .01$). There was no interaction or effect of ambiguity (Fs < 1).

Discussion

The critical result in this experiment was the relative difficulty at the embedded verb in the implausible ambiguous condition relative to its unambiguous control, coupled with the lack of a difference between the two plausible conditions.

In a deterministic serial model, there should have been no difference between the size of the ambiguity effects for the plausible and implausible conditions, because in all four conditions, the processor should have computed only the sentence complement interpretation: In the unambiguous conditions, the relative clause interpretation was ruled out in advance, and in the ambiguous conditions, the sentence complement should be preferred, so the relative clause should never have been computed. (Experiment 2 provides direct evidence about the preference for the sentence complement interpretation.) Thus the relative clause interpretation's plausibility should have had no influence on processing.

A ranked-parallel model (competitive or not), on the other hand, accounts for the embedded verb pattern straightforwardly: Although the sentence complement interpretation is preferred, the relative clause is also computed. In the plausible conditions, both interpretations are plausible, so the system has no difficulty. In the implausible ambiguous condition, however, the relative clause interpretation becomes implausible at the embedded verb, and this results in some difficulty. Thus these results indicate that the relative clause interpretation must have been considered on at least some of the trials in the experiment. However, they do not establish whether the relative clause interpretation was considered on a separate set of trials from the sentence complement interpretation, as a probabilistic serial model would predict, or whether both alternatives were considered on the same trials, as a ranked-parallel model would predict. Experiment 1B examined this issue.

Experiment 1B

Experiment 1A provided evidence against deterministic serial parsing models, which always choose the same single interpretation for a given structural ambiguity. However, a serial processor might instead choose a single interpretation probabilistically (e.g., Ferreira & Henderson, 1990; Mitchell, 1987; van Gompel et al., 1999), selecting the preferred sentence complement interpretation most of the time, but occasionally selecting the relative clause interpretation instead. This class of serial model could account for the ambiguity effect at the implausible embedded verb in Experiment 1A by attributing it to the minority of trials in which the relative clause interpretation was chosen instead of the sentence complement: On these trials (only), the implausibility of the relative clause will be noticed and will create some difficulty. This class of models could also account for the effect of ambiguity at disambiguation in the same way: In a minority of trials, the relative clause will have to be dropped and the sentence complement interpretation constructed in its place, and this reanalysis will happen at the disambiguation.

To examine probabilistic serial models, we therefore considered an important additional prediction, which derives from the fact that in any serial model, the sentence complement and relative clause interpretations must be in complementary distribution. That is, given that the sentence complement interpretation will be selected on some proportion of trials for a given stimulus, the relative clause must be selected on the rest of the trials for that stimulus. Consequently, if some biasing factor predicts how often the sentence complement interpretation is selected, it must also necessarily predict how often the relative clause interpretation is chosen.

One biasing factor which most ranked-parallel and probabilistic serial models predict will influence the choice or ranking among alternatives is the frequency with which the initial, ambiguity-triggering noun takes a sentence complement. This argument structure frequency bias is like that shown to influence ambiguity resolution (either immediately or eventually) in a variety of other ambiguities, although verb-triggered cases have been examined far more extensively than noun-based ones (e.g., Boland, Tanenhaus, Garnsey, & Carlson, 1995; Ferreira & Henderson, 1990; Garnsey et al., 1997; Holmes, Stowe, & Cupples, 1989; MacDonald, 1994; MacDonald et al., 1994; Mitchell, 1987, 1989; Schmauder & Egan, 1998; Trueswell et al., 1993).

To examine the influence of argument structure frequency in the sentence complement versus relative clause ambiguity, we measured sentence complement bias (hereafter *SC bias*) for each of the ambiguity-triggering nouns in the Experiment 1A stimuli, and then correlated it with difficulty at disambiguation (the size of the ambiguity effect) in the plausible conditions of Experiment 1A (4a and 4b). In ranked-parallel models and in probabilistic serial models which make use of lexical information, the amount of difficulty experienced at disambiguation should depend on the strength of the noun's SC bias. This is because SC bias should predict the relative strength of the two alternatives in a ranked-parallel model, and it should predict the frequency with which the sentence complement is chosen as opposed to the relative clause in a probabilistic serial model.

If SC bias does have an influence on difficulty at disambiguation in the plausible conditions, then we can consider the predictions of various models about its influence at the embedded verb in the implausible conditions. The size of the ambiguity effect at the embedded verb in these conditions measures the degree to which the relative clause interpretation (and therefore its implausibility) is noticed.

As noted above, the critical property of serial models is that they construct only a single interpretation on a given trial, and thus if there are only two alternatives to be considered, any factor which increases the chance of constructing one alternative necessarily decreases the chance of constructing the other. Thus if SC bias determines how often the sentence complement interpretation is constructed, it also must determine how often the relative clause interpretation is constructed. Because the likelihood of constructing the relative clause alternative determines whether or not its implausibility is noticed, stimuli for which the relative clause is often constructed should show a larger difference between the ambiguous and unambiguous implausible conditions at the embedded verb than stimuli for which the relative clause is not often constructed. As a result, the correlation between SC bias and ambiguity effect size at the embedded verb in the implausible conditions should be comparable to the correlation between SC bias and ambiguity effect size at disambiguation in the plausible conditions: In both cases, as SC bias increases, the size of the ambiguity effect should decrease. At the disambiguation, this is because the sentence complement will have been initially chosen more often for nouns with stronger SC biases, resulting in less reanalysis difficulty. At the embedded verb, the correlation should arise because SC bias also determines how often the relative clause alternative is chosen, and it will have been chosen less often for nouns with stronger SC biases, resulting in less attention to its implausibility at the embedded verb.

This same pattern of correlations is also predicted by competitive ranked-parallel models, because the effect of competition in such models is to cause the two alternatives to tend to be in complementary distribution, just as they necessarily must be in probabilistic serial models. Thus to the extent that two alternatives compete, any factor which increases the support for one alternative will, directly or indirectly, decrease the support for the other. Assuming that the relative clause interpretation's implausibility causes a slowdown to the extent that that interpretation is supported, the result of competition will be that an increase in support for the sentence complement causes a decrease in support for the relative clause and a corresponding decrease in the effect of its implausibility. This will yield the same pattern of correlations as for probabilistic serial models.

However, because ranked-parallel models have multiple interpretations available on a single trial, they can be differentiated in terms of a parameter not available to serial models, which is how (or to what extent) the interpretations interact with each other during processing. In a competitive ranked-parallel model, the interpretations are tightly linked, and a change in the support for one directly affects the support for the other. Eliminating this link yields a non-competitive ranked-parallel model, in which changes to the support for one interpretation do not necessarily have any impact on the other interpretation. In such a model, even though SC bias should control the preference for the sentence complement interpretation (assuming a lexically-driven model), this biasing factor will not control the support for the relative clause interpretation. Thus, for example, in stimuli with a strong SC bias, this will result in increased support for the sentence complement interpretation, but it will not decrease support for the relative clause possibility. Thus SC bias should correlate with the disambiguation ambiguity effect in the plausible conditions, but it should not correlate with the embedded verb ambiguity effect in the implausible conditions.

To examine the above predictions, we first used a sentence completion survey to measure SC biases. We then examined the two critical correlations involving SC bias and Experiment 1A reading times: first, with ambiguity effect size at disambiguation in the plausible conditions, and, second, with ambiguity effect size at the embedded verb in the implausible conditions. Probabilistic serial models and competitive ranked-parallel models predict that the two correlations should be similar; non-competitive ranked-parallel models predict that only the first correlation should be reliable.

Method

Participants. Fifty-nine Northeastern University undergraduates participated, but one participant's data were excluded because of a very high rate of ungrammatical responses (34%).

Materials and Design. The 35 nouns from the Experiment 1A stimuli were normed (one noun, *information*, was used twice in Experiment 1A), along with three additional nouns which could also take a sentence complement. Each noun was presented as part of the beginning of a sentence, preceded by *The* and followed by *that*, as in (7). Four 3-page lists were created, each with a different random order of items. The order of the pages was randomized separately for each participant.

(7) The complaint that

Procedure. Participants were instructed to complete each beginning to form a full sentence.

Data Coding. The responses were coded for the presence of a sentence complement or a relative clause following the noun. Responses were excluded if they were ambiguous between the two types (3.7% of all responses) or if they were ungrammatical (3.5%).

Results

Across the set of nouns, the mean percentage of sentence complement completions out of all sentence complement and relative clause completions (SC bias) was 48 ($SD = 25$). Individual noun SC bias values are presented in Appendix A.

Figure 2 shows a scattergram for the correlation between noun SC bias and the size of the ambiguity effect (ambiguous − unambiguous) for the Experiment 1A plausible conditions at the disambiguating region (the direct object). This correlation was reliable ($r = .48$, $p < .01$) and accounted for a substantial proportion of the variance in ambiguity effect size (23%), indicating that the ambiguity-triggering noun's argument structure frequency bias predicts readers' preference for a sentence complement versus a relative clause.

Figure 3 shows the corresponding scattergram for the correlation between noun SC bias and the size of the ambiguity effect for the Experiment 1A implausible conditions at the embedded verb. As is evident from the figure, this correlation did not approach reliability ($r = .10$, $p > .55$), and a test of whether it differed from the correlation in Figure 2 above revealed that it did ($z = -1.68$, $p < .05$, one-tailed). Thus SC bias predicted difficulty at the disambiguation in the plausible conditions, but it did not predict difficulty with the relative clause's implausibility at the embedded verb. This indicates that whereas the ambiguity-triggering noun's argument frequency bias predicts the amount of support for the sentence complement interpretation, it does not influence the amount of support for the relative clause possibility.

Because this lack of a correlation relative to the reliable one in Figure 2 is critical for differentiating probabilistic serial, competitive ranked-parallel, and non-competitive ranked-parallel models, we also considered some methodological explanations for it. Comparison of Figures 2 and 3 suggests that the embedded verb reading time difference variable (the y-axis of Figure 3) is not limited in its range ($SD = 58$, vs $SD = 38$ for the disambiguation reading time difference variable in Figure 2), and of course the SC bias variable is exactly the same in the two correlations. It was also not the case that the embedded verb difference variable failed to correlate with anything, as it did

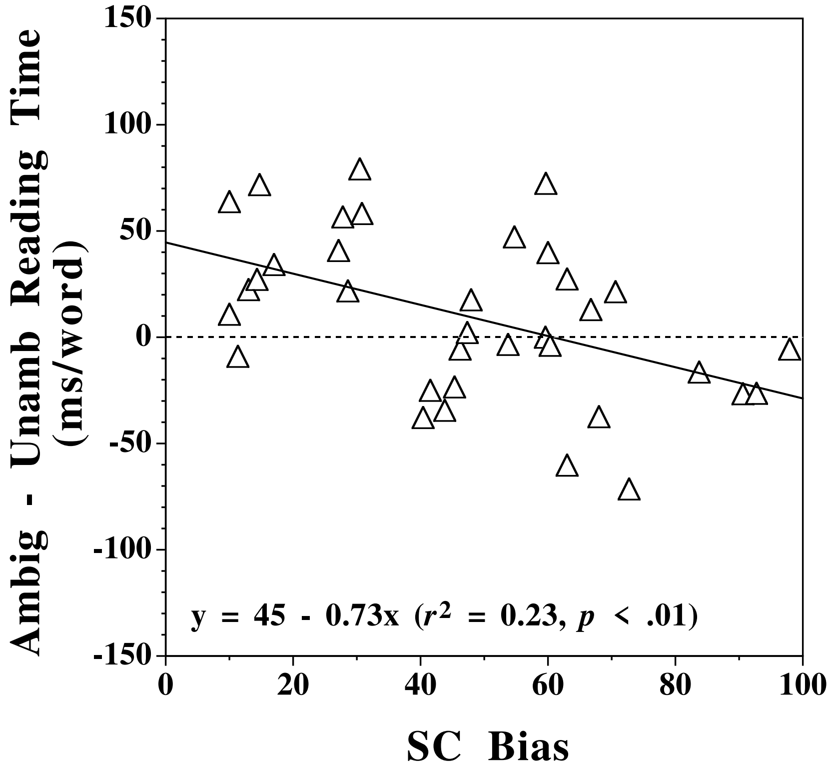

Figure 2: Experiment 1 correlation across items between SC bias and residual reading time ambiguity effect (ambiguous − unambiguous) at disambiguation in the plausible conditions.

correlate reliably with the log-transformed frequency (Francis & Kučera, 1982) of the embedded verb ($r = .35$, $p < .05$).

An additional alternative is that the pattern of correlations might result from effects in the unambiguous controls rather than in the ambiguous stimuli. This was also not the case, as the correlation at disambiguation in the plausible conditions was reliable for the ambiguous stimuli ($r = .38$, $p < .05$) but not the unambiguous stimuli ($r = .01$), whereas the correlation at the embedded verb in the implausible conditions was not reliable for either (ambiguous: $r = .01$; unambiguous: $r = .10$). The implausible condition correlation with embedded verb log frequency also arose from the ambiguous rather than the unambiguous conditions (ambiguous: $r = .34$, $p < .05$; unambiguous: $r = .01$).

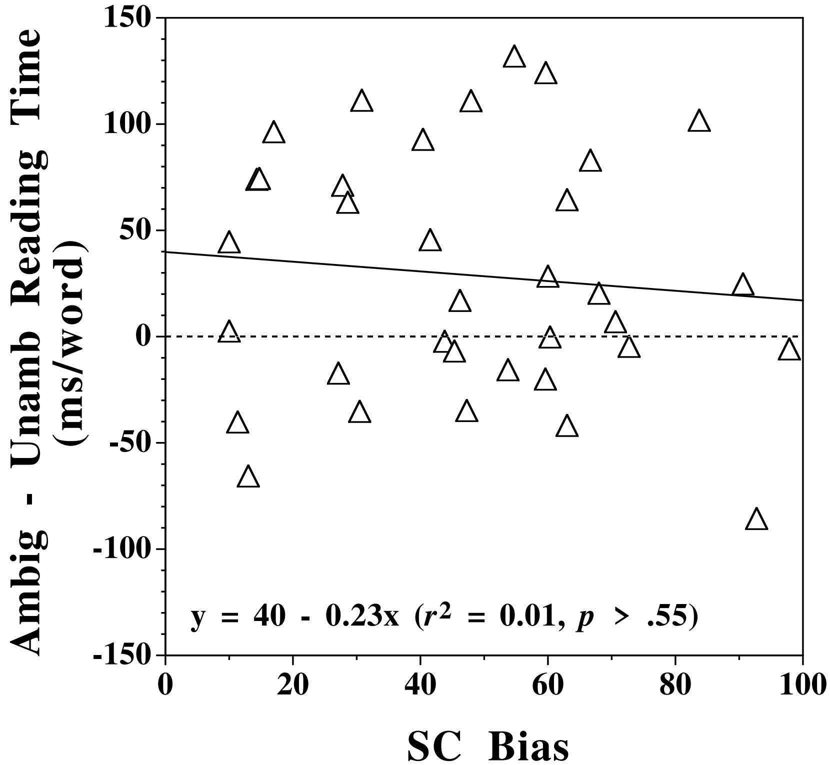

Figure 3: Experiment 1 correlation across items between SC bias and residual reading time ambiguity effect (ambiguous − unambiguous) at the embedded verb in the implausible conditions.

Discussion

The pattern of correlations in Figures 2 and 3 provides evidence against both probabilistic serial and competitive ranked-parallel models. As described above, these models predict that SC bias should correlate with the disambiguation ambiguity effect (in Figure 2) and the embedded verb ambiguity effect (in Figure 3) to about the same degree, but this was not the case: SC bias correlated with the disambiguation ambiguity effect, but it did not correlate with the embedded verb ambiguity effect. The critical lack of a correlation in the latter case could not be the result of a lack of variance to be accounted for in the embedded verb ambiguity effect variable, nor could it be the result of unrelated interference from the unambiguous conditions involved in the ambiguity effects.

This pattern does match that predicted by a non-competitive ranked-parallel model, in which the strengths of the two alternatives in an ambiguity can vary independently, and it therefore suggests (1) that, at least in some situations, the human sentence processing system does handle ambiguities by keeping track of multiple possible interpretations simultaneously (in parallel), and (2) that the interpretations do not necessarily compete with each other but can instead be maintained relatively independently. We will consider these points further and discuss how such a model might operate in the General Discussion. First, however, it is important to recognize that these conclusions rely on the assumption that the sentence complement is the preferred interpretation for the ambiguity we have been examining. If instead the relative clause interpretation were preferred in the Experiment 1 stimuli, both serial and ranked-parallel models would be able to account for the results straightforwardly. Experiment 2 was therefore conducted to test for the preferred interpretation of the ambiguity.

Experiment 2

For the Experiment 1 results to provide evidence about parallelism, comprehenders' preference for our stimuli must be for the sentence complement interpretation over the relative clause interpretation. Many sentence comprehension theories predict this preference, either because of minimal attachment[1] (e.g., Frazier, 1979, 1987), because sentence complements are arguments whereas relative clauses are adjuncts (e.g., Abney, 1989; Pritchett, 1988; also Frazier & Clifton, 1996: their primary/non-primary distinction is equivalent to an argument/adjunct one for this ambiguity), or because the relative clause interpretation creates more discourse complexity (e.g., Altmann et al., 1992; Altmann & Steedman, 1988; Crain & Steedman, 1985). Lexical constraint-based theories (e.g., MacDonald et al., 1994; Trueswell & Tanenhaus, 1994) do not make an overall prediction about the preferred interpretation for the ambiguity, but they instead predict that the preferred interpretation will be driven by lexical properties, and particularly by how often the initial noun occurs with a sentence complement. The correlation in Experiment 1B, between SC bias and plau-

[1] On Frazier's (1990) assumptions, argument or modifier attachment (as in the sentence complement interpretation) and true adjunction (as in the relative clause interpretation) into a simple NP such as *the claim* each require the addition of a single node to the NP's structure. In the case of arguments and modifiers, an intermediate node must be inserted between N and NP. In the case of adjunction, an additional NP is required (e.g., Chomsky, 1981). However, sentence complements and relative clauses also differ in internal clausal structure, in that relative clauses require a specifier position for the relative pronoun or operator, whereas complementizers head sentence complements and require no specifier position (see, e.g., Chomsky, 1986). Minimal attachment will thus favor the sentence complement interpretation.

sible condition ambiguity effect size at the disambiguation (see Figure 2), supported this prediction, but it is still important to show that the sentence complement is preferred overall for these stimuli, if the Experiment 1 results are to provide evidence about the handling of a secondary interpretation for an ambiguity.

To our knowledge, despite the various theoretical predictions, this ambiguity has not previously been examined experimentally. Thus in order to see whether the relative clause is non-preferred, we had comprehenders read examples of the ambiguity from Experiment 1 resolved either with the sentence complement interpretation or with the relative clause interpretation. The plausible conditions from Experiment 1 were used again, and they were compared to (plausible) relative clause conditions created by dropping the direct object from the plausible condition embedded clauses.

Method

Participants. Eighty-four Northeastern University undergraduates participated. To balance the number of participants across lists, 4 participants with poor comprehension performance (below 76% overall) were excluded, leaving 80 participants.

Materials and Design. The stimulus sets in this experiment were constructed from those in Experiment 1 by dropping the two implausible conditions and replacing them with two conditions (ambiguous and unambiguous) in which the ambiguity was resolved with the relative clause interpretation, as shown in (8).

(8) a. The claim that the cop ignored the informant might have affected the jury.

b. The claim alleging that the cop ignored the informant might have affected the jury.

c. The claim that the cop ignored might have affected the jury.

d. The claim which the cop ignored might have affected the jury.

The first two versions in (8) are the same sentence-complement-resolved stimuli as in (4), but the second two versions are resolved with the relative clause interpretation, which does not include the direct object following the embedded verb (*the informant* in (8a) and (8b)). The unambiguous relative clause version (8d) was always created by inserting *which* in place of *that*. Because *which* cannot be a complementizer and must be a relative pronoun, it forces the relative clause interpretation. The plausible embedded verb was used in all four versions, so both interpretations were plausible prior to disambiguation. As in Experiment 1, the direct object disambiguated the sentence complement interpretation. The matrix verb immediately following the embedded verb disambiguated the relative clause interpretation, because the embedded verbs were transitively biased (see Method section of Experiment 1A). The relative clause versions of the stimuli can be derived from Appendix A.

The 36 experimental stimuli were combined with 64 fillers to form four 100-item lists, as in Experiment 1A. The fillers had a variety of different syntactic structures, including 26 items which began with sentence-complement-taking nouns like those in the experimental items but did not continue with a similar embedded clause. Each experimental stimulus and filler also had a Yes/No comprehension question. Unlike in Experiment 1, the question did ask about the ambiguity for nearly all of the experimental items. For example, for (8), the comprehension question was *Did the police officer ignore the claim?*, which should have been answered No for the sentence complement versions (8a and 8b), but Yes for the relative clause versions (8c and 8d).

Apparatus and Procedure. These were identical to those in Experiment 1A.

Results

Comprehension Question Performance. Comprehension performance was 72% correct in both sentence complement conditions, 75% correct in the ambiguous relative clause condition, and 81% correct in the unambiguous relative clause condition. This pattern yielded reliable effects in the analysis by participants, but patterns in the analysis by items were marginal or non-significant (interpretation main effect: $F_1(1,79) = 16.05$, $MS_e = 199$, $p < .001$; $F_2(1,35) = 1.21$, $MS_e = 1111$, $p > .25$; ambiguity main effect: $F_1(1,79) = 6.06$, $MS_e = 141$, $p < .05$; $F_2(1,35) = 3.84$, $MS_e = 88$, $p < .10$; interaction: $F_1(1,79) = 5.53$, $MS_e = 154$, $p < .05$; $F_2(1,35) = 3.28$, $MS_e = 103$, $p < .10$). Paired comparisons revealed that the two sentence complement conditions did not differ (Fs < 1), but the two relative clause conditions did ($F_1(1,79) = 14.43$, $MS_e = 118$, $p < .001$; $F_2(1,35) = 5.27$, $MS_e = 128$, $p < .05$).

Reading Times. Residual times were computed and trimmed at 2.5 *SD*, affecting less than 2.3% of the data, as in Experiment 1. Trials for which the comprehension question was answered incor-

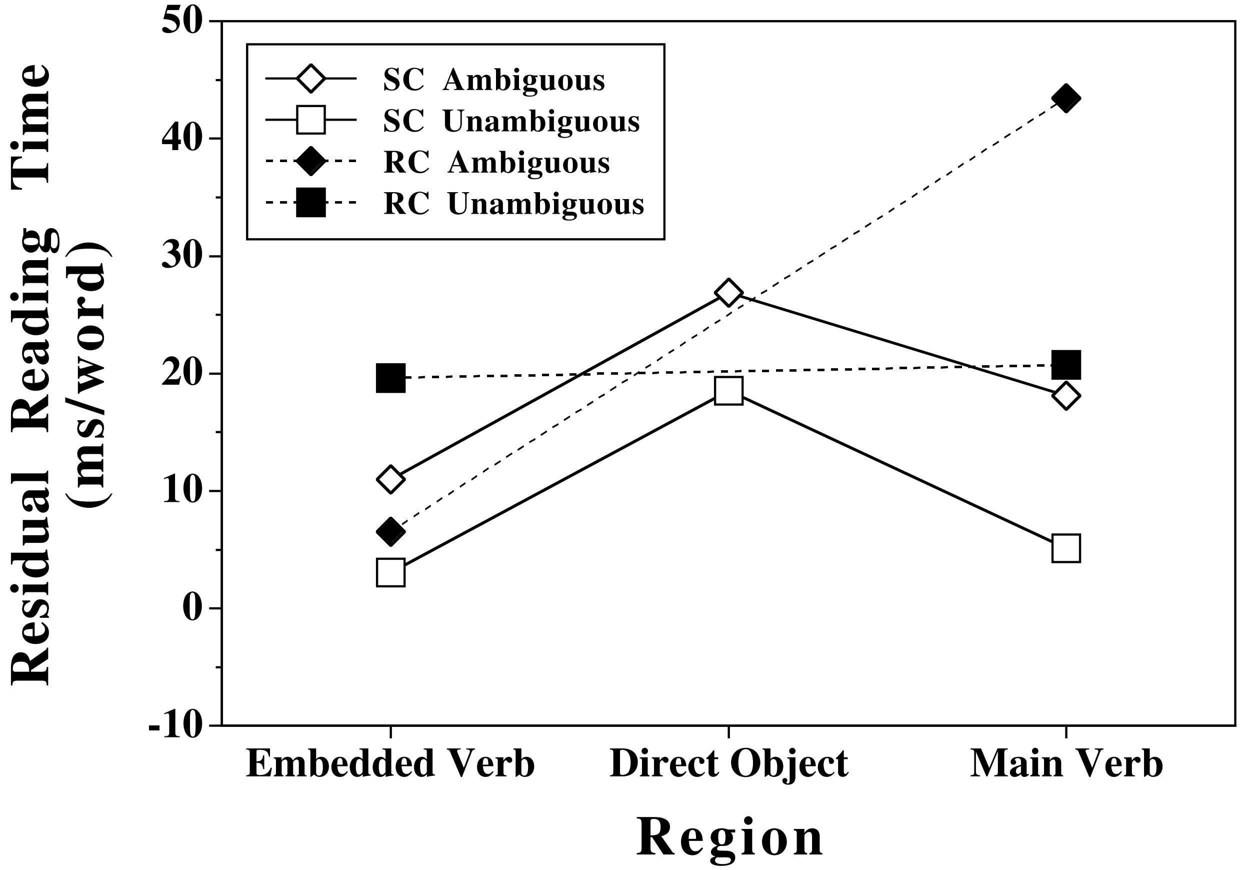

Figure 4: Experiment 2 residual reading time per word by condition and region.

rectly were also dropped. Figure 4 shows residual reading times by condition for the embedded verb, the direct object (present only for the sentence complement conditions, where it forces that interpretation), and the first two words of the main verb, which disambiguate the relative clause versions. Correspondingly-trimmed raw times are shown in Appendix B.

At the embedded verb, there were no reliable main effects (Fs < 2, ps $> .20$), nor was there an interaction (Fs < 3, ps $> .10$). Similarly, at the direct object, the two sentence complement conditions did not differ (Fs < 1).

At the first two words of the main verb phrase, main effects of interpretation ($F_1(1,79) = 16.63$, $MS_e = 2190$, $p < .001$; $F_2(1,35) = 5.76$, $MS_e = 2948$, $p < .05$) and ambiguity ($F_1(1,79) = 8.71$, $MS_e = 3202$, $p < .01$; $F_2(1,35) = 9.95$, $MS_e = 1139$, $p < .01$ were present. There was no interaction (Fs < 2, ps $> .25$), because the ambiguity effect for both the sentence complement and relative clause interpretations was in the same direction. However, the relative size of these ambiguity effects differed, as the comparison for the sentence complement conditions was not reliable ($F_1(1, 79) = 1.87$, $p = .17$; $F_2(1, 35) = 2.47$, $p = .12$), whereas that for the relative clause conditions was ($F_1(1, 79) = 8.88$, $p < .01$; $F_2(1, 35) = 5.93$, $p < .05$).

Discussion

Following the predictions of most theories, readers displayed difficulty relative to an unambiguous control when the relative clause interpretation was required at disambiguation (the main verb region), and they did not show difficulty relative to an unambiguous control when the sentence complement interpretation was forced, either at the direct object or in the following main verb region. This same pattern was visible in comprehension performance as well: Performance was not affected by ambiguity in the two sentence complement conditions; but in the relative clause conditions, performance was better in the unambiguous than in the ambiguous condition. (Although comprehension performance tended to be lower overall in the sentence complement conditions, this difference was confounded with differences in the questions used for the two interpretations. The questions for the sentence complement interpretations may simply have been harder overall.) The pattern in both reading times and comprehension performance thus indicates that the sentence complement is generally the preferred interpretation for the sentence complement versus relative clause ambiguity, at least for the specific set of stimuli in the current experiments.

General Discussion

These experiments examined an ambiguity not previously considered in any detail in the literature, the noun-triggered sentence complement versus relative clause ambiguity, and used it to investigate whether the human sentence comprehension system can maintain multiple interpretations for an ambiguity in parallel, or whether it is instead limited to considering only a single possibility at a time. Experiment 2 indicated that comprehenders overall preferred the sentence complement interpretation to the relative clause interpretation: Readers displayed significant difficulty in resolving the ambiguity toward the relative clause, but they showed no reliable difficulty in reading the sentence complement resolution. Experiment 1A manipulated the plausibility of the non-preferred (relative clause) interpretation while holding constant the plausibility of the preferred (sentence complement) interpretation. The sentence complement possibility thus was preferred initially, was plausible throughout, and was forced at disambiguation. Readers nevertheless showed effects of the plausibility of the non-preferred interpretation during the ambiguity: Compared to controls, readers slowed down when the relative clause possibility became implausible, whereas they did not slow down in the cases where the relative clause remained plausible. Experiment 1B further examined these effects of plausibility by considering the influence of argument structure bias (SC bias, the preference to be followed by a sentence complement over a relative clause, associated with each ambiguity-triggering noun). SC bias did predict the amount of difficulty (ambiguous versus unambiguous) readers displayed in reading the disambiguation toward the sentence complement interpretation in Experiment 1A, when both interpretations remained plausible throughout the ambiguity. However, it did not predict the amount of difficulty with the implausibility of the relative clause possibility at the embedded verb within the ambiguity (at the point where the relative clause became implausible).

To apply these results to the serial versus parallel issue, we divided serial and parallel models into two classes each. Serial models can be either deterministic or probabilistic. In deterministic serial models (e.g., Frazier, 1979, 1987; Frazier & Rayner, 1982; Frazier & Clifton, 1996, for the current ambiguity), the comprehension system always selects the same interpretation for a given sequence of grammatical categories, regardless of argument structure frequency biases and other probabilistic sources of information. The preferred interpretation for the current ambiguity is the sentence complement alternative. Thus deterministic serial models should select the sentence complement initially for the ambiguity, and because in Experiment 1 the sentence complement remains plausible throughout and turns out to be correct, no reanalysis should be necessary. The relative clause alternative should therefore never be considered and its (im)plausibility should have no influence on processing. Deterministic serial models thus cannot account for the effect of the relative clause's implausibility in Experiment 1A.

For probabilistic serial models (e.g., Ferreira & Henderson, 1990; Mitchell, 1987, 1989; van Gompel et al., 1999), just as for deterministic ones, the results of Experiment 2 indicate that the sentence complement interpretation is preferred. However, probabilistic serial models are more flexible than deterministic ones in that they are not limited to always selecting the same alternative for a given structural ambiguity. In particular, they can account for the overall preference for the sentence complement interpretation (Experiment 2) as well as the small ambiguity effect at the disambiguation toward the sentence complement (unreliable in Experiment 2, reliable as a main effect but not in paired comparisons in Experiment 1A) by assuming that the sentence complement is usually, but not always, the selected alternative. This assumption also accounts for the effect of the relative clause interpretation's implausibility in Experiment 1A: The effect arises from the minority of trials in which the relative clause is chosen instead of the sentence complement. The pattern of correlations found in Experiment 1B, however, presents a problem for probabilistic serial models. These models are compatible with the effect of SC bias on difficulty at the disambiguation toward the sentence complement interpretation. However, because the sentence complement and relative clause interpretations are necessarily in complementary distribution in such models, they also predict a corresponding effect of SC bias on the degree to which the implausibility of the relative clause causes difficulty, and this effect (measured at the embedded verb) was not present.[2] As a result, neither deterministic nor probabilistic serial models appear compatible with the full set of results.

To see the implications of these results for ranked-parallel models, we divided them into two classes as well: competitive and non-competitive. In competitive ranked-parallel models (e.g., Boland, 1997; Gibson, 1991; Jurafsky, 1996; Just & Carpenter, 1992; MacDonald et al., 1994; Spivey & Tanenhaus, 1998; Stevenson, 1994; Tabor et al., 1997), multiple interpretations can be maintained, but usually one is preferred, and, critically, as support for one interpretation increases, support for other interpretations decreases. The source of this trade-off in support varies with the model (see discussion below). Such models can account for most of the Experiment 1 and 2 results because they predict that the sentence complement interpretation is typically preferred, that this preference varies at least in part as a function of SC bias, and that the relative clause

[2]Lewis (in press) suggests that a probabilistic serial model might account for the difference between the SC bias correlations if the correlation at the disambiguating region reflects ease of reanalysis rather than likelihood of selecting the sentence complement interpretation prior to disambiguation — no correlation would then be expected at the embedded verb. However, this model requires that argument structure frequency biases only influence processing when reanalysis is triggered for other reasons, and it will therefore fail to account for a variety of results indicating that such biases influence processing within at most a word or two of their theoretically-earliest availability (e.g., Boland et al., 1995; Ferreira & Henderson, 1990; Garnsey et al., 1997, Holmes et al., 1989; McElree & Griffith, 1995; Mitchell, 1989; Schmauder & Egan, 1998; Trueswell & Kim, 1998; Trueswell et al., 1993).

interpretation is also maintained: The overall preference for the sentence complement yields the greater difficulty for the relative clause resolution in Experiment 2, and the small ambiguity effect at disambiguation for the sentence complement interpretation arises because the preference for the sentence complement is not absolute. Because the relative clause interpretation is maintained as a secondary alternative, the system will be able to detect its implausibility, accounting for the Experiment 1A results. However, competitive ranked-parallel models have the same difficulty as probabilistic serial models in handling the correlations from Experiment 1B, and for essentially the same reason: The effect of the competition between alternatives in a parallel model is the same as the enforcement of a selection between alternatives in a serial model. To the extent that one alternative gains, others lose. Thus a competitive ranked-parallel model straightforwardly accounts for the correlation of SC bias with difficulty at the sentence complement disambiguation, but like a probabilistic serial model, it predicts the same correlation between SC bias and difficulty at the embedded verb due to the relative clause's implausibility.

Unlike serial models, however, ranked-parallel models do not necessarily require the alternatives for an ambiguity to be in complementary distribution. Although purely competitive models create this complementarity effect, degree of competition can be treated as a continuous property (e.g., across ambiguities), and models in which multiple (ranked) alternatives are maintained in parallel and do not compete against each other for support are thus possible. Such a model decouples the two SC bias correlations in Experiment 1B: SC bias can control the support for the sentence complement interpretation (thus correlating with how difficult it is to end up with that interpretation at its disambiguation) without controlling the support for the relative clause interpretation (predicting no substantial correlation with difficulty at the embedded verb due to the relative clause's implausibility). A non-competitive model will account for the other Experiment 1 and 2 results just as a competitive ranked-parallel model will. Thus the overall pattern of results suggests that the human sentence comprehension system does maintain multiple interpretations in parallel, at least in some situations, and that the interpretations do not necessarily compete with each other for support.

Although these results are compatible with non-competitive ranked-parallel models as a class, this leaves unspecified the details of how or why a ranked-parallel model might be non-competitive. Three other related issues also deserve attention, including the relationship between competitive and non-competitive models, the effect of handling multiple simultaneous interpretations on processing difficulty, and the effect of implausibility on the maintenance of an ambiguity's interpretations. We discuss these questions below, beginning with some further detail about our proposed model.

Compatibility and Competition

The conclusion that competitive ranked-parallel models cannot account for the Experiment 1B pattern is of some concern given that all current parallel models rely on either direct or indirect competition in handling alternatives for an ambiguity. In direct competition cases (e.g., MacDonald et al., 1994; Spivey & Tanenhaus, 1998; Stevenson, 1994), ambiguities are resolved by interactive activation processes, so alternatives effectively attempt to turn each other off until one of them wins, either by reaching a threshold or by eliminating the other alternatives. In indirect cases (e.g., Gibson, 1991; Jurafsky, 1996; Just & Carpenter, 1992), capacity limitations create competition between alternatives, and only the most strongly supported possibilities will be retained. Although the alternatives do not directly inhibit each other, increasing support for one alternative will tend to result in decreased support for other alternatives, as resources become scarcer. In addition, recent empirical results suggest that alternatives do compete with each other in at least some cases of ambiguity resolution (e.g., McRae, Spivey-Knowlton, & Tanenhaus, 1998; Richardson & Spivey, 1999; Spivey & Tanenhaus, 1998; cf. Traxler, Pickering, & Clifton, 1998; van Gompel et al., 1999).

The contrast between our results and the existing evidence for competition can be explained by a proposed additional property of ranked-parallel models: *compatibility*, which controls the degree to which alternatives for an ambiguity compete, such that highly compatible alternatives will tend not to. Compatibility between alternatives varies across ambiguities and can be conceived of as a ratio of the amount of overlap in the representations for the alternatives, relative to the total representational material needed.[3] Whether the representations involved are lexical (e.g.,

[3]Compatibility may in fact depend on the degree to which elements of the representations for the alternatives specifically conflict with each other, rather than just on the amount of extra material needed in one or the other representation. We adopt this latter, extra-material, view, but it does not differentiate between cases in which the extra material needed for one alternative is simply left out of the second alternative, and cases in which the extra material in one alternative specifically conflicts with the material needed in the other alternative. We therefore note conflicting versus non-conflicting elements where relevant. Additional data will be needed to distinguish these possibilities.

grammatical categories, agreement features, semantic features), syntactic (e.g., chunks of phrase structure), interpretive components (e.g., propositions, mental model components), or some combination of these is an open question. We will focus on syntactic and semantic/discourse interpretive representations, as they appear to be most relevant for the current ambiguity. However, if syntactic and lexical ambiguities involve the same processing mechanisms, as lexicalist constraint-based approaches suggest, we would expect that compatibility of lexical representations will be relevant as well.

For the noun-triggered sentence complement versus relative clause ambiguity, the proportion of shared representational elements for the sentence complement and relative clause is quite high while the embedded verb is being processed: In terms of syntactic structure, both alternatives share the same subtree for the embedded clause subject (*the cop* in (9)), the structure connecting the embedded clause subject to the embedded verb head (*ignored*) and its verb phrase, and possibly a postulated direct object position (given that our embedded verbs were strongly transitively biased). In addition, most of the clausal structure required to support a finite embedded clause (constructed while *that* and perhaps the next word are processed) will also be shared, although the details of this structure vary across syntactic theories. In terms of interpretive components, the representation of the embedded subject and its relationship to the embedded verb, as well as the appropriate senses of the words in the ambiguity, are all shared between the representations.

(9) The claim that the cop ignored…

In contrast to the overlapping material, there are relatively few representational components needed for only one of the alternatives. Prior to disambiguation, the two alternatives differ only in the relationship between the matrix subject head (*claim*) and the embedded clause itself, which in the syntax may involve no more than the addition of an operator position, needed at the beginning of the representation for the relative clause but not needed for the sentence complement. In some syntactic theories, the details of the structure for the attachment of the clause into the matrix subject noun phrase also vary across the two alternatives (e.g., Chomsky, 1981, 1986). In terms of interpretive components, the relative clause may involve additional discourse complexity, expressed as the need for a set of claims from among which the relative clause can select (Altmann & Steedman, 1988; Crain & Steedman, 1985). Thus although there are (necessarily) some differences between the representations involved in the two alternatives for the ambiguity, these differences are relatively limited compared to the amount of overlap between the two alternatives. It is also worth noting that with the exception of the possible differences in internal matrix subject noun phrase structure, none of the extra elements used in only one alternative directly conflict with anything in the representation for the other alternative. Thus, for example, the extra operator position required by the relative clause to control the interpretation of the direct object gap is an addition to the clausal structures that are required in both alternatives; there is no conflicting node attached at that point in the sentence complement alternative. Likewise, if the relative clause interpretation for (9) requires the postulation of a set of claims over which the relative clause can restrict, this set is an addition to what is required by the sentence complement interpretation (a single claim) and does not conflict with that alternative's discourse representation.

The degree of compatibility in this ambiguity can be contrasted with that in other ambiguities commonly examined in the sentence processing literature. For example, the two alternatives have a much smaller amount of overlap in the main verb versus reduced relative ambiguity (e.g., *The horse raced past the barn (fell).*), which Spivey and colleagues (McRae et al., 1998; Richardson & Spivey, 1999; Spivey & Tanenhaus, 1998) have used to argue for a competition-based ambiguity resolution mechanism. The relevant sense of the ambiguity-triggering verb in the two cases is shared (but cf. approaches to the passive like Pinker, 1989), and in some versions of the ambiguity, the prepositional phrase following the verb (e.g., *past the barn*) receives the same structural and semantic interpretation in the different alternatives (e.g., MacDonald, 1994; Pearlmutter & MacDonald, 1995). However, a variety of the triggering verb's (e.g., *raced*) lexical properties directly conflict with each other (MacDonald et al., 1994): The main verb interpretation requires active voice, whereas the reduced relative requires passive voice; the main verb requires a past tense verb form, whereas the reduced relative requires a past participle form; and in some cases, the necessary argument structures vary (e.g., MacDonald, 1994; Pearlmutter & MacDonald, 1995). Furthermore, in the syntax, the reduced relative interpretation requires clausal structure which is not needed for the main verb possibility, and it requires a passive verb phrase, which conflicts with the active structure required for the main verb interpretation. The semantic/discourse elements needed for the two interpretations are also substantially different: Because of the active/passive difference, the relationship between the subject noun phrase (*the horse*) and the verb is typically conflicting in the two alternatives (e.g., Trueswell et al., 1994; cf. Pearlmutter & MacDonald, 1995), and as with the relative clause interpretation in the ambiguity in the present experiments, the reduced relative clause interpretation involves additional discourse elements not needed for the main verb interpretation. Thus this ambiguity will generally involve much less compatibility between its alternatives than the ambiguity we have examined, and as a result, its alternatives will compete much more strongly.

Other ambiguities investigated in the literature are also likely to display less compatibility between alternatives than the current one: In the direct object versus sentence complement ambiguity (e.g., *John knew the answer.* vs *John knew the answer was wrong.*), the syntactic and semantic/discourse representations for the ambiguous noun phrase (e.g., *the answer*) will be shared, but the required argument structures for the ambiguity-triggering verb (*knew*) will be incompatible, as will the required semantic relationships between the verb and the ambiguous noun phrase, and none of the clausal syntactic material needed for the sentence complement interpretation will be used in the direct object interpretation. Although the degree of overlap between the alternatives in this ambiguity will be greater than in the main verb versus reduced relative case, it will still clearly be less than in the sentence complement versus relative clause case, and thus it is likely to show stronger effects of competition between alternatives. Prepositional phrase attachment ambiguities (e.g., *The girl hit the boy with the bookbag.*) will be similarly intermediate in degree of competition: The internal structure of the prepositional phrase (e.g., *with the bookbag*) will be shared between the two alternatives, but the semantic relationships among the prepositional phrase, the verb (*hit*), and the object noun phrase (*the boy*), as well as the syntax required to attach the prepositional phrase into the earlier structure, will differ between the two alternatives for the ambiguity.

This consideration of compatibility in ambiguity resolution thus reveals that work on sentence-level ambiguity resolution has focused on cases in which the alternatives tend to compete with each other. These cases are the ones which are most likely to yield patterns of results suggesting serial processing, because they are cases in which the preferred alternative (regardless of the reason it is preferred) will tend to eliminate other alternatives. In the noun-triggered sentence complement versus relative clause ambiguity we have examined, on the other hand, the alternatives are much less likely to compete. Thus compatibility can explain why we were able to find results supporting a non-competitive ranked-parallel model in this case, while competition between alternatives appears to be an important component in the handling of many other ambiguities.

Compatibility in a Ranked-Parallel Processor

Although the presentation of compatibility above may seem ad hoc, it arises naturally in a processing system which constructs representations (e.g., syntactic trees, semantic/discourse representations) within an activation framework, and has two additional important properties: First, individual parts of representations have potentially independent activation levels which can decay over time, so that while one part of a syntactic tree (for example) is highly active, another portion of the same tree might be only weakly active. This is similar to the approach of Stevenson (1994) for syntax processing and of Kintsch and colleagues (e.g., Kintsch, 1988; Kintsch, Welsch, Schmalhofer, & Zimny, 1990) and Dell, McKoon, and Ratcliff (1983) for discourse processing, with an activation decay function associated with each portion of a representation, though without enforced competition. An activation decay function in such a system would also straightforwardly yield recency effects (e.g., Gibson, 1998; Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; Pearlmutter, in press; Pearlmutter & Gibson, 1999; Stevenson, 1994), assuming that attachment is easier to more active representations.

The second important property of the system we are proposing is that it shares representational components among the alternatives for an ambiguity, so while it maintains multiple alternatives in parallel, the alternatives are not independent representations. They are linked to the extent that they share components. Thus the representational components needed for both alternatives for an ambiguity (e.g., the syntactic subtree for the embedded subject noun phrase in the sentence complement vs relative clause case) are only represented once. It is this property in particular which yields compatibility, because for ambiguities with a high percentage of material overlapping between the alternatives, factors which increase the activation of one alternative will also tend to increase the activation of (at least parts of) the other alternative. Thus as alternatives share more and more material, they will tend to reinforce each other rather than competing.

While this structure-sharing property is critical in accounting for the current results, it is also the most novel component of this model. Other proposed activation-based models (e.g., Stevenson, 1994; Spivey & Tanenhaus, 1998) rely extensively on direct competition for ambiguity resolution, with the result that the processor will settle, either immediately or within a few words, on a single, strongly preferred alternative for the ambiguity. Thus the assumption is that the driving force behind ambiguity resolution is the pressure to resolve conflicts between mutually exclusive alternatives in favor of whichever alternative is most strongly supported by the available evidence, as it accrues over time. On our view, the pressure to resolve conflicts is clearly still present and still drives processing, but the processor can tolerate a limited degree of inconsistency across the network of activated representations. The degree to which active conflicting alternatives can be sustained may be a fixed universal property of the human processor, or it may be subject to variation across languages or individuals (e.g., Gibson et al., 1996; Just & Carpenter, 1992; MacDonald et al., 1992; Pearlmutter & MacDonald, 1995), but, critically for our account, it requires enough overlap (and therefore shared representations) between the alternatives to offset the conflicting material. A more exact specification will require further investigation, both in terms of theoretical modeling and empirical investigation. As noted above, it may turn out that specifically conflicting material creates competition even if a large amount of representational overlap is also present; the processor's ability to tolerate multiple non-competing alternatives may be most apparent only in cases in which the representational material needed for one but not the other alternative has no competing component (e.g., the extra structure required for an operator position in the relative clause alternative, which conflicts with nothing in the sentence complement representation).

Although some details obviously remain to be firmly established, this kind of ranked-parallel processor, with shared, independently-activated representations, helps to provide an account of two other issues raised by the current work: (1) the relationship between working memory capacity, degree of parallelism, and reading difficulty; and (2) the effect of implausibility on the maintenance of representations for an ambiguity.

The first issue, relating working memory, parallelism, and reading difficulty, arises because the pattern of results in Experiment 1A (and Experiment 2) indicates that the processor's consideration of multiple alternatives during the ambiguity was not by itself sufficient to cause a slowdown. That is, Experiment 1A reading times at the embedded verb in the plausible ambiguous condition did not differ from those in the corresponding unambiguous control. This fits with the great majority of other results in the literature showing that reading times during an ambiguity are not elevated over those to the same region of an unambiguous control (e.g., Altmann et al., 1992; Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Trueswell et al., 1994). Frazier and Rayner (1982; see also Lewis, in press) suggested that this pattern was evidence for a serial over a parallel system, but there are several possible accounts of such a result within a parallel framework. Gibson and Pearlmutter (in press) provide one: Reading time might not be directly related to available memory or processing capacity, or to the number of representations being considered. Comprehenders might simply maintain as many interpretations in parallel as possible, up to some resource capacity, without cost, pruning alternatives when necessary to stay within the capacity.

Our proposal suggests an alternative possibility, which is that for highly compatible alternatives, because of structure-sharing, the cost associated with maintaining a secondary interpretation should be quite minimal, resulting in little or no additional difficulty. In our stimuli, for example, most of the processing necessary to maintain the relative clause alternative will have to be done anyway to maintain the sentence complement alternative. Structure-sharing prevents it from having to be done twice. For more competitive cases, the alternatives will trade off against each other, so that as one alternative draws more resources, other alternatives will tend to draw fewer, with the result that overall resource use will not be greatly increased. On this account, a substantially increased demand for computational resources might still slow down processing, as might cases in which two alternatives are incompatible but roughly equally supported (the traditional cases of competition-induced slowdowns). Parallel consideration of highly compatible alternatives would not itself create much difficulty, however.

The other issue raised by these results is the effect of implausibility on the maintenance of an ambiguity's interpretations. The central question is why the implausibility of the relative clause alternative in Experiment 1A resulted in a slowdown in reading the embedded verb. An obvious alternative outcome is that the processing system could simply have noted that the non-preferred interpretation was implausible and dropped it. That this did not happen is indicated by a comparison of the plausible condition and implausible condition ambiguity effects at the disambiguation in Experiment 1A. There was no interaction at the disambiguation, and the two ambiguity effects were nearly identical; if the implausibility of the relative clause had caused it to be dropped, readers should have had less difficulty at disambiguation than in the plausible conditions, because the implausibility of the relative clause would have provided disambiguation in advance.

To explain this, we propose a straightforward extension, to ambiguous structures handled in parallel, of Murray and Rowan's (1998) claim about unambiguous cases: Computation of plausibility is fast and mandatory (see also, e.g., Garnsey et al., 1997; Marslen-Wilson & Tyler, 1987;

Pickering & Traxler, 1998). When multiple interpretations are being considered simultaneously, the processor necessarily attempts to construct semantic representations for each of them. Furthermore, the computation of a semantic representation itself varies in duration with plausibility: Constructing a representation for an implausible input is more difficult than constructing one for a plausible input. Thus the delay at the embedded verb in the implausible conditions of Experiment 1A results from the attempt to interpret the relative clause possibility rather than from the action taken by the processor once it determines that the relative clause is relatively implausible (though nothing prevents such actions from having an effect as well).

Of course, this still leaves open the question of why the implausibility of the relative clause did not make it easier to resolve the ambiguity at the disambiguation, as has been seen with other ambiguities (e.g., Garnsey et al., 1997; McRae et al., 1998; Pickering & Traxler, 1998; Trueswell et al., 1994). Structure-sharing provides one possible explanation here: Because there is so much overlap between the relative clause and sentence complement, the cost of maintaining the former in addition to the latter is minimal (indeed, across both Experiments 1 and 2, the size of the ambiguity effect at the disambiguation to the sentence complement is only 10 ms), so even when the relative clause becomes implausible, the processor might maintain it as a secondary possibility. However, two other possible explanations might apply: Because the disambiguating region occurs immediately after the embedded verb, either (1) the processor may simply not yet have had the chance to use the plausibility information to eliminate or weaken the relative clause alternative, or (2) the potential ease of handling the disambiguation may be offset by a spillover of implausibility-induced processing difficulty from the embedded verb. We are examining these possibilities in ongoing work by inserting additional material between the embedded verb and the disambiguation.

One additional point regarding the effect of the relative clause's implausibility is also worth noting: Our effect is perhaps surprising given the findings of Boland et al. (1995), who examined gap-postulation ambiguities in long-distance dependencies (e.g., *Which client did the salesman visit while in the city?*), where the processor can predict a direct object gap following the first verb (*visit*) or can guess that the gap will appear later (e.g., *Which client did the salesman visit with while in the city?*). Boland et al. used the detection of implausibility as a measure of whether a gap had been postulated or not, comparing cases like that above to cases in which the initial noun was implausible as a direct object (e.g., *Which prize did the salesman visit...*). When the verb was biased to be transitive, readers detected an anomaly at the verb more often in the implausible cases than in the plausible cases, indicating that they had postulated a direct object gap and had noticed that it resulted in an odd interpretation. However, with both object control verbs (e.g., *remind*) and dative verbs (e.g., *read*), whose argument structures include an additional constituent that can contain the gap (e.g., *Which child/movie did Mark remind them to watch?*), Boland et al. found no effect of plausibility on anomaly detection until the gap was forced to the direct object position. They also found that when the initial noun was plausible as a direct object, readers did postulate the direct object gap. Thus readers seemed to use both argument structure and plausibility to make decisions about gap-postulation, and readers apparently were not bothered by the implausibility of one possible gap as long as another possibility was available. This latter effect is the surprising one in comparison to the current results, because in our stimuli, readers slowed down in response to the implausibility of one alternative even though another plausible alternative was available (and generally preferred).

A variety of explanations, which are not mutually exclusive, might account for this difference: For example, Boland et al. (1995) used a self-paced word-by-word anomaly-detection task rather than a self-paced reading task. Readers were to press one button at a word which rendered a sentence implausible and a different button if the sentence remained plausible. In principle, readers might have adopted a strategy of pressing the "plausible" button as long as they could find at least one plausible interpretation, resulting in an effect of implausibility at the verb for the transitive verbs, where no other analysis was made available by the verb's argument structure, and no effect of implausibility for the other verbs, where their argument structures made available a still-plausible alternative.

Another possibility is that readers in our experiments may have had a stronger commitment to the interpretations for the ambiguity than in the Boland et al. case, because the ambiguity has already been maintained for several words before the relevant plausibility computation occurs. In the long-distance dependency cases, although the need for a gap (and uncertainty about its location) is present well before the verb, the verb is the first position where some differential evaluation of alternatives can be made. Thus the computation of plausibility in the implausible cases may never make much progress in the long-distance dependency construction, because there is no prior commitment to the implausible possibility and another plausible possibility is available. Although strength of prior commitment to an alternative has not previously received much direct attention in ambiguity resolution, it may turn out to be an important factor to consider in understanding parallel models involving continuous constraint satisfaction (e.g., McRae et al., 1998; Spivey & Tanenhaus, 1998).

In addition to these two possible explanations, compatibility and the effects of structure-sharing may again come into play here. In particular, the mandatoriness of plausibility computation described earlier is likely to be influenced by degree of compatibility: When alternatives are highly compatible, as in our stimuli, computing an interpretation for one alternative will result in much of the work of the computation already being done for the other alternative as well, so the cost of the additional computation for the secondary interpretation will be limited. In more competitive cases, on the other hand, the alternatives will be more distinct, and resources may be devoted more heavily to computations (and thus alternatives) which are making better progress. This should in turn result in less attention being paid to computing less plausible possibilities if a more promising competing alternative is available. This will likely be the case for Boland et al.'s (1995) long-distance dependency cases, because the alternatives involve little relevant structure-sharing: The alternative semantic relationships between the gap and the verb conflict, and, especially at the verb itself, where the syntactic representations have not yet been constructed, there is little or no overlapping material. Thus on this compatibility-based explanation, the processor will be more likely to thoroughly analyze even the relatively implausible secondary alternative in our stimuli, compared to the construction used by Boland et al. (1995).

While the details of some of the proposals we have discussed above still remain to be further established, the results of the experiments provide specific and strong evidence against serial models as a class, as well as against ranked-parallel models which rely purely on competition to determine rankings. Readers notice when a secondary interpretation becomes implausible, and the strength of support for the secondary interpretation is independent of the strength of support for the preferred interpretation. The combination of these results indicates that multiple interpretations can be simultaneously considered by the sentence processor for an ambiguity, and it suggests that the nature of processing is in part controlled by the degree to which alternatives overlap, or are compatible, with each other. We proposed an interactive-activation, constraint-based model of processing in which (1) components of representations can be activated, can decay, and can compete separately;

and (2) components which are shared among alternatives are represented only once. Such a model naturally accounts for a variety of results in the literature and continues to make use of competition between mutually exclusive possibilities as a mechanism for ambiguity resolution. However, it also delimits the range of influence of competition, suggesting that as the alternatives for an ambiguity increase in degree of representational overlap, competition will decrease, and the alternatives will tend to support each other instead of competing for computational resources. These influences of compatibility can be seen in the effects, both of working memory limitations and of interacting constraints such as plausibility, on the maintenance of interpretations during processing.

References

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, **18**, 129–144.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, **31**, 685–712.

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

Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, **12**, 423–484.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, **34**, 774–806.

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

Chomsky, N. (1986). *Barriers.* Cambridge, MA: MIT Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, **12**, 335–359.

Connine, C. M., Ferreira, F., Jones, C., Clifton, C., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, **13**, 307–319.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320–358). Cambridge, England: Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in comprehending and producing words in context. *Journal of Verbal Learning and Verbal Behavior*, **19**, 450–466.

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior*, **22**, 121–132.

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

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 555–568.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton-Mifflin.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Ph.D. dissertation, University of Connecticut. Bloomington, IN: Indiana University Linguistics Club.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Lawrence Erlbaum.

Frazier, L. (1990). Parsing modifiers: Special purpose routines in the human sentence processing mechanism? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 303–330). Hillsdale, NJ: Lawrence Erlbaum.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, **24**, 437–468.

Frazier, L., & Clifton, C., Jr. (1996). *Construal.* Cambridge, MA: MIT Press.

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

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, **37**, 58–93.

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

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, **68**, 1–76.

Gibson, E., & Pearlmutter, N. J. (in press). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research.*

Gibson, E., Pearlmutter, N. J., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, **59**, 23–59.

Gorrell, P. (1987). *Studies in human syntactic processing: Ranked-parallel versus serial models.* Unpublished Ph.D. dissertation, University of Connecticut, Storrs, CT.

Hickok, G. (1991). *Gaps and garden-paths: Studies on the architecture and computational machinery of the human sentence processor.* Unpublished Ph.D. dissertation, Brandeis University, Waltham, MA.

Holmes, V. M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language*, **28**, 668–689.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, **20**, 137–194.

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

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, **111**, 228–238.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, **95**, 163–182.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, **29**, 133–159.

Kurtzman, H. S. (1985). *Studies in syntactic ambiguity resolution.* Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Lewis, R. L. (in press). Serial and parallel parsing: Empirical conundrums and some overlooked evidence. *Journal of Psycholinguistic Research.*

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, **9**, 157–201.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, **24**, 56–98.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, **101**, 676–703.

Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 37–62). Cambridge, MA: MIT Press.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **21**, 134–157.

McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 1219–1228.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, **38**, 283–312.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 601–618). Hillsdale, NJ: Erlbaum.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, **4**, SI123–154.

Murray, W. S., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research*, **27**, 1–22.

Pearlmutter, N. J. (in press). Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research.*

Pearlmutter, N. J., & Gibson, E. (1999). Recency in verb phrase attachment. Manuscript submitted for publication.

Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, **34**, 521–542.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 940–961.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: MIT Press.

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

Richardson, D. C., & Spivey, M. J. (1999). On the resolution of syntactic alternatives in parallel: Race or arm wrestle? Paper presented at the 12th Annual CUNY Sentence Processing Conference, New York, NY.

Schmauder, A. R., & Egan, M. C. (1998). The influence of semantic fit on on-line sentence processing. *Memory & Cognition*, **26**, 1304–1312.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential complexity and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 1521–1543.

Stevenson, S. (1994). *A competitive attachment model for resolving syntactic ambiguities in natural language parsing* (Tech. Rep. RuCCS TR-18). New Brunswick, NJ: Rutgers Univ., Center for Cognitive Science.

Swinney, D., Ford, M., Frauenfelder, U., & Bresnan, J. (1988). On the temporal course of gap-filling and antecedent assignment during sentence comprehension. In B. Grosz, R. Kaplan, M. Macken, & I. Sag (Eds.), *Language structure and processing.* Stanford, CA: CSLI.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, **12**, 211–271.

Traxler, M. J., Pickering, M. J., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, **39**, 558–592.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, **39**, 102–123.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Towards a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, **33**, 285–318.

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**, 528–553.

van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J.. (1999). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. Manuscript submitted for publication.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **8**, 385–399.

West, R. F., & Stanovich, K. E. (1986). Robust effects of syntactic structure on visual word processing. *Memory & Cognition*, **14**, 104–112.

Wright, B., & Garrett, M. (1984). Lexical decision in sentences: Effects of syntactic structure. *Memory & Cognition*, **12**, 31–45.

Appendix A

The stimuli for both experiments are shown below. The word in parentheses in each item was included only in the unambiguous sentence complement versions. The plausible and implausible embedded verbs for each item are separated by a slash, and the two-word direct object (included only in the sentence complement versions) is in boldface. The unambiguous relative clause version was always created by replacing *that* with *which*. The values in parentheses after each item are, in order, the plausibility rating (1–7, 1 = least plausible) of the plausible sentence complement, the implausible sentence complement, the plausible relative clause, the implausible relative clause, the plausible subject-verb fragment, the implausible subject-verb fragment, the plausible unambiguous fragment, and the implausible unambiguous fragment; the transitivity bias (%) of the plausible and implausible embedded verbs; and the sentence complement bias (%) of the ambiguity-triggering noun. The stimulus set for which data were excluded from Experiment 1 (see Results) is marked with an asterisk.

1. The message (saying) that Bill's mother delivered/baked **the pizza** was surprising to his friends. (6.22, 6.22, 6.5, 1.56, 5.95, 6.15, 6.29, 6.12; 100, 94; 14)

2. The complaint (establishing) that the athlete ridiculed/offended **the umpire** was played up by the media. (6.39, 6.39, 4, 2.6, 5.05, 5.55, 4.68, 4.41; 100, 100; 40)

3. The reason (proving) that the comedian ridiculed/imitated **the newscaster** was not kept a secret. (6.33, 6.29, 4.11, 2.94, 5.45, 6.8, 3.46, 4.18; 100, 100; 93)

4. The revelation (showing) that the administrator belittled/fired **the employee** had to be completely ignored. (5.5, 6.56, 4.83, 2.28, 5.2, 5.6, 4.55, 4.45; 100, 96; 48)

5. *The perception (suggesting) that the photographer exploited/harassed **the model** was not an issue at the agency. (6.33, 6.06, 4.61, 2.11, 5.55, 5.4, 4.14, 3.77; 100, 100; 59)

6. The verification (confirming) that the ranger checked/fixed **the canoe** came over the new walkie-talkie. (5.56, 6.28, 5.28, 3.17; 5.1, 4.95, 5.45, 4.68; 75, 100; 60)

7. The conclusion (reporting) that the spy reached/left **the checkpoint** has not yet been confirmed. (5.61, 5.06, 5.34, 2.83, 4.45, 5.1, 5.41, 4.88; 100, 84; 44)
8. The pledge (saying) that the volunteers would obtain/prevent **the injunction** was a shock to the lobbyists. (3.83, 3.44, 5.39, 3.44, 4.9, 4.8, 5.27, 5.25; 100, 100; 17)
9. The belief (asserting) that the terrorists held/injured **the hostage** was depressing to the negotiators. (6.61, 6.61, 4.11, 2.28, 5.75, 6.1, 4.29, 4.52; 96, 100; 91)
10. The proof (confirming) that the arsonist hid/lit **the matches** was highlighted by the prosecution. (6.06, 6.67, 5.78, 4.44, 5.35, 6.25, 5.56, 5.95; 24, 100; 60)
11. The hope (suggesting) that the teenager encouraged/tutored **the boy** was shared by both parents. (5.89, 6.06, 4.22, 2.11, 5.25, 5.45, 3.14, 2.91; 100, 92; 84)
12. The allegation (stating) that the senator confirmed/attended **the meeting** was leaked to the press. (6.17, 6.67, 5.61, 2.94, 5.75, 6.3, 5.45, 5.83; 100, 100; 60)
13. The theory (claiming) that the chief defended/escorted **the tribe** was proven to be untrue. (6.78, 5.78, 4.89, 2.22, 5.3, 4.55, 5.38, 4.23; 100, 100; 60)
14. The claim (alleging) that the cop ignored/shot **the informant** might have affected the jury. (5.28, 4.61, 5.17, 2.17, 4.6, 6.2, 5.44, 5.17; 100, 100; 68)
15. The gossip (insinuating) that the neighbor hated/kicked **the cat** had spread around the neighborhood. (6.5, 5.89, 5.5, 2.5, 5.2, 4.95, 5.91, 4.84; 83, 100; 27)
16. The comment (asserting) that the hotel posted/reduced **the prices** was encouraging to the traveler. (6.06, 6.33, 5.17, 2.17, 4.75, 4.6, 3.8, 4.54; 100, 100; 11)
17. The confession (confirming) that the congressman wrote/cashed **the check** did not explain the murder. (5.89, 6.17, 4.78, 1.94, 6.1, 5.6, 6.05, 6.14; 92, 100; 31)
18. The discovery (revealing) that the physicist publicized/canceled **the project** should be the top story. (5.67, 5.89, 6.11, 2.94, 4.95, 5, 5.45, 5.68; 100, 100; 55)
19. The argument (suggesting) that the couple heard/attended **the show** did not persuade their friends. (5.39, 6.61, 6.56, 2.94, 5.9, 6.45, 5.04, 5.55; 88, 100; 42)

20. The accusation (implying) that the guard denied/granted **the request** was unfounded and completely ludicrous. (6.17, 5.61, 6.39, 2.94, 5.9, 4.6, 5.24, 5.29; 92, 100; 63)

21. The criticism (alleging) that the general debated/implemented **the decision** was ignored by the Pentagon. (5.39, 5.28, 5.39, 3.5, 5, 5.4, 3.58, 4.27; 36, 100; 28)

22. The story (stating) that the journalist used/married **the celebrity** was intriguing to the media. (5.06, 6.22, 6.33, 2.06, 6, 5.05, 5.52, 6.14; 100, 87; 15)

23. The implication (stating) that the philosopher refuted/mesmerized **his opponents** had to be considered illogical. (4.94, 4.39, 5.06, 2.33, 5.4, 4.7, 4.82, 4.36; 100, 100; 71)

24. The threat (suggesting) that the outlaws disregarded/broke **the laws** might have scared the sheriff. (6.39, 6, 5.56, 3.17, 6.2, 5.95, 4.56, 3.64; 100, 82; 46)

25. The rule (stating) that the preschoolers should follow/chase **the clown** was part of a game. (4.44, 4.11, 5.44, 2, 5.95, 4.65, 6.17, 4.88; 100, 100; 63)

26. The confirmation (indicating) that the parent made/sewed **the costumes** was sent to the PTA. (5.89, 5.28, 5.5, 2.33, 5.65, 5.55, 5.68, 4.5; 100, 91; 67)

27. The acknowledgment (mentioning) that the fireman got/parked **the truck** did not reach the station. (5.06, 6.33, 5.94, 1.83, 5.85, 5.95, 3.88, 4.08; 100, 58; 54)

28. The statement (showing) that the company issued/traded **the bonds** was published in the brochure. (5.89, 5.89, 6.44, 2.78, 5.5, 5.9, 6.32, 6.14; 100, 100; 29)

29. The danger (implying) that the contractor overlooked/violated **the regulation** was a concern for investors. (6.06, 6.39, 4.72, 2.67, 5.25, 5.7, 3.68, 3.6; 100, 100; 30)

30. The point (noting) that the teacher emphasized/changed **the homework** was discussed at the meeting. (6.61, 6.44, 6.28, 4.61, 6.4, 5.7, 4.64, 3.5; 96, 75; 13)

31. The concern (stating) that the instructor raised/misplaced **the grades** was discussed by the board. (5.33, 6.17, 5.44, 2.56, 4.85, 5.75, 4.82, 4.82; 100, 100; 47)

32. The information (showing) that the agent divulged/foiled **the plot** was examined and discussed immediately. (4.28, 4.83, 5.44, 4.28, 4.4, 4.95, 4.68, 5.67; 100, 100; 10)

33. The opinion (recognizing) that the runner challenged/beat **his opponent** was good for his reputation. (3.94, 3.61, 3.78, 1.72, 4.25, 3.35, 3.04, 3.68; 95, 87; 45)

34. The assumption (implying) that the girl made/ripped **the dress** had misled the sewing teacher. (6.56, 6.61, 5.78, 1.5, 5.85, 5.4, 5.5, 4.91; 100, 95; 73)

35. The fact (establishing) that the pupil remembered/failed **the class** might have made a difference. (5.78, 6.28, 6.06, 2.78, 6.05, 6.25, 4.67, 5.08; 68, 78; 98)

36. The information (indicating) that the assassin stole/drove **the car** might worry the CIA immensely. (4.83, 5.72, 5.61, 2.22, 4.55, 4.45, 5.5, 6.17; 100, 45; 10)

Appendix B

Table B.1: Experiment 1 Trimmed Raw Reading Time (ms/word)

	Region		
Condition	Emb Verb	Dir Obj	Main Verb
Plausible			
Ambiguous	469	436	426
Unambiguous	479	429	427
Implausible			
Ambiguous	472	437	412
Unambiguous	455	431	416

Note. Reading times were trimmed at 2.5 *SD*. Emb Verb = embedded verb. Dir Obj = direct object.

Table B.2: Experiment 2 Trimmed Raw Reading Time (ms/word)

	Region		
Condition	Emb Verb	Dir Obj	Main Verb
Sentence Complement			
Ambiguous	467	433	420
Unambiguous	451	423	404
Relative Clause			
Ambiguous	446	—	449
Unambiguous	465	—	426

Note. Reading times were trimmed at 2.5 *SD*. Emb Verb = embedded verb. Dir Obj = direct object.

Author Note

This research was supported by NSF Grant SBR-9729007 and a Northeastern University Research Scholarship and Development Fund award to the first author. We greatly appreciate the comments and advice, on earlier presentations of this work, of Gerry Altmann, Susan Garnsey, Edward "Fletch" Gibson, Rick Lewis, Maryellen MacDonald, Brian McElree, Joanne Miller, Martin Pickering, Suzanne Stevenson, and Matt Traxler. We also thank Vered Argaman, Collin Chang, Sydney Goddard, Sahoko Hama, Chad Loder, Joanne Malinka, Kim Mitchell, Patricia Pimentel, Janet Randall, Shelagh Scollin, Mei-Ling Smith, and Jasmine Whitney for their aid in constructing stimuli, running participants, and coding data. Portions of this work were presented at the March 1998 CUNY Sentence Processing Conference (New Brunswick, NJ), the November 1998 Psychonomic Society Conference (Dallas, TX), and the February 1998 Bovine Conference Digest (The Demise of the Cereal Processor, New Brunswick, NJ). Address correspondence to either author: Psychology Dept., 125 NI, Northeastern University, Boston, MA 02115. E-mail: pearlmutter@neu.edu, amendels@lynx.dac.neu.edu.