The use of thematic role information in parsing: Syntactic processing autonomy revisited

Charles Clifton Jr.,[a],[*] Matthew J. Traxler,[b] Mohamed Taha Mohamed,[a]
Rihana S. Williams,[c] Robin K. Morris,[c] and Keith Rayner[a]

[a] *Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA*
[b] *University of California, Davis, CA 95616, USA*
[c] *University of South Carolina, SC 29208, USA*

Abstract

Two eye-movement experiments examined the processing of sentences containing reduced relative constructions. In the first experiment, animacy of the sentential subject, structural ambiguity, and parafoveal preview of syntactically disambiguating material were manipulated. Evidence of disruption was found in temporarily ambiguous sentences, regardless of animacy or preview. In the second experiment, readers with high versus low verbal working memory capacity read the sentences from Experiment 1. High and low-span readers exhibited very similar patterns of processing. As in the first experiment, evidence for disruption was found in temporarily ambiguous sentences whether the sentential subject was animate or inanimate. Sentences with animate subjects were hard to interpret, and relatively late measures of processing indicated that an animate subject made ambiguity especially hard to overcome. We interpret the findings as being consistent with serial, depth-first models of parsing.
© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Syntax; Parsing; Reduced relative clause; Garden-pathing; Eyetracking

One theoretical approach to sentence parsing gives logical priority to the grammatical structure of a sentence. The reader's or listener's knowledge of the grammar of his or her language provides a structure that constrains how the meanings of the words in the sentence are combined into a sentence interpretation. In some versions of this approach (e.g., Ferreira & Clifton, 1986; Frazier, 1977, 1987, 1990; Frazier & Rayner, 1982), the reader/listener initially constructs a single grammatical analysis of the sentence and then interprets it, revising the analysis if necessary. While constructing a single analysis is not an obligatory part of a model that gives grammar logical priority in sentence comprehen-

sion (e.g., Gibson, 1991, 1998), we will concentrate on serial "depth-first" models (Frazier, 1995) that do assume that a single analysis is initially chosen, and in fact that assume that the initial analysis is simply the first one that is completed (see Frazier, 1987; Frazier & Clifton, 1996; cf. Traxler, Pickering, & Clifton, 1998).

A contrasting approach to sentence parsing views grammatical structure as one of multiple interacting constraints on sentence interpretation (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995). In such a constraint-based model, grammatical structure may carry a heavy weight in determining the interpretation of a sentence, but it does not necessarily take precedence over such factors as plausibility and contextual appropriateness.

Serial, depth-first models have stimulated a great deal of research, including some that gave them surprisingly

---
[*] Corresponding author. Fax: 1-413-545-0996.
E-mail address: cec@psych.umass.edu (C. Clifton Jr.).

strong support (e.g., Ferreira & Clifton, 1986; Frazier & Rayner, 1982; Pickering & Traxler, 1998; Rayner, Carlson, & Frazier, 1983). However, much of the later research that these models stimulated was designed to show that non-grammatical factors could affect the difficulty of comprehending sentences, even obscuring or eliminating the apparent contribution of grammatical factors (see Clifton & Duffy, 2001; Mitchell, 1994, and Rayner & Clifton, 2002, for reviews of this literature). Judging from a haphazard sample of recent textbooks in psychology and in cognitive psychology, it seems fair to conclude that many cognitive psychologists judge that constraint-based models have carried the day. Several factors may contribute to this judgment. First, the constraint-based models appealed to the connectionist metaphor, which was very popular when much of the research was conducted (and in fact, in some cases have been implemented as connectionist models; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey & Tanenhaus, 1998). Second, experimental psychologists have been extremely cautious about basing cognitive processes on grammatical rules, which appear to change frequently with seemingly-arbitrary theoretical changes in linguistics. Third, some results provide dramatic and apparently-convincing evidence that factors of meaning and plausibility can completely override the grammatical factors that take precedence in a depth-first model of sentence comprehension.

Two of the most convincing and often-cited examples of such evidence are those provided by Trueswell, Tanenhaus, and Garnsey (1994) and by Just and Carpenter (1992), following up an earlier report by Ferreira and Clifton (1986). Ferreira and Clifton used eye-movement monitoring techniques to show that reading was disrupted when a sentence fragment that could be understood as a subject plus a main verb was disambiguated toward the normally unpreferred reduced relative clause analysis. This disruption, or 'garden-pathing,' was observed even when the main verb analysis was rendered implausible by using an inanimate noun as the grammatical subject of the sentence that was an unlikely agent of the verb. As one concrete example, Ferreira and Clifton showed disruption of reading the region *by the lawyer* in both (1a) and (2a), compared to their unambiguous counterparts (1b, 2b).

(1a) The defendant examined by the lawyer turned out to be unreliable.
(1b) The defendant that was examined by the lawyer turned out to be unreliable.

(2a) The evidence examined by the lawyer turned out to be unreliable.
(2b) The evidence that was examined by the lawyer turned out to be unreliable.

The initial portion of sentences like (1a), *The defendant examined...* is thought to be given a main clause analysis

in which *the defendant* is subject and typically agent of *examined* (following the Minimal Attachment strategy, Frazier, 1987, in a depth-first account, or honoring the high frequency of main clause analyses in some versions of a constraint-based account; cf. Bever's, 1970, initial discussion of the infamous *The horse raced past the barn fell* garden path). The agentive phrase *by the lawyer*, and the following main verb phase *turned out...*, are inconsistent with the main clause analysis, forcing it to be revised. This slows reading compared to (1b), in which the phrase *that was examined by...* is unambiguously a relative clause that modifies the first noun, *defendant*.

Ferreira and Clifton (1986) reasoned that if readers used plausibility and meaning, not just grammatical knowledge, to guide their initial interpretation of a sentence, then the disruption of reading would disappear in (2a), with an inanimate initial noun. *Evidence* is not a plausible subject of the verb *examined*, and this implausibility should block the analysis in which it is the subject. Rather, since it is a plausible theme (or patient; the entity affected by the action) of *examined*, it will be initially taken as the head of the (reduced) relative clause *examined by the...* In this case, no revision will be needed when the eyes reach the *by the...* phrase and no disruption will take place. However, Ferreira and Clifton (1986) observed very comparable amounts of disruption in (1a) and (2a). They argued that this meant that *evidence* was initially taken to be the subject of the verb *examined*, contradicting an apparent prediction of a constraint-based account and supporting a serial, depth-first model in which grammatical constraints are honored first even when they clash with plausibility.

Trueswell et al. (1994) challenged the Ferreira and Clifton (1986) demonstration by noting that Ferreira and Clifton selected their implausible subject sentences on intuitive grounds and arguing that several of their sentences were not in fact fully implausible on the main verb analysis. Trueswell et al. improved upon Ferreira and Clifton's materials and used (in their Experiment 2) items that Burgess (1991) had developed using sentence-completion norms, so that a fragment consisting of an inanimate noun plus a verb was seldom completed as a subject plus an object. They further had participants rate the typicality of each initial noun as an agent and as a theme of the first verb and selected inanimate items that were rated as atypical agents but typical themes. Trueswell et al. reported that, with these improved materials, little or no trace remained of the disruption that Ferreira and Clifton (1986) had reported for the sentences with an initial inanimate noun. Trueswell et al. concluded that their more adequate manipulation of the plausibility of the initial subject + verb analysis did override any structural preference for the subject + verb analysis, and argued that this supported a constraint-based account over a serial depth-first account.

Just and Carpenter (1992) also noted that some of the materials used by Ferreira and Clifton (1986) were not adequately biased, but were more concerned about individual differences in parsing strategies caused by differences in the cognitive resources brought to the parsing task. Just and Carpenter replicated the Ferreira and Clifton results (even after changing a few of the sentences) when the participants had reading spans of 2.5 or less as measured by the Daneman and Carpenter (1980) reading span task, but reported that beginning a sentence with an inanimate initial noun (implausible as agent of the first verb) did benefit participants whose reading spans were 4.0 or higher. They suggested that only the high-span participants had the processing capacity needed to take plausibility into account.

The experiments to be reported here revisit the question of whether readers make use of semantic information (including animacy) to avoid syntactic misanalysis of reduced relative clauses. The Just and Carpenter (1992) results have some peculiarities that make it worthwhile attempting to replicate them. In particular, the advantage that an inanimate initial noun gave their high-span participants was equally as large for their unambiguous (unreduced relative clause) sentences as for their temporally ambiguous sentences, raising a question about whether animacy guided parsing decisions or simply affected the ease of interpreting the test sentences. Further, although it was not mentioned in the published report, most of the participants in Ferreira and Clifton's (1986) research were graduate students of psychology at the University of Massachusetts, and thus not likely to be a group of low-span individuals. Finally, as Waters and Caplan (1996a) noted, even Just and Carpenter's high-span readers showed evidence of disrupted reading upon encountering syntactically disambiguating information in sentences with inanimate initial nouns.

The results of Trueswell et al. (1994) appear to be very convincing, but research conducted by Burgess (1991; reported in Spivey-Knowlton, Trueswell, & Tanenhaus, 1993) raises an interesting potential reason for the difference between their results and those of Ferreira and Clifton (1986). Burgess conducted experiments measuring the self-paced reading of sentences like those used by Trueswell et al. The sentences were presented either in two-word chunks so that the preposition *by* appeared together with the initial verb (*The defendant/examined by/...*) or word-by-word so that the verb would be presented separately from the preposition. In this latter case, the results were very similar to those reported by Ferreira and Clifton. When the preposition was presented together with the verb, however, inanimacy of the initial noun eliminated reading disruption, as reported by Trueswell et al. It may be that the grammatical information presented

(perhaps parafoveally) by the preposition *by* is needed to guide the sentence processing system in using the information that the initial inanimate noun is not a suitable subject for the verb that follows (see also the discussion of the *by*-constraint in McRae et al., 1998). Alternatively, preview of the preposition *by* may trigger syntactic reanalysis while readers are still fixating the preceding word.

How does this bear on the difference between Trueswell et al. and Ferreira and Clifton? The latter researchers used a display with only 40 characters across the screen, so that they had to divide the *by the NP* phrase in the middle (as noted by Trueswell et al.). This, together with the relatively coarse display on the point-plotting oscilloscope they used for a display device, may have reduced the value of *by the* as a parafoveal preview in the Ferreira and Clifton experiment (and, as noted by Trueswell et al., the line divisions used by Ferreira and Clifton resulted in unnaturally short first lines in the ambiguous conditions, which could have slowed reading time on the disambiguating second lines).

The first experiment to be reported here tests the importance of parafoveal preview of the *by* phrase. It used the boundary change technique (Rayner, 1975) so that in half the conditions, the reader was provided with preview of the *by* phrase, but in the other half of the conditions, no valid information was provided about *by the*, which became legible only after the eye had landed on it. It thus tested the suggestion (McRae et al., 1998; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993) that parafoveal information restricting the syntactic analysis of a sentence can combine with animacy information to guide the parsing process. The experiment also used materials used by Trueswell et al. (1994), so that the no-boundary-change condition would be largely a replication of their research. The second experiment to be reported did not manipulate preview, but instead tested high vs. low reading span participants from a different population using the same experimental materials. It was designed to evaluate Just and Carpenter's (1992) claims about limitations that low-span readers may have on using semantic information to guide parsing.

Experiment 1

Method

Participants

Twenty-four University of Massachusetts undergraduates participated in the experiment for cash or course credit. They were all native English speakers and had normal uncorrected vision. Each was tested in a single session approximately 45 min long.

Table 1
An example of the four versions of one pair of sentences

Sentence form 1	
Animate reduced	The man/paid/by the parents/was unreasonable
Animate unreduced	The man/who was/paid/by the parents/was unreasonable
Inanimate reduced	The ransom/paid/ by the parents/was unreasonable
Inanimate unreduced	The ransom/that was/paid/ by the parents/was unreasonable
Sentence form 2	
Animate reduced	The man/paid/by the parents/saved their/son's life
Animate unreduced	The man/who was/paid/by the parents/saved their/son's life
Inanimate reduced	The ransom/paid/ by the parents/saved their/son's life
Inanimate unreduced	The ransom/that was/paid/by the parents/saved their/son's life

Materials

Twenty-four of the sentences from Experiments 1 or 2 of Trueswell et al. (1994) served as the experimental items.[1] To increase the number of experimental items each participant could see, we made two versions of each sentence by writing a new ending for each of the original 24 sentences. The two versions differed only after all critical material, and permitted each participant to read sentences with both the inanimate and the animate subjects from the original sentence. This resulted in a total of 48 sentences. Four versions of each sentence were constructed, created by the factorial combination of animate vs. inanimate initial noun phrase (typical vs. atypical as agent, and atypical vs. typical as theme, in the ratings provided by Trueswell et al.), crossed with reduced vs. unreduced relative clause. One full set of examples for a single sentence appears in Table 1, together with indications of the regions to be analyzed. All sentences appear in the Appendix. These 48 sentences were combined with four practice sentences and 72 filler sentences of a variety of syntactic forms, including 24 that were temporarily ambiguous between main clause and reduced relative clause analyses but resolved in favor of the main clause analysis.

Each of the experimental sentences had a version in which the *by the* portion of the agentive phrase was replaced with a string of six of letters (e.g., *hlekde*). A different string of six letters was used for each experimental sentence, and in each string, each letter was randomly chosen with a probability equal to its frequency of occurrence in printed text. In the no parafoveal preview condition, these strings were displayed in place of *by the* until the eye moved into the space before *by*. When a very fast display change (5 ms maximum) is employed, this procedure produces minimal disruption in reading. Most readers report no awareness of the display change, but some readers report that a flicker occurred on some items. They almost never can report the content of the preview material (unless it consists of a string of repeated *x*'s) (Rayner, 1998). Further, Inhoff, Starr, and Liu (1998) presented data suggesting that the effects of display changes were not related to degradation of visual quality but "derived solely from the masking of useful text" (p. 101). Each of the filler sentences also had one word or one short sequence of words replaced by random letters until the eye entered the word or sequence. The word or sequence was chosen approximately equally often from all positions of the sentence.

A two-choice question was made up for one-half of all the sentences, including 24 of the 48 experimental sentences. Two-thirds of the questions of the experimental sentences concerned some aspect of the thematic relations involved in the reduced relative clause.

Design

Each participant received six experimental sentences in each of eight conditions. The conditions were defined by the factorial combination of animate vs. inanimate initial NP, reduced vs. unreduced relative clause, and presence vs. absence of parafoveal preview of the agentive *by* phrase. Counterbalancing procedures were used to ensure that, across all participants, each sentence was tested equally often in each of the eight conditions. If a given sentence was tested with an inanimate first noun using one of the two endings, it was tested with the animate first noun using the other ending (and similarly, if the sentence was tested as an unreduced relative clause with one ending, it was tested as a reduced relative clause with the other ending, and if it was tested with parafoveal preview with one ending, it was tested without preview with the other).

Apparatus and procedure

The stimuli were presented on a NEC 4FG monitor through a VGA video board operating at a 200 Hz refresh rate and controlled by a 486 PC. An A to D

---

converter interfaced the computer with a Fourward Technologies Generation V Dual Purkinje eyetracker. The eyetracker monitored movements of the right eye, although viewing was binocular. Letters were formed from a 7×8 array of pixels, using the Borland C default font. Participants sat 61 cm away from a computer screen and silently read single line sentences while their head position was stabilized by a bite bar. At this viewing distance, 3.8 letters equaled one degree of visual angle. At the beginning of the experiment, the eyetracking system was calibrated for the participant. At the start of each trial, a check calibration screen appeared, and participants who showed a discrepancy between where their eye fixated and the location of the calibration squares were recalibrated before the next trial. A trial consisted of the following events: The check calibration screen appeared and the experimenter determined that the eye-tracker was correctly calibrated. The participant was instructed to look at the calibration square on the far left of the screen, then the experimenter presented the sentence. The participant read the sentence silently and at his/her own pace, then clicked a response key to make the sentence disappear. Following half of the trials, a comprehension question appeared on the screen. The participant responded by pressing the response key that corresponded with the position of the correct answer. Then the check calibration screen appeared before the next trial. The experiment was completed in one session of approximately 45 min.

When a sentence involved a display change, the computer initiated the change from random letters to the correct letters as soon as the eye had entered the space before the critical word (by in the experimental sentences) and completed the change within one raster sweep, 5 ms. This occurred while the eye was still moving and thus vision was suppressed. The relatively few trials in which the display change was initiated at the very beginning or end of a saccade were discarded. While participants were aware that some display changes had taken place, they were uniformly unable to report anything about the details of the change.

On the first four trials of an experimental session, four practice sentences were presented. Following that, the order of presentation of sentences was individually randomized for each participant.

Results and discussion

The analyses were conducted by dividing each sentence into regions as indicated by the / character in (3):

(3) The evidence—defendant$_1$/(That—who was)$_2$/examined$_3$/by the lawyer$_4$/turned out to be$_5$/unreliable$_6$.

Region 1 contained the head of the subject NP, Region 2 contained the relative pronoun plus the auxiliary in the unambiguous versions (but was empty in the ambiguous versions), and Region 3 contained the verb of the relative clause. Region 4 contained the phrase *by the NP* which effectively disambiguates an ambiguous sentence. Region 5 contained the matrix verb, which fully disambiguates the sentence, plus one or two following words if they completed the sentence or (as in the present example) went up to a line break. Region 6 contained any remaining material in the sentence.

Following standard procedure in eye movement experiments (Rayner, 1998; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989) a variety of measures of eye movements were analyzed. The informative measures include (a) first fixation duration, the duration of the initial fixation (if any) in a region, not counting trials in which a fixation past the region had occurred previously; (b) first pass time, the summed duration of fixations in a region from first entering it to first leaving it, to the left or to the right (not counting trials on which there was no initial fixation in the region; note, this measure is commonly called "gaze duration" when the region is a single word); (c) regressions out, the percentage of trials in which a first pass fixation in a region was followed by a regression to an earlier region; (d) regression path duration, the summed duration of all fixations from first entering a region to first going past the region, including regressive fixations to earlier regions (Konieczny, Hemforth, Scheepers, & Strube, 1997; Liversedge, Paterson, & Pickering, 1998; cf. Rayner & Duffy, 1986, for an early use of the measure); and (e) second pass time, the summed duration of all fixations made in a region after going past the region and then regressing back into it (counted as zero if no such regression was made). Table 2 presents means of these eye movement measures for Regions 1–5.[2] Fig. 1 presents the regression path durations, which we consider to be the most inclusive reflection of early disruption in reading. Analyses of variance with three within-participant factors (initial noun animacy, ambiguity, and presence vs. absence of preview) were conducted separately on Regions 3, 4, and 5 (and for second pass times, on Region 1). All significant results, and particularly informative non-significant results, will be reported.

Region 3: relative clause verb

Disambiguation (presence vs. absence of the *who/that was* sequence before the verb) decreased first fixation durations on the verb ($F1(1, 23) = 10.32$, MSe $= 1455$, $p < .01$; $F2(1, 47) = 7.96$, MSe $= 3248$, $p < .01$). The first

---

[2] Measures of initial reading times are not reported for Region 1, because the eyetracking software was configured to begin recording only 200 ms after the presentation of a sentence, on the assumption that any eye movements made during this time might simply reflect the display onset, not normal reading processes.

Table 2
Mean scores on five dependent measures by condition and region, Experiment 1

	Region											
	Preview						No preview					
	1	2	3	4	5	6	1	2	3	4	5	6
First fixation time												
Animate ambiguous	—	—	253	246	257	252	—	—	268	253	269	243
Animate control	—	220	247	235	239	268	—	211	239	241	254	245
Inanimate ambiguous	—	—	253	214	242	247	—	—	252	258	238	254
Inanimate control	—	232	234	226	247	250	—	224	235	250	244	246
First pass time												
Animate ambiguous	—	—	295	553	497	447	—	—	336	541	470	449
Animate control	—	297	297	458	482	490	—	280	302	504	482	468
Inanimate ambiguous	—	—	305	469	445	445	—	—	316	509	462	452
Inanimate control	—	313	269	427	483	431	—	298	315	486	430	411
Second pass time												
Animate ambiguous	84	—	106	118	76	1	74	—	165	91	112	3
Animate control	45	98	105	82	102	3	83	109	105	108	99	11
Inanimate ambiguous	44	—	45	44	74	2	68	—	91	27	56	0
Inanimate control	48	95	35	59	72	8	31	55	46	48	82	11
Regression path duration												
Animate ambiguous	—	—	342	651	645	753	—	—	373	763	600	796
Animate control	—	327	376	590	591	878	322		387	688	595	833
Inanimate ambiguous	—	—	366	540	509	682	—	—	411	663	498	670
Inanimate control	—	356	377	478	533	672	—	328	369	559	513	628
Regressions out												
Animate ambiguous	—	—	6	12	12	22	—	—	6	23	6	26
Animate control	—	3	16	11	10	25	—	8	15	20	6	30
Inanimate ambiguous	—	—	6	5	7	25	—	—	11	18	3	27
Inanimate control	—	6	20	4	4	15	—	4	14	10	9	30

Example sentence: The defendant (evidence)/(who (that) was)/examined/by the lawyer/turned out to be/unreliable ("/" marks indicate region boundaries).

fixation in the initial verb region averaged 239 ms following *who/that was* phrase (which disambiguated the relative clause) and 257 ms when this phrase had been absent. However, there were more initial regressions out of Region 3 when the verb followed the disambiguating phase that when the verb immediately followed the initial NP (16.25 vs. 7.25%; $F1(1, 23) = 9.07$, MSe $= 433$, $p < .01; F2(1, 47) = 19.82$, MSe $= 310$, $p < .001$). It appears that readers were likely to return quickly to the early part of the sentence upon reading the verb when that part of the sentence included a disambiguating *who/that was* phrase.

Preview vs. no preview of the *by the* sequence (Region 4) affected Region 3 reading times. Presence of preview did not affect first fixation duration, but speeded first pass reading times of Region 3 by 22 ms ($F1(1, 23) = 4.89$, MSe $= 4936$, $p < .05$; $F2(1, 47) = 5.01$, MSe $= 6653$, $p < .05$). The interaction of ambiguity and preview was non-significant ($p > .25$ in both first fixation and first pass measures), indicating that the effect of ambiguity was not mediated by any difficulty of processing a par-

afoveal *by the* phrase. It seems most likely that reading was simply disrupted by the absence of preview.

Second pass reading time in Region 3 reflected disruption caused by ambiguity, lack of preview, and animacy of the initial noun. Second pass time was faster for unambiguous than for ambiguous items ($F1(1, 23) = 9.02$, MSe $= 4555$, $p < .01$; $F2(1, 47) = 8.51$, MSe $= 9111, p < .01$). Further, second pass time was faster when there was preview than when no preview was present ($F1(1, 23) = 8.60$, MSe $= 4850$, $p < .01$; $F2(1, 47) = 6.25$, MSe $= 15015$, $p < .02$). Inanimate sentences had shorter second pass reading times than animate sentences ($F1(1, 23) = 22.97$, MSe $= 9170$, $p < .001$; $F2(1, 47) = 18.95$, MSe $= 22000$, $p < .001$). The only interaction that approached significance involved ambiguity and preview ($F1(1, 23) = 2.61$, MSe $= 9867$, $p = .12$; $F2(1, 47) = 3.69$, MSe $= 9012$, $p < .07$). This interaction, if it can be trusted, reflects a second pass time of 129 ms in the ambiguous, no preview condition, compared with second pass times ranging from 70 to 77 ms in the other conditions.
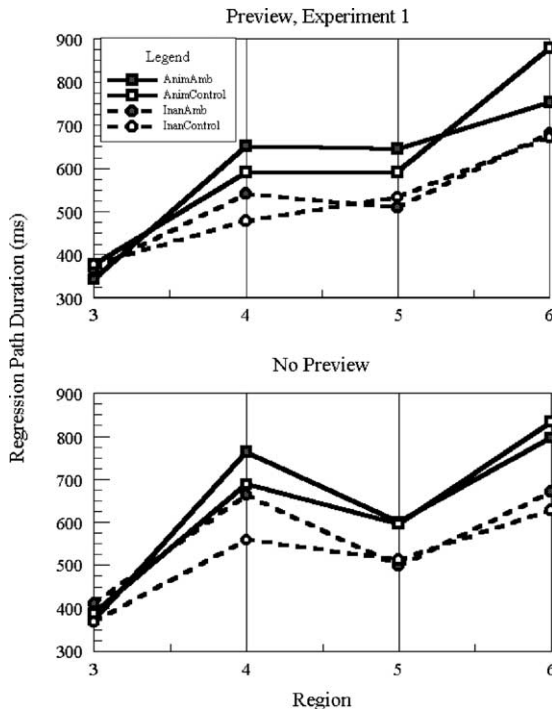
Fig. 1. Regression path durations, Experiment 1.

Region 4: ''by the NP''

Region 4 is critical in that it effectively disambiguates the sentence toward a relative clause analysis. Syntactic misanalysis would be indicated by disruption in the ambiguous conditions compared to the unambiguous conditions. First fixation duration was not affected by ambiguity, but first fixations were shorter when there was preview than when no preview had been provided, 231 vs. 251 ms ($F1(1, 23) = 14.64$, MSe = 1366, $p < .001; F2(1, 47) = 12.56$, MSe = 3238, $p < .001$). The benefit of preview on first fixation duration was significantly greater for inanimate than animate initial nouns, 33 vs. 7 ms (F1(1, 23) = 5.45, MSe = 1582, $p < .03$; $F2(1, 47) = 4.62$, MSe = 2639, $p < .04$). However, since this difference did not interact with ambiguity, it presumably reflects an effect of interpretation difficulty, not an effect of parsing decisions.

Similar to what was observed in Region 3, first pass reading time in Region 4 reflected disruption caused by ambiguity, lack of preview, and animacy of the initial noun. First pass time was faster for unambiguous (control) than for ambiguous sentences, 469 vs. 519 ms ($F1(1, 23) = 19.34$, MSe = 6078, $p < .001; F2(1, 47) = 14.20$, MSe = 16118, $p < .001$). It was also faster when preview was provided, 478 vs. 510 ms ($F1(1, 23) = 5.56$, MSe = 9156, $p < .03; F2(1, 47) = 5.56$, MSe = 16,399, $p < .03$) and faster for inanimate than animate initial nouns, 478 vs. 514 ms ($F1(1, 23) = 8.23$, MSe = 9996,

$p < .01; F2(1, 47) = 8.90$, MSe = 21,685, $p < .01$). The ambiguity effect was numerically smaller for inanimate than for animate initial nouns. The overall advantage of unambiguous sentences was 34 ms vs. 66 ms, for inanimate and animate initial nouns, respectively. When preview had been present, the advantages of unambiguous sentences were 42 vs. 95 ms for inanimate and animate initial nouns. The difference between the latter two effects is quite similar to the difference that was reported by Trueswell et al. (1994), Experiment 2 (29 vs. 128 ms). However, the crucial interaction between animacy and ambiguity was not significant in our data ($F1(1, 23) = 2.38, p > .13; F2(1, 47) = 0.45, p > .49$) (and we note that the interaction was not significant in either experiment in Trueswell et al., 1994, considered separately). Follow-up contrasts indicated that the effect of ambiguity was significant ($p < .05$) for the inanimate as well as the animate initial noun conditions, which was not the case in Trueswell et al. (1994).

Questions of statistical significance aside, the present first pass data are not too different from those reported in Trueswell et al. (1994). However, we do not believe that first pass reading time results tell the whole story This measure reflects the initial time spent looking at a region. Difficulty of integrating the material in that region into a developing sentence analysis can appear as lengthened fixations in that region, but it can also appear as quick regressions out of the region (Rayner & Sereno, 1994). The regression path duration measure (Konieczny et al., 1997; Liversedge et al., 1998) is sensitive to both these effects of processing. It is also sensitive to the difficulty of recovering from misanalysis (not unlike first pass time), but we submit that the existence of an effect in the regression path duration measure unambiguously indicates the presence of some processing difficulty even if the magnitude of the effect reflects in part the process of recovery.[3]

Regression path durations showed much the same pattern of effects as first pass times, except that the effect of ambiguity was fully as large for sentences with initial inanimate nouns as for those with animate nouns. Regression path durations were shorter for control than for ambiguous sentences, 580 vs. 655 ms ($F1(1, 23) = 10.16$, MSe = 26,640, $p < .01; F2(1, 47) = 12.83$, MSe = 40,485, $p < .001$). They were also shorter for sentences with preview than for sentences with no preview, 565 vs. 669 ms ($F1(1, 23) = 27.43$, MSe = 18,821, $p < .001$; $F2(1, 47) = 22.14$, MSe = 44,902, $p < .001$), and for sentences with inanimate rather than with animate initial nouns, 560 vs. 674 ms ($F1(1, 23) = 24.18$, MSe = 25,395,

---

[3] We note that this measure was not in general use when Ferreira and Clifton (1986) was published or when Trueswell et al. (1994) appeared (although it had been used as early as Rayner & Duffy, 1986).

$p < .001;$ $F2(1,47) = 18.74,$ $MSe = 69,183,$ $p < .001).$ No interaction approached significance (all $F < 1$). The effect of ambiguity was 83 ms for inanimate initial nouns and 68 ms for animate initial nouns (note, non-significantly greater for inanimate nouns).

Regressive eye movements out of the *by the NP* region were less frequent when there had been preview than when no preview had been provided, 8.2 vs. 18.8% $(F1(1,23) = 18.04,$ $MSe = 273,$ $p < .001; F2(1,47) = 15.76,$ $MSe = 656,$ $p < .001),$ and when the initial noun had been inanimate rather than animate, 9.6 vs. 16.9% $(F1(1,23) = 12.40,$ $MSe = 209,$ $p < .002; F2(1,47) = 15.76,$ $MSe = 656,$ $p < .001).$ The numerical difference between ambiguous and control conditions, 14.9 vs. 11.6%, was not significant $(p > .15).$

The only Region 4 effect that approached significance in second-pass times was animacy of the initial noun (45 ms for inanimate initial nouns, 100 ms for animate initial nouns; $F1(1,23) = 7.97,$ $MSe = 18,408,$ $p < .01;$ $F2(1,47) = 21.76,$ $MSe = 15,905,$ $p < .001).$ Since this effect did not interact with ambiguity, it can be considered to be an effect of the difficulty of interpreting sentences with initial noun phrases that are good agents but not good patients of the relative clause verb.

Region 5: main verb

Preview and ambiguity had no significant effect in Region 5, but there was some evidence of reading difficulty when the initial noun was animate. Animacy of that noun marginally slowed first fixations $(F1(1,23) = 3.28,$ $MSe = 4912,$ $p < .08; F2(1,47) = 7.18,$ $MSe = 952,$ $p < .02)$ and first pass times $(F1(1,23) = 5.91,$ $MSe = 6190,$ $p < .03; F2(1,47) = 3.61,$ $MSe = 30,994,$ $p < .07)$ and significantly increased regression path durations $(F1(1,23) = 13.81,$ $MSe = 31,078,$ $p < .001; F2(1,47) = 13.47,$ $MSe = 93,674,$ $p < .001).$ Second pass times were also marginally longer for animate than for inanimate sentences in Region 5 $(F1(1,23) = 4.08,$ $MSe = 8212,$ $p < .06; F2(1,47) = 6.22,$ $MSe = 19,409,$ $p < .02).$

Region 1: initial NP

Examination of second pass reading times for Region 1 provides some understanding of how readers coped with factors that made comprehension of animate and ambiguous sentences difficult. The only effect that was significant in both the participants and the items analysis was a three-way interaction of preview, ambiguity, and animacy $(F1(1,23) = 6.97,$ $MSe = 3456,$ $p < .02;$ $F2(1,47) = 6.14,$ $MSe = 10,548,$ $p < .02).$ Examination of Table 2 indicates a complex interaction: When preview had been present, re-reading times were long only in the animate initial noun, ambiguous sentence condition. However, when preview had been absent, re-reading times were long in all conditions except the easiest inanimate noun, unambiguous sentence condition. It appears that preview of the *by the* phrase did ease the

interpretation of sentences that contained either helpful semantic constraints or syntactic disambiguation. When preview was absent, readers needed both syntactic and semantic information to interpret sentences without difficulty.

Conclusions from Experiment 1

Experiment 1 offers no support for the suggestion (see Spivey-Knowlton, Trueswell, & Tanenhaus, 1993) that parafoveal preview of the *by the* phrase is needed for effective use of semantic biasing information provided by the initial NP. The effects of ambiguity and animacy were essentially the same, regardless of parafoveal preview. To the extent that there were differences between the first pass time results of Ferreira and Clifton (1986) and Trueswell et al. (1994), the differences were not likely due to inadequate preview in the former study.

Some aspects of the data from Experiment 1 are consistent with the principal finding of Trueswell et al. (1994): the effect of ambiguity on first pass reading time was numerically (but not significantly) reduced when the initial NP was inanimate. However, unlike the dramatic result of Trueswell et al.'s (1994) Experiment 1, in which NP inanimacy essentially eliminated the effect of ambiguity, the effect of ambiguity was significant even with an inanimate initial NP. Further examination of the data, especially the regression path duration measure, indicated evidence of substantial disruption in the processing of temporarily ambiguous sentences with inanimate initial NPs, not reduced in magnitude compared to the disruption observed with animate sentences. We submit that this additional examination of the data provides evidence that inanimacy of the initial noun did not eliminate garden-pathing.

Even though it did not eliminate garden-pathing, an inanimate initial noun slowed reading for both ambiguous and unambigous items, as had been observed in high-span participants by Just and Carpenter (1992). The general lack of interactions involving animacy and ambiguity indicates that animacy did not play a role in guiding parsing decisions, but instead affected the ease of constructing an interpretation of a sentence. However, the three-way interaction in second-pass times of Region 1 does suggest that inanimacy and parafoveal preview of the *by the* region can combine to help recovery from a garden-path. When both an inanimate initial noun and parafoveal preview of *by the* were present, syntactic ambiguity resulted in only minimal re-reading of the first region of the sentence, but when either was absent, syntactic ambiguity resulted in more re-reading (and when both were absent, substantial re-reading occurred regardless of syntactic ambiguity).

In short, we do not claim that our results are dramatically different from those reported by Trueswell et al., insofar as we used the measures that were in general use at the time Trueswell et al. reported their

data. However, we do claim that extended analyses of our data provide evidence that contradicts Trueswell et al's claim that ''only animate nouns showed clear signs of difficulty'' (p. 285).

Experiment 2

It is possible that limitations of the Experiment 1 participants, such as low reading span, were responsible for their failure to use information about the plausibility of the initial noun phrase as agent vs. theme of the first verb. Just and Carpenter (1992) reported that low reading span individuals are limited in their use of plausibility information. This suggestion that Experiment 1 participants were limited in this way is called into question by the fact that they did show effects of initial noun phrase animacy (or plausibility) in both ambiguous and unambiguous sentences, much like Just and Carpenter's high-span participants. Nonetheless, we directly examined the effect of reading span in Experiment 2, since it has been commonly suggested that the difference between the results reported by Trueswell et al. (1994) and those reported by Ferreira and Clifton (1986) is attributable to differences in their participants' reading ability.

The question of whether reading span affects parsing performance has attracted interest in part because existing accounts of any span effects differ in how easily they are accommodated by parallel constraint-based vs. serial depth-first accounts of parsing. A 'shared resource' account proposes that the lack of verbal working memory resources may limit comprehenders' ability to deploy probabilistic constraints in general and semantic plausibility information in particular (Just & Carpenter, 1992; Just, Carpenter, & Keller, 1996; cf. Traxler, 2002). On this account, readers who lack sufficient verbal working memory capacity (low-span readers) do not use relative plausibility to rank alternative syntactic analyses, whereas readers who have greater verbal working memory resources (high-span readers) would use this and other probabilistic information to rank alternatives (Pearlmutter & MacDonald, 1995). Low-span comprehenders may consider only one analysis of an ambiguous string, or they may adopt heuristics that minimize demands on working memory resources. A 'dedicated resource' account differs from the shared resource account (Caplan & Waters, 1999; Waters & Caplan, 1996a, 1996b, 1997; Waters, Caplan, & Rochon, 1995). On the dedicated resource account, differences in reading behavior emerge because of language interpretation processes operating on parsed input that has been assigned a standard sense-semantic interpretation, not because of differences in the parsing process itself. The dedicated resource account is neutral regarding the existence of differences in the comprehension of sentences

with animate versus inanimate sentential subjects, such as those tested in Experiment 1, but high- and low-span readers should not differ in the parsing strategies they use to interpret the sentences. This account predicts that verbal working memory capacity (high vs. low span) will not interact with sentence type and animacy. Hence, Experiment 2 may produce evidence that bears on the shared versus dedicated resource question.

Method

Participants

Forty-four University of South Carolina undergraduates participated in the experiment for cash or course credit. Half of the participants were high-span readers and half were low-span readers. All were native English speakers with normal uncorrected vision.

Working memory test

To select the 44 participants for the experiment, a large number of participants were screened using a variant of the Daneman and Carpenter (1980) sentencespan test (La Pointe & Engle, 1990). Participants were presented with groups of sentences to read aloud. Each sentence was followed by an unrelated target word that the participants were to remember. After reading all the sentences in a group, participants were asked to recall all of the target words in order. The number of sentences and the number of target words in a group increased from two to six as the participant proceeded through the experiment. Participants saw three groups of the same size before they saw the next larger group. They were initially given three groups of two sentences as practice.

Span was calculated as the largest sentence-group size for which the participant could successfully recall all the target words for at least two of the three groups of that size (following Pearlmutter & MacDonald, 1995). If a participant remembered all of the words of one trial at the next higher group of sentences perfectly, but could recall no more than one trial perfectly at that next group, then half credit (0.5 points) was given (cf. La Pointe & Engle, 1990; MacDonald, Just, & Carpenter, 1992). The possible range of scores was thus 0–6. Participants whose span scores were 3.5 or higher were classified as high-span readers; participants whose span scores were 2 or lower were classified as low-span readers. The mean span was 4.0 ($SD = .64$) for high-span readers and 1.8 ($SD = .58$) for low-span readers.

This version of the sentence-span task was used, rather than the original version developed by Daneman and Carpenter (1980) and used by Just and Carpenter (1992) because it has been extensively investigated by Engle and his colleagues and, in our opinion, its psychometric properties are better understood. The convergent validity among variants of the working memory

span procedure is well-documented (e.g, Engle, Tuholski, Laughlin, & Conway, 1999; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Kail & Hall, 2001), as is the power of the span task to predict performance on a variety of higher level cognitive tasks including reading comprehension (e.g., Engle et al., 1999). Additionally, participants classified by the version of the sentence-span test used here have shown group differences in processing of subject- and object-extracted relative clauses (Traxler, Morris, & Williams, 2003), so there is good reason to believe that it is measuring largely the same construct of memory span that the original Daneman and Carpenter task was designed to measure.

Stimuli

The test and filler sentences were identical to those in Experiment 1. One version of each test sentence was randomly assigned to one of four lists for presentation to the participants. The sentences were assigned such that exactly one version of each item appeared on each

of the four lists and such that no participant saw more than one version of each item.

Apparatus and procedure

The eye-movement apparatus and procedure was very similar to the previous experiment, except that preview was not manipulated so that readers had full preview of all regions at all times. A Fourward Technologies DPI eyetracker was used, and was interfaced with a 486 PC that displayed sentences in the same font used in Experiment 1, using the same software as in that experiment.

Results

Table 3 presents mean scores of the dependent measures for each scoring region for high- and low-span readers. Fig. 2 presents regression path duration by region and condition for high- and low-span readers. The data from each region were submitted to 2 (span: high vs.

Table 3
Mean scores on five dependent measures by condition and region for high- and low-span readers, Experiment 2

	Region											
	High-span readers						Low-span readers					
	1	2	3	4	5	6	1	2	3	4	5	6
First fixation time												
Animate ambiguous	—	—	309	298	315	337	—	—	287	286	311	298
Animate control	—	287	290	280	297	328	—	271	276	272	299	295
Inanimate ambiguous	—	—	305	279	308	339	—	—	295	279	284	320
Inanimate control	—	272	288	294	307	365	—	259	275	270	286	292
First pass time												
Animate ambiguous	—	—	373	535	511	500	—	—	354	505	533	460
Animate control	—	325	338	479	487	497	—	338	325	453	497	434
Inanimate ambiguous	—	—	351	492	479	520	—	—	371	473	469	470
Inanimate control	—	376	311	461	502	501	—	400	327	450	471	454
Second pass time												
Animate ambiguous	87	—	152	200	168	50	72	—	135	116	106	39
Animate control	111	136	125	164	147	81	74	80	65	98	82	35
Inanimate ambiguous	72	—	72	62	87	28	81	—	85	79	69	31
Inanimate control	74	100	60	84	103	34	65	80	41	48	71	35
Regression path duration												
Animate ambiguous	—	—	401	601	595	685	—	—	413	609	614	640
Animate control	—	349	367	536	563	690	—	369	383	529	549	591
Inanimate ambiguous	—	—	378	558	538	639	—	—	446	549	516	594
Inanimate control	—	384	340	517	544	667	—	415	372	476	534	582
Regressions out												
Animate ambiguous	—	—	4.2	11.4	14.7	20.1	—	—	9.1	14.5	9.8	21.6
Animate control	—	2.7	6.1	9.0	11.3	20.8	—	4.2	9.9	9.6	6.4	20.5
Inanimate ambiguous	—	—	5.3	6.9	9.4	15.3	—	—	11.8	8.9	6.4	16.7
Inanimate control	—	6.1	5.2	6.6	5.7	19.5	—	8.0	6.9	4.2	7.6	14.8

Example sentence: The defendant (evidence)/(who (that) was)/examined/by the lawyer/turned out to be/unreliable ("/" marks indicate region boundaries).
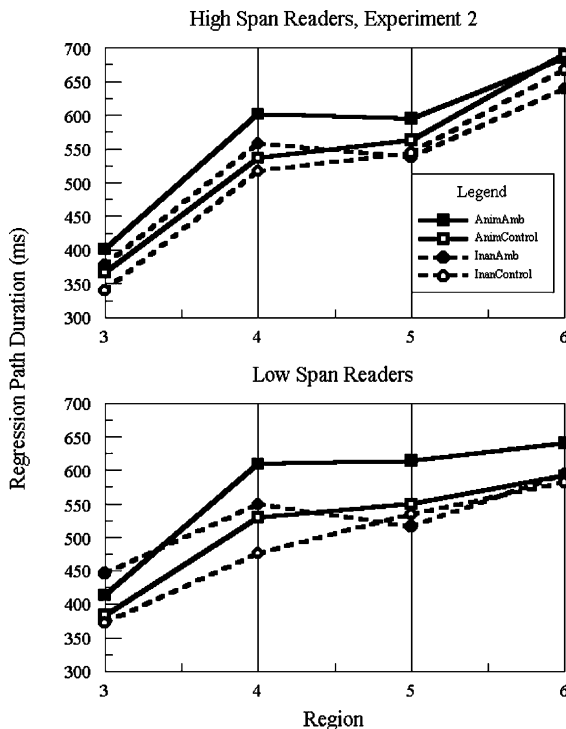
Fig. 2. Regression path durations, Experiment 2.

low span) $\times$ 2 (animacy: sentential subject is animate vs. inanimate) $\times$ 2 (ambiguity: temporarily ambiguous vs. control) ANOVAs. All factors were treated as within participants and items, except for span, which was treated as a between-participants factor. All effects that were significant at the .05 level or beyond are reported here.

Region 3: relative clause verb region

All of the dependent measures except frequency of regressions out of Region 3 indicated that ambiguous sentences were more difficult to process than controls in Region 3. (First-fixation time 298 vs. 282 ms; $F1(1,42) = 8.19, p = .01$, MSe = 1469; $F2(1,47) = 8.08$, $p = .01$, MSe = 3517; First-pass time 502 vs. 461 ms; $F1(1,42) = 12.3, p < .001$, MSe = 6043; $F2(1,47) = 24.1, p < .0001$, MSe = 5661; Regression path duration: 410 vs. 365 ms; $F1(1,42) = 10.9, p < .01$, MSe = 7802; $F2(1,47) = 15.8, p < .001$, MSe = 10,799; Second Pass time: 111 vs. 73 ms; $F1(1,42) = 23.4, p < .0001$, MSe = 2750; $F2(1,47) = 10.3, p < .01$, MSe = 11,948.) These effects extend the effect observed in Experiment 1 for first fixation duration to several other measures. However, Experiment 2 did not replicate the Experiment 1 finding of an increased number of regressions from the relative clause verb when a disambiguating *who/that was* phrase had been present (the corresponding effect in Experiment 2 was numerically in the opposite direction and statistically non-significant).

Sentences with inanimate sentential subjects had shorter second pass times than sentences with animate sentential subjects (65 vs. 199 ms; $F1(1,42) = 14.5$, $p < .001$, MSe = 9144; $F2(1,47) = 23.9, p < .0001$, MSe = 13,432). The effect is consistent with the animacy effects in Experiment 1, and may reflect general difficulty assigning an interpretation to sentences with animate sentential subjects modified by a relative clause rather than some aspect of parsing.

Two span-related effects appeared in the relative clause verb region. First-pass regresion data from the past participle region produced a main effect of span, with low-span readers having more regressions out of Region 3 than high-span readers (9.4 vs. 5.1%; $F1(1,42) = 9.78, p < .01$, MSe = 79.1; $F2(1,47) = 20.3$, $p < .0001$, MSe = 92.0). The absence of interactions in the regresion data indicates that the regressions were more or less evenly distributed across conditions. Hence, the best interpretation of this effect is that low-span readers may generally experience greater difficulty processing the sentences. Additionally, span interacted with ambiguity in second-pass reading time ($F1(1,47) = 5.57, p < .05$, MSe = 2750; $F2(1,47) = 4.23, p < .05$, MSe = 9952). This interaction occurred because the difference between the ambiguous and control conditions was greater for low-span readers than high-span readers (57 vs. 19 ms). This difference suggests that low-span readers were having greater difficulty than high-span readers in resolving syntactic ambiguity. Note, however, that the lack of a three-way interaction between span, ambiguity, and animacy (with $F1 = 0.07$ and $F2 = 0.13$) shows that the greater difficulty in processing the relative clause verb region by the low-span readers was not moderated by the animacy of the sentential subject.

No other main effects or interactions were significant by both participants and items, although there was a trend toward an interaction of animacy and ambiguity in frequency of regressions out of the relative clause verb region, with $F1(1,42) = 3.73, p = .06$, MSe = 43.6, and $F2(1,47) = 2.80, p = .10$, MSe = 109. Inanimacy of the initial noun increased the frequency of regressions in ambiguous sentences, but decreased it in control sentences. This effect, if significant, might suggest that preview of the disambiguating *by* phrase combined with inanimacy of the initial noun triggered reanalysis prior to readers fixating directly on the *by* phrase. However, note that no similar tendency toward an interaction was observed in Experiment 1, even in the preview condition.

Region 4: ''by the NP''

First fixation produced a nearly significant interaction of animacy and ambiguity ($F1(1,42) = 4.12$, $p < .05$, MSe = 1027; $F2(1,47) = 3.87, p = .06$, MSe = 2357): The effect of ambiguity was 16 ms for animate and 3 ms in the opposite direction for inanimate initial

nouns. First pass reading time, by contrast, indicated that reading was faster overall for unambiguous than for ambiguous sentences, 461 vs. 502 ms ($F1(1, 42) = 12.3$, $p < .001$, MSe $= 6043$; $F2(1, 47) = 12.3, p < .001$, MSe $= 11,464$). Reading was also faster following inanimate than animate initial nouns, 470 vs. 493 ms ($F1(1, 42) = 4.45, p < .05$, MSe $= 5317$; $F2(1, 47) = 4.98, p < .05$, MSe $= 10,975$). The interaction of animacy and ambiguity was not significant ($F1(1, 42) = 2.93$, $p = .10$, MSe $= 2553$; $F2(1, 47) = 2.07, p = .16$, MSe $= 11,072$). The simple effect of ambiguity was significant for animate initial NPs ($t1(43) = 3.58$, $p < .001$; $T2(1, 47) = 2.75, p < .01$) and nearly significant for inanimate initial NPs ($t1(43) = 2.26, p < .03$; $T2(1, 47) = 1.85, p = .071$).

Regression path duration and percentage of regressions also indicated that reading was disrupted by ambiguity and by animacy of the initial noun, with little or no sign of an interaction between these variables. Regression path time was faster in the inanimate conditions than in the animate conditions, 525 vs. 569 ms ($F1(1, 42) = 8.76, p < .01$, MSe $= 9623$; $F2(1, 47) = 14.5, p < .001$, MSe $= 14,305$), and faster in unambiguous than ambiguous sentences, 515 vs. 579 ms ($F1(1, 42) = 18.8, p < .0001$, MSe $= 9977$; $F2(1, 47) = 26.4, p < .0001$, MSe $= 15,065$). The interaction of ambiguity and animacy was non-significant ($F < 1$). The effect of ambiguity was significant for animate initial NPs ($t1(43) = 3.48, p < .001; t2(47) = 3.31, p < .002$) and nearly significant for inanimate initial NPs ($t1(43) = 1.81, p < .08; t2(47) = 2.42, p < .02$). Animate conditions also evoked more regressions from the *by the NP* region than inanimate conditions, 11.1 vs. 6.7 % ($F1(1, 42) = 7.56, p < .01$, MSe $= 116$; $F2(1, 47) = 10.1$, $p < .01$, MSe $= 180$). More regressions occurred in the ambiguous conditions than the unambiguous conditions, 10.4% vs. 7.3% ($F1(1, 42) = 9.54, p < .01$, MSe $= 44.5$; $F2(1, 47) = 7.42, p < .01$, MSe $= 140$) but there was no interaction between ambiguity and animacy, ($F < 1$).

Effects of span occurred in the items analysis in first fixation time ($F2(1, 47) = 12.9, p < .001$, MSe $= 1698$), first pass time ($F2(1, 47) = 15.0, p < .001$, MSe $= 6184$), regression path duration ($F2(1, 47) = 6.32, p < .05$, MSe $= 8979$), and second pass time ($F2(1, 47) = 23.3, p < .001$, MSe $= 12,670$). In all cases, reading times were longer for high than for low-span participants. However, none of these effects were significant in the participants analysis (all $p > .16$). No other effects involving span approached significance.

Region 5: main verb

Animacy of the initial noun disrupted reading in the main verb region, as indicated by first fixation duration, first pass time, regression frequency, regression path duration, and second pass time (first fixation 296 vs. 305 ms;

$F1(1, 42) = 6.76, p < .01$, MSe $= 1619$; $F2(1, 47) = 5.47, p < .05$, MSe $= 2568$; first pass 480 vs. 507 ms; $F1(1, 42) = 7.34, p < .01$, MSe $= 4323$; $F2(1, 47) = 6.41, p < .01$, MSe $= 10,356$; regressions 7.3 vs. 10.6%; $F1(1, 42) = 7.19, p < .01$, MSe $= 68.3$; $F2(1, 47) = 8.27$, $p < .01$, MSe $= 180$; regression path duration 533 vs. 580 ms; $F1(1, 42) = 14.4, p < .01$, MSe $= 6882$; $F2(1, 47) = 13.8, p < .01$, MSe $= 16,647$; second pass 82 vs. 126 ms; $F1(1, 42) = 10.6, p < .01$, MSe $= 7818$; $F2(1, 47) = 9.92$, $p < .01$, MSe $= 16,045$). The only other significant effect was an interaction of animacy and ambiguity in regression path duration ($F1(1, 42) = 4.01, p = .05$, MSe $= 9983$; $F2(1, 47) = 5.05, p < .05$, MSe $= 21,346$), suggesting that some residual difficulty was experienced in the animate ambiguous condition, which differed from its control (605 vs. 533 ms; $F1(1, 42) = 5.26, p < .05$, MSe $= 9983$; $F2(1, 47) = 7.42, p < .01$, MSe $= 21,346$) while there was no significant difference between the inanimate ambiguous condition and its control (527 vs. 539 ms; $F1$ and $F2 < 1$). No other main effects or interactions involving regression path duration were significant in the main verb region, although the main effect of span was significant by items ($F2(1, 47) = 14.5, p < .001$, MSe $= 9753$) but not by participants ($F1(1, 42) = 2.25$, NS, MSe $= 37,289$).

Discussion

The main effect of ambiguity in the relative clause verb region (Region 3) indicates that readers had greater difficulty processing this region when the sentence was ambiguous, whether the initial noun was animate or inanimate. A similar effect was observed in first fixation time in Experiment 1, and (non-significantly) in first pass time. No such effect was reported by Trueswell et al. (1994). The effect, however, is predicted by both types of theories discussed in the Introduction. A constraint satisfaction theory would predict slowed reading time in the ambiguous condition because of conflict between the competing main clause and relative clause analyses, especially when the relative clause analysis was highly plausible. A serial depth-first theory could predict the effect if syntactic reanalysis of temporarily ambiguous sentences were triggered by preview of the *by* phrase. Alternatively, the effect of ambiguity in the relative clause verb region may reflect a spillover effect. In the unambiguous conditions, readers have just encountered two short, frequent function words. In the other conditions, there may be some residual effect of having read two less frequent, longer content words.

As in Experiment 1, regression path duration in the *by the NP* region showed comparable disruption of processing of ambiguous sentences whether the sentential subject was animate or inanimate (73 vs. 57 ms). This ambiguity effect extended into the next region, Region 5,

only for animate initial NPs. This suggests that readers are less able to revise an initially plausible (initial animate NP) analysis. Note this interaction of animacy and ambiguity in Region 5 was numerically present in Experiment 1, but not significant. It is possible that the increased number of participants in Experiment 2 allowed us to detect this effect, whereas only hints were available in the Experiment 1 data. However, it is not clear why Experiment 2 failed to detect the second pass time interaction in Region 1 (involving animacy and ambiguity, when preview was present) that had been obtained in Experiment 1.

One striking result in Experiment 2 is the complete absence of three-way interactions of span, animacy, and ambiguity. This pattern of results is consistent with those previous studies that have also failed to show that low-span readers are disproportionately affected by syntactic complexity or less sensitive to semantic plausibility (e.g., Caplan & Waters, 1999; Waters & Caplan, 1996a,b, 1997). The entire pattern of results suggests that low-span readers may generally have greater difficulty interpreting sentences, but the data do not support the more specific predictions made by the shared resource account.

Analysis of combined experiments

In an attempt to explore further what factors do affect the comprehension of reduced relative clause sentences, we combined the data from the preview-present conditions of Experiment 1 and from all conditions of Experiment 2. We did this to permit a more powerful test of the interaction of ambiguity and animacy (just as Trueswell et al., 1994, did; p. 301), and to permit an examination of the extent to which the ratings of agent- and theme-appropriateness provided in Trueswell et al. could predict the size of our reading time effects (once again, following Trueswell et al.).

The analyses of first pass time, based on all 68 participants, indicated that a Region 3 verb following a *who/ that was* phrase was read 31 ms. faster than the verb of a temporarily ambiguous sentence $(F1(1,67) = 13.75,$ $MSe = 5005, \ p < .001; \ F2(1,47) = 38.82, \ MSe = 1498,$ $p < .001)$, and 22 ms faster following an animate than an inanimate initial NP $(F1(1,67) = 8.31, \ MSe = 1206,$ $p < .005; F2(1,47) = 4.67, \ MSe = 2035, \ p < .04)$. The interaction of ambiguity and animacy approached significance in the by-participants analysis $(F1(1,67) = 3.87, MSe = 2540, p = .054)$, suggesting that in this potentially-ambiguous region, ambiguity exacted a larger penalty for animate NP sentences than for inanimate NP sentences (44 vs. 20 ms), but the difference was not significant by items $(F2(1,47) = 1.25, MSe = 2061, p > .25)$.

First pass reading times for the disambiguating *by*-phrase region, Region 4, were faster by 51 ms for unambiguous than for ambiguous items $(F1(1,67) = 28.61, \ MSe = 5699, \ p < .001; \ F2(1,47) = 16.86, MSe = 5176, p < .001)$ and by 49 ms for animate than for inanimate items $(F1(1,67) = 13.95, \ MSe = 5677, p < .001; F2(1,47) = 11.63, MSe = 3827, p < .001)$. The 35 ms effect of ambiguity for inanimate NP items was only non-sigificantly smaller than the 65 ms effect for animate NP items $(F1(1,67) = 2.89, \ p < .10; F2(1,47) = 1.51, p > .20)$, and was significantly greater than zero $(p < .01$ by participants, $p < .05$ by items$)$.

Analyses of regression path duration for the combined data were consistent with the data reported for the individual experiments. No effect was fully significant in Region 3, and the effects of both ambiguity and animacy were significant in Region 4 (all $Fs > 16.5$) while the interaction of these factors was non-sigificant (all $ps > .20$).

As Trueswell et al. (1994) did, we correlated our measures of reading time for the disambiguating region of the inanimate ambiguous items with the ratings of theme and agent typicality ratings. Trueswell et al. found that reading times were generally longer as rated theme typicality decreased, both in simple correlations and in hierarchical multiple regressions in which the reading time for the unambiguous version of each inanimate NP sentence was forced in first. We did not obtain such results. Theme typicality ratings correlated a non-significant −.221 with first pass times, and −.24 for regression path durations. Neither contributed significantly to the multiple regression equation.[4] The difference between our results and those of Trueswell et al. (1994) could possibly be attributed to a restriction of range of theme typicality ratings in our materials. However, the standard deviation of our theme ratings was 1.14, only moderately smaller than the standard deviation for Trueswell et al.'s Experiment 1 (1.38) and larger than their Experiment 2 (0.834) (where, we note, the coefficients were generally not significant). We note further, however, that the presence vs. absence of these correlations has no theoretical consequences. Correlations are expected from a constraint satisfaction perspective, where clearer information supporting the final analysis should speed its selection. However, they are also compatible with a serial, depth-first perspective, if clear information about lexical preferences speeds revision of an initial, erroneous analysis.

---

[4] We note that agent typicality rating did contribute a significant $(t = 2.09, p < .05)$ negative coefficient $(−30.45)$ to the prediction of Region 4 regression path duration. However, the simple correlation between agent typicality rating and Region 4 regression path duration was $+0.072$. Examination of the data indicated the operation of a suppressor variable, since agent typicality rating was correlated $+.364$ with the regression path duration of unambiguous inanimate NP items, which was forced in as the first predictor variable.

General discussion

Experiments 1 and 2 both showed disruption of processing during the *by*-phrase portion of the reduced relative clause, whether the grammatical subject was animate or inanimate. This effect emerged in first pass time, in regression path duration, and in frequency of first-pass regressions (the last measure was non-significant in Experiment 1). Experiment 2 also produced an interaction of animacy and ambiguity in regression path duration in the main verb region and Experiment 1 indicated an interaction of these two factors in re-reading times for the first region when preview had been present. These effects demonstrate that animacy affected how quickly readers recovered from a syntactic ambiguity, but they do not support the claim that readers avoid syntactic misanalysis in sentences with inanimate grammatical subjects.

Experiment 1 showed that preview (or the lack thereof) affected reading, but failed to show that it affected parsing decisions. Processing time on the by-phrase decreased when preview of that phrase was available. This effect may have been greater for sentences with inanimate grammatical subjects, but there was no clear evidence that lack of preview was especially disturbing for ambiguous sentences. First pass reading time on the relative clause verb (e.g., *examined*) increased when there was no preview of the following *by the* region. It is possible that this means that readers normally make use of parafoveal information to make early structural decisions about the relative clause verb, but it seems equally likely that they were simply disrupted by the unusual parafoveal letter sequences in the no-preview conditions. Note, however, that full preview of the *by* phrase did not prevent syntactic misanalysis. In both experiments, significant disruption of processing was observed in the prepositional phrase portion of the reduced relative clause even when the *by* phrase had been present throughout each trial.

Experiment 2 provided evidence about verbal working memory and parsing of reduced relative clauses. High span and low-span readers had similar patterns of results, with comparable early effects of ambiguity. Low-span readers may have had greater difficulty early in the sentence, as evidenced by their greater number of regressions from the relative clause verb region. The only other significant span-related effect in Experiment 2 was the interaction of span and ambiguity in the second-pass reading times of the relative clause verb region. The time that low-span readers spent re-reading this region was affected by ambiguity. However, this effect did not interact with animacy of the initial noun; an inanimate initial noun reduced the time spent recovering from a misanalysis equally for high- and low-span readers. In the disambiguating *by*-phrase region and the following main verb region, high-span participants seemed to read

more slowly than low-span participants, but this effect was significant only in the items analysis, and did not involve an interaction with ambiguity.

On balance, these data favor the dedicated resource view over the shared resource account. Critically, there was no interaction involving working memory span, animacy, and ambiguity. If high-span readers adopted different parsing strategies than low-span readers, they should have reacted to disambiguating information differently. In fact, both high- and low-span readers showed evidence of very comparable disruption following syntactic disambiguation. Further, they did not appear to be differentially sensitive to semantic information. High span readers as well as low-span readers showed evidence of disruption while processing the *by* phrase in sentences with inanimate grammatical subjects. This finding makes it hard to argue that high-span readers in the current Experiment 2 overrode syntactic constraints favoring the main clause interpretation when animacy information pointed toward the alternative relative clause analysis.

Although the current findings point to a different conclusion than that reached by Trueswell et al. (1994), they are actually consistent with the bulk of what has been reported in the literature. The current first pass times are very similar to what was reported by Ferreira and Clifton (1986), using admittedly less adequate materials.[5] The current results are generally consistent with the self-paced reading data from McRae et al. (1998), who used all animate grammatical subjects, manipulating whether the grammatical subject was a good thematic agent or a poor thematic agent. McRae et al. presented the word *by* with the relative clause verb in a single presentation region. Their next presentation region included the determiner and noun of the relative clause. The final region corresponded to our main verb region. They found a main effect of ambiguity in their *examined by* region, such that the region was difficult to process when the sentences were ambiguous (similar to our effect of ambiguity on the initial verb, except that the effect was especially large when the initial noun was a good patient). Their noun-phrase region produced only a main effect of ambiguity, with ambiguous sentences taking longer to process than controls, which corresponds to similar findings in the current experiments in the *by the NP* region. Finally, their main verb region produced an interaction of ambiguity and thematic fit, as ambiguous sentences with good agents as their grammatical subjects proved especially difficult. The current

---

[5] In contrast to the present results, Ferreira and Clifton did not observe speeded second pass times for inanimate initial nouns. This may have been due to the inadequate materials they used or to the early division into two presentation lines, especially in the ambiguous conditions, discussed earlier.

Experiment 2 produced a similar pattern in the regression path duration measure.

The current findings are also largely compatible with Just and Carpenter's (1992) results. These researchers reported eye-movement data on a region corresponding to our *by the NP* region. They observed a main effect of ambiguity in first pass reading times on this region, similar to what was found in the present experiments. Although Just and Carpenter interpreted their data as showing that only high-span readers use animacy information to block a syntactic misanalysis, their data actually do not show the interaction of working memory span, animacy, and ambiguity that would be needed to support such an interpretation. Rather, while their data indicate that only high-span readers were facilitated by inanimate initial nouns, this facilitation held true regardless of sentence ambiguity. The data from the current Experiment 2 show that an inanimate initial noun facilitates reading for both high- and low-span readers. We do not know why our low-span readers appeared to use animacy in interpreting sentences while Just and Carpenter's did not, but in our view, there is no evidence in either our data or in Just and Carpenter's data that either high- or low-span readers used animacy to guide their initial parsing decisions.

Finally, we acknowledge that, if we restrict ourselves to examining just the measures reported by Trueswell et al. (1994), our actual findings may not differ greatly from theirs. There are differences, to be sure; for instance, we obtained significant or near-significant ambiguity effects in the first pass time measure for inanimate initial NPs, whereas Trueswell et al. did not. These differences might be traced to our repetition of sentences (each sentence appeared twice, albeit with different initial NPs and different endings). We do not believe that this is actually the source of any differences. We analyzed the first occurrence of each sentence in the Experiment 1 data, and found that the sizes of the ambiguity effect in regression path duration were 60 and 99 ms for animate and inanimate initial NPs, respectively. These values are closely comparable to the overall effect sizes of 68 and 83 ms. A second reason for the differences comes from a possible difference in the makeup of the filler items between Trueswell et al. and the present study. We explicitly included 24 filler items that had a temporary ambiguity between main clause and relative clause interpretations, resolved in favor of the main clause. Trueswell et al. simply indicated that an unspecified number of their filler items had past tense main clause verbs, without indicating how many or whether these were temporarily ambiguous. We suggest that this may have enabled some of their participants to detect that a temporary ambiguity was relatively likely to be resolved as a relative clause. A third possible reason is that Trueswell et al. followed up each experimental sentence with a second sentence, resulting in two-sentence mini-discourses. This difference may be associated with the differences in results; clearly, further research on the effects of relative clause reduction in full, natural discourses is warranted.

Regardless of these possible bases for the differences between our results and the results of Trueswell et al., we suggest that the most important conclusion to take from the current data is not that our data are different from theirs. Rather, it is that some, but not all, measures of processing difficulty indicated that the presence of an inanimate initial NP did not eliminate or even reduce the immediate disruption caused by a temporary reduced relative clause ambiguity. The claim made by Trueswell et al. (1994) that it did eliminate disruption is thus contradicted by a more exhaustive analysis of the data.

Conclusions

Our findings do not provide a basis for choosing between serial, depth-first parsing models and parallel, constraint satisfaction models. However, they do cast substantial doubt on Trueswell et al.'s (1994) frequently-cited conclusion that semantic factors can override syntactic processing biases. A favorite argument against serial models is that the syntactic preferences they propose can in fact be completely eliminated by semantic and pragmatic factors, and Trueswell et al. is frequently cited in support of this argument. While it is certainly possible that syntactic preferences can be eliminated in some domains, it appears that the reduced relative clause domain studied by Trueswell et al. is not one of them.

We would like to suggest that our findings are most straightforwardly understood in terms of serial, depth-first parsers. On such an account, readers should adopt the main verb analysis initially for sentences with either animate or inanimate grammatical subjects. This initial analysis is ruled out when readers encounter the *by* phrase, and hence syntactic reanalysis is triggered. This account explains the ambiguity effects in the current experiments, along with the absence of animacy by ambiguity interactions in measures of early processing. However, animacy can affect the speed with which reanalysis is completed, and it can affect various measures of the difficulty of interpreting sentences.

Although the present data are easily interpreted in terms of a serial, depth-first model, they cannot rule out a constraint-based processing architecture. A constraint-based model might rank the structural simplicity constraint higher than the semantic plausibility constraint or take into account the relative frequency with which lexical items occur in different constructions (Trueswell, 1996). Hence, the parser might favor the main clause analysis most of the time, with weaker commitment to that analysis when the grammatical subject was inanimate. Encountering the by-phrase would cause the

parser to increase the activation of the previously disfavored reduced relative analysis, leading to competition and increased processing time (cf. Lewis, 2000; Spivey-Knowlton & Sedivy, 1995). This would account for globally elevated processing times in the ambiguous sentences, with especially long processing times in reduced relative clauses modifying animate grammatical subjects.

In our opinion, the real challenge is not to decide whether current serial or parallel parsing theories are correct. Rather, it is to develop better theories and better ways of testing them. Existing serial, "garden-path," theories (e.g., Frazier, 1987, 1990; Frazier & Clifton, 1996) have developed some interesting and provocative claims about how parsers build grammatical structures for sentences, but have not adequately developed claims about how these structures are evaluated and revised (see Fodor & Ferreira, 1998; and Rayner, Carlson, & Frazier, 1993, for some attempts). On the other hand, existing parallel parsing theories have developed primarily as procedures for choosing among existing structures, overlooking the problem of how the structures are built (McRae et al., 1998; Spivey & Tanenhaus, 1998) or accepting patently-inadequate procedures for building structure (cf. Frazier, 1995). There are encouraging signs new directions in the development of parsing models that attend both to the task of creating structure and the task of evaluating the structure that is created (e.g., Gibson, 1998; Jurafsky, 1996; Sturt, Costa, Lombardo, & Frasconi, submitted), but these models have yet to be subjected to stringent experimental tests.

The other challenge is to develop more adequate ways of evaluating models. Simply demonstrating that some semantic or pragmatic manipulation reduces the effect of syntactic ambiguity carries little force. Serial models attribute such effects to reanalysis; parallel models to semantic guidance. Demonstrating the effective disappearance of any effect of syntactic ambiguity (as Trueswell et al., 1994, claimed to do) has often been taken to be more convincing evidence against an initial, obligatory, modular syntactic analysis, but as the present experiments demonstrate, it can be misleading to accept a conclusion of "no effect." More diagnostic experimental approaches are needed. One potentially-useful avenue would be to identify manipulations where structural preferences apparently can be completely overcome. Although the present data indicate that relative clause reduction is not such a manipulation, potential instances do exist. We suggest that they may include verb structure and its influence on the preference for interpretation of prepositional phrases (Spivey-Knowlton & Sedivy, 1995), question focus and its influence on recency preferences (Altmann, van Nice, Garnham, & Henstra, 1998), and argument vs. adjunct structure and their impact on referential discourse effects (Britt, 1994). Simply identifying these manipulations and

giving them a "heavy weight" in quantitative models of sentence interpretation is, however, inadequate. What is needed are explanations of why some manipulations can overcome structural preferences while others (including the manipulation examined in the present paper) cannot. Another potentially-informative approach is to examine the processing of normally-preferred structures, where Frazier (1995) claims strong tests of competing models are possible (and where, we note, Binder, Duffy, & Rayner, 2001, have provided direct evidence against current implemented constraint-satisfaction models of parsing). A third approach would be to look for effects of competition in an ambiguous region of a sentence, apparently predicted by parallel models (cf. Lewis, 2000) but notably difficult to find.

While the present research cannot pretend to solve the major challenges facing parsing theories, we do submit it has value: It reminds psycholinguistic theorists that the influential findings of Trueswell et al. (1994) cannot legitimately be taken as conclusive evidence that semantic factors guide the analysis of syntactic structure. The interesting questions of how we parse and interpret sentences remain unsolved.

Acknowledgments

Appendix A

Note: Alternative initial nouns (animate vs. inanimate) divided by |; alternative sentence endings indicated by {.}

1. The speaker|solution proposed by the group {would work perfectly for the program. turned out to be disastrous.}
2. The man|ransom paid by the parents {was unreasonable. saved their son's life.}
3. The lawyer|package sent by the governor {arrived late. was neglected by the officials in the town hall.}
4. The student|award accepted by the school {was very pleased|pleasing. indicated its good reputation and credibility.}
5. The woman|portrait sketched by the artist {was very beautiful. was very famous.}
6. The defendant|evidence examined by the lawyer {turned out to be unreliable. did not make a good case.}
7. The specialist|equipment requested by the hospital {had finally arrived. was very good for the patients.}
8. The artist|painting studied by the historian {was relatively unknown. was famous only in the mid 1800s.}

9. The man|van recognized by the spy {took off down the street. was on a secret mission.}

10. The man|message recorded by the secretary {could not be understood. was informative and straightforward.}

11. The author|book read by the student {was very difficult to understand. addressed some sensitive issues.}

12. The director|building watched by the cop {was in a bad part of the town. looked strange and suspicious.}

13. The scientists|alternatives considered by the committee {each had limitations. had been proposed by the chairperson.}

14. The student|paper graded by the professor {was very interesting. deserved a good grade.}

15. The mailman|package expected by the secretary {arrived too late. was delayed at the post office.}

16. The man|car towed by the garage {was parked illegally. had been in an accident.}

17. The prisoner|gold transported by the guards {was closely watched. was locked in the van.}

18. The teacher|textbook loved by the class {was very easy to understand. taught them a lot.}

19. The contestant|recipe selected by the judges {did not deserve to win. was the best of the group.}

20. The thief|jewelry identified by the victim {was held for questioning|as evidence. was photographed by the police.}

21. The troops|power plant attacked by the terrorists {suffered heavy losses. was on the army base.}

22. The boy|necklace described by the lady {was quite handsome|beautiful. arrived in time for the party.}

23. The woman|sofa scratched by the cat {was badly injured|damaged. was not noticeably hurt|torn.}

24. The client|account wanted by the advertiser {was worth a lot of money. requires a lot of work.}

References

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

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York: Wiley.

Binder, K. S., Duffy, S. A., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language, 44*, 297–324.

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.

Burgess, C. (1991). The interaction of syntactic, semantic, and visual factors in syntactic ambiguity resolution. Unpublished PhD Dissertation, University of Rochester.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77–126.

Clifton, C., Jr., & Duffy, S. (2001). Sentence comprehension: Roles of linguistic structure. *Annual Review of Psychology, 52*, 167–196.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, shortterm memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163–184.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, shortterm memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309–331.

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

Fodor, J. D., & Ferreira, F. (Eds.). (1998). *Sentence reanalysis*. Dordrecht: Kluwer Academic Publishers.

Frazier, L. (1977). On comprehending sentences: Syntactic parsing strategies. Unpublished doctoral dissertation, University of Connecticut.

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

Frazier, L. (1990). Exploring the architecture of the language system. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 409–433). Cambridge, MA: MIT Press.

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

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

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

Gibson, E. A. F. (1991). A computational theory of human linguistic processing: Memory limitations and processing breakdown. Unpublished doctoral dissertation, Carnegie Mellon University.

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

Inhoff, A. W., Starr, M. S., & Liu, W. (1998). Eye-movement-contingent display changes are not compromised by flicker and phosphor persistence. *Psychonomic Bulletin & Review, 5*, 101–106.

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

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

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers and evidence of arguments. *Psychological Review, 103*, 773–780.

Kail, R., & Hall, L. K. (2001). Distinguishing shortterm memory from working memory. *Memory & Cognition, 29*, 19.

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1997). The role of lexical heads in parsing: Evidence from German. *Language and Cognitive Processes, 12*, 307–348.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1118–1133.

Lewis, R. L. (2000). Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research, 29*, 241–248.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance while reading and watching dynamic scenes* (pp. 55–75). Oxford: Elsevier.

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

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

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

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 375–410). New York: Academic Press.

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

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

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358–374.

Rayner, K., & Clifton, C., Jr (2002). Language comprehension. In D. L. Medin (Ed.), *Stevens' handbook of experimental psychology* (Vol. X, pp. 261–316). New York: Wiley.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of ord frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition, 22*, 281–285.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*, SI21–50.

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

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context and syntactic ambiguity resolution. *Canadian Journal of Psychology, 47*, 276–309.

Sturt, P., Costa, F., Lombardo, V., & Frasconi, P. (submitted). Learning first-pass structural attachment preferences with dynamic grammars and recursive neural networks. *Cognition*.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller, & P. Eimas (Eds.), *Handbook of perception and Cognition: Speech, Language, and Communication* (Vol. 11, (2nd ed., pp. 217–262). San Diego: Academic Press.

Traxler, M. J. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from selfpaced reading. *Quarterly Journal of Experimental Psychology, 55A*, 75–96.

Traxler, M. J., Morris, R. K., & Williams, R. S. (2003). Working memory and semantics in the processing of relative clauses. Poster presented to the CUNY Conference on Human Sentence Processing. Boston, MA.

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

Trueswell, J. C. (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.

Waters, G. S., & Caplan, D. (1996a). The capacity theory of sentence comprehension: Critique of Just and Carpenter. *Psychological Review, 103*, 761–772.

Waters, G. S., & Caplan, D. (1996b). Processing resource capacity and the comprehension of garden path sentences. *Memory and Cognition, 24*, 342–355.

Waters, G. S., & Caplan, D. (1997). Working memory and on-line sentence comprehension in patients with Alzheimer's disease. *Journal of Psycholinguistic Research, 26*, 377–400.

Waters, G. S., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology, 12*, 1–30.