Journal of Memory and Language **46,** 267–295 (2002)
doi:10.1006/jmla.2001.2808, available online at http://www.academicpress.com on IDEAL®

Syntactic Complexity in Ambiguity Resolution

Daniel Grodner, Edward Gibson, and Susanne Tunstall

Massachusetts Institute of Technology

This article presents two self-paced reading experiments which investigate the role of storage costs associated with maintaining incomplete syntactic dependencies in structural ambiguity resolution. We argue that previous work has been equivocal regarding syntactic influences because it has examined ambiguities where there is little or no resource differential between competing alternatives. The candidate structures of the ambiguities explored here incur substantially different storage costs. The results indicate that storage-based biases can be sufficiently powerful to create difficulty for a structural alternative even when it is promoted by nonsyntactic factors. These findings are incorporated into a model of ambiguity resolution in which structural biases operate as independent graded constraints in selecting between structural alternatives.

Key Words: sentence comprehension; ambiguity resolution; syntactic complexity.

Perhaps the quintessential computational property of natural language is the ability to productively combine linguistic units into larger linguistic units. These computations are particularly prominent at the level of syntax. This, in part, is why syntactic descriptions have traditionally played a privileged role in the study of structural ambiguity resolution. For instance, representations at the level of syntax are typically thought to determine, and distinguish between, the structural alternatives compatible with an input word string. It is not then surprising that a number of investigators have sought syntactic explanations for the resolution of syntactic-level ambiguity. These theorists have posited the existence of specialized syntactic parsing principles that operate over phrase markers (e.g., Frazier, 1979, 1987; Gorrell, 1995), transformations (Fodor, Bever, & Garrett, 1974), thematic grids (Pritchett, 1988; Gibson, 1991), or other formal linguistic representations of a sentence to explain why one structural alternative is favored over another. While such approaches differ in many respects, they all emphasize the role of intrinsic architectural constraints arising from the combinatorial properties of language.

In recent years a great deal of evidence has accumulated demonstrating that nonstructural influences are integral to explaining syntactic ambiguity resolution (see Gibson & Pearlmutter, 1998; Tanenhaus & Trueswell, 1995, for summaries of relevant evidence). This work has shifted the focus of research away from syntactic parsing principles toward lexical and contextual factors. To give one illustration, Trueswell, Tanenhaus, and Garnsey (1994) demonstrated that thematic information heavily influences reading patterns for the main verb/reduced relative clause (MV/RR) ambiguous sentences as in (1).

(1) a. The evidence examined by the lawyer turned out to be unreliable.

b. The defendant examined by the lawyer turned out to be unreliable.

The NP "the evidence" is a poor agent of "examined" because inanimate objects do not nor-

We thank the following people for helpful discussions on presentations of this work: the audience at the 10th CUNY Human Sentence Processing Conference, Neal Pearlmutter, Colin Phillips, Carson Schütze, Lyn Frazier, and two anonymous reviewers. We also acknowledge the assistance of Andrew Nevins and Anna Lane, who helped compile head bias statistics. All errors are our own. Funding for this work was provided by a grant provided by the National Science Foundation to the second author, award SBR-9729037. The first author was supported by the Research Training Grant "Language: Acquisition and Computation" awarded by the National Science Foundation to the Massachusetts Institute of Technology (DIR 9113607). Additional funding for the project was provided by the JST/MIT joint international research project "Mind/Articulation."

Address correspondence and reprint requests to Daniel Grodner, Department of Brain and Cognitive Sciences, NE20-437D, MIT, Cambridge, MA 02139. Fax: (617) 258-8654. E-mail: dgrodner@psyche.mit.edu

0749-596X/01 $35.00

mally examine things. At the same time, evidence is something that might plausibly undergo examination. This comports well with the thematic structure of the RR reading, but is inconsistent with a MV interpretation. In contrast, "the defendant" in (1b) could easily engage in either examining something or being examined by something. This is consistent with the thematic relationships required by either structural alternative. Eye-tracking measures reveal that individuals do not experience measurable difficulty relative to unambiguous controls in processing sentences like (1a), where thematic fit promotes the RR reading, but they do experience difficulty in sentences like (1b). Other nonsyntactic factors such as verb argument structure frequency (MacDonald, 1994; Pearlmutter & MacDonald, 1995), the frequency that the verb appears as a past tense or past participle (Trueswell, 1996), the availability of postambiguity constraints (MacDonald, 1994; Trueswell et al., 1994), the semantics of the determiner in the NP preceding the ambiguous verb (Ni, Crain, & Shankweiler, 1996), and referential context (Britt, Perfetti, Garrod, & Rayner, 1992; Ferreira & Clifton, 1986; Rayner, Garrod, & Perfetti, 1992; Spivey-Knowlton & Tanenhaus, 1994) also contribute to reading patterns in processing the MV/RR ambiguity. It can be concluded that parsing preferences are to a certain extent governed by nonstructural influences.

Lexical and discourse factors contribute to the resolution of several other ambiguities as well. Referential information and lexically specific frequency and/or argument structure information are sufficient to account for prepositional phrase attachment preferences (Schütze & Gibson, 1999; Spivey-Knowlton & Sedivy, 1995). Verb bias and thematic plausibility play primary roles in determining preferences in the direct object/sentential complement ambiguity (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell, Tanenhaus, & Kello, 1993). Lexical frequency factors and plausibility explain much of the resolution of the noun–noun/noun–verb ambiguity (Frazier & Rayner, 1987; MacDonald, 1993).

Because preferences for the MV/RR and other ambiguities are so susceptible to manipulations of lexical factors, it may be that structural biases play only a modest role in ambiguity resolution. Some investigators have endorsed a particularly strong version of this idea: "Reinterpreting syntactic ambiguity resolution as a form of lexical ambiguity resolution obviates the need for special parsing principles to account for syntactic interpretation preferences. . ." (MacDonald, Pearlmutter, & Seidenberg, 1994, p. 676) and "There does not appear to be a role for purely structurally defined parsing principles (i.e., minimal attachment)" (Spivey-Knowlton & Sedivy, 1995, p. 227).

Even those few constraint-based lexicalist models that include a structural component rarely specify the nature of this constraint. For instance, referring to a bias in their model to interpret an initial *NP-Ved* sequence as a MV or RR, Spivey and Tanenhaus (1998) acknowledge the following: "We remain agnostic about whether this configurational bias is best characterized at the structural level or emerges from other more local constraints" (p. 1524).

The nature and importance of structural biases is clearly a murky issue. Though most researchers acknowledge a role for structural complexity, it is not generally thought to arise from any resource limitations imposed on language-specific computations. Rather, structural complexity is often conceived of as a by-product of domain general mechanisms. One such domain general mechanism is locality (Gibson, 1998, 2000; see also Frazier's, 1979, late closure; Kimball's, 1973, right association; and Gibson's, 1991, recency). Locality specifies that the cost associated with integrating new material into a partial interpretation of a sentence increases monotonically with the distance [measured in words (e.g., Hawkins, 1994) or discourse referents (e.g., Gibson, 1998)] between the site of integration and the new element. In ambiguity resolution, this cost induces a graded preference for new input material to attach to more recent potential attachment sites. The clearest evidence for this bias comes from investigations of ambiguous modifier attachment (Altmann, van Nice, Garnham, & Henstra, 1998; Gibson, Pearlmutter, & Torrens, 1999; Pearlmutter & Gibson, 2000). No known combination of lexical and

contextual factors can account for the array of phenomena that has been explained by locality. Thus, locality appears to be a structurally based bias, but not necessarily a bias specific to *syntactic* structure. Locality arises because linguistic units are, to a first approximation, perceived in succession. As mental representations degrade over time, more recently encountered sentential material will be more accessible as target sites of attachment. Construed in this way, locality could be reflective of a domain general principle applicable to any serialized information processing. It would not then derive from any intrinsic property of syntactic combination.

This article presents evidence supporting the use of a second resource-based structural bias that derives from the combinatorial aspect of language. We hypothesize that maintaining a partial syntactic representation incurs a storage cost for each syntactic head necessary to complete a (partial) input string as a grammatical utterance. For instance, after encountering a sentence-initial NP (e.g., "Mary"), only a predicate is required to construct a sentence. One unit of storage cost is associated with this input because the role of predicate can be fulfilled by a single verbal head (e.g., "sleeps"). Such syntactic predictions serve to guide the interpretation of further, possibly ambiguous, lexical material in light of the partial structural context. Storage costs are specific to language qua syntax because they apply to intrasentential dependencies between syntactic heads. No other level of linguistic or nonlinguistic representation makes explicit these sorts of intrasentential relations.

Storage costs and locality-based integration costs form the core of the Dependency Locality Theory (DLT) of unambiguous sentence complexity (Gibson, 1998, 2000). The hypothesis pursued presently is that the same graded resource constraints operate in structural ambiguity resolution. In this framework, the DLT acts as an independent factor, as summarized in (2) (adapted from Gibson):

(2) Ambiguity Resolution Hypothesis

In choosing among structures consistent with an ambiguous input, two of the factors that the processor uses to evaluate its choices are DLT storage and integration costs (in addition to lexical frequencies, plausibility, and discourse context). When conflicts arise between minimizing storage cost and integration cost, storage costs should be minimized.

Note that storage costs are assessed for incomplete dependencies. These dependencies represent lexical heads that will need to be integrated into the current structure at some point. By minimizing storage costs at the current word, the parser can limit the number of potentially long-distance integrations to be made later in the sentence. This provides the motivation for ordering storage costs before integration costs. As a consequence, perceivers should associate fillers with gaps as soon as grammatically possible in order to disburden themselves the cost of a gap site prediction. This is in accordance with evidence supporting the Active Filler Strategy (Clifton & Frazier, 1989; Frazier, 1987; Frazier & Flores d'Arcais, 1989). The DLT therefore provides a theory of resource allocation that explicates this strategy, much in the same spirit as the Minimal Chain Principle (De Vincenzi, 1991).

For expository purposes, we discuss the effects of storage cost within a framework where multiple analyses of an ambiguous string compete for activation in parallel. Given the assumption of a framework in which alternative structures are constructed in parallel, storage costs for these partial structures can be compared from the onset of ambiguity. When the difference in storage costs is large relative to the influence of other factors, the candidate requiring the least storage cost will be more accessible than other candidates. When this difference is smaller, lower ranked alternatives will be more accessible, depending on the strength of other constraints. Within a parallel framework, the accessibility of the correct interpretation is a central factor in determining how difficult it will be to recover from misanalysis. In cases where all factors support a particular alternative, the subordinate alternative will be effectively inaccessible.[1]

The DLT predicts that the bias from structural complexity will be small for the MV/RR construction in (1). Integration costs at the point where the ambiguity is introduced [at "examined" in (1)] are identical across the two readings. For the MV reading, the verb is integrated thematically and structurally to the subject NP. This link crosses no intervening lexical material. The RR interpretation also requires linking the verbal head to the preceding NP. All immediate integrations are local for both readings. Storage costs are also comparable for the structural alternatives. The MV reading requires only an object NP to complete the input as a grammatical string. This requirement could be satisfied by a single lexical head (e.g., "papers"). Under the RR reading, the input string requires a matrix predicate and perhaps an embedded modifier to be grammatical.[2] Thus the storage-cost difference between the structural alternatives is zero or one predicted lexical head. This difference may be small enough to be overwhelmed by nonstructural factors in many cases. Other ambiguous constructions where lexical and discourse manipulations are known to strongly affect parsing preferences have similarly small resource cost differentials. This is true for the prepositional phrase attachment ambiguity, the noun–noun/noun–verb ambiguity, and the direct object/sentential complement ambiguity.

In sum, previous work has been equivocal regarding the role of structural influences other than locality because it has focused on ambiguities with little or no resource differential between alternatives. Here we present two reading experiments on ambiguities in which substantial syntactic storage cost differences between the candidate interpretations are predicted. In each of these experiments nonstructural constraints were aligned in favor of the more costly construction, and thereby, pitted against the structural bias. The results indicate that storage-based biases can be sufficiently powerful to create difficulty for a structural alternative even when it is promoted by nonsyntactic factors.

EXPERIMENT 1

Though the structural bias for the MV/RR is small in a matrix sentence context, consider the same ambiguity embedded within a relative clause, as in (3):

(3) a. The witness who the defendant examined was lying.

b. The witness who the defendant examined by the lawyer implicated was lying.

Storage costs for the two readings at "examined" diverge more than they do for the respective readings of the matrix version of the ambiguity (1). The MV reading (3a) requires only one unrealized syntactic head to become grammatical: a matrix predicate ("was lying"). The RR reading (3b) entails three (possibly four) unsatisfied dependencies: (1) a matrix predicate ("was lying"), (2) an embedded predicate (satisfied by "implicated"), (3) an embedded gap site associated with the filler "who" (satisfied by the object of "implicated"),[3] and (4) possibly a modifier (see footnote 2). The storage-cost difference in (3) for the MV/RR embedded within a relative clause is therefore at least two dependencies (possibly three dependencies)—more

[1] Our results are also consistent with a framework where structures are generated in parallel but only a single alternative is maintained. One way to account for the graded nature of storage-cost biases within a serial framework is to probabilistically choose among the alternatives. Under such a model, the likelihood of choosing a particular structure would be related to its level of activation. An incorrect reading will elicit a greater misanalysis effect the more frequently it is chosen.

[2] RRs like (1) are slightly odd without a modifier following the embedded verb:

(i) ?The evidence examined turned out to be unreliable.

(ii) ?The money accepted was from an unscrupulous investor.

Condoravdi (1989) suggests that modifiers are required in middle constructions to occupy the role of nuclear scope (e.g., "?This bread cuts" vs "This bread cuts easily"). There may likewise be a grammatical requirement for modification of the verb in a reduced relative clause.

[3] Empty categories are not crucially assumed here. Any linguistic theory that involves dependency positions makes the right predictions (cf. Pickering & Barry, 1991; Gibson & Hickok, 1993).

than a difference of zero (possibly one) dependency in the matrix sentence context.

Note that there is a conflict between minimizing storage and integration costs at the point of encountering the verb "examined." Under the MV analysis (3a), the filler "who" is linked to the object role assigned by "examined." This integration crosses two new discourse referents, corresponding to the object indicated by the NP "the defendant" and the event indicated by the verb "examined." In contrast, integrations at the point of "examined" are local in the embedded RR (3b), just as they are for the unembedded RR (1a) above. Fulfillment of the filler-gap dependency is postponed in the RR analysis until several words (and two discourse referents) later. Hence, whereas storage costs favor the MV over the RR at this point, integration costs favor the RR over the MV. The Ambiguity Resolution Hypothesis (2) favors minimizing storage costs over minimizing immediate integration costs—thus favoring the MV resolution in (3)—because this will minimize integration costs across the sentence. When the RR filler-gap dependency is eventually satisfied, at "implicated," a larger integration cost will be incurred than for the MV analysis, in which the gap site is encountered at "examined."

Experiment 1 investigates reading behavior for temporarily ambiguous RRs in an embedded context. Items were created using those from Trueswell, Tanenhaus, and Garnsey (1994) where the combined effects of thematic typicality and frequency information reduced difficulty with the RR reading. For instance, in (4), as in (1a), the MV interpretation of "examined" is implausible, because "evidence" is not a good agent for "examined":

(4) The witness who the evidence examined by the lawyer implicated was lying.

Unlike (1) syntactic storage costs are strongly biased against the RR toward the MV in (4). Therefore the RR reading should be less available when the MV/RR is embedded within a relative clause in the present items than in Trueswell et al.'s unembedded items.

A control condition was included in the study, in which the MV/RR ambiguity was contained within a sentential complement of a verb as in (5):

(5) The witness thought that the evidence examined by the lawyer implicated his next-door neighbor.

DLT complexity costs for this ambiguity are the same as in the matrix MV/RR case in (1a). Thus, there is little or no structural bias in (5) for the MV, and reading behavior is expected to approximate that in the matrix context. Though Trueswell et al. found no garden path effect in matrix versions of the items employed here, they used an eye-tracking methodology which permitted preview of the disambiguating preposition at the point of the embedded verb. Presenting the disambiguating by-phrase together with the ambiguous verb has been shown to enhance the availability of the RR analysis (Burgess, 1991; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; MacDonald, 1994). In the present items no such preview was available, so it is possible that some difficulty might arise at disambiguation in the sentential complement condition. Nevertheless, assuming that the magnitude of the misanalysis effect is determined in part by the degree of commitment to the incorrect reading, this effect is predicted to be smaller than that for relative clause items like (4). The sentential complement condition also served to ensure that any difficulty observed with the RR in the relative clause context could not be attributed simply to the complexity of processing a multiclause sentence.

In summary, plausibility factors favor the RR in each of the two embedding conditions. Syntactic storage costs favor the MV reading in the relative clause embedding, but have a small or negligible bias in the sentential complement embedding. If syntactic storage costs influence ambiguity resolution, then more difficulty should arise when a temporarily ambiguous sentence is resolved toward a RR in the relative clause condition (4) than in the sentential complement condition (5).

Method

Participants. Sixty participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and naive as to the purpose of the study.

Materials. The items were constructed using the poor agent conditions from Trueswell et al.'s (1994) Experiment 1 as a base, so that plausibility factors favored the RR interpretation of the ambiguous word [e.g., "examined" in (1a)]. The sequence *The N Ved by the N* was embedded in two contexts: within a relative clause [RC, "the witness who. . ." in (6a) and (6b)] and after a verb which took a sentential complement [SC, "the witness thought that. . ." in (6c) and (6d)]. In this way, we created sentences in which there were large and small storage cost differences, respectively, between the MV and RR interpretations of the ambiguous verb. All items were resolved as RRs. Unambiguous versions of the sentences were formed by inserting the string "that was" prior to the ambiguous verb. The verb following the by-phrase ["implicated" in (6)] was the same in all conditions. According to the DLT, an interaction between structure (RC and SC) and ambiguity was anticipated at the by-phrase, with a larger effect of ambiguity in the high-cost RC conditions:

(6) a. RC (large storage cost difference), ambiguous

The witness who the evidence examined by the lawyer implicated seemed to be very nervous.

b. RC (large storage cost difference), unambiguous

The witness who the evidence that was examined by the lawyer implicated seemed to be very nervous.

c. SC (small storage cost difference), ambiguous

The witness thought that the evidence examined by the lawyer implicated his next-door neighbor.

d. SC (small storage cost difference), unambiguous

The witness thought that the evidence that was examined by the lawyer implicated his next-door neighbor.

There were 16 target items, which were combined with 56 fillers of various types. These included sentences similar to the target items except that the ambiguous verb was resolved as a matrix verb. The items were counterbalanced across four lists using a Latin-square design. In half of the targets, the initial noun ["witness" in (6a)–(6b)] was animate in all four conditions. In the other half, the initial noun was inanimate in the RC conditions and animate in the SC conditions. This animacy factor was balanced across lists. Eight different embedding verbs like "thought" were used in the SC conditions. Each verb was used twice in the materials but only once per list. Appendix A provides a complete list of the Experiment 1 stimuli.

Plausibility norm. As stated above, target items were constructed from Trueswell et al.'s (1994) items in which the NP preceding the ambiguous verb (NP_2) was a poor agent of that verb in order to support the RR reading. However, the addition of a relativized NP (NP_1) in the RC conditions of the present items calls for further consideration of plausibility factors. Under the MV interpretation, this initial NP plays the patient role for the embedded ambiguous verb, whereas in the RR interpretation NP_2 plays this role. As a result, the competition between the MV and RR readings is, in part, a competition between the first and second NPs acting as patient for the embedded verb. In order to establish that thematic plausibility did not support the MV reading (i.e. that plausibility was not aligned with the syntactic complexity bias), a survey was conducted to assess whether NP_1 or NP_2 was a better patient for the embedded verb. Twenty participants from the MIT community who did not take part in the main experiment were asked to respond to questions like the following: "Which is more typically examined?" Ratings were given on a 7-point scale with NP_1 (e.g., "a witness") at one end of the scale and NP_2 (e.g., "some evidence") at the other. Values ranged from 3 at either extreme, indicating a strong preference for

the NP corresponding to that end of the scale, and converged at 0 at the middle, indicating zero bias between the two options. Mean ratings for each item are given in Appendix A. As desired, there was an overall bias toward NP_2 as a better patient, with a mean of 0.84 ($SD = 0.32$ by participants, $SD = 1.25$ by items); this mean was significantly different from zero [$t1(19) = 11.7$, $p < .001$; $t2(15) = 2.67$, $p < .01$]. Note that at the point of encountering the ambiguous verb, readers are comparing the likelihood of Verb + NP_2 (the RR role assignment) to the likelihood of NP_2 + Verb + NP_1 (the MV role assignment). The set of events denoted by the latter (e.g., "evidence examining a witness") are included in the set of events denoted by Verb + NP_1 (e.g., "examining a witness"), the relation evaluated in this survey. As a result, the survey was a conservative filter because the actual likelihood of the MV reading is much lower than that for Verb + NP_1 alone.[4]

Procedure. The task was self-paced word-by-word reading with a moving window display (Just, Carpenter, & Woolley, 1982). Participants fell into two groups. One group was tested using a Macintosh computer running software developed in our lab. The second group was tested using an IBM-compatible computer running Micro-Experimental Laboratory (MEL) software. Each trial began with a series of dashes marking the length and position of the words in the sentences. Participants pressed the spacebar to reveal each word of the sentence. As each new word appeared, the preceding word disappeared. The amount of time a participant spent reading each word was recorded as the time between key-presses. After the final word of each item, a comprehension question appeared which asked about information contained in the preceding sentence. Participants pressed one of two keys to respond "yes" or "no." After an incorrect answer, the word "INCORRECT" flashed briefly on the screen. No feedback was given for correct responses. Participants were asked to read sentences at a natural rate and to be sure that they understood what they read. They were told to answer the questions as quickly and accurately as they could and to take wrong answers as an indication to read more carefully.

Up to 80 characters could appear on each line for the MEL display, and up to 100 characters could appear on each line for the Macintosh display. Each item spanned from one to one-and-one-half lines. The disambiguating by-phrase for all target items appeared on the first line. Items were pseudorandomized separately for each participant, with at least one filler sentence preceding each target.

Before the main experiment, a short list of practice items and questions was presented in order to familiarize the participant with the task. Participants took approximately 25 min to complete the experiment. For most participants, this experiment was combined with an unrelated self-paced reading experiment (using the same procedure), resulting in a session about 45 min long. Participants were given short breaks between the two experiments.

Results

Comprehension question performance. Questions for experimental items were answered correctly on 75.5% of trials. For the RC conditions, accuracy rates were 71.4 and 72.1% for the ambiguous and unambiguous items, respectively. For the SC conditions, accuracy rates were 80.0% for ambiguous items and 78.3% for unambiguous items. A two-factor ANOVA crossing structure (RC and SC) with ambiguity (unambiguous and ambiguous) on the question-answering data revealed that responses to SC questions were significantly more accurate than responses to RC questions [$F1(1,58) = 7.01$, $MSE = 0.047$, $p = .01$, $F2(1,15) = 4.96$, $MSE = 0.032$, $p < .05$]. This difference is unsurprising because the RC structure involve an extra level of nested structure and are therefore more difficult to understand than the SC structure. Of the 16 item questions, 14 queried role

[4] For two individual items, NP_1 was judged a significantly better patient of the verb than NP_2: "a textbook" was a less typical patient of "love" than "a scientist," and a "a sofa" was less likely to be "scratched" than "a table." However, the full MV thematic assignments for these items were highly implausible (i.e., "a table scratching a sofa" and "a textbook loving a scientist"). In any case, when analyses of reading times were performed without these two items, the results reported below were unaffected.

assignments to the embedded verb. Accuracy rates on these items closely resembled overall accuracy rates.

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 derived 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. All items were analyzed, regardless of how the comprehension question was answered. Residual reading times beyond 3 *SD* from the mean for a given condition and position were excluded from analyses. This adjustment affected less than 2% of the data. Appendix B reports the raw and residual reading times trimmed at 3 *SD*.

Figure 1 shows residual reading times by condition. For analysis purposes, the words were grouped into regions. This was done in order to reduce the number of comparisons and focus on effects of interest. This technique also facilitates comparison between the results of the present study and those of others that used similar methodologies (e.g., Trueswell, 1996). Regions are given in Table 1. Region 1 included the initial subject NP. Region 2 consisted of the relative pronoun ("who" or "which") in the RC conditions and the matrix verb plus "that" in the SC conditions. Region 3 consisted of the following determiner and noun. Region 4 consisted of the segment "that was," present only in the unambiguous conditions. Region 5 consisted of the ambiguous verb. Region 6 consisted of the by-phrase—the critical disambiguation point in the ambiguous conditions. Region 7 consisted of the next three words (at least the first word was the same across all conditions). Finally, region 8 consisted of the remainder of the sentence.

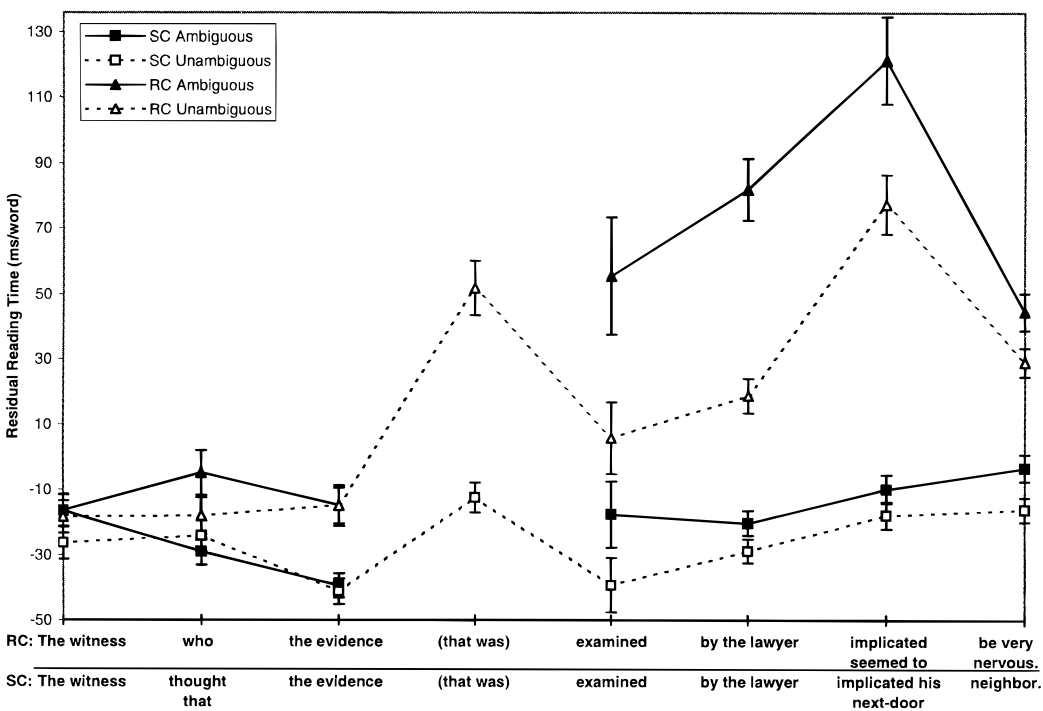

FIG. 1. Residual reading times in Experiment 1. Error bars represent 1 *SE* of the Grand Mean.

TABLE 1

Analysis Regions for Experiment 1

	Region							
	1	2	3	4	5	6	7	8
RC	The witness	who	the evidence	(that was)	examined	by the lawyer	implicated seemed to	be very nervous.
SC	The witness	thought that	the evidence	(that was)	examined	by the lawyer	implicated his next-door	neighbor.

There were no reliable effects or interactions related to group (PC vs Macintosh presentation). This factor is therefore omitted in the analyses reported below. Length-adjusted reading times over regions 2–8 (excepting region 4, which was not present in all conditions) were submitted to a $2 \times 2 \times 6$ ANOVA crossing ambiguity by structure by region. All possible main effects and interactions were reliable for both the participants and items analyses (all Fs $>$ 2.5, ps $< .05$).

The first prediction tested in the current experiment was that greater storage costs in the RC embedding context would produce elevated reading times over the embedded clause. This was evaluated by comparing reading times for the unambiguous RC and SC conditions across regions 3–6 where these conditions were lexically and structurally identical. As anticipated, the embedded clause was read significantly faster when embedded within a SC than when embedded within a RC [$t1(1,59) = 65.1$, $MSE = 5900$, $p < .001$; $t2(1,15) = 123$, $MSE = 800$, $p < .001$].

The second prediction bears directly on the Ambiguity Resolution Hypothesis. If increased storage costs promote the incorrect MV analysis more in the RC condition than in the SC condition, then there should be a more pronounced effect of ambiguity in the RC conditions over the disambiguating region. To test this, two-factor ANOVAs crossing structure and ambiguity were performed individually over the region just prior to disambiguation (region 5), the disambiguating region (region 6), and the region immediately following disambiguation (region 7).

On the ambiguous verb (region 5) there were reliable main effects of both structure and ambiguity, with RC items read more slowly than SC items [$F1(1,59) = 20.8$, $MSE = 9966$, $p < .001$; $F2(1,15) = 16.9$, MSE $= 3353$, $p < .001$] and ambiguous items read more slowly than unambiguous items [$F1$ $(1,59) = 4.99$, $MSE = 14{,}608$, $p < .05$; $F2(1,15) = 20.6$, $MSE = 995.7$, $p < .001$], but these factors did not interact significantly.

At the principal region of interest, the disambiguating by-phrase in region 6, RC conditions were again read more slowly than SC conditions [$F1(1,59) = 61.4$, $MSE = 5727$, $p < .001$; $F2(1,15) = 121$, $MSE = 742$, $p < .001$] and ambiguous items were read more slowly than unambiguous items. The ambiguity effect was greater for the RC conditions than for the SC conditions, as evidenced by a reliable interaction between structure and ambiguity [$F1(1,59) = 20.2$, $MSE = 2369$, $p < .001$; $F2(1,15) = 11.4$, $MSE = 1044$, $p < .005$].

In region 7, the main effect of structure persisted [$F1(1,59) = 58.9$, $MSE = 13{,}703$, $p < .001$; $F2(1,15) = 92.1$, $MSE = 2242$, $p < .001$] as did the main effect of ambiguity [$F1(1,59) = 10.06$, $MSE = 4257$, $p < .005$; $F2(1,15) = 9.79$, $MSE = 1144$, $p < .01$]. There was also an interaction of structure and ambiguity in this region, but this effect was marginal in the items analysis [$F1(1,59) = 5.23$, $MSE = 3729$, $p < .05$; $F2(1,15) = 3.35$, $MSE = 1628$, $p < .10$].

It is clear from Fig. 1 that the ambiguous items were read more slowly in the RC conditions both on the ambiguous verb and over the ensuing disambiguating by-phrase relative to the unambiguous control. Two plausible accounts are consistent with this pattern. One possibility is that structural complexity strongly supports the MV reading at the ambiguous verb. This promotes the implausible

MV argument structure, resulting in elevated reading times over region 5. The ambiguity effect over region 6 then arises from misanalysis when the sentence is disambiguated to the RR reading. Under this explanation, the effects of elevated ambiguous reading times over regions 5 and 6 should be independent of one another, or possibly, inversely related: more support for the MV should decrease difficulty over the ambiguous verb and increase misanalysis difficulty over disambiguation. The second possibility is that the introduction of ambiguity at region 5 in itself results in difficulty, perhaps due to competition between the MV and RR alternatives. This difficulty could be exacerbated by the high storage costs of the RC conditions and spill over into the disambiguating region, thus engendering the observed interaction between structure and ambiguity. Under the second account we would expect the slowdown in the RC ambiguous items prior to disambiguation (region 5) to be positively correlated with the slowdown over disambiguation (region 6), whereas we would expect no such correlation under the first account.

To distinguish these possibilities, difference scores were calculated for each item by subtracting unambiguous reading times from ambiguous reading times over regions 5 and over region 6. Correlations were then performed to see if the ambiguity effects before and after disambiguation were related. For the RC conditions the effects exhibited a nonsignificant negative correlation ($r = -.26$, $p < .33$). Note that difference scores across contiguous regions of a condition are likely to be positively correlated because reading times within each individual trial are usually highly correlated. As a result, the lack of a correlation strongly suggests that the pre- and postdisambiguation effects have independent origins. To correct for the correlated effects of contiguous regions, an additional analysis was performed. Difference scores were calculated for every region of each experimental item. Each score was regressed against the score for the previous region in order to model the influence of difference scores across consecutive regions. Using the equation derived from this regression, predicted scores for regions 5 and 6 of the RC conditions were computed and then subtracted from actual difference scores. A correlation was performed on the residual variance. There was a marginally reliable inverse relationship between difference scores over region five and region six [$r = -.48$, $F(1,14) = 4.12$, $p = .06$]. This supports the view that the pre- and postdisambiguation ambiguity effects stem from different sources.

It is not clear whether subjects had difficulty with the RR resolution in the ambiguous low-storage-cost difference conditions. The slowdowns for ambiguous SC items pictured in regions 6, 7, and 8 are consistent with the pattern observed by Burgess (1991), who used word-by-word presentation on temporarily ambiguous RRs in a matrix context. Burgess interpreted his data as evidence of a garden path effect arising when no parafoveal preview of the disambiguating preposition is available. In the present data there was also an ambiguity effect prior to disambiguation (region 5), so it is not clear whether the effect over the disambiguating region can be attributed to misanalysis as in Burgess' work. Difference scores over regions 5 and 6 were reliably correlated ($r = .68$, $p < .01$), suggesting perhaps that participants were slowed by the presence of ambiguity prior to disambiguation and that this effect persisted over the disambiguating by-phrase. As with the RC conditions, an analysis was conducted to adjust for the linear contribution of difference scores across consecutive regions. The difference score at each region was used to predict the difference score in the succeeding region across all experimental items. The residual variance in difference scores was calculated for regions 5 and 6 of each item. These values were then regressed against one another. The positive relationship observed before the correction disappeared after this adjustment was made ($F < 1$). Thus, there was no clear evidence of particular difficulty over the disambiguating region in the SC condition.

Correlational analyses. The SC conditions were expected to be processed like Trueswell et al.'s poor agent conditions. In their experiments, the typicality of thematic relations associated with both the MV and the RR analyses influ-

enced the effect of ambiguity during the ambiguous verb and over disambiguation. These results suggest that both the MV and RR analyses were partially activated. Likewise, in our low-storage-cost SC conditions, the MV and RR analyses should both be accessible. Thematic fit with both the MV and RR argument structures might then be expected to contribute to the resolution of the ambiguity. For the RC conditions the MV should be more active than the RR because of a strong syntactic storage-cost bias toward the MV. Thus, the thematic fit with the MV argument structure is more likely to affect ambiguous reading times than the thematic fit with the RR argument structure. Factors that promote the MV reading should decrease difficulty at the verb when the MV analysis is still consistent with the input. In contrast, these factors should make it more difficult to arrive at the RR reading after disambiguation.

A set of correlational analyses was performed to examine the influence of thematic typicality on processing the ambiguous conditions. Thematic typicality was assessed via two surveys which evaluated the plausibility of NP_2 (e.g., "evidence") as agent or patient of the embedded verb (e.g., "examined"). In the first of these, 33 participants who did not take part in the main experiment answered questions regarding NP_2's patient typicality (e.g., "How typical is it for someone/something to examine evidence?"), and in the second, 33 additional participants answered questions regarding NP_2's agent typicality (e.g., "How typical is it for evidence to examine someone/something?"). These questions were identical to those from surveys performed by Trueswell et al (1994). A 7-point scale was used in both questionnaires, where 1 represented a maximally atypical judgement and 7 a maximally typical judgement. The mean ratings for each item are reported in Appendix A. Our patient NP_2 ratings were comparable to Trueswell et al.'s (ours: mean = 3.7, SD = 1.0, range = 1.5 to 5.1; Trueswell et al.'s: mean = 4.7, range = 1.8 to 6.5; correlation: $r = .56$, $df = 14$, $p < .05$), but our agent ratings were not (ours: mean = 2.6, SD = 0.7, range = 1.2 to 3.9; Trueswell et al.'s: mean = 1.4, range = 1.0 to 2.2; correlation: $r = .12$, $p > .12$, $p > .10$). This discrepancy may be related to method: Our ratings were gathered using a single list, with no fillers, resulting in a larger range of ratings than was obtained by Trueswell et al. Because our values exhibited greater variance for the patient ratings, and comparable variance on the agent ratings, they were entered into the analyses below.

Thematic typicality for each stimulus was regressed against difference scores for the RC and SC items over the ambiguous verb, the by-phrase, and the following three word region (regions 5, 6, and 7). The results of these analyses are given in Table 2. To ensure that the effects reported below were due to the manipulation of ambiguity, these statistics were also evaluated as predictors of unambiguous reading times, but there were no significant correlations. Hence none of the correlations with difference scores reported below are attributable to trends in the unambiguous conditions.

In the SC conditions, the agenthood of NP_1 was a marginally significant contributor to the ambiguity effect over the verb (region 5) and after disambiguation (region 7). Patienthood of NP_1 also significantly affected reading times over region 7. This is consistent with the hypothesis that both the MV and RR argument structures influenced the processing of the ambiguity in the SC conditions. In contrast, only the agenthood of NP_1 affected difference scores in the RC conditions—better agents led to an increased ambiguity effect at the disambiguating by-phrase. This is in accord with the view that

TABLE 2

Correlations between Difference Scores and Thematic Typicality [df = (1,14)]

	Region		
	Verb	By-phrase	Next region
RC-agency	−.36	.53*	.24
RC-patiency	.04	.26	.26
SC-agency	.43+	.13	.45+
SC-patiency	−.31	.02	−.61*

+$p < .10$.
*$p < .05$.

only the MV argument structure was active when storage cost differences were high.

Discussion

The results of Experiment 1 demonstrate that the manipulation of syntactic environment alone can have a significant impact on the availability of structural alternatives in the MV/RR ambiguity. Individuals had more difficulty with RRs in the syntactic context of a relative clause than in the syntactic context of a sentential complement clause. This was evident in the interaction observed over the disambiguating by-phrase. Both during this region and after there was a significantly larger effect of ambiguity in the RC conditions than in the SC conditions. The SC conditions did not show unequivocal evidence of a garden path. To be sure, the ambiguous condition was read slower than its unambiguous control over the disambiguating region. However, this was significantly correlated with a larger effect in the region prior to disambiguation. The introduction of the ambiguity appears to have created difficulty in the SC conditions, obscuring any indication of a garden path that may have arisen. In contrast, the pronounced garden path effect in the RC conditions was unrelated to the effect of ambiguity at the preceding verb.

The hypothesis under investigation here is that the difference in the configurational complexity of the structural candidates biases the parser to favor the simpler alternative in the RC conditions. An alternative resource-based explanation of the present results is that the computation of frequency and plausibility constraints becomes more difficult as working memory load increases (e.g., MacDonald, Just, & Carpenter, 1992; Pearlmutter & MacDonald, 1995).[5] This could lead to the attenuation of nonstructural factors in the high storage-cost (RC) conditions. Any structural factor that favors the MV might therefore be more influential in the high storage-cost conditions because nonstructural factors are relatively weaker. Further, contextual constraints like thematic typicality may be more difficult to compute than lexical constraints like morphological frequency (Pearlmutter & MacDonald, 1995). Thus typicality in particular might have a stronger influence in the low storage-cost conditions. Indeed, though the both the agent and patient typicality of NP_2 affected reading times in the SC conditions, individuals were only sensitive to agent typicality in the RC conditions. Nevertheless, the Pearlmutter and MacDonald account makes no predictions about which contextual constraints should be harder to compute. For instance, that account does not distinguish between patient and agent typicality. Therefore, the differential sensitivity to patient and agent typicality in the high-cost conditions is not explained. Under the account offered here, the difference is anticipated because the RR construction is not accessible when storage-cost differences weigh heavily against this reading. As a result, thematic relations associated with the RR interpretation, such as the patient typicality of NP_2, are not sufficiently activated to impact reading behavior. Still, the items above were not designed to compare the MacDonald et al. explanation with the Ambiguity Resolution Hypothesis. Experiment 2 addresses this issue more directly by examining an ambiguity in a matrix context, where extrinsic memory load is minimal and remains constant across conditions.

EXPERIMENT 2

Experiment 2 explored another ambiguity in which structural and nonstructural constraints were contraposed: the noun–noun (NN)/relative clause (RC) ambiguity, exemplified in (7) (from Marcus, 1980).

(7) a. The cotton clothing is made of cotton from Mississippi.

b. The cotton clothing is made of grows in Mississippi.

(cf. The cotton which clothing is made of grows in Mississippi.)

In the NN reading (7a) "cotton" acts as a modifier of the head noun "clothing" to form the compound nominal "cotton clothing." In the RC alternative (7b) "clothing" is the subject of a RC which modifies "cotton." There is a strong tendency to analyze the ambiguous string "the cot-

[5] We are indebted to Neal Pearlmutter for this observation.

ton clothing" as a NN, resulting in unproblematic processing of the NN continuation in (7a). In contrast, a great deal of difficulty arises in processing the dispreferred RC continuation (7b) at the point of encountering the verb "grows."

To our knowledge, the NN/RC ambiguity has not been studied in any on-line experiments. Nevertheless, this ambiguity is common. It occurs, for instance, whenever a NN with a plural or mass head noun is specified by a determiner with ambiguous number marking (e.g., "some," "no," or "the"). The parsed Brown Corpus of just over 1 million words (Kučera & Francis, 1967; Marcus, Santorini, & Marcinkiewicz, 1993) contains 5395 plural NNs, 13.2% of which include ambiguous determiners. The fact that this arrangement occurs frequently is not surprising because NNs are highly productive and are generally comprehended without noticeable effort (Clark, 1983; Clark, Gelman, & Lane, 1985).

For our present purposes, the NN/RC ambiguity is interesting because, unlike ambiguities examined in most on-line studies, there is a large disparity between the storage costs incurred by its candidate readings. The ambiguity is introduced when the word "clothing" is input to the parser. For both readings, integrating "clothing" into the existing structure is a local operation. In the NN, the head noun, "clothing," is attached directly to "cotton," the modifier. In the RC, "clothing" is attached as the subject of a RC modifying "cotton." Neither of these integrations spans a new discourse referent. Although integration costs are balanced, storage costs are biased toward the NN. For the NN reading, only a matrix predicate is required to complete the input grammatically. In contrast, the RC interpretation requires satisfying two predictions in addition to the matrix predicate: an embedded verb and an embedded NP-gap site. The differential between one and three unsatisfied dependencies for the NN/RC is comparable to that of the embedded MV/RR ambiguity.

In Experiment 2, we examined the contributions of syntactic and nonsyntactic factors in processing the NN/RC ambiguity. Nonstructural factors were manipulated to support the RC reading. In the absence of a structural bias, the NN reading should be rendered inaccessible, and there should be no difficulty with temporarily ambiguous RC sentences. Consequently, any evidence of difficulty with the RC reading relative to its unambiguous control indicates the influence of a storage-cost bias favoring the NN.

Method

Participants. Forty-eight participants from MIT and the surrounding community were paid for their participation. All were native speakers of English and naive as to the purposes of the study.

Materials. Forty temporarily ambiguous RC sentences were prepared. For twenty of these, the NN interpretation was less plausible than the RC throughout the ambiguous region. For ease of exposition, we refer to this set of items as the implaus-NN items. (Note that only the NN reading was implausible in these items: The RC reading was highly plausible.) The other half of the items—the plaus-NN items—contained NN readings that were more plausible than their RC interpretations. Each stimulus item had an unambiguous control. Sample implausible and plausible items are given in (8) and (9) respectively.

(8) Implaus-NN

a. Ambiguous:

The tool plumbers need to have is a good monkey wrench for loosening rusty pipes.

b. Unambiguous:

The tool which plumbers need to have is a good monkey wrench for loosening rusty pipes.

(9) Plaus-NN

a. Ambiguous:

The alley mice run rampant in is damp and dimly lit but relatively clean.

b. Unambiguous:

The alley which mice run rampant in is damp and dimly lit but relatively clean.

Note that the second noun in each condition is plural. Plural nouns are generally infelicitous as prenominal modifiers.[6] Thus the second noun is strongly biased to be interpreted as a head and not a modifier of a third noun yet to come. To avoid the possibility of the noun–noun sequences being misconstructed as noun–verb sequences and introducing a third interpretation of the present ambiguity, only lexically unambiguous nouns were used as the second noun (e.g., "plumbers" and "mice").

The ambiguous conditions are consistent with a NN analysis through the sixth word of the sentence. The point of disambiguation occurs at the seventh word, when the matrix verb is encountered. At this point there are two inflected verbs, requiring one level of embedding. Unambiguous RC conditions were created from the ambiguous conditions by inserting a relative pronoun (e.g., "which") between the first and second nouns.

Because of their intuitive difficulty, the ambiguous plaus-NN items were expected to generate a strong garden path effect. These items were intended as controls to establish the qualitative effect of a garden path in this ambiguity. Because of the large storage cost bias, an analogous misanalysis effect in the crucial implaus-NN conditions should be observed. It should follow a similar qualitative pattern and arise at the same word positions. In this way the plaus-NN items served to diagnose whether difficulty with the RC resolution in the ambiguous implaus-NN items was due to misanalysis difficulty or to some alternative factor.

In addition to the 40 target items, there were 80 fillers, some of which were items from an unrelated experiment. Four lists were prepared to balance all factors in a Latin-square design. Each subject saw five versions of each of the above four conditions. There were also an identical number of temporarily ambiguous and unambiguous sentences which contained NN constructions. These were constructed using the implaus-NN and plaus-NN items above as a base. The presence of NN items served to discourage participants from developing strategic response biases to the RCs. No subject saw two versions of a single stimulus. Appendix C provides a complete list of Experiment 2 stimuli including both RC and NN versions.

[6] There are some counterexamples (e.g., "bonds market") but these are rare and used in specialized contexts (Kiparksy, 1982; Seidenberg, Haskell, & MacDonald, 1999). Plural modifiers are not generally productive.

Controls for Lexical and Contextual Influences

The focus of this experiment is the implaus-NN items in which structural and nonstructural factors are pitted against one another. It was therefore important to ensure that nonconfigurational influences did not support the NN reading in these items. The following sections detail the ways in which plausibility, lexical frequency, and referential parsimony were evaluated for the present materials.

Plausibility norms. The predilection for the NN in (7) is probably strengthened by the fact that "cotton clothing" refers to an extremely plausible real-word concept. It has been argued that noun–noun compound conceptual combination derives from a thematic relation between the constituent concepts (Downing, 1977; Gagne & Shoben, 1997; Levi, 1978). There is reason to believe that the results of such thematic combinations are available early enough to affect initial parsing decisions (Trueswell et al., 1994, Garnsey et al., 1997). We controlled for thematic plausibility by constructing NN/RC items with NN conceptual interpretations that were implausible relative to their RC alternatives.

In order to verify that the NN interpretations of the implaus-NN items were less plausible than the RC interpretations, two surveys were conducted. For the first of these, 40 individuals from the MIT community who did not participate in the main experiment were asked to judge the naturalness of the unambiguous versions of the NN and RC readings as sentence onsets (e.g., "The tool which plumbers. . ." corresponding to (8b) and "Tool plumbers. . ." corresponding to the unambiguous NN reading). The 20 norming stimuli were pseudorandomly mixed with 56 filler items to form two counterbalanced lists, each of which contained either the NN or RC version of each item. The partici-

pants were evenly divided into two groups and given one or the other of the lists. The RC versions of implausible items received a mean plausibility rating of 2.91 ($SD = 1.16$) on a 7-point scale where 7 indicated maximal implausibility.[7] The NN conditions were judged more implausible, with a mean rating of 4.87 ($SD = 1.04$). This difference was reliable [$t1(1,39) = 82.0$, $MSE = 0.942$, $p < .001$; $t2(1,19) = 52.1$, $MSE = 0.741$, $p < .001$]. Eighteen of the 20 items were numerically more plausible in the RC reading. Fourteen of these comparisons were significantly different ($ts > 5.5$, $ps < .05$), three tended toward significance ($ts > 1.5$, $ps < .20$), and one did not differ reliably. The final two items were numerically more plausible in the NN reading ($ts < 2.25$). These last two items were omitted from analyses of reading times below. Item by item ratings are provided in Appendix C.

One confound inherent to the first survey is that it compared different structures; the NN conditions were complete NPs at the point of comparison, whereas the RC conditions were partial NPs. It is possible that the observed differences in plausibility may have been caused by this structural difference. In particular, the incompleteness of the RCs may have led to more plausible ratings. If so, then the ensuing predicates might have made the RC conditions less plausible than the NNs prior to disambiguation. A second survey examined this possibility for the remaining 18 implaus-NN items. One hundred nine additional members of the MIT community rated the plausibility of the disambiguated NN and RC forms of the items on a 7-point scale. In this survey, the full RC was provided (e.g., "The tool which plumbers need to have") and the NN was as before (e.g., "Tool plumbers"). Again, two lists were constructed and the participants were divided into two groups, one of 57 and another of 52. The RCs were rated more plausible, 2.23 ($SD = 0.83$), than the NNs, 4.39 ($SD = 1.00$). This difference was highly significant [$t1(1,108) = 380$, $MSE = 0.673$, $p < .001$, $t2(1,17) = 201$, $MSE = 0.215$, $p < .001$]. All item by item t tests were also highly significant ($ts > 15$, $ps < .01$). These results suggest that the RC verbal continuations actually improve the plausibility of the RC relative to the NN.

It is significant to note that this pattern of judgements must be due to nonstructural influences. Under all existing theories of structural complexity, the NN is favored over the RC (e.g., Frazier, 1979, 1987; Gorrell, 1995; Pritchett, 1988; Gibson, 1991, 1998). Since the results of the surveys indicate a clear bias toward the RC, syntactic complexity cannot be contributing to this effect.

A separate questionnaire with a design identical to that of the first survey described above was prepared to ensure that the NN interpretations were more plausible than the RC interpretations for the plaus-NN items. Another 50 participants completed this survey. Overall, the NN onsets were judged to be significantly more plausible, 2.26 ($SD = .94$), than the RC onsets, 4.23 ($SD = 1.19$) [$t1(1,49) = 191$, $MSE = 0.508$, $p < .001$; $t2(1,19) = 68.8$, $MSE = 0.564$, $p < .001$]. As with the implaus-NN items, pairwise comparisons were also conducted. Fifteen of the items reliably favored the NN ($ts > 7$, $ps < .05$). Another four approached significance ($ts > 2.2$, $ps < .15$). The remaining item numerically favored the RC, but this difference was not significant ($t < 1$).

Lexical frequency. Lexical frequencies have been shown to play an important role in ambiguity resolution (e.g., MacDonald 1993, 1994; MacDonald et al., 1994; Trueswell, 1996). However, there are many ways to tabulate lexical frequencies, and it is not yet known which of these might be relevant to the NN/RC ambiguity. MacDonald (1993) studied the statistical factors involved in a related ambiguity, the NN/NV ambiguity, illustrated in (10):

(10) The warehouse fires . . .

a. Noun–noun continuation: . . . can be dangerous.

b. Noun–verb continuation: . . . twenty employees every month.

[7] For all surveys in this experiment, higher numbers indicated more implausible values. To make the presentation of the correlations with reading times below more perspicuous, these values were subtracted from zero so that higher values imply greater plausibility.

In the NN reading (10a), the first noun "warehouse" acts a modifier of the second noun, "fires." In the NV reading (10b) "warehouse" acts as the head of a NP. MacDonald found that two statistics related to the first noun in the NN construction affected the processing of examples like these. Because the factors involved in this ambiguity could plausibly play a role in the resolution of the NN/RC ambiguity, an effort was made to control for both of these statistics in the current experiment. It is unlikely that any statistical tendencies associated exclusively with the second noun could affect the processing of the NN/RC because the lexical category and modifier/head status of the second noun are identical under both structural alternatives.

First, MacDonald observed that co-occurence bias, the rate at which lexical items appear together as a NN, influenced the likelihood with which participants will conjoin them in on-line reading. For instance, "cotton" might be readily conjoined with "clothing" because it is a lexically encoded feature of "cotton" that it often modifies "clothing." MacDonald found that NNs with constituents that frequently co-occur in a head–modifier sequence promoted a NN reading, while NNs with low co-occurrence rates supported the alternative NV structural interpretation. To avoid a co-occurrence bias toward the NN, implaus-NN items in the present experiment were constructed using novel NN combinations.

Values for co-occurrence bias were determined by searching through three electronic corpora of written English obtained from the Linguistics Data Consortium at the University of Pennsylvania: the Wall Street Journal (WSJ) corpus, which contains over 1 million words taken from articles in the WSJ from 1987 to 1989; the Brown corpus, which contains approximately 1 million words collected from various sources (Kučera & Francis, 1967); and the AP corpus, which contains about 40 million words from Associated Press newswire articles published in 1989. The WSJ and Brown corpora are parsed (Marcus, Santorini, & Marcinkiewicz, 1993), permitting a search on instances where the words used in our stimuli occurred together within a noun phrase. Searches for both plural and singular variants of the head noun yielded no instances of modifier–head sequences matching those in our items. Because the AP corpus is unparsed, it was impossible to constrain the search to noun phrases. Instead, a search was conducted for sentences in which the modifier string was followed by a plural or singular variant of the head noun. Again no instances of the noun–noun strings used in the experimental items were discovered.

In an effort to find co-occurrence rates in a larger and more naturalistic corpus, an additional set of searches was conducted on the World Wide Web (WWW) using the Wired Digital Incorporated Hotbot^{TM} search engine. This search engine covered 57.5% of the indexicable web (approximately 15 billion words in 1998) distributed across 110 million documents representing a tremendous variety of topics and conversational registers (Lawrence & Giles, 1998). Hotbot was employed to find only web pages containing modifier–head sequences that matched plural and singular variants of the noun–noun sequences in our items. The output of these searches was normally compact enough to search by hand. When over 100 web pages were returned, 50 of them were randomly chosen for coding. This yielded no exact matches for most of the stimulus items queried.[8]

Co-occurrence frequencies for the plaus-NN items took on a broad range of values. Following previous work (e.g., MacDonald, 1993; Spivey-Knowlton & Tanenhaus, 1994) co-occurrence frequency was normalized against the frequency of the initial noun (N_1) in isolation according to the following formula: log(co-occurrence frequency)/ log(N_1 frequency). The results of these calculations were used in the correlational analyses performed below.

A second factor which MacDonald found to be influential in processing the NN/NV ambi-

[8] There were some counterexamples to this trend. However, no compound from an implaus-NN item appeared in more than 0.00001% of occurrences of the first noun. Further, all such co-occurrences arose in obscure and specialized contexts. As a result, participants almost certainly did not use instances like these to tune preferences for the implaus-NN items.

guity was the first noun's head bias—the ratio of instances in which it was used as a head relative to all instances of the noun. More concretely, co-occurrence frequency alone cannot explain the intuitive ease with which "cotton parkas" is analyzed as a NN compound. This intuition can be partially explained by the fact that "cotton" has a weak head bias (or, equivalently, a strong modifier bias); it is used as a head in only 44.2% of usages. MacDonald found that facilitation for the NN interpretation varied directly with the proportion of modifier usages of the first noun. Just as with the NN/NV, the first noun in the NN/RC ambiguity can play the role of a head or modifier; it is a head in the RC reading and a modifier in the NN. Reasoning by analogy, a modifier-biased first noun should vie in favor of the NN reading, whereas a head-biased first noun should support the RC reading. Corpus searches were conducted to ensure that the first nouns in the implaus-NN items below were head-biased, so that they were biased against the syntactically less complex NN interpretation.

All sentences containing instances of N_1 were collected from the AP corpus. When more than 200 sentences were returned, 150 of them were randomly chosen for coding. All sentences were coded if less than 150 were available.[9] Two judges independently coded the head/modifier status of each instance of the target word string in the retrieved sentences. When the composite meaning of a hyphenated or proper NP was not entirely attributable to the meanings of its constituent parts—an idiomatic usage—the NP was excluded from head/modifier counts. A total of over 2600 sentences were coded. The judges agreed on over 92% of all classifications. A third judge settled all disputes. The proportion of modifier and head uses in this sample was tallied relative to all nominal instances of the word string. The results of all of these searches are provided with the items in Appendix C.

In MacDonald's (1993) items, initial nouns that biased the reader toward a NN compound appeared as heads on average 58.8% of the time. In contrast, those that did not bias toward a NN occurred as heads in 85.5% of instances. The average head bias (computed as an average of head biases for each item) for implaus-NN items in the present study was 85.3% (*SD* = 15.1%). Thus, N_1s from implaus-NN items resemble MacDonald's head-biased nouns. Similar counts were performed for the plaus-NN items. The average head bias for these items was also 85.3% (*SD* = 8.1%).

Referential theory. A number of theorists have proposed that discourse complexity is a major determinant of structural preferences in language comprehension (e.g., Crain & Steedman, 1985; Altmann & Steedman, 1988). According to referential theory, the sentence processing mechanism follows the principle of Parsimony, choosing the candidate reading associated with the fewest presuppositions that are not supported by the discourse context (Crain & Steedman, 1985). Complex NPs are dispreferred relative to unmodified NPs because they presuppose a contrast set. As with postnominal modifiers such as relative clauses, prenominal modifiers such as compound nominals can implicitly indicate a contrast set by highlighting a relevant property of interest (Markman, 1989). Because both the NN and the RC readings potentially entail a contrast set, Parsimony does not straightforwardly favor one reading over the other. Indeed, the existence of a contrast set has been shown to reduce difficulty both in processing certain postnominally modified NPs (Altmann, Garnham, & Dennis, 1992; Altmann & Steedman, 1988; Britt, 1994) and in processing certain prenominally modified NPs (Sedivy, 1999).

Indeed, if there is any general semantic-level bias in the NN/RC ambiguity, it arguably favors the RC. Whereas RC modification and adjectival modification unambiguously highlight a feature of interest, the semantic interpretation of NNs is variable and subject to the vagaries of pragmatic context: The feature or constellation of features picked out by a modifying noun can be highly dependent on the situation in which the noun is interpreted. For instance, in certain circumstances, the NN "winter book" might refer either to a book to be read in the winter or

[9] One N_1, "fusebox," did not appear in the AP corpus. Fifty sentences containing this noun were obtained from the WWW.

a book about winter. The relative complexity of NN compounding compared to adjectival modification has been confirmed by Murphy (1990). Individuals were quicker to judge that an adjectivally modified NP was meaningful than one containing a novel NN compound. This was true even for anomalous adjectives (e.g., "the pregnant boy").

In sum, a number of nonsyntactic factors were controlled in the implaus-NN items to favor the RC reading of the ambiguous string. None of these factors supported the structurally simpler NN interpretation.

Procedure

The procedure was the same self-paced word-by-word moving window paradigm as used in Experiment 1. This experiment was done using Macintosh computers running software developed in our lab. Each stimulus item spanned from one to one-and-one-half lines, with up to 100 characters on each line. All critical material analyzed below was contained in the first line.

Participants took approximately 25 min to complete the experiment. For most participants, this experiment was combined with two unrelated experiments: one using an identical self-paced reading procedure and a second employing a questionnaire methodology. The combined session was roughly 45–60 min. Participants took a short break between experiments.

Results

Comprehension question performance. Questions for experimental items were answered correctly 93.4% of the time. A 2 × 2 ANOVA crossing ambiguity and plausibility revealed that questions following implaus-NN items were answered correctly more frequently (95.6%) than those for plaus-NN items (91.2%). This was reliable in the participants analysis [$F1(1, 47) = 8.6$, $MSE = 0.01$, $p < .01$] but not in the items analysis ($F < 2$). This difference probably arose because the questions for the implaus-NN items were easier than those for the plausible items. No other reliable effects or interactions were observed. Because of high accuracy rates, no data were excluded on the basis of comprehension performance.

Predictions. There were three predictions for the present experiment. One of these relates to the plaus-NN items. In this condition, both structural and nonstructural factors were aligned in support of the NN reading. As a result, the parser should pursue the NN analysis in the ambiguous condition and experience difficulty only when this analysis becomes untenable. Thus we expected reading times over disambiguation to be slower for the plaus-NN ambiguous condition relative to its unambiguous control.

In the implaus-NN conditions, the Ambiguity Resolution Hypothesis (2) posits that the NN analysis should also be accessible, despite its implausibility. This stems from a strong syntactic storage-cost bias in favor of the NN. Two effects on reading times were predicted. First, just as with the plaus-NN items, we anticipated that the active NN interpretation would cause difficulty when the sentences were disambiguated toward a RC analysis. This should lead to a slowdown on ambiguous items over the same word positions as for the plaus-NN conditions though the magnitude of this garden path effect was expected to be less than that for the plaus-NN conditions where all factors are heavily biased in support of the NN. Second, the NN reading is active from the point of the second noun where the ambiguity is introduced. Because this reading is extremely implausible relative to the RC interpretation, an early slowdown on implaus-NN ambiguous items was expected over the second noun and ensuing verb. There should be no similar difficulty due to implausibility in the plaus-NN conditions. Hence, we predicted an interaction between ambiguity and plausibility over the first several words of the ambiguous region; the early ambiguity effect for the implaus-NN items should be significantly greater than that for the plaus-NN items. In sum, storage-cost biases predict that readers should be penalized twice for maintaining the NN analysis in the ambiguous NN-implaus conditions. The first penalty arises because the the NN is highly implausible, and a second penalty arises when this implausible reading turns out to be incorrect.

Reading times. As in Experiment 1, reading time analyses were performed on length-adjusted residual reading times for each word. Val-

ues that were more than 3 SDs away from the mean of a given condition and word position were excluded from analysis. This correction eliminated 2.0% of the data. Reading times for three implaus-NN items were omitted: two because the NN reading was judged more plausible than the RC reading (as mentioned above), and one because of typographical error that went uncorrected in the self-paced session. Appendix D reports the raw and residual reading times.

Figure 2 depicts residual reading times per word for each condition. A $12 \times 2 \times 2$ ANOVA crossing word position by plausibility by ambiguity was conducted. All possible effects and interactions were significant (Fs > 7, ps $< .01$).

The ambiguous plaus-NN conditions were anticipated to elicit a large misanalysis effect over the disambiguating region. This pattern is evident in comparing the filled and unfilled triangles of Fig. 2 and was confirmed by a significant interaction between position and ambiguity over the final word of the ambiguous region ("have") and the first two words of the disambiguating region ("is" and "a") [$F1(2,94) = 31.6$, $MSE = 32000$, $p < .001$; $F2(2,38) = 49.3$, $MSE = 8090$, $p < .001$]. This implies that ambiguous plaus-NN sentences were initially misanalyzed as NNs, resulting in elevated reading times when they were resolved as RCs.

The critical question in the present experiment was whether the implaus-NN conditions would also precipitate difficulty when ambiguous items were disambiguated toward an RC interpretation. Figure 2 demonstrates that the pattern of reading times in the implaus-NN items follows a profile similar to that of the plaus-NN items. Ambiguous reading times (the filled squares) become increasingly separated from unambiguous reading times (the unfilled squares) at the second word of the disambiguating region. Just as for items where the NN was more plausible than the RC, there was a significant interaction of ambiguity by region at the onset of the disambiguating region [$F1(2,94) = 5.15$, $MSE = 4050$, $p < .01$; $F2(2,32) = 5.32$, $MSE = 3180$, $p < .01$].

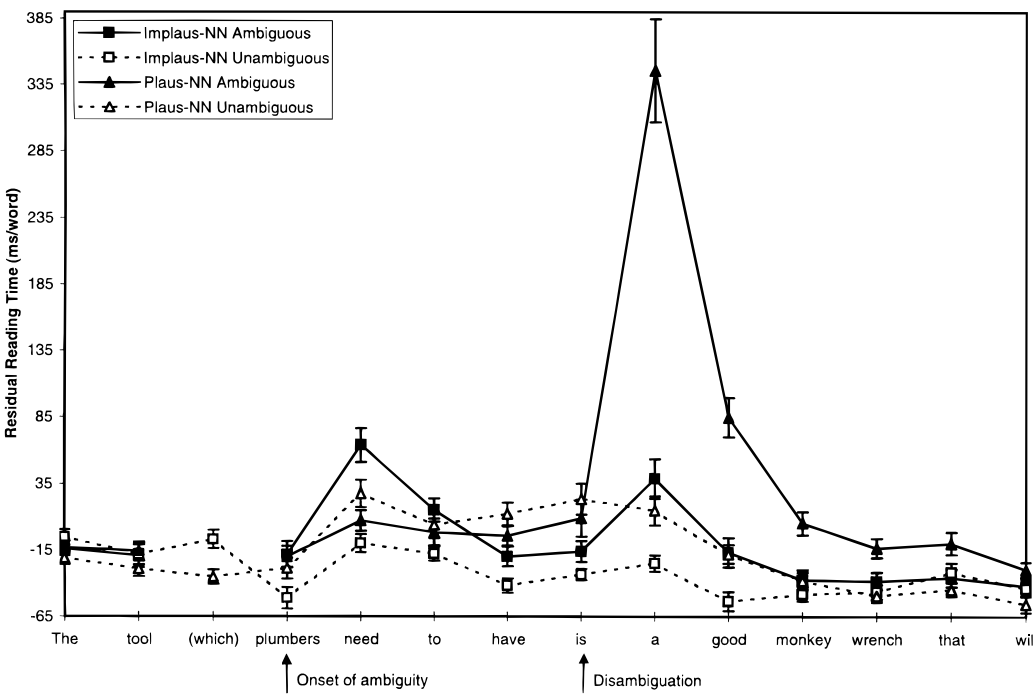

FIG. 2. Residual reading times in Experiment 2. Error bars represent 1 *SE* of the Grand Mean.

This indicates that the NN interpretation made it difficult for readers to arrive at the correct RC interpretation at disambiguation, even for items which were plausibility biased toward the RC reading.

A second pattern of separation between the implaus-NN conditions is visible in Fig. 2 at the onset of the ambiguous region. A 3 × 2 ANOVA crossing word position by ambiguity confirmed that there was a reliable increase in ambiguous reading times from the last word before the ambiguous region (the first noun) to the first two words of the ambiguous region (the second noun and the following word) [$F1(2,94) = 6.35$, $MSE = 5260$, $p < .01$; $F2(2,32) = 12.6$, $MSE = 980$, $p < .001$]. A 2 × 2 ANOVA crossing amibiguity and plausibility over the first two words of the ambiguous region confirmed the prediction that difficulty on ambiguous items was reliably larger in the implaus-NN conditions than in the plaus-NN conditions [$F1(1,47) = 14.9$, $MSE = 3040$, $p < .001$; $F2(1,35) = 30.3$, $MSE = 535$, $p < .001$]. These results imply that the implausibility of the NN analysis was responsible for elevated reading times at the onset of the ambiguity.

Correlations. Under the account offered above, increased reading times over the first few words of the ambiguous region in the implaus-NN condition are attributable to the peculiarity of the NN interpretation. Therefore we would expect factors that bolster the NN to be inversely correlated with the magnitude of the ambiguity effect. In contrast, the ambiguity effect at disambiguation is allegedly due to competition from the NN with the correct RC reading. In this case, factors that attenuate support for the NN reading should result in reduced interference with the RC resolution. Thus, there should be a direct relationship between factors that support the NN reading and the ambiguity effect size over disambiguation.

To further establish the independence of the effect over the initial portion of the ambiguous region from the effect over disambiguation, a set of correlational tests was performed. In order to isolate the effects of ambiguity from differences between unambiguous conditions, ambiguous reading times over the first three words of the ambiguous region and the first three words of the disambiguating region were regressed against unambiguous times at those word positions for each item. Critical analyses were conducted on the residual variance in ambiguous times to assess the contributions of plausibility, co-occurrence bias, and head bias to reading behavior at the introduction of the ambiguity and over disambiguation. A summary of these correlational analyses is given in Table 3.

All trends over the initial portion of the ambiguity reversed in direction in the disambiguating region. For the early effect, factors that promoted the NN reading led to a decreased ambiguity effect. Over disambiguation, the same factors led to increased reading times. The fact that nonsyntactic factors contributed in opposite ways to the pre- and postdisambiguation effects lends credence to the hypothesis that the ambiguity effects are distinct rather than parts of a single overarching slowdown. The pattern of correlations observed here is expected if the structurally simpler NN reading was strongly activated in the ambiguous RC conditions. On this view, factors that enhance the NN reading—e.g., high co-occurrence bias and plausible NN readings—caused the ambiguous items to be

TABLE 3

Coefficients for Stepwise Regressions between Four Measures and Residual Ambiguous Reading Times for RC Conditions after Unambiguous Times Have Entered [$df = (1,35)$]

	Region	
	Onset of ambiguity	Disambiguation
	The tool **plumbers need to** have	**is a good** monkey wrench . . .
NN plausibility	−.52***	.63***
RC plausibility	.05	−.23
Co-occurrence bias	−.29+	.28+
Head bias	−.01	−.01

Note. Relevant reading times are from regions in boldface.

+$p < .10$.

***$p < .001$.

read faster, thereby reducing the slow-down effect prior to disambiguation. In the disambiguating region, these same factors strengthened the misanalysis effect when the sentence was resolved as a RC.

Discussion

The reading time data demonstrate that structural complexity plays a significant role in resolving the NN/RC ambiguity. Even when nonsyntactic factors were heavily biased toward the RC interpretation, the NN structural reading was sufficiently activated to affect reading behavior. This was evident from a pattern of two independent slowdowns for ambiguous implaus-NN items relative to their unambiguous controls. The first of these was inversely related to the plausibility of the NN reading and occurred early in the ambiguous region. This indicates that the NN reading was highly activated in ambiguous cases. Items with implausible NN readings then created difficulty prior to disambiguation, but items with plausible NN interpretations did not. For the unambiguous conditions, the implausible NN analysis was not permitted, so no analogous difficulty arose. A second slowdown occurred in the disambiguating region. Again, this is explained if the NN analysis was sufficiently active over the ambiguous region to disrupt resolution toward the RC reading. Unlike the first slowdown, the second effect increased with the plausibility of the NN interpretation, implying that the NN reading was more strongly activated for ambiguous items with plausible NN readings, resulting in a stronger garden path.

In sum, there is evidence here for a powerful configurational bias at work in the resolution of the NN/RC. Lexical frequency, combinatorial plausibility, and discourse reference did not promote the NN reading in our stimuli, yet participants exhibited difficulty with the RC reading. The forces pushing the experimental participants to the NN interpretation cannot be ascribed to any peculiarities of the particular lexical items used. These results are consistent with a view of the parser in which syntactic storage costs are weighed in evaluating the viability of structural alternatives. At the point where the ambiguity is introduced, there is a substantial discrepancy between the number of syntactic dependencies entailed by the NN and RC analyses. If the parser attempts to minimize storage costs, then the simpler NN reading should be promoted in the present ambiguity.

GENERAL DISCUSSION

In both of the ambiguities explored above, individuals exhibited difficulty with the more complex structural alternative even when lexical and contextual constraints supported this alternative. In Experiment 1, temporarily ambiguous RR sentences that cause little or no difficulty in matrix contexts elicited significantly more difficulty when embedded in a relative clause. Manipulating structural environment alone was sufficient to affect the pattern of preferences in the MV/RR ambiguity. This effect was not idiosyncratic to the MV/RR. Experiment 2 tested sentences that were temporarily ambiguous between a NN and a RC structure. Readers exhibited difficulty in processing sentences disambiguated toward the RC even though this interpretation was judged more plausible in off-line surveys and was supported by lexical statistics. Thus individuals activated the syntactically simpler NN reading despite the fact that nonsyntactic factors were aligned against this alternative. In both experiments, correlational analyses showed that ambiguous reading times were affected by lexical and contextual properties associated with the structurally simpler reading. These results refute theories that explicitly exclude the influence of structurally based biases as well as theories that permit the influence of syntactic regularities only when lexical-level constraints are uncommitted with respect to structure (e.g., Ford, Bresnan, & Caplan, 1982).

Though the nature of the structural biases observed here cannot be conclusively determined, the results are consistent with a constraint that favors the candidate reading associated with the fewest unsatisfied syntactic dependencies. Certain other metrics of syntactic complexity also predict that the MV is simpler than the RR and that the NN is structurally simpler than the RC (e.g., Frazier, 1979, 1987;

Gorrell, 1995). However, these measures only yield ordinal value judgements, which rank structural candidates in terms of complexity. As a result, they cannot distinguish ambiguities in which syntactic biases are substantial, such as the RC embedded MV/RR and the NN/RC, from ambiguities in which they are less influential, such as the matrix MV/RR and the SC embedded MV/RR. In contrast, the DLT outputs interval value judgements of syntactic complexity. This allows the DLT to quantify the contribution of syntactic complexity to resolving any given ambiguity.

Certain theories of ambiguity that employ an ordinal metric of syntactic complexity appeal to diagnosis or repair mechanisms of reanalysis in order to explain asymmetric misanalysis difficulty. Reanalysis difficulty is thought to be a function of the cues to disambiguation, the syntactic characteristics of the target structure, and syntactic characteristics of the incorrect structure (see Fodor & Ferreira, 1998, for a review of several prominent approaches to reanalysis). In Experiment 1, the alternative structures and cues to disambiguation were constant across the high- and low-memory-cost difference conditions. Yet readers experienced differential difficulty with the RR continuation across the two embedding environments. In order to account for the present results it is crucial to consider how committed the perceiver is to the target and alternative structures. We know of no theory of reanalysis that incorporates this type of information.

The experiments presented here necessitate an ambiguity resolution mechanism that makes reference to structural biases. We have proposed that these biases derive from resource-based constraints which can be characterized by the DLT. However, a statistical explanation of these results is still logically possible. Such an account would have to employ coarse-grained, statistical records which are nonlexical (Mitchell, Cuetos, Corley, & Brysbaert, 1995). There are two challenges faced by any statistical account of the structural biases observed above.

First, because of the large number and complexity of potential linguistic contexts, the number of contextual frequencies that might be compiled is enormous. Processing theories that make extensive use of contextual frequencies must place a principled bound on those that might be important. Otherwise these theories cannot be falsified; any reading time data that are not explained by known contextual frequencies might be influenced by an as yet undiscovered statistic. Statistical accounts also need to address when and why certain probabilistic constraints are more powerful than others, otherwise they have little predictive value. Recent advances in computational modeling might provide a means of bounding statistical information. Neural networks are thought to be capable of attending to only those contingent frequencies that have predictive value (Elman, 1991; Tabor, Juliano, & Tanenhaus, 1997). These approaches are still in their infancy, however. It remains to be seen whether they can describe a system with anything close to the complexity of natural language. Moreover, one still has to specify a grain size indirectly by choosing the primitives in the system (e.g., whether certain semantic features should be included in the input), and by fixing various other parameters. More importantly, none of these approaches fully solves the problem of theory underspecification. If frequency information is conditionalized on structural contexts, it must be specified which contexts are important and when. Until such a theory can be elaborated, there is no way to apply these statistical approaches to the present ambiguities.

The second challenge a statistical account of the present results faces is that statistical correlations could reflect the effects of nonstatistical constraints on reading behavior. Parallels between perceivers' preferences and the distribution of constructions in the linguistic environment might arise either because readers tune their expectations based on experience or because similar constraints operate in comprehension and production. The probabilistically tuned model views structural frequencies as historical artifacts or as the consequence of independent constraints on production. On the other hand, a resource-based theory like the DLT offers the opportunity to explain existing

syntactic distributions (Gibson & Pearlmutter, 1994) in a framework that presumes similar cognitive mechanisms for constructing the syntactic structure of a sentence in comprehension and production. If individuals avoid producing resource intensive constructions, then the absolute complexity of a given structure is predicted to determine its real-world frequency, modulo the influence of other communicative demands. In comprehension, where the parser compares alternatives directly or indirectly, the *relative* complexity of each candidate will be the primary determinant of structural biases. In principle then, a resource-based theory might be able to explain syntactic distributions just as well as a statistical tuning model.

CONCLUSION

The present experiments indicate that a specialized metric of linguistic resource complexity acts as an independent constraint in structural ambiguity resolution. This metric biases the parser toward the structural candidate that incurs the least integration and storage costs. The strength of this bias is monotonically related to the size of the discrepancy in resource costs between structural alternatives. Construed in this way, resource complexity is defined over configurations of lexical items and cannot be reduced to costs associated with individual words. The present experiments therefore reaffirm the importance of nonlexical information in ambiguity resolution. They also begin to articulate the way in which nonlexical biases operate.

APPENDIX A

Items for Experiment 1

The ambiguous forms of the RC and SC versions of the items in Experiment 1 are given below. The unambiguous forms are created by adding the words "that was" immediately before the embedded ambiguous MV/RR verb (e.g., before "examined" in item 1). Mean typicality ratings for NP_1 and/or NP_2 combined with the embedded verb are listed after each item in the following order: NP_1 versus NP_2 as patient; NP_1 as patient; NP_1 as patient combined with NP_2 as agent; NP_2 as agent; NP_2 as patient.

1. RC: The witness who the evidence examined by the lawyer implicated seemed to be very nervous.
   SC: The witness thought that the evidence examined by the lawyer implicated his next-door neighbor.
   (0.94, 5.38, 1.76, 1.97, 4.09)
2. RC: The student who the paper graded by the professor was written by was unhappy with the grade.
   SC: The student said that the paper graded by the professor was written by another student.
   (1.9, 6.06, 2.15, 2.3, 4.15)
3. RC: The young artist who the painting studied by the historian had motivated has a lot of raw talent.
   SC: The young artist thought that the painting studied by the historian had motivated him in his work.
   (2.15, 1.9, 3.63, 1.79, 3.18, 4)
4. RC: The bystander who the van recognized by the spy had hit got up and then ran down the street.
   SC: The bystander reported that the van recognized by the spy had hit a woman who then got up and ran down the street.
   (−0.15, 4.38, 1.79, 1.24, 5.06)
5. RC: The woman who the recipe selected by the judges had made famous was an excellent cook.
   SC: The woman believed that the recipe selected by the judges had made her famous.
   (0.9, 4.9, 1.7, 2.12, 4.55)
6. RC: The construction worker who the bricks lifted by the crane had brushed against was not hurt.
   SC: The construction worker noticed that the bricks lifted by the crane had brushed against the ladder and knocked it down.
   (1.95, 3.13, 1.73, 1.67, 3.79)
7. RC: The executive who the account wanted by the advertiser depended upon was about to be promoted.
   SC: The executive believed that the account wanted by the advertiser depended upon a number of intangible factors.
   (1.7, 4.13, 1.97, 2.30, 4.48)
8. RC: The scientist who the textbook loved by the class was about discovered that the text had some inaccuracies.
   SC: The scientist discovered that the textbook loved by the class was inaccurate in many places.
   (−1.5, 4.94, 1.52, 1.7, 3.06)
9. RC: The mansion which the gold transported by the guards had paid for had nineteen rooms.
   SC: The criminal knew that the gold transported by the guards had paid for a mansion with nineteen rooms.
   (2.5, 1.44, 1.42, 2.58, 2.52)

APPENDIX A—*Continued*

10. RC: The trees which the power plant attacked by the terrorists damaged are going to take a long time to recover.
SC: The mayor reported that the power plant attacked by the terrorists damaged some trees in the nearby forest.
(0.2, 2.84, 2.45, 2.09, 1.67)
11. RC: The drugs which the jewelry identified by the victim had been traded for had a high street-value.
SC: The punk indicated that the jewelry identified by the victim had been traded for drugs with a higher street-value.
(0.15, 1.91, 3.88, 3.79)
12. RC: The factory which the equipment requested by the hospital is manufactured at is in the middle of Iowa.
SC: The supplier said that the equipment requested by the hospital is manufactured at a factory in the middle of Iowa.
(2.25, 2.22, 1.76, 2.82, 4.39)
13. RC: The documents which the package expected by the secretary contained were critical for the meeting.
SC: The manager indicated that the package expected by the secretary contained documents that were critical for the meeting.
(1.55, 5.13, 1.61, 2.09, 4.48)
14. RC: The earrings which the necklace described by the lady matched were found in the trunk of her car.
SC: The investigator discovered that the necklace described by the lady matched earrings which were found in the trunk of her car.
(0.1, 4.5, 2.18, 2.85, 3.91)
15. RC: The river which the valley captured by the enemy contains has its source at a glacier.
SC: The commander knows that the valley captured by the enemy contains a river that has its source at a glacier.
(0.2, 2.09, 3.03, 3.03, 1.52)
16. RC: The table which the sofa scratched by the cat was near was made of bamboo.
SC: Jason noticed that the sofa scratched by the cat was near a table made of bamboo.
(−1.45, 4.84, 2.85, 3.79, 3.36)

APPENDIX B

Residual and Raw Reading Times per Word (in Milliseconds) for Experiment 1

	Condition			
	RC		SC	
Region	Ambiguous	Unambiguous	Ambiguous	Unambiguous
1	−18.0 (340)	−18.9 (339)	−18.9 (341)	−26.1 (335)
2	−3.0 (339)	−15.1 (327)	−29.8 (327)	−23.2 (334)
3	−14.4 (334)	−16.1 (332)	−39.4 (308)	−41.8 (307)
4		54.8 (392)		−13.2 (324)
5	53.4 (434)	7.2 (389)	−16.7 (364)	−40.2 (340)
6	84.5 (424)	18.7 (358)	−20.4 (320)	−29.7 (311)
7	123.9 (474)	79.2 (431)	−10.1 (342)	−18.8 (333)
8	44.3 (388)	29.6 (373)	−0.5 (348)	−14.6 (333)

APPENDIX C

Items for Experiment 2: The Following Items Were Fillers in the Main Experiment

The ambiguous forms of the RC and NN versions of each item in Experiment 2 are given below. Items 1–20 are the implausible NN items; items 21–40 are the plausible NN items. Norms are listed after each item in the following order: plausibility rating for the NN; plausibility rating for the RC; head bias; and co-occurrence bias, calculated as log (co-occurrence frequency)/log (initial noun frequency). The unambiguous version of a NN item is formed by removing the initial determiner "the." The unambiguous version of a RC item is formed by inserting the wh-pronoun "which" between the two initial nouns in the item. Items 18–20 were excluded from analyses for reasons cited in the text.

1. RC: The country enemies may soon attack decided to increase military spending and tighten restrictions on immigration.
NN: The country enemies may soon attack innocent people despite the harsh penalties the country would levy against them as a result.
(3.96, 3.84, 0.947, 0)

APPENDIX C—*Continued*

2. RC: The park dogs can play in is beside the pond near the railway station.
   NN: The park dogs can play in the pond while their owners relax on a bench nearby.
   (4.3, 3.2, 0.805, 0)
3. RC: The appliance cats tend to damage is a dishwasher that vibrates too much.
   NN: The appliance cats tend to damage old dishwashers when the mechanical vibrations start to bother them.
   (6.2, 4.65, 0.658, 0)
4. RC: The chemical refrigerators could run on is freon or a closely related man-made substitute.
   NN: The chemical refrigerators could run on freon gas but they usually run on carbon dioxide gas which is cheaper.
   (3.95, 2.8, 0.381, 0.139)
5. RC: The warranty televisions usually come with is for one year but includes no labor.
   NN: The warranty televisions usually come with a contract since the companies that make them have many legal responsibilities toward their customers.
   (5, 3.5, 0.659, 0)
6. RC: The cafeteria students frequently eat at serves both vegetarian and kosher meals.
   NN: The cafeteria students frequently eat at cheap diners although the cafeteria will feed them for free.
   (4.2, 1.65, 0.758, 0)
7. RC: The doctor women will rely on is a skilled physician who listens carefully to their problems.
   NN: The doctor women will rely on the nurses despite the reputation for incompetence among the hospital staff.
   (5.5, 2.35, 0.98, 0.185)
8. RC: The newspaper neighbors frequently argue over is the local community paper delivered on Wednesday afternoon.
   NN: The newspaper neighbors frequently argue about editorial columns but the arguments rarely get violent.
   (6.05, 2.7, 0.869, 0)
9. RC: The fusebox electricians have to service contains many old and frayed wires.
   NN: The fusebox electricians have to service the wiring while all the painters have to work on the ceilings of the old building.
   (4.3, 2.2, 0.885, 0)
10. RC: The stroller mothers prefer to push has large rubber wheels and a good breaking system.
    NN: The stroller mothers prefer to push baby carriages rather than carry their children in their arms.
    (5.4, 2.9, 0.968, 0)
11. RC: The outfit girls like to wear is not always the one that they look best in.
    NN: The outfit girls like to wear high heels although most woman prefer more comfortable shoes.
    (5.45, 2.95, 1, 0)
12. RC: The award authors want to receive is the Pulitzer Prize in literature.
    NN: The award authors want to receive lucrative contracts since they know publishers make a lot of money from their books.
    (4.35, 2.1, 0.868, 0.169)
13. RC: The program lawyers love to watch is a courtroom drama called Law and Order on Tuesday nights.
    NN: The program lawyers love to watch hospital dramas although courtroom shows are usually televised at the same time on other channels.
    (4.4, 2.2, 0.95, 0.142)
14. RC: The doll children like to have is a Barbie doll of one kind or another.
    NN: The doll children like to have doll houses because dolls should live in homes just like real people.
    (5.15, 2.35, 0.925, 0.246)
15. RC: The nation hunters like to visit is in central Africa where wild game are plentiful.
    NN: The nation hunters like to visit exotic countries when the weather at home gets cold and wet.
    (4.75, 3.95, 0.965, 0)
16. RC: The tool plumbers need to have is a good monkey wrench that will loosen rusted pipes.
    NN: The tool plumbers need to have big toolboxes because unforeseen problems often arise on the job.
    (5.6, 1.85, 0.899, 0)
17. RC: The duty priests have to fulfill is to be faithful to God and sympathetic to their congregation.
    NN: The duty priests have to fulfill a mission when they graduate from the seminary.
    (4.9, 2.5, 0.884, 0)
18. RC: The stuffing pillows are filled with is usually a blend of synthetic fibers because natural fibers can cause terrible allergies.
    NN: The stuffing pillows are filled with synthetic fibers because natural fibers can cause terrible allergies.
    (2.52, 3.8, N/A, N/A)

APPENDIX C—*Continued*

19. RC: The customer receptionists are nice to is courteous and doesn't demand too much attention.
    NN: The customer receptionists are nice to considerate clients but also treat mean clients well if they are wealthy.
    (2.72, 4.4, N/A, N/A)
20. RC: The inn businessmen can stay at provides bagels for breakfast along with orange juice and coffee.
    NN: The inn businessmen can stay at lousy hotels when they are unable to find a better place to stay.
    (4.5, 2.6, 0.95, 0)
21. RC: The alley mice run rampant in is damp and dimly lit but relatively clean.
    NN: The alley mice run rampant in dark streets because there isn't much traffic there to scare them away.
    (2, 2.72, 0.859, 0)
22. RC: The kitchen lamps shine brightest in is one with lots of white tile and little dark wood.
    NN: The kitchen lamps shine brightest in the corner because the paint there reflects their light to the rest of the room.
    (1.4, 5.12, 0.717, 0.219)
23. RC: The shirt hooks tend to rip is made of fine silk and is quite delicate.
    NN: The shirt hooks tend to rip silk garments regardless of the way in which they are hung.
    (2.08, 3.84, 0.912, 0)
24. RC: The coat shops are now advertising is being marketed to young professionals
    NN: The coat shops are now advertising winter coats despite the warm weather we have been having.
    (1.96, 5.72, 0.859, 0.173)
25. RC: The river kayaks float slowly down is broad and contains a large volume of water.
    NN: The river kayaks float slowly down shallow streams since the current isn't very strong in shallow waters.
    (2.84, 3.8, 0.791, 0.369)
26. RC: The track horses can run around needs to be cleaned often because horse manure accumulates very rapidly even when the horses are racing.
    NN: The track horses can run around muddy fields if they are fitted with special racing shoes.
    (2.24, 3.2, 0.937, 0.24)
27. RC: The egg boxes will not crush possesses a shell that is half an inch thick.
    NN: The egg boxes will not crush most eggs although they might crush some with especially weak shells.
    (3.48, 5.36, 0.814, 0.426)
28. RC: The highway billboards are placed along becomes extremely congested during rush hour.
    NN: The highway billboards are placed along major freeways so that many potential customers can see them.
    (1.32, 3.36, 0.76, 0.382)
29. RC: The jacket pockets are sewn on is good for keeping your hands warm though it isn't very fashionable.
    NN: The jacket pockets are sewn on the fabric although certain adhesives are more durable than thread.
    (1.76, 4.64, 0.959, 0.563)
30. RC: The wall tiles are falling from has started to leak but it should still be sturdy enough to support the roof.
    NN: The wall tiles are falling from the bathroom but the tiles on the floor remain fixed in place.
    (1.76, 4.56, 0.983, 0.542)
31. RC: The sidewalk stones are piled near will be torn up when workmen from the city come to dig up electric wires and water mains.
    NN: The sidewalk stones are piled near the curb whenever workmen from the city come to dig up electric wires and water mains.
    3.32, 4.72, 0.832, 0.135)
32. RC: The sink water always stagnates in requires a good cleaning to get rid of its awful odor.
    NN: The sink water always stagnates in the kitchen until someone is willing to unclog the drain.
    (2.36, 3.92, 1, 0.391)
33. RC: The house shingles are nailed to is not worth as much as the brick houses in the neighborhood.
    NN: The house shingles are nailed to wood boards because glue simply cannot hold them in place.
    (2.12, 5.04, 0.946, 0.182)
34. RC: The beach trucks are driven on is less than a mile from the beach where people swim.
    NN: The beach trucks are driven on wet sand since it offers more traction than dry sand.
    (3.2, 4.6, 0.889, 0.0989)
35. RC: The commander pilots receive orders from wears two stars to display his high rank.
    NN: The commander pilots receive orders from mission control while other pilots receive orders from officers in the field.
    (2.48, 2.2, 0.821, 0.195)

APPENDIX C—*Continued*

36. RC: The restaurant tables are placed behind is trying to gain more business with outside seating.
NN: The restaurant tables are placed behind a fence so people driving by won't see them.
(1.52, 4.6, 0.848, 0.489)
37. RC: The desk pens write best on has a hard flat surface and plenty of space to spread papers out.
NN: The desk pens write best on legal pads rather than typing paper or post-it notes.
(3.16, 5.16, 0.868, 0.363)
38. RC: The school computers help to organize is running smoothly because administrators have less paperwork than they used to.
NN: The school computers help to organize class schedules before the term but few students take advantage of them.
(1.56, 4.12, 0.665, 0.484)
39. RC: The beeper clips are attached to sells very well because it is easy to carry around.
NN: The beeper clips are attached to belt loops so that pockets can be left free to carry other things.
(2.52, 3.32, 0.87, 0.0679)
40. RC: The juice blenders are corroded by is highly acidic and can also cause stomach problems.
NN: The juice blenders are corroded by acidic liquids while more alkaline liquids do very little to damage them.
(2.04, 4.52, 0.711, 0.174)

APPENDIX D

Residual and Raw Reading Times per Word (in Milliseconds) for Experiment 2

	Condition			
	Plausible		Implausible	
Word	Ambiguous	Unambiguous	Ambiguous	Unambiguous
1. The	−13.3 (319)	−21.0 (311)	−14.0 (316)	−5.4 (327)
2. tool	−15.8 (345)	−29.2 (329)	−19.6 (346)	−18.3 (353)
3. which		−35.2 (316)		−7.0 (344)
4. plumbers	−20.0 (342)	−28.9 (333)	−18.5 (359)	−51.1 (328)
5. need	7.0 (346)	27.4 (367)	63.9 (414)	−10.1 (344)
6. to	−1.7 (359)	3.9 (360)	15.2 (342)	−17.9 (312)
7. have	−4.4 (337)	12.2 (350)	−20.0 (328)	−41.8 (304)
8. is	9.1 (347)	23.4 (358)	−16.1 (314)	−33.3 (299)
9. a	345.6 (689)	14.6 (356)	38.6 (367)	−25.0 (304)
10. good	84.9 (434)	−18.2 (326)	−17.0 (339)	−53.6 (309)
11. monkey	5.3 (352)	−37.9 (312)	−37.2 (315)	−48.0 (304)
12. wrench	−13.5 (330)	−49.2 (294)	−38.3 (309)	−46.3 (304)
13. that	−10.0 (341)	−44.5 (299)	−35.6 (313)	−31.6 (318)
14. will	−30.0 (313)	−55.8 (298)	−41.9 (307)	−43.7 (306)
15. loosen	−8.9 (338)	−38.2 (313)	−33.2 (337)	−46.5 (317)

REFERENCES

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.

Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J. (1998). Late closure in context. *Journal of Memory and Language*, **38,** 459–484.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language*, **33,** 251–283.

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

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

Clark, H. H. (1983). Making sense of nonce sense. In G. B. Flores d'Arcais & R. Jarvella (Eds.), *The process of understanding language*. New York: Wiley.

Clark, H. H., Gelman, S. A., & Lane, N. N. (1985). Compound nouns and category structure in young children. *Child Development*, **56,** 84–94.

Clifton, C., Jr., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, **4,** SI77–SI103.

Clifton, C., Jr., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht, The Netherlands: Kluwer.

Condoravdi, C. (1989). The Middle: Where semantics and morphology meet. In *MIT Working Papers in Linguistics* (Vol. 11). Cambridge, MA: MIT Press.

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

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht, The Netherlands: Kluwer.

Downing, P. (1977). On the creation and use of English compound nouns. *Language*, **53,** 810–842.

Elman, J. L. (1991). Distributed representations, simple recurrent networks and grammatical structure. *Machine Learning*, **7,** 195–225.

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

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language*. New York: Holt, Rinehart & Winston.

Fodor, J. D., & Ferreira, F. (1998). *Reanalysis in sentence processing*. Boston: Kluwer.

Ford, M., Bresnan, J., & Kaplan, R. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727–796). Cambridge, MA: MIT Press.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–585). Hillsdale, NJ: Erlbaum.

Frazier, L., & Flores d'Arcais, G. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, **28,** 331–344.

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, **26,** 505–526.

Gagne C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23,** 71–87.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, B. (1997). The relative 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 doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

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

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain*. Cambridge, MA: MIT Press.

Gibson, E., & Hickock, G. (1993). Sentence processing with empty categories. *Language and Cognitive Processes*, **8,** 147–161.

Gibson, E., & Pearlmutter, N. (1994). A corpus-based analysis of psycholinguistic constraints on PP attachment. In C. Clifton Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 181–198). Hillsdale, NJ: Erlbaum.

Gibson, E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Science*, **2,** 262–268.

Gibson, E., Pearlmutter, N., & Torrens, V. (1999). Recency and lexical preferences in Spanish. *Memory & Cognition*, **27,** 603–611.

Gorrell, P. (1995). *Syntax and perception*. Cambridge, UK: Cambridge Univ. Press.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge, UK: Cambridge Univ. Press.

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

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

Kiparksy, P. (1982). *Explanation in phonology*. Dordrecht, The Netherlands: Foris.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*, Providence, RI: Brown Univ. Press.

Lawrence, S., & Giles, C. L. (1998). Searching the World Wide Web. *Science*, **280,** 98–100.

Levi, J. N. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.

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

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

Marcus, M. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, **19,** 313–330.

Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, **24,** 469–488.

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, **29,** 259–288.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden-paths: Assessing the contribution of syntax, semantics, and plausibility in resolving ambiguities. *Language and Cognitive Processes*, **11,** 283–334.

Pearlmutter, N. J., & Gibson, E. (2001). Recency in verb phrase attachment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **27,** 574–590.

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

Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, **6,** 229–259.

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

Rayner, K., Garrod, S., & Perfetti, C. A. (1992). Discourse influences during parsing are delayed. *Cognition*, **45,** 109–139.

Schütze, C., & Gibson, E. (1999). Argumenthood and English Prepositional phrase attachment. *Journal of Memory and Language*, **40,** 409–431.

Sedivy, J. (1999). *Investigating the discourse-based properties of adjectives in on-line semantic interpretation*. Paper presented at the 12th CUNY sentence processing conference.

Seidenberg, M. S., Haskell, T., & MacDonald, M. C. (1999). *Constraints on plurals in compounds: Some implications of compounds research*. Talk given at the 40th Annual Meeting of the Psychonomic Society.

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

Spivey-Knowlton, M. J., & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, **55,** 227–267.

Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 415–439). Hillsdale, NJ: Erlbaum.

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

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,** 211–272.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language and communication* (pp. 217–262). San Diego, CA: Academic Press.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, **35,** 566–585.

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.

(Received August 26, 1999)
(Revision received March 6, 2001)
(Published online November 9, 2001)