Published in final edited form as:
Q J Exp Psychol (Hove). 2011 September ; 64(9): 1797–1820. doi:10.1080/17470218.2011.582127.

Processing temporary syntactic ambiguity: The effect of contextual bias

Mohamed Taha Mohamed and
Ain Shams University, Cairo and Emirates College, Abu Dhabi

Charles Clifton Jr.
University of Massachusetts Amherst

Abstract

This paper reports two experiments using sentences with a temporary ambiguity between a direct object and a sentence complement analysis that is resolved toward the normally-preferred direct object analysis. Post-verbal noun phrases in these sentences could be ambiguously attached as either a direct object or the subject of a sentence complement whereas in unambiguous versions of the sentences the subcategorization of the verb forced the direct object interpretation. Participants read these sentences in relatively long paragraph contexts, where the context supported the direct object analysis ("preferred"), supported the sentence complement analysis ("unpreferred"), or provided conflicting evidence about both analyses ("conflicting"). Self-paced reading times for ambiguous post-verbal noun phrases were almost equivalent to the reading times of their unambiguous counterparts, even in unpreferred and conflicted context conditions. However, time to read a following region, which forced the direct object interpretation, was affected by the interaction of verb subcategorization ambiguity and contextual support. The full pattern of results do not fit well with either an unelaborated single-analysis ("garden path") model or a competitive constraint-satisfaction model, but are consistent with a race model in which multiple factors affect the speed of constructing a single initial analysis.

Keywords

syntactic ambiguity; context; ambiguous region

The past few decades have seen a large increase in our understanding of how people use grammatical structure in understanding sentences (see Clifton & Duffy, 2001, for a review). Much of this increase was arguably stimulated by the experimental analyses of the comprehension of sentences with temporary syntactic ambiguities. The logic behind these studies was that one could identify the principles governing decisions about what a sentence meant by studying what happens when decisions lead a reader or listener astray (as in Bever's, 1970 infamous "the horse raced past the barn fell"). In a seminal article, Frazier and Rayner (1982) measured eye movements while people read sentences like (1), and found slowed reading times in the region following the postverbal noun phrase (*belongs to her* in the example).

(1) The second wife will claim the entire family inheritance belongs to her.

Frazier and Rayner argued that readers initially interpreted the postverbal noun phrase *the entire family inheritance* as the direct object of the verb, and were slowed when later

Correspondence should be addressed to: Mohamed Taha Mohamed, Emirates College for Advanced Education, Abu Dhabi City, P.O. Box: 126662 Abu Dhabi, United Arab Emirates, taham2003@hotmail.com.

material forced it to be revised as the subject of a complement sentence. They attributed this effect to a simple parsing principle, Minimal Attachment, which claimed that readers initially commit themselves to the first available analysis, which will generally be the syntactically simplest one (the direct object analysis in this case). Readers are claimed to consider other analyses only if later-processed material forces them to (see also Rayner, Carlson, & Frazier, 1983).

Since this early work, researchers have explored how a multitude of factors affect a reader's or listener's understanding of a sentence: structural factors (Clifton & Duffy, 2001), lexical factors (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994), prosody (Frazier, Carlson, & Clifton, 2006), discourse (Altmann & Steedman, 1988), pragmatics (Altmann, 1998), statistical constraints (Jurafsky, 1996), and on and on. One could easily argue that the basic way in which each of these factors affects sentence comprehension is fairly well understood, and the problem now is to understand how they interact, how different sources of information interface with each other.

We submit, however, that a fundamental question of the parser's architecture has not yet been fully answered. The early theorists (e.g., Frazier & Rayner, 1982; Frazier, 1987; cf. Crocker, 1995; Forster, 1979; Kimball, 1983; for convenience, we will refer to them as "serial garden path" theorists) adopted a modular (J. A. Fodor, 1982) position, suggesting that some sources of information (e.g., syntax, morphology) took logical and temporal priority over others (e.g. meaning, discourse context, world knowledge). Other theorists adopted a connectionist position (e.g., Rumelhart et al, 1986; McClelland et al. 1986) and have proposed a "constraint satisfaction" approach[1], according to which a reader or listener uses all available sources of information in a single stage to select the most satisfactory analysis of a sentence. Such theories have most often been developed with an "activation" metaphor, where the different possible analyses of a sentence are activated to varying degrees and compete with one another (e.g., McRae, Spivey-Knowlton, & Tanenhaus, 1998). It is clear that the constraint satisfaction approach has stimulated researchers to investigate how a wide variety of factors affect language comprehension, which has vastly increased our knowledge of the topic. But this does not mean that its theoretical claims are correct.

In this paper, we examine a three-way distinction between garden-path models, competitive constraint satisfaction models, and unrestricted race models (Traxler, Pickering, & Clifton, 1998; van Gompel et al., 2001; 2005), and argue that the models make distinct predictions for patterns of reading times in ambiguous and unambiguous regions (see also Farmer et., 2007, for an independent presentation of the same theoretical distinctions). Consider a sentence like (2). The region *the entire family inheritance* is ambiguous between a direct object (transitive) analysis and the beginning of a sentence complement analysis. Imagine that this sentence occurs in a discourse context that provides a pragmatic bias toward the sentence complement analysis (without making the direct object analysis anomalous), and further imagine that the sentence is being compared with one that contains an unambiguously transitive verb, e.g., *take*. The garden path and race models predict that reading time for the ambiguous region will not be slowed by the ambiguity of the verb (it may even be speeded, according to the race model) or by the biasing context in which the sentence appears. The first analysis that can be constructed will be accepted. However, a parallel competitive model (as will be argued below) will generally predict slowed reading time in the noun phrase that follows an ambiguous verb, due to competition between possible analyses.

[1]The "constraint satisfaction" position has earlier roots in Tyler and Marslen-Wilson (1977) and in the "detective model" advanced by Clark and Clark (1977) and Fodor, Bever, & Garrett (1974).

(2) The second wife will claim the entire family inheritance for herself.

Next, consider reading time in the disambiguating region, *for herself*, which forces a transitive analysis of the preceding noun phrase. Given a discourse context that favors the sentence complement analysis, a parallel competitive model would predict slowed reading time here following the ambiguous verb for the same reason the ambiguous region was read slowly: competition between alternatives. A simple serial garden path model would predict no effects of ambiguity and no interaction with context. It claims that the reader will generally have adopted the simplest analysis of the ambiguous noun phrase, the direct object analysis, and will experience no difficulty when the disambiguating region supports this analysis. (Any difficulty of integrating the sentence with an inconsistent context should appear for both unambiguous and ambiguous verb sentences.) In contrast, an unrestricted race model can permit factors other than syntactic simplicity to affect which analysis is constructed most quickly. In particular, when context favors a sentence complement analysis, the ambiguous region will more frequently receive such an analysis, and this will result in more frequent need for reanalysis in the disambiguating region, slowing reading.

Table 1 presents a synopsis of these predictions, which we will go on to justify. The garden-path theorists' account claims that a reader uses syntactic principles to construct a quick-and-easy verb+direct object analysis of *claim the entire family inheritance.* It is initially sensitive only to grammatical (phonological, morphological, syntactic) information and accepts the most quickly-built analysis, which in general is the analysis that requires application of the fewest syntactic rules. The existence of other possible analyses will not slow reading time, and could in principle speed it, since the model posits a race between alternative analyses (Frazier & Fodor, 1978). Thus, ambiguity will not slow reading. Since the direct object analysis is simpler than the sentence complement analysis (Rayner & Frazier, 1987), it is the initially-constructed analysis (and we will refer to it throughout as the 'normally-preferred' analysis). Extra-grammatical factors do not affect the choice of an initial analysis, although they may trigger or guide revision of the analysis (Fodor & Ferreira, 1998; Rayner et al., 1983). However, if the disambiguated analysis is the same as the initially-constructed analysis (e.g., the normally-preferred direct object analysis), there is no reason to expect slowed reading (apart from effects of other factors such as plausibility, which Clifton et al., 2003 showed to affect ambiguous and unambiguous sentences alike).

While a serial garden-path theorist must posit two different processes in ambiguous and disambiguating regions, a constraint-satisfaction theorist can provide a uniform account of how both regions are read. A competitive constraint-satisfaction theory claims that the different possible analyses of *the entire family inheritance* are activated when that phrase is read, and various sources of information (frequency, plausibility, lexical biases, etc.) select among them. When multiple analyses are activated, there is a period of competition among them during which one becomes dominant. If later information (e.g., the final verb phrase *belongs to her* in the example) forces a different analysis (the sentence complement analysis) to become dominant over the earlier-favored analysis, these analyses also compete, slowing reading time.

Constraint satisfaction approaches, as they have generally been developed, claim that the same competitive process that slows reading in a disambiguating region applies in the ambiguous region itself. When only one analysis of the "ambiguous" region is possible (e.g., if the verb *claim* were replaced with *take* in (1)), there should be little or no competition between alternative analyses, and thus reading should be faster than if two or more analyses are initially activated.

Competition is not the only way that a parallel constraint-satisfaction theory can choose among alternative analyses. There are parallel models (e.g., Gibson, 1991; cf. Gibson &

Pearlmutter, 1998) that simultaneously consider constraints such as lexical, contextual, and computational resources constraints in a non-competitive way. Recently-developed statistically-based models (Hale, 2001; Jurafsky, 1996; Levy, 2008) are able to account for some of the phenomena that motivate parallel models, but their implications for the effect of ambiguity per se are not clear (it may be that an ambiguity results in less reduction of uncertainty than an unambiguous phrase, predicting that ambiguity will speed reading) and they have not been developed to deal with discourse context effects. In any event, competitive constraint satisfaction is the most widely proposed way of adjudicating among multiple possible analyses, and is the way that is used in the most influential explicit model, McRae et al's. (1998) 'competition-integration' model. Elman, Hare, & McRae (2005) applied this model to sentences with direct object/sentence complement ambiguities like (1) (an ambiguity that was examined in the present experiments). The model made quantitative predictions of longer reading times both in the postverbal noun phrase region (corresponding to *the entire family* inheritance in (1)) when it was ambiguous than when it was unambiguous and in the disambiguating region (corresponding to *belongs to her* in (1)) following an ambiguously-attached noun phrase compared to when the complementizer *that* forced a sentence complement reading.

Other explicit parallel models that employ competition to satisfy multiple grammatical and extra-grammatical constraints have been developed. These include early models by Bates and MacWhinney (1982) and Cottrell (1985), as well as models by MacDonald, Pearlmutter, and Seidenberg (1994), Stevenson (1994), Tabor and Tanenhaus (1999), and Vosse and Kempen (2000; cf. Vosse & Kempen, 2009, for an alternative development). Most of the presentations of these models explicitly state that resolving an ambiguity slows processing. Further, the field seems to have accepted that ambiguity does slow processing in a way that is explained by parallel constraint satisfaction models, to the extent that it is the position taken in standard textbooks (e.g., Harley, 2008, pp 305–306).

As mentioned earlier, parallel models that take multiple information sources simultaneously into account do not have to rely on time-consuming competition to reach an interpretation (see Gibson & Pearlmutter, 2000, and Lewis, 2000, for some discussion). Further, Green and Mitchell (2006) have challenged the view that competitive constraint satisfaction models do predict slowing in an ambiguous region. They show that such models (in particular, the McRae et al., 1998, competition-integration model) do not predict slower processing in ambiguous than in unambiguous regions when one of the possible analyses is strongly activated before entering a region. In this case, ambiguity merely allows the existing strong activation to be maintained.

Green and Mitchell's valid insight is that when context permits a particular structural analysis to be predicted and activated, an ambiguous instance of that structure is taken to be consistent with the prediction and its reading will slowed modestly or not at all, compared to an unambiguous instance. But in many cases of interest, there is no reason to predict one particular analysis. In one construction that Green and Mitchell discuss extensively (the relative clause attachment structure, *the daughter of the colonel who was on the balcony*; Cuetos & Mitchell, 1988), a reader has no reason to predict that a relative clause will follow the initial noun phrase *the daughter of the colonel*, much less to predict that which of the two nouns it will modify. In the constructions studied in the present research (the NP/S complement ambiguity, illustrated in Examples 1 and 2), there is little reason for a reader to predict or activate either an NP or S complement analysis before the ambiguously-attached phrase-initial noun is read. Further, we explicitly manipulate the extent to which context biases toward assigning one particular structural analysis when the noun is read, resulting in conditions that, following Green and Mitchell, should permit ambiguity-based competition to appear. Finally, and concretely, Elman et al. (2005) explicitly used the competition-

integration model to derive quantitative estimates of reading time of sentences with the direct object/sentence complement ambiguity presented in biasing contexts. They showed that the model predicted that reading would be slowed in the ambiguous noun phrase region (see their Figure 4). However, their experimental data showed that reading was actually slowed nonsignificantly if at all, whereas it was substantially slowed (as predicted) in a following disambiguating region (again, see Elman et al's Figure 4; see Clifton & Staub, 2008, for further discussion).

A third type of model, an unconstrained race model (Traxler et al., 1998; van Gompel et al., 2005) shares features with both garden-path and parallel competitive models. The garden-path model we have been considering is in fact a race model (Frazier & Fodor, 1978), but the winner of the race is determined only by syntactic simplicity considerations. An unrestricted race model permits extragrammatical factors such as context and construction frequency to affect which analysis wins the race. The existence of multiple possible analyses does not slow reading, as it does in the competition model. In fact, the existence of different possible analyses of an ambiguous sentence can speed its reading, since the mean of the fastest analysis (the "winner of the race") should be faster than the mean completion time of any particular analysis (Traxler et al., 1998) On the other hand, since extragrammatical factors can affect which analysis wins the race, an unrestricted race model can predict that such factors will affect reading time of a disambiguating region. If discourse context results in one interpretation winning the race, reading time will be slowed in a following region that disambiguates toward the other interpretation.

What are the facts about reading times? There is abundant evidence (see, e.g., Clifton, Staub, & Rayner, 2007, for a review) that reading is slowed in a disambiguating region when the disambiguation clashes with syntactic preferences. Reading in a disambiguating region is also sometime slowed when the disambiguation clashes with extragrammatical factors (although this is not always observed, at least in the face of strong syntactic preferences; Binder, Duffy, & Rayner, 2001, for example, showed that an extragrammatical context that was sufficient to speed reading of a region that disambiguated in favor of the unpreferred reduced relative clause construction did not slow reading when the region disambiguated to the preferred direct object interpretation). However, there is little compelling evidence that reading is slowed in an ambiguous region compared to a matched unambiguous region. Clifton and Staub's (2008) analysis of the sentence-processing literature makes it clear that there has been rather little research targeting such an increase. Most potentially-relevant reports focus on the disambiguating region, and either show no effects in the ambiguous region or do not present the data. The few instances of apparent slowing in an ambiguous region are generally confounded with implausibility of one analysis (e.g., Ferreira & Clifton, 1987), different lexical items (e.g., Ni, Crain, & Shankweiler, 1998),. The apparently most adequate test of whether reading is slowed in an ambiguous region was performed by Hare, McRae, and Elman (2003). As discussed above, Elman et al. (2005) fit the competition-integration model to these data. Their model clearly predicted that reading should be slowed in the ambiguous region compared to a disambiguated region. However, judging from the figures presented by Hare et al. (and by Elman et al., 2005), no such slowing was observed.

The current experiments

One goal of the present research is to attempt to determine the pattern of how reading is affected by a temporary direct object/sentence complement ambiguity both in the ambiguous region and the disambiguating region. A second goal of the present research is to explore the effects of discourse context on the response to syntactic ambiguity and its resolution (as was also done by Hare et al., 2003). As outlined in Table 1, the three classes of theories that have

been discussed make different predictions about how ambiguity and discourse context will affect processing in the ambiguous and disambiguating regions.

We report two experiments, measuring self-paced reading time of sentences that are temporarily ambiguous between direct object and sentence complement interpretations, the ambiguity illustrated in (1) and (2) (see Ferreira & Henderson, 1990; Garnsey et al..1997; Hare et al., 2003; Kennison, 2001; Pickering, Traxler, & Crocker, 2000; Rayner & Frazier, 1987; Trueswell, Tanenhaus, & Kello, 1993; Wilson & Garnsey, 2009 for some other studies of this construction). An example of one of our sentences appears in (3), in its temporarily ambiguous form in (3a) (where the verb supports both direct object and sentence complement analysis) and in an unambiguous form in (3b), where the verb is obligatorily transitive. Many studies of this pair of constructions use control sentences with the complementizer *that*, which makes a sentence unambiguously have a sentence complement. In contrast, we used control sentences with unambiguously transitive verbs, which require the direct object analysis (see Staub, 2007, for convincing evidence that readers honor transitivity in online processing). We did this because we wanted to study disambiguation toward the analysis which is normally preferred, for reasons discussed below.

(3) **a.** Kent announced the details of the murder and explained why they had to be made public.

b. Kent presented the details of the murder and explained why they had to be made public.

Our participants read these sentences in the context of long paragraphs whose content biased them toward the preferred direct object analysis of the postverbal noun phrase (*the details of the murder* in the example), or toward the unpreferred sentence complement analysis, or provided conflicting information supporting leaving both analyses possible. We believed that this context manipulation sets the stage so that ambiguity effect in ambiguous region, if exists, can be easily detected. While most discourse contexts that have been used in the sentence processing literature have manipulated how context provides support for noun phrase reference, our contexts manipulated the plausibility of the propositions expressed by direct object vs. sentence complement analyses. The experiments were designed to determine first, whether reading time for the postverbal noun phrase was affected by whether its relation to the verb it followed was ambiguous or not, and second, whether any differences in reading time of the ambiguous and disambiguating regions were modulated by the plausibility of the possible analyses in the discourse context.

The following disambiguating region of the experimental sentences (beginning *and explained* in the example) always resolved the ambiguity in favor of the direct object analysis, which is the normally preferred analysis in a null or neutral context (Rayner & Frazier, 1987). . We concentrated on this resolution rather than a resolution in favor of the normally unpreferred, sentence complement, analysis, because all models under consideration can predict slowing of the disambiguating region when it forces analysis of the ambiguous noun phrase as subject of a complement sentence (Frazier, 1995) and thus, such disambiguation does not test between the models. Our use of normally preferred sentences in the current research is similar to their use in a recent study by Wilson and Garnsey (2009) who used them to show that a verb bias toward the sentence complement analysis slowed the reading of temporarily-ambiguous sentences when they were disambiguated toward the direct object analysis, when compared to similarly-disambiguated sentences whose verb was biased toward a direct object analysis. Thus, for example, in their Experiment 1 they found 14 ms longer reading times in the (italicized) disambiguating region in the clause-biased (4) below than in the direct object-biased (5) (Wilson & Garnsey, 2009).

(4) The ticket agent admitted the mistake *because she* had been caught.

(5) The CIA director confirmed the rumor *when he* testified before Congress.

Wilson and Garnsey (2009) thus reported a result that is inconsistent with a prediction of the garden-path theory. We extend their research by comparing reading time for a region that disambiguates a temporarily ambiguous sentence in favor of the direct object analysis with the time for the same region in an unambiguous sentence, a comparison that was not included in the Wilson and Garnsey design. We will also go beyond examining effects in the disambiguating region and report effects found in the ambiguous region, the post-verbal noun. The Wilson & Garnsey design did not allow them to make secure conclusions about this region, since in their unambiguous condition, the critical noun phrase was preceded by the complementizer *that*, a highly frequent word which can be expected to influence the reading of the following word.

To summarize our predictions, a garden path model and an unrestricted race model do not predict that a temporarily ambiguous region will be read slowly, compared to a matched unambiguous region. A competitive constraint satisfaction model predicts that, in the absence of prior expectations, ambiguity will slow reading. In contrast, the unrestricted race model and the constraint satisfaction model predict that contextual bias can affect the initial analysis of a temporary ambiguity, slowing reading time when a following region disambiguates against the bias. Without additional assumptions, a garden-path model would not predict that contextual bias could slow reading time of a region that disambiguates in favor of the normally-preferred, minimal attachment (here, direct object) analysis.

Experiment 1

The first experiment was designed to measure reading time of a postverbal noun phrase and a following region that forced a direct object analysis of this noun phrase. Critical sentences were either ambiguous or unambiguous in terms of how the postverbal noun phrase could be related to the verb. They were presented following contexts that supported the normally-preferred direct object analysis of the noun phrase, or that supported the normally-unpreferred subject of sentence complement analysis, or were conflicted between supporting the two analyses. Thus, the experimental design had three context conditions (preferred, conflicted, and unpreferred) crossed with two ambiguity conditions (ambiguous and unambiguous verb subcategorization).

Method

Materials—Temporarily ambiguous normally-preferred sentences like (6a), with the ambiguously subcategorizing verb italicized in the examples, were compared to unambiguous but otherwise equivalent versions of the same sentences as in (6b). These unambiguous sentences contained pure transitive verbs with similar meanings to those in the ambiguous sentences.

(6)

a. The independent inspector *proved* the charges against Mr. John and issued his final report accordingly.

b. The independent inspector *supported* the charges against Mr. John and issued his final report accordingly.

In preparing the material, ambiguous verbs were chosen to permit both direct objects and sentence complements based on the fragment completion norms presented by Garnsey et al (1997) and Kennison (2001) (see Table 2; note, some verbs were used in as many as three experimental passages). Since these (and other) norms show modest-to-good agreement in terms of the relative frequencies of direct object (NP) vs. sentence complement (SC), and since a measure of this relative frequency seems best to predict behavioral data (Gahl, Jurafsky, & Roland, 2004), the mean percentage of NP/(NP+SC) completions was

calculated across the two sets of norms (when verbs were available in both norms). The resulting bias percentages appear in the final column of Table 2, and averaged 46.7%. Unambiguous verbs were chosen to permit only a direct object analysis and block a sentence complement analysis, and to have approximately the same meaning as their counterparts in the ambiguous sentences in the judgment of the authors[1].

These sentences were used in the three previously mentioned context conditions. Each context introduced a noun phrase denoting some object or event (*theory*, *solution*, *decision*, *murder*, etc.) and making some assertions about it. In the preferred context, the assertions were such that the object or event would most plausibly be taken as the theme of the verb of the critical sentence. The assertions of the unpreferred context most plausibly supported a state of affairs in which the verb would take a complement making some assertion (generally some kind of denial) about the object or event. Conflicted contexts contained some assertions that supported each expectation.

Six versions of each of 24 passages were constructed. Each version contained either a conflicted, preferred, or unpreferred context and the critical sentences contained either a pure transitive verb (unambiguous) or a verb that permitted either a noun phrase or a sentence complement (ambiguous). In all experimental passages, the critical sentence was resolved to its normally preferred structure (direct object rather than sentence complement). Table 3 gives an example of the three contexts followed by the ambiguous and unambiguous versions of the critical sentence. The table also gives an example of the question that was asked after the passage. Questions were typically rather demanding, in that they required comprehension of the passage and often some inferencing. All experimental materials for Experiment 1 are available at https://udrive.oit.umass.edu/xythoswfs/webui/_xy-5032239_1 To avoid the possibility that participants might learn to expect a normally-preferred sentence, a set of 12 passages was used as fillers. Each of these passages contained a critical ambiguous sentence resolved toward the unpreferred sentence complement analysis. Four of these passages had a conflicted context, 4 passages had a preferred context, and 4 passages had an unpreferred context. Thus, passages that contained a temporarily ambiguous attachment of the noun phrase to the preceding verb (half of the experimental passages and all of the fillers) were resolved equally often to direct object and sentence complement interpretations. Two norming studies were conducted to test the validity of the experimental materials.

Norming study 1—The main goal of this norming study is to make sure that each context evokes the intended interpretation(s). In the conflicted condition, there should be roughly equal preference for the direct object and the subject of sentence complement interpretations. For the preferred and unpreferred contexts, there should be a preference for the preferred (direct object) and unpreferred (sentence complement) interpretations, respectively. Participants were presented with experimental items up through the noun phrase of the critical sentence, following an ambiguous verb. Each critical sentence fragment was presented in the three context conditions (see Table 4). Items were identical except for one or two sentences which were meant to set the context. Even these were of

[1]The ambiguity manipulation in this experiment compares unambiguous direct object verbs with verbs that ambiguously take a direct object or a sentence complement. There is a substantial literature about the effects of relative frequency of direct object vs. sentence complement use of ambiguous verbs, but the literature is somewhat unclear about such effects. Some studies examine whether relative frequency of usage affects the size of any garden path effects observed (Ferreira & Henderson, 1990; Kennison, 2001; Trueswell, Tanenhaus, & Kello, 1993; Wilson & Garnsey, 2009) and others examine the interaction of relative frequency and plausibility (Garnsey et al., 1997; Pickering, Traxler, & Crocker, 2000). Perhaps the safest conclusion to draw from these studies is that large biases toward one analysis or the other do not always block the simpler direct object analysis, but that evidence for that the direct object analysis is entertained and evaluated for plausibility is most reliably obtained when the direct object and sentence complement analyses are roughly equally frequent, as they are in our materials.

similar length and shared most of their words. Participants were given two alternative continuations, one which supported the preferred direct object analysis, and the other the unpreferred sentence complement analysis. The participants' task was to choose one of these two alternatives as the best completion of the sentence. Three groups of 8 participants completed the completion choice task (the data from one clearly uncooperative participant was eliminated). Each group completed 24 passages of the 72 possible passages (24 items * 3 context conditions). The construction of the three sets of passages followed a Latin square so that for each sentence, one passage representing one condition was chosen to be included in each set. The order of presenting the interpretations to be chosen from was alternated, so half of the items had NP complement first and the other half of items had the sentence complement first.

The results (Table 5) indicated that the contexts work in the expected direction. A simple ANOVA showed that there was a significant difference among contexts in terms of their preferences for the type of completion that follows the verb in the critical sentences ($F1 (2, 44) = 20.696$, $p < .001$, $F2 (2, 46) = 26.095$, $p < .001$). Pairwise comparisons among contexts showed that they were significantly different at the .01 level. The conflicted conditions elicited nearly equal percentages of choices of the preferred and the unpreferred analyses. The majority of participants chose the preferred analysis in the preferred context. Similarly, the majority of participants chose the unpreferred analysis in the unpreferred context. To be sure, the three context conditions were not as sharply separated as would be ideal. The inherent bias in favor of a direct object interpretation was difficult to overcome, and the complexity of the contexts permitted variation in how they were interpreted. Nonetheless, the differences among the context conditions were statistically significant, and as we will demonstrate in the following experiments, significantly affected the time taken to read the text that made the actual interpretation clear.

Norming study 2—The second norming study was designed to make sure that the critical sentences in each passage are plausible. The passages, up through the critical sentence, were presented to 12 participants and they were asked to rate the plausibility of the underlined critical sentence in each passage on a 7-point scale ranging from extremely plausible (7) to extremely implausible (1). Six filler passages were constructed to be extremely implausible. The averages of evaluation rates for each condition are presented in Table 6 below.

The results indicated that ratings of plausibility of the critical sentences in each condition is high (none was less than 5.5) and a simple ANOVA showed that there was no difference among the six conditions in terms of participants' rating of their plausibility ($F1(5, 55) = 1.897$, $p = .11$, $F2 (5, 115) < 1$.

Taken together, the results of the two norming studies indicate the validity of the material. In particular, they show that (1) context conditions evoke the analyses that they are supposed to evoke, and (2) the critical sentences in different conditions were acceptable to readers from the same pool from which participants of the two experiments were recruited.

Participants—Fifty-eight participants took part in the main experiment. They were recruited from UMass undergraduate psychology classes. They participated for credit or were paid $8. The data of 10 participants with an accuracy level below 70% in answering questions that followed each passage were eliminated, a number that appears high but perhaps is not unexpected given the complexity of the passages and the difficulty of the questions. Moreover, it was observed that some participants seemed to be reading very fast and carelessly, not looking at all presentation regions long enough to really read them. These participants were identified as those who read 100 or more experimental sentence presentation regions (approximately 10% of all regions) in less than 400 ms each. Since

nearly all presentation regions had multiple words, it was assumed that these participants were not always reading carefully. Seven more participants were eliminated by this criterion, leaving the data of 41 participants to be analyzed.

Procedures—A 3 (context conditions) * 2 (ambiguity conditions) design was used in the experiment. The six versions of the 24 passages were assigned to six different counterbalanced lists, so that each passage was tested in each version in one list and each list contained four passages in each version. A pharase-by-phrase self-paced reading technique was used to measure the reading times of presentation regions. The presentation regions were chosen to be coherent short phrases. This procedure was chosen over word-by-word presentation to encourage readers to comprehend each region before moving on, thus avoiding the delayed effects and spillover effects that sometimes appear with word-by-word presentation (Mitchell, 2004) and permitting reading times to be interpreted as comprehension times. Participants were presented with the paragraphs on a computer screen. Before each paragraph, participants saw the sentence "*press thumb button to see next passage".* On pushing a lever with the right hand, they were presented with a preview display. In this preview display, each letter was replaced by a dot but the spaces and punctuation were preserved. To see the first region of the text, participants had to pull a trigger with the right hand. After reading a particular presentation segment, participants had to pull the same trigger again to bring up the next one, and simultaneously, to turn the previous presentation segment into dots again. Reading time for each region was recorded. Participants continued pulling the trigger after reading each presentation segment until the end of the paragraph.

Each participant read 24 experimental paragraphs plus 12 fillers that contained temporarily ambiguous critical sentences resolved in favor of the unpreferred sentence complement analysis. Participants were instructed to read the paragraphs according to their normal reading speed as they would do in any natural reading situation. The paragraphs were presented in a constrained-randomized mode. Paragraphs were randomized over the six experimental conditions according to a Latin square so that every critical sentence appeared equally in the six conditions and no participant saw a single sentence in more than one condition. However, two filler passages appeared at the beginning of the session for each subject. These two paragraphs were used to familiarize participants with the paradigm. Each paragraph was followed by a simple comprehension question. The total time needed to finish the experimental session varied from 35 to 45 minutes.

Data analysis—Reading times in four presentation regions of the critical sentence were analyzed. The first three regions are illustrated in Table 7, and all four are indicated by ~ marks in Table 4. Region 1 contained the subject and verb of the critical sentence, Region 2 contained the ambiguous postverbal noun phrase, and Region 3 contained the disambiguating region. For 18 of the 24 items, the disambiguating region was a conjoined verb phrase (conjunction + verb + argument); for the remaining items, it was a conjoined sentence (including a different subject). Region 4 contained the next phrase of the sentence, and was analyzed simply to see if there was any spillover from the disambiguating region. Reading times under 300 ms (1.2% of the data) were eliminated, since the regions were too long to have been read carefully that quickly. Then boxplots of the reading times for each region were constructed, and clear outliers were identified as data points that were at least several hundred ms longer than the bulk of the RT distributions. This resulted in eliminating all times greater than 3500 ms in Region 1 (subject + verb) and 4000 ms in each of the other regions (0.9% of the data).

The resulting times were analyzed as a linear mixed model, as described in Baayen, Davidson, & Bates, 2008. This analysis was chosen because it permitted simultaneous

evaluation of both participants and passages as random effects factors, because it permitted any effects of variables such as region length and verb bias to be taken into account, and because it permitted analysis of the imbalanced design necessitated by eliminating participants. The R statistical programming language was used for all analyses (The R Project for Statistical Computing; http://www.r-project.org/). Reading times in each of the four analysis regions were analyzed with two fixed effect factors and their interaction, ambiguity of the verb (ambiguous vs. pure transitive unambiguous), and context (preferred, conflicted, and unpreferred), plus a third fixed effect scalar variable of region length in characters. Models that included the mean percentage of NP verb bias (last column of Table 1) were also evaluated (see Crawley, 2007, for discussion of model evaluation). Including this scalar variable as an additive term improved the model fit significantly only for analysis of Region 1, the initial noun + verb region; including it as an interacting term did not result in further significant model improvements. The reported analyses therefore will include NP verb bias only for Region 1. There were two random factors, participant and passage (item) intercepts. The verb ambiguity factor was analyzed using sum (or effect) coding (Crawley, 2007; Myers & Well, 2003). This results in the test of the two-level verb ambiguity factor evaluating the overall difference between sentences with ambiguous and unambiguous verbs (the factor's beta weight as reported in the analysis is half the difference between the means). It also results in the intercept of the ambiguity factor being the grand mean of that factor, so that the effects of context are evaluated as the marginal means over ambiguous and unambiguous verbs (as in an ordinary ANOVA, or when a multi-level factor is centered in a linear model). The 2-df context factor was analyzed using two single df *a priori* contrasts, for theoretical reasons. The first contrast compared preferred contexts to the average of conflicted and unpreferred contexts, and effectively tested whether reading was slowed when the context did not provide clear support for the normally-preferred direct object analysis. The second contrast compared conflicted to unpreferred contexts, to see whether a strong bias against the normally-preferred analysis slowed reading. (Because these two contrasts exhausted the available df, it was inappropriate to perform other contrasts of possible interest, e.g., the contrast of preferred and unpreferred contexts.) Because of the recognized difficulty of estimating df in designs like the present one, significance was evaluated using Markov Chain Monte Carlo estimation (Baayen et al., 2008).

Results

Reading times for the critical regions (subject + verb, ambiguous, and disambiguating regions), gathered from 41 subjects, are shown in Figure (1). FIGURE 1 ABOUT HERE

In the *subject + verb region*, there was a marginal main effect of ambiguity (t = 1.62, pMCMC <.11), reflecting possibly-longer times to read the unambiguous pure transitive verbs than the ambiguous verbs. Further, there was a significant interaction between ambiguity and the first context contrast (preferred context vs. the other contexts) (t = 1.98, pMCMC < .05). Times to read unambiguous noun plus verb were particularly long in the preferred context compared to the other contexts. This result was unexpected, and was not replicated in Experiment 2. There was a highly significant main effect of region length in this and in all other analyses. However, since all analyses are consistent with finding that reading time is affected by region length, the effect of region length will not be discussed further. Subject + verb reading speed was significantly affected by verb NP vs SC bias (pMCMC < .01; effect coefficient = −4.32), indicating faster reading when the verb was more strongly biased toward taking an NP complement. However, since allowing this effect to interact with the other effects did not improve the model fit (reflecting the nonsignificance of the interactions), we disregard the effect; it may be due to unintended differences among lexical items, so we do not attempt to interpret it.

The ambiguous region (the noun phrase following the subject + verb region) had no significant differences in reading times. All values of t were < 1.0, except for the first context contrast, preferred context vs. the other conditions, where reading time was marginally faster in the preferred condition (t = 1.93, pMCMC < .09). In particular, there was no sign that reading was slowed when the verb ambiguously permitted a direct object vs. a sentence complement, as opposed to being unambiguously transitive (pMCMC > 0.95). The coefficient of the verb ambiguity effect in the linear mixed model analysis (estimated from the MCMC analysis) was 1.14 with HPD95% lower and upper bounds of −28.21 and 30.370.

As for the *disambiguating region*, there was a significant main effect of ambiguity (t = 2.52, pMCMC < .02) when tested at the pooled context conditions; the disambiguating region was read slower in the ambiguous than in the unambiguous condition (1375 vs. 1314 ms). More importantly, the interaction of ambiguity with the first contrast of context (preferred vs the unpreferred + control) was significant (t = 2.38, pMCMC < .03). While reading times were numerically faster following disambiguation of an ambiguous than an unambiguous verb in the preferred context (1304 vs 1353 ms), they were slower in the pooled conflicted and unpreferred contexts (1411 vs. 1295 ms). The coefficient of this interaction was −25.13 ms, with HPD95 bounds of −46.238 and −3.779. This interaction shows that the cost of ambiguity increased when the discourse context was not biased in favor of the actual resolution of the ambiguity (that is, not biased to favor the normally preferred, direct object analysis). The penalty for ambiguity appeared to have been larger for conflicted contexts than for unpreferred contexts (penalties of 166 vs. 65 ms), but the interaction of this contrast with ambiguity was not significant. (t = 1.5, pMCMC < .15, with a coefficient of −27.08 and rather large HPD95 bounds of −65.024 and 7.604). No effects (apart from the ever-present effect of region length) were significant in the region following the disambiguation (all p > . 15).

Discussion

The core results are straightforward: the pattern of data matched the predictions of the race model, and not of serial garden-path or competition models. There was no sign that reading of a noun phrase was slowed when it could be interpreted as either a direct object or as the subject of a sentence complement, compared to when it was unambiguously a direct object. This lack of effect held regardless of the bias induced by the preceding context – namely, whether it supported the preferred direct object reading of the noun phrase, or the sentence complement interpretation, or provided ambiguous (conflicting) support for both readings. On the other hand, reading time for the disambiguating region showed a clear interaction of ambiguity of the noun phrase and preceding context. The disambiguating region, which forced the noun phrase to be interpreted as the direct object of the verb, was read more slowly following an ambiguous than an unambiguous verb when the context failed to provide clear support for the required direct object interpretation.. However, there was no substantial effect of verb ambiguity when the context supported the preferred direct object interpretation (reading times were slightly, but not reliably, slower following an unambiguously transitive verb).

We suggest that the lack of effect of ambiguity in the ambiguous region is inconsistent with parallel competitive models, but consistent with both serial garden-path and race models. We further suggest that the interaction of ambiguity with context in the disambiguating region is inconsistent with a serial garden-path model, which would have predicted no difficulty with the direct object analysis (apart from late-stage difficulties of interpretation, which should have appeared for both the ambiguous and the unambiguous conditions). The slowed reading of the disambiguating region following a sentence with an ambiguous verb in a context that did not favor the direct object analysis indicates frequent selection of the

sentence complement analysis. According to parallel competitive models, such an analysis would result in competition with the actual transitive analysis; according to a race model, the initial sentence complement analysis would have to be revised in favor of a direct object analysis. In either case, reading would be slowed.

The pattern of results across the ambiguous and disambiguating region is what was predicted by the race model, and is inconsistent with straightforward serial garden-path and parallel competitive models. To be sure, the lack of an obtained effect of ambiguity in the ambiguous region does not prove that there is no true difference, but the observed effect is very small in absolute terms, and its 95% bounds come very close to excluding the observed effect of the interaction of ambiguity and context in the disambiguating region. To increase confidence that ambiguity affects only the disambiguating region, we report an extended replication of Experiment 1 below.

There are some unexpected patterns in the data of Experiment 1. These include the fact that reading time was longer for the noun plus unambiguous verb in the preferred context than for the noun plus ambiguous verb in this context, and the fact that reading times for the disambiguating material were numerically (if not significantly) faster following an unambiguous verb than following an ambiguous verb in the conflicted and unpreferred contexts but not in the preferred context. We are inclined to attribute these effects to unintended effects of the plausibility of particular passages (despite the nonsignificant effects of passage type on rated plausibility), but rather than attempting to interpret them, we will present a second experiment, partly to see which if any of these effects replicate, but also to test one alternative account of the reading time effects in the disambiguating region, as explained below.

Experiment 2

The main purpose of the second experiment was to exclude the possibility that the difference between ambiguous and unambiguous sentences in the disambiguating region is due to a delayed effect of competitive interpretations of the ambiguous region which spilled over to the following disambiguating region. This was unlikely, because of the use of fairly long presentation regions, but nonetheless a new phrase modifying the noun phrase of the ambiguous region was added after the ambiguous region and before the disambiguating one. This region will be referred to as the extended ambiguity region. If the effect of ambiguity and context in the disambiguating region of Experiment 1 was really a spillover effect from the ambiguous region, and not a response to the disambiguation, the effect should appear in the extended ambiguity region.

Method

Material—The same materials used in Experiment 1 were used in Experiment 2 except for adding the extended ambiguity region, an adjunct phrase which modified the ambiguously-attached postverbal noun phrase. In 19 of the 24 items, the extended ambiguity region was a prepositional phrase. In the remaining cases, it was a relative clause (as in (7)) or an adjectival phrase. Thus, the critical sentence in the earlier example became (7), with an extended ambiguity region (*that were being investigated*). Similar extended ambiguity regions were added to the sentences in the filler paragraphs.

(7) The independent inspector proved/supported the charges against Mr. John that were being investigated and issued his final report accordingly.

Participants—65 University of Massachusetts undergraduates participated in the experiment. They participated for course credit or were paid $8. Seven participants with an

accuracy level below 70% in answering comprehension questions that followed each passage were eliminated. Moreover, 13 participants were excluded for having more than 100 reading times below 400 ms.

Procedure—The procedures were the same as in Experiment 1, except for the addition of the extended ambiguity region.

Data analysis—Reading times in four regions of the critical sentence were analyzed. As in Experiment 1, Region 1 contained the subject and verb of the critical sentence and Region 2 contained the ambiguous postverbal noun phrase. Region 3 now contained the extended ambiguous region, and Region 4 was the disambiguating region. The analysis of the following region is not reported (as in Experiment 1, it showed no significant effects of the manipulations). Reading times under 300 ms (1.1% of the data) were eliminated, and boxplots were used to identify outliers. This resulted in eliminating times greater than 3500 ms in Regions 1 through 3 and 4000 ms in Region 4 (1.5% of the data).

The data were analyzed using linear mixed models and contrasts as described for Experiment 1. In each analysis, percentage of NP usage of ambiguous verbs was included as a scalar additive or interacting factor. As in Experiment 1, including the additive factor improved the model only for Region 1 (Noun plus Verb), and including it as an interacting factor resulted in no further improvement. Therefore, models with the NP frequency factor will be reported only for Region 1.

Results and Discussion

Reading times of the critical regions (subject + verb, ambiguous, extended ambiguity, and disambiguating regions), collected from 45 subjects, are summarized in Figure 2.

For the *subject + verb region*, no effect or interaction (apart from the ubiquitous effect of region length) approached significance (all pMCMC > .15). The puzzling anomaly of long reading times for unambiguous verbs in the preferred context did not appear in Experiment 2. As in experiment 1, reading speed increased as for verbs that were more strongly biased toward direct object usage (pMCMC < .02, effect size = −4.11, very similar to what was observed in Experiment 1).

An anomaly did appear, however, in the *ambiguous region.* Reading times were notably fast for the noun phrase that followed an ambiguously-transitive verb in the preferred context condition (mean times of 1316 ms for the ambiguous preferred condition, 1422 for its unambiguous counterpart) This effect resulted in an interaction between verb subcategorization ambiguity and the first context contrast (preferred context vs. the others) (t = 2.51, pMCMC < .02); the mean times for the pooled conflicted and unpreferred conditions were 1405 ms for the ambiguous items vs. 1390 ms for the unambiguous items). The fast times in the ambiguous condition were not observed in Experiment 1 (identical to Experiment 2 up through the ambiguous region), and in any event go counter to the analysis in which ambiguity should result in slow reading times. No other effects (apart from length) approached significance. In particular, the overall effect of ambiguity was nonsignificant (pMCMC > .64). The estimated effect size was 13.02 with 95% HPD interval of −42.13 to 70.84; note that the nonsignificant effect was in the direction of faster, not slower, reading for the ambiguous than the unambiguous region.

The anomalously fast reading times for the ambiguous/preferred context condition also appeared in the *extended ambiguity region*, resulting in significance of the interaction between verb subcategorization ambiguity and the first Context contrast (t = 2.58, pMCMC < .01) (for the preferred condition, ambiguous items took 1048 ms while unambiguous items

took 1161 ms; for the pooled conflicted and unpreferred conditions, the means were 1173 and 1140 ms) Again, no other effects of interest approached significance. The overall effect of ambiguity had a $p > .19$, and an estimated effect size of 30.3341 (95% HPD interval −15.489 to 76.253), again trending toward faster reading in the ambiguous than the unambiguous region.

The *disambiguating region* showed essentially the same pattern of results as in Experiment 1. The interaction between verb ambiguity and the first Context contrast (preferred vs. conflicted and unpreferred contexts) was significant ($t = 2.13$, $pMCMC < .04$). The coefficient of the interaction was 46.65 (95% HPD interval 2.876 to 88.697) As in Experiment 1, there was only a small and nonsignificant difference in reading time between disambiguations following ambiguous vs. unambiguous verbs in the preferred contexts (1563 vs 1578 ms), but a substantial penalty for reading the disambiguating material after ambiguous verbs existed in the conflicted and unpreferred contexts (ambiguous, 1654 ms; unambiguous, 1520 ms). The interaction between verb ambiguity and the second context contrast (conflicted vs. unpreferred) once again was nonsignificant ($t < 1.0$). As in Experiment 1, no effect (apart from region length) approached significance in the post-disambiguating region.

Discussion

The main novel contribution of Experiment 2 was to show that the slowed reading times following ambiguous verbs in the conflicted and unpreferred contexts could not be attributed to spillover from the ambiguous region. The extended ambiguity region (which followed the head portion of the ambiguous noun phrase) showed the same pattern of results as the ambiguous region, which were very different from the results in the disambiguating region.

Experiment 2 confirmed the primary findings of Experiment 1. There was no slow-down in reading of an ambiguously-attached noun phrase following a verb that supports both direct object and sentence complements, compared to following a verb that unambiguously supports only a direct object complement, nor was there any interpretable effect of context on reading times in this region. A similar description applies to the extended ambiguity region, which modifies the noun phrase that began in the ambiguous region. We have no plausible account of the fast reading time following ambiguous verbs in the preferred contexts, as compared to following unambiguous verbs. Examination of the distributions of reading times for these regions indicated that there were an unusually large number of times in the 2000 to 3000 ms range for the unambiguous preferred context condition, as compared to the ambiguous preferred context condition, while the distributions were otherwise very similar. However, these long times were not associated with any particular subset of the passages or the subjects, and we have no account of them.

Considering the disambiguating region, Experiment 2 confirmed the interesting results of Experiment 1. Compared to the ambiguous verb condition, there was a cost to reading the phrase that forced a direct object interpretation of the preceding noun phrase when the attachment of that noun phrase to the verb was ambiguous (direct object vs. subject of sentence conflict). This cost appeared only following conflicted and unpreferred contexts, and disappeared entirely for preferred contexts.

Once again, as in Experiment 1, we must acknowledge that reading times for disambiguating phrases following unambiguous verbs were particularly fast in the conflicted and unpreferred context. We do not have a good account of why these times were apparently as fast as or faster than the times to read in the preferred context condition, where conflict is

essentially nonexistent. The finding is unexpected under any of the theories we have discussed.

General Discussion

The two experiments that have been presented explored the effect of syntactic ambiguity and propositionally-biasing contexts on the reading of different regions of sentences: regions that could be temporarily ambiguous, and regions that resolved this ambiguity. The syntactic ambiguity that was examined was the often-studied direct object/sentence complement ambiguity (e.g., Rayner & Frazier, 1987). Its ambiguity was blocked in the unambiguous condition by using a verb that permitted only a direct object noun phrase complement. Material that followed the ambiguity always forced the direct object analysis (the strongly-preferred analysis out of context; Rayner & Frazier, 1987). Context was manipulated by preceding the critical sentence with a relatively long paragraph whose content provided information that favored the direct object interpretation (“preferred context”), favored the sentence complement interpretation (“unpreferred context”), or favored each interpretation (“conflicted context”). Reading time was measured in a phrase-by-phrase self-paced reading task.

Analysis of three types of parsing models that have been proposed (serial garden-path models, competitive constraint satisfaction models, and race models) indicated that the three types of models make differing patterns of predictions for reading times (see Table 1; see also Farmer et al., 2007). Straightforward versions of garden-path models predict no cost of ambiguity in the ambiguous region, and no interaction of context and ambiguity in the disambiguating region. Competitive constraint-satisfaction models predict that both effects should appear, reflecting a uniform underlying mechanism of competition. Race models were argued to share the prediction of garden-path models in the ambiguous region and the prediction of constraint-satisfaction models in the disambiguating region.

The pattern of results of both experiments supported the predictions of the race model. The first finding is that reading time in the ambiguous region was not slowed, compared to the same region in which the syntactic ambiguity was blocked by unambiguous verb subcategorization. The second finding is that reading of the following region, which forced the direct object interpretation of any ambiguity, was substantially slowed when the earlier verb permitted a sentence complement analysis and the preceding discourse context failed to provide strong evidence favoring a direct object interpretation. The current findings are interestingly similar to those obtained by Wilson and Garnsey (2009). That is, our second finding is consistent with their findings in emphasizing the frequently mentioned effect of non-syntactic factors on processing the disambiguating region. However, the current findings go beyond their results by permitting a comparison of temporarily ambiguous and unambiguous sentences that are resolved toward the normally-preferred direct object interpretation (and, as discussed earlier, the first finding addresses a topic that the Wilson and Garnsey experiment was not designed to address). The lack of ambiguity or contextual effects in this region indicates that the processor constructs a single analysis in early stage. Accordingly, while Wilson and Garnsey’s results in the disambiguating region were taken as evidence to support constraints satisfaction model, our results in both ambiguous and disambiguating region support a more complicated pattern. On one hand, the first finding is inconsistent with the predictions of competitive constraint satisfaction models (both as generally understood, as argued earlier, and as explicitly predicted by the most adequate implemented model; Elman et al., 2005). It is, however, consistent with both garden-path and race models. The second finding, on the other hand, is consistent with competitive constraint-satisfaction and race models, but not with unelaborated garden-path models.

The present data thus indicate that the process of building an initial analysis may not be guided solely by grammatical information, as claimed by early serial garden-path models. Moreover, a core conclusion of the research presented here is that the conceptually-attractive claim of competitive constraint-satisfaction models – that the same mechanism of competition is responsible for both the disruption seen in recovery from a syntactic misanalysis (a garden-path) and the presumed disruption seen during the ambiguity itself – may not be correct. We realize that our current analyses cannot establish that there is no effect of ambiguity in the ambiguous region, but we do note that the pattern of results was replicated in two experiments and extended to a second region in Experiment 2. We further note that the experiments were sensitive enough to detect effects of context in the ambiguous region and ambiguity effects in the disambiguating region. Moreover, this absence of ambiguity effects in the ambiguous region seems to be consistent with ambiguous region data presented Hare et al (2003), as discussed in the Introduction., and with the absence of a verb bias effect in the ambiguous region reported by Wilson and Garnsey (2009).

We note that a race model could support the expectation of faster reading times in the postverbal noun phrase region when two analyses are available than when only one is available (as was found for different constructions by Traxler et al., 1998 and by van Gompel et al., 2005). The only sign of such an effect was the fast reading times in the ambiguous and the extended-ambiguity regions of the preferred-context condition of Experiment 2, but given the absence of this effect in Experiment 1, we are not inclined to give it much credence.

The manipulation of context in the present experiments has some potentially-interesting properties. As mentioned earlier, most manipulations of context have varied the felicity of definite noun phrase reference (e.g., whether or not there is a unique referent for a definite noun phrase; Altmann & Steedman, 1988) and less often the discourse pragmatics (e.g., whether or not a phrase satisfies a preceding indirect question; Altmann et al., 1998). Here, we manipulated whether or not the propositional content of the preceding paragraph supported the eventual interpretation of a postverbal noun phrase as the theme of that verb or as the subject of a sentential complement. The manipulation clearly impacted the reading of the phrase that disambiguated in favor of the normally-preferred direct object interpretation. When context supported the direct object interpretation, reading time was not affected by temporary ambiguity. This is consistent with all the models we have considered.

However, when context provided some support for the normally-unpreferred sentence complement analysis, temporary ambiguity resulted in slowed reading of the disambiguating region, an apparent garden path. This finding poses difficulty for the garden-path theory. At the least, it requires a garden-path theorist to claim that, faced with conceptual conflict, the processor has to go back and consider how the verb could have supported an analysis other than the initial direct object analysis (cf. Rayner et al., 1983, for some suggestions in this direction). The finding, however, is easily accommodated by a constraint satisfaction approach (the conflict between verb subcategorization and context, on the one hand, and the unambiguous content of the disambiguating region, on the other hand, slows reading) and a race model which permits contextual factors to affect which analysis ‘wins the race.’)

We acknowledge that some matters are yet to be resolved. One is that our finding that context did affect the reading time of a region that disambiguated in favor of the normally-preferred analysis seems to conflict with the findings of Duffy et al. (2001). These authors studied the reduced relative/main verb ambiguity (Bever, 1970) and found that the region that disambiguated toward the main verb analysis was read equally fast when referential context supported the reduced relative analysis as when it supported the main verb analysis.

We suggest that this may simply reflect the fact that the reduced relative vs. main verb construction used by Duffy et al. may be much more strongly biased toward one analysis than the direct object/sentence complement construction we used (cf. the results of Altmann et al., 1998, who had to reduce the strength of a late closure or local attachment bias by lengthening intervening material to allow pragmatic context to affect processing of the normally-preferred local attachment). We think that the argument that Duffy et al. mounted against competition accounts is still valid, but acknowledge that context may still have an effect on the choice of an initial analysis when the out-of-context bias is not extreme.

A second unresolved matter is the lack of any observed difference between the conflicted and unpreferred contexts in the present experiments. A competitive constraint satisfaction position would seem to predict more conflict in the region that disambiguates toward the direct object analysis when it followed the unpreferred conflict than when it followed the conflicted context, and a race model would seem to predict that the normally-unpreferred sentence complement analysis would have been chosen more often following the unpreferred than the conflicted context. However, there was no evidence for more disruption following an ambiguous verb in the unpreferred context than in the conflicted context. As long as we accept that our off-line norming results are valid estimates of on-line interpretive preferences, we have no satisfying account of the lack of a difference between the contexts (or, indeed, of the notably fast reading of the disambiguating region following an unambiguous verb in conflicted and unpreferred contexts, as compared to the preferred context). We must acknowledge that our rather complex manipulation of the propositional content of preceding context has some effects that are not yet understood.

Conclusions

The observed pattern of how ambiguity and context affected reading time in syntactically-ambiguous and disambiguating regions is consistent with unrestricted race models of parsing, but not with simple serial garden-path models or with competitive constraint satisfaction models as they have been developed. Ambiguity per se did not slow reading an ambiguous region, but ambiguity in concert with discourse bias did affect reading of a disambiguating region. Our data argue against the attractively-simple claim that competition between alternative analyses underlies the frequently-reported slowing of reading following a syntactic misanalysis. They also argue against the attractively-simple claim that an initial analysis of a phrase is built with reference only to grammatical (especially syntactic) information and only later interpreted in the light of extragrammatical information. They are best understood in terms of an unrestricted race model, in which multiple sources of information are used in a non-competitive fashion to construct an initial analysis, and where the first-constructed analysis is then interpreted in the light of all available information and, if necessary, revised.

Acknowledgments

This research was reported as a part of a Masters thesis submitted by Mohamed Taha Mohamed to the University of Massachusetts. It was supported in part by Grant Number HD18708 from NICHD to the University of Massachusetts. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of NICHD or NIH. We thank Adrian Staub for his comments on earlier versions of this manuscript, especially for advocating the race model we present.

References

Altmann GTM. Ambiguity in sentence processing. Trends in Cognitive Sciences. 1998; 2:146–152. [PubMed: 21227111]

Altmann G, Steedman M. Interaction with context during human sentence processing. Cognition. 1988; 30:191–238. [PubMed: 3215002]

Altmann GTM, van Nice KY, Garnham A, Henstra JA. Late closure in context. Journal of Memory and Language. 1998; 38:459–484.

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language. 2008; 59:390–413.

Bates, E.; MacWhinney, B. Functionalist approaches to grammar. In: Wanner, E.; Gleitman, L., editors. Language acquisition: The state of the art. New York: Cambridge University Press; 1982.

Bever, TG. The cognitive basis for linguistic structures. In: Hayes, JR., editor. Cognition and the development of language. New York: Wiley; 1970. p. 279-352.

Binder K, Duffy S, Rayner K. The effects of thematic fit and discourse context on syntactic ambiguity resolution. Journal of Memory and Language. 2001; 44:297–324.

Clark, HH.; Clark, EV. Psychology and language. New York: Harcourt Brace Jovanovich; 1977.

Clifton, C., Jr; Staub, A.; Rayner, K. Eye movements in reading words and sentences. In: Gompel, RV.; Fisher, M.; Murray, W.; Hill, RL., editors. Eye movement research: Insights into mind and brain. New York: Elsevier; 2007. p. 341-371.

Clifton C Jr, Staub A. Parallelism and competition in syntactic ambiguity resolution. Language and Linguistics Compass. 2008; 2:234–250.

Clifton C Jr, Traxler M, Mohamed MT, Williams RS, Morris RK, Rayner K. The use of thematic role information in parsing: Syntactic processing autonomy revisited. Journal of Memory and Language. 2003; 49:317–334.

Cottrell, GW. Connectionist parsing. Paper presented at the Seventh Annual Meeting of the Cognitive Science Society; Irvine, CA. 1985.

Crawley, MJ. The R Book. Chichester, England: John Wiley & Sons; 2007.

Cuetos F, Mitchell DC. Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. Cognition. 1988; 30:73–105. [PubMed: 3180704]

Tomasello, M.; Slobin, D. Cues, constraints, and competition in sentence processing. In: Elman, JL.; Hare, M.; McRae, K., editors. Beyond nature-nurture: Essays in honor of Elizabeth Bates. Mahwah, NJ: Lawrence Erlbaum Associates; 2005. p. 111-138.

Farmer TA, Cargill SA, Hindy NC, Dale R, Spivey MJ. Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. Cognitive Science. 2007; 31:889–910. [PubMed: 21635321]

Ferreira F, Henderson J. The 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. 1990; 16:555–568.

Fodor, JA.; Bever, TG.; Garrett, MF. The psychology of language: An introduction to psycholinguistics and generative grammar. New York: McGraw-Hill; 1974.

Fodor, JD.; Ferreira, F., editors. Sentence reanalysis. Dordrecht: Kluwer Academic Publishers; 1998.

Forster, KI. Levels of processing and the structure of the language processor. In: Cooper, WE.; Walker, ECT., editors. Sentence Processing. Hillsdale, NJ: Erlbaum Press; 1979. p. 27-86.

Frazier, L. Exploring the architecture of the language system. In: Altmann, G., editor. Cognitive models of speech processing: Psycholinguistic and computational perspectives. Cambridge, MA: MIT Press; 1990. p. 409-433.

Frazier L, Calson K, Clifton C Jr. Prosodic phrasing is central to language comprehension. Trends in Cognitive Sciences. 2006; 10:244–249. [PubMed: 16651019]

Frazier L, Fodor JD. The sausage machine: A new two-stage parsing model. Cognition. 1978; 6:291–326.

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

Gahl S, Jurafsky D, Roland D. Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. Behavior Research Methods, Instruments, & Computers. 2004; 36:432–443.

Garnsey SM, Pearlmutter NJ, Myers E, Lotocky MA. The contributions of verb bias and plausibility to the comprehension of temporarily-ambiguous sentences. Journal of Memory and Language. 1997; 37:58–93.

Gibson, E. Doctoral dissertation. Carnegie Mellon University; 1991. A computational theory of human linguistic processing: Memory limitations and processing breakdown.

Gibson E, Pearlmutter NJ. Constraints on sentence comprehension. Trends in Cognitive Science. 1998; 2:262–268.

Gibson E, Pearlmutter NJ. Distinguishing serial and parallel processing. Journal of Psycholinguistic Research. 2000; 29:231–240. [PubMed: 10709187]

Green MJ, Mitchell DC. Absence of real evidence against competition during syntactic ambiguity resolution. Journal of Memory and Language. 2006; 55:1–17.

Hale J. A probabilistic Earley parser as a psycholinguistic model. Proceedings of NAACL. 2001; 2:159–166.

Hare M, McRae K, Elman JL. Sense and structure: Meaning as a determinant of verb subcategorization preferences. Journal of Memory and Language. 2003; 48:281–303.

Harley, TA. The psychology of language: From data to theory. 3. Hove, UK: Psychology Press; 2008.

Jurafsky D. A probabilistic model of lexical and syntactic access and disambiguation. Cognitive Science. 1996; 20:137–194.

Kennison SM. Limitations on the use of verb information during sentence comprehension. Psychonomic Bulletin & Review. 2001; 8:132–137. [PubMed: 11340858]

Lewis RL. Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. Journal of Psycholinguistic Research. 2000; 29:241–248. [PubMed: 10709188]

MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. Pyschological Review. 1994; 101:676–703.

McClelland, JL.; Rumelhart, DE. the PDP Research Group. Psychological and Biological Models. Vol. 2. Cambridge, MA: MIT Press; 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition.

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

Mitchell, DC. On-line methods in language processing: Introduction and historical review. In: Carreiras, M.; Clifton, C., Jr, editors. The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond. Brighton, UK: Psychology Press; 2004. p. 15-32.

Myers, J.; Well, AD. Research design and statistical analysis. 2. Mahwah, NJ: Lawrence Erlbaum Associates; 2003.

Ni W, Crain S, Shankweiler D. Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibilty in resolving ambiguities. Language and Cognitive Processes. 1996; 11:283–334.

Pickering MJ, Traxler MJ, Crocker MW. Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. Journal of Memory and Language. 2000; 43:447–475.

Rayner K, Carlson M, Frazier L. The interaction of syntax and semantics during sentence processing. Journal of Verbal Learning and Verbal Behavior. 1983; 22:58–374.

Rayner K, Frazier L. Parsing temporarily ambiguous complements. Quarterly Journal of Experimental Psychology. 1987; 39A:657–673.

Rumelhart, DE.; McClelland, JL. the PDP Research Group. Foundations. Vol. 1. Cambridge, MA: MIT Press; 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition.

Staub A. The parser doesn't ignore intransitivity, after all. Journal of Experimental Psychology: Learning, Memory and Cognition. 2007; 33:550–569.

Stevenson S. Competition and recency in a hybrid network model of syntactic dependencies. Journal of Psycholinguistic Research. 1994; 23:295–322.

Tabor W, Tanenhaus MK. Dynamical models of sentence processing. Cognitive Science. 1999; 23:491–516.

Traxler M, Pickering M, Clifton C Jr. Adjunct attachment is not a form of lexical ambiguity resolution. Journal of Memory and Language. 1998; 39:558–592.

Trueswell, JC.; Tanenhaus, MK. Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In: Clifton, C.; Rayner, K.; Frazier, L., editors. Perspectives on sentence processing. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994. p. 155-179.

Trueswell JC, Tanenhaus MK, Kello C. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1993; 19:528–553.

Tyler L, Marslen-Wilson W. The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior. 1977; 16:683–692.

van Gompel RPG, Pickering M, Pearson J, Liversedge SP. Evidence against competition during syntactic ambiguity resolution. Journal of Memory and Language. 2005; 52:284–307.

van Gompel RPG, Pickering M, Traxler M. Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. Journal of Memory and Language. 2001; 45:225–258.

Vosse T, Kempen G. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and lexicalist grammar. Cognition. 2000; 75:105–143. [PubMed: 10771275]

Vosse T, Kempen G. In defense of competition during syntactic ambiguity resolution. Journal of Psycholinguistic Research. 2009; 38:1–9. [PubMed: 18521754]

Wilson M, Garnsey S. Making simple sentences hard: Verb bias effect in simple direct object sentences. Journal of Memory and Language. 2009; 60:368–392. [PubMed: 20160997]

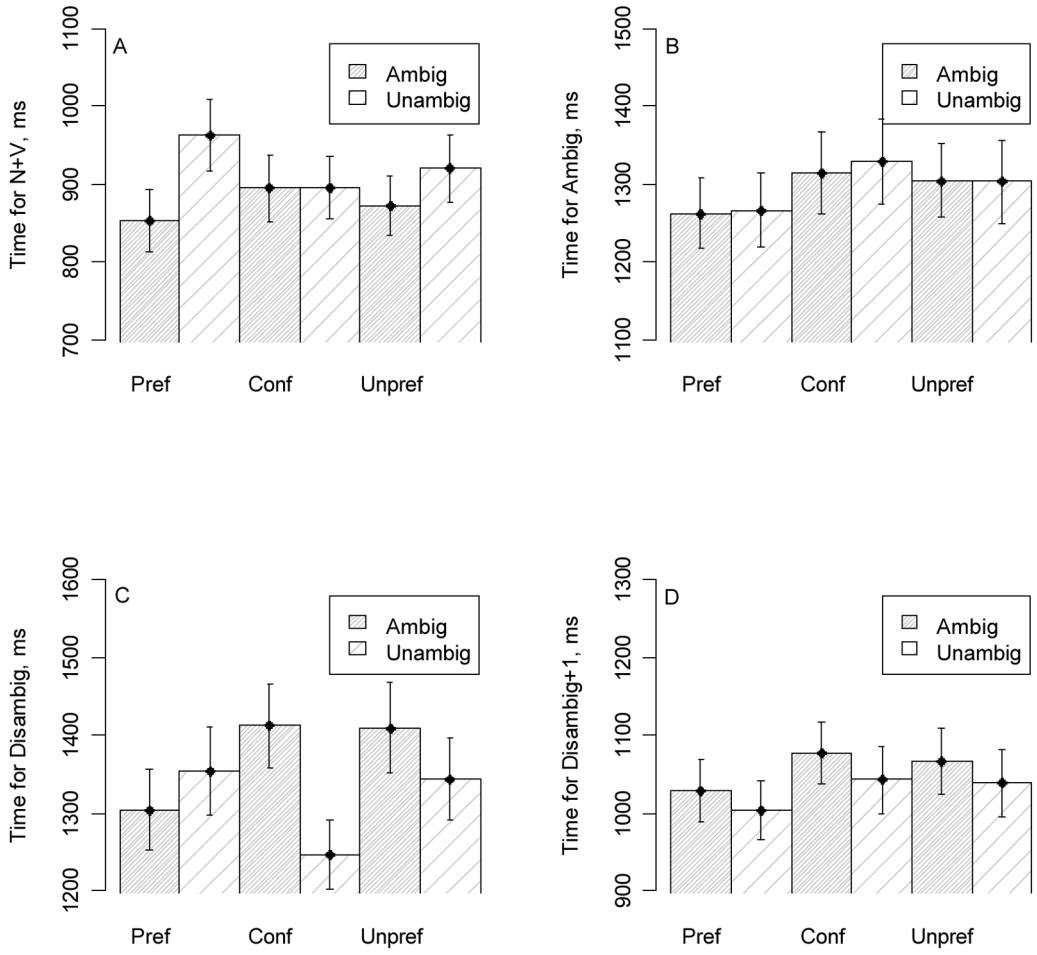

Figure 1.
Average Reading Times, Experiment 1 (with standard error bars). Panel A: Noun+verb region. Panel B: Ambiguous region. Panel C: Disambiguating region. Panel D: Post-disambiguating region.

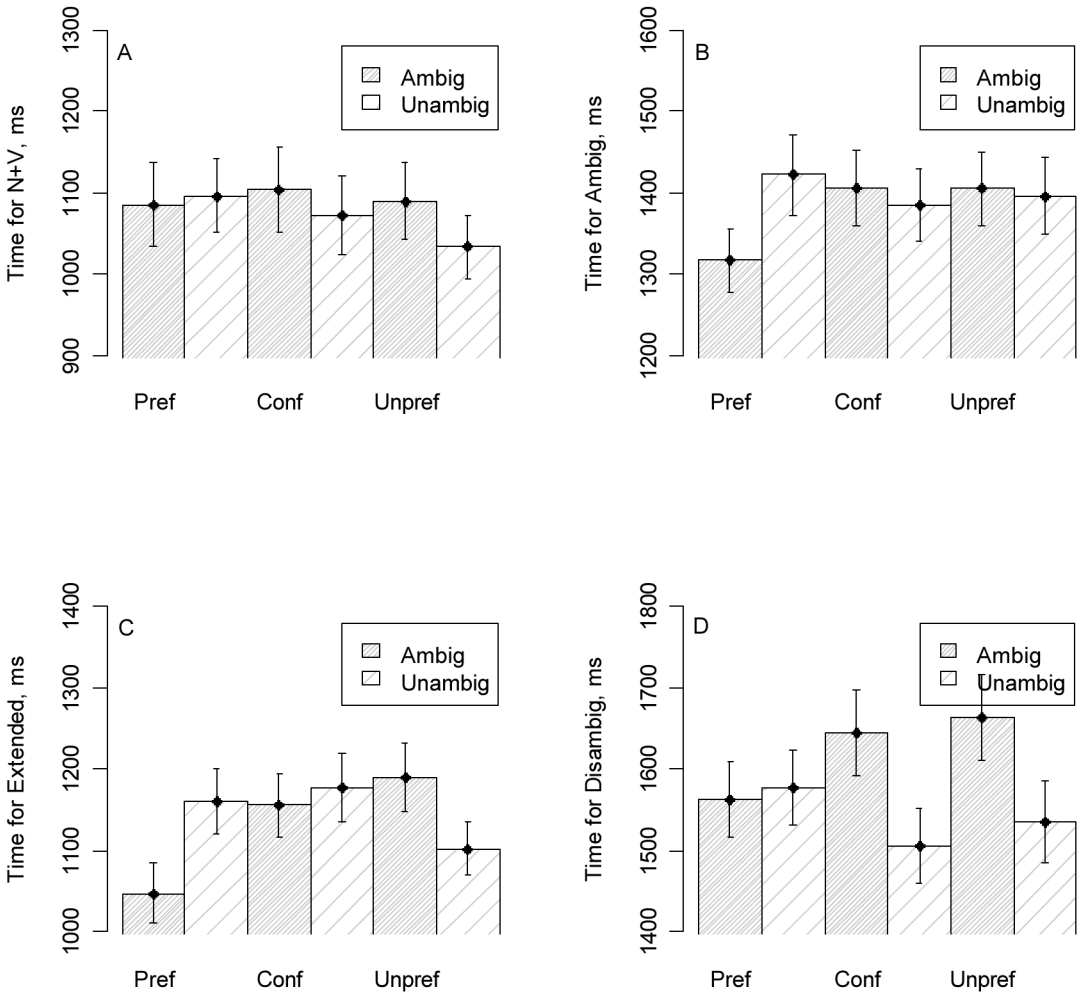

Figure 2.
Average Reading Times, Experiment 2 (with standard error bars). Panel A: Noun+verb region. Panel B: Ambiguous region. Panel C: Extended ambiguous region. Panel D: Disambiguating region.

Table 1

Synopsis of predictions of serial garden-path, parallel competitive, and unrestricted race models, for effects of verb ambiguity and its interaction with biasing context, assuming biasing context toward sentence complement analysis and disambiguation to simpler transitive analysis. “slowed” in Biasing context x ambiguity = reading time following ambiguous verb particularly slowed in biasing context.

Model	**Region**	
	Ambiguous	**Disambiguating**
Garden-path model		
Verb ambiguity	no effect	no effect
Biasing context x ambiguity	no effect	no effect
Parallel competition model		
Verb ambiguity	slowed	slowed
Biasing context x ambiguity	slowed	slowed
Unrestricted race model		
Verb ambiguity	no effect (or faster)	slowed
Biasing context x ambiguity	no effect	slowed

Table 2

Verb Bias (Percentage) of Verbs Used in the Materials. Also Shows Percentage of NP/(NP+SC) for Both Norms, Pooled when Available

Verb	Garnsey et al.		Kennison		Mean %
	NP	SC	NP	SC	NP
Admit	9	60	14	43	18%
Announce	49	48	75	19	65%
Believe	14	50	39	29	40%
Discover	69	30	59	25	70%
Emphasize	75	19	17	71	51%
Expect			26	10	28%
Feel	12	11	14	7	59%
Figure	8	46			15%
Guess	39	25	45	9	71%
Indicate	21	70	21	67	23%
Know	31	46			40%
Prove	23	61	61	20	51%
Recommend			55	45	55%
Regret	17	9	82	6	87%
Suggest	18	61	33	59	30%

Q J Exp Psychol (Hove). Author manuscript; available in PMC 2012 July 05.

Table 3

Example of the Material. Critical sentence is **boldfaced**. Alternative verbs separated by /. Analysis regions delimited by ~. Extended ambiguity region noted by { } Relevant differences are *italicized*. Common ending to all texts appears at the bottom.

Preferred Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money collected. Many people began to talk about serious charges against him. *Over Mr. John's opposition and his suspicious resistance,* an independent judicial inspector was appointed to investigate the whole issue.
Conflicted Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money he collected. Many people began to talk about serious charges against him. *Whether he was honest or not was a serious controversial issue. In search for the truth,* an independent judicial inspector was appointed to investigate the whole issue.
Unpreferred Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money collected. Many people began to talk about serious charges against him. *As a result of Mr. John's confident insistence on proving his innocence,* an independent judicial inspector was appointed to investigate the whole issue.
Common Ending: Finally, **~ the independent inspector proved/supported ~ the charges against Mr. John ~{that were being investigated~} and issued his final report accordingly.~ Mr. John is about to resign~** and may be sued. Many of his advocates were astonished to learn that. Others, though, were skeptical about this conclusion.
Question:
The inspector found Mr. John is guilty? True False

Table 4

Example of material used in norming study 1. Common ending to all texts and response options appear at the bottom.

Preferred Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money collected. Many people began to talk about serious charges against him. Over Mr. John's opposition and his suspicious resistance, an independent judicial inspector was appointed to investigate the whole issue.
Conflicted Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money collected. Many people began to talk about serious charges against him. Whether he was honest or not was a serious controversial issue. In search for the truth, an independent judicial inspector was appointed to investigate the whole issue.
Unpreferred Context: Mr. John is the president of our local association for animal rights. He is responsible for fund raising activities and promoting the contributions of members and non-members to support the activities of the association. Last year, some questions were raised about the honesty of Mr. John and how he dealt with the money collected. Many people began to talk about serious charges against him. As a result of Mr. John's confident insistence on proving his innocence, an independent judicial inspector was appointed to investigate the whole issue.
Common Ending: Finally, the independent inspector proved the charges against Mr. John
(1) and issued his final report accordingly.
(2) were false and issued his final report accordingly.

Table 5

Results of Norming Study 1

	Completion	
Context	**Preferred Completion**	**Unpreferred Completion**
Conflicted Context	59.78%	40.22%
Preferred Context	71.20%	28.80%
Unpreferred Context	34.78%	65.22%

Table 6

Results of Norming Study 2 (7 = “extremely plausible”)

Experimental condition	Ratings
Ambiguous sentence in preferred context	6.2
Unambiguous sentences in preferred context	6.3
Ambiguous sentences in conflicted context	6.5
Unambiguous sentences in conflicted context	5.8
Ambiguous sentences in unpreferred context	6.5
Unambiguous sentence in unpreferred context	5.5

Table 7

Regions in Critical Sentences, Experiments 1 and 2

Experiment	Region			
1	Subject+Verb They proved	Ambiguous the theory under discussion	Disambiguating and their results were confirmed	
2	Subject + Verb They proved	Ambiguous the theory under discussion	Extended ambiguity by the audience	Disambiguating and their results were confirmed