Constituent Attachment and Thematic Role Assignment in Sentence Processing: Influences of Content-Based Expectations

ROMAN TARABAN AND JAMES L. MCCLELLAND

Carnegie Mellon University

The experiments reported here use individual word reading times in a self-paced word-by-word reading task to examine the processing of prepositional phrase constituents in sentences like "The spy saw the cop with a revolver but the cop didn't see him." In Experiment 1 we show that reading times for words immediately following the prepositional phrase ("with a revolver") are predictable from the consistency of subjects' expectations for the attachment of such prepositional phrases with the attachment dictated by the content of the prepositional phrase itself. These expectations vary from sentence to sentence with the content of the material preceding the prepositional phrase and do not appear to reflect the syntactic principle of Minimal Attachment. Experiment 1 conflated violations of expectations for prepositional phrase attachment with violations of role and filler expectations; Experiment 2 examined the contribution of each of these three types of expectation violations to the slowing of reading times. Violations of filler expectations that did not change expected role or attachment produced a small but significant slowdown in processing the words just following the prepositional phrase. Violations of thematic role expectations and filler expectations produced a much larger slowdown, but violation of attachment expectations as well as filler and role expectations did not produce any additional slowing of processing. We interpret these results as supporting models of sentence processing in which thematic role expectations for upcoming constituents play a role in guiding the interpretation of these constituents as they are initially processed. © 1988 Academic Press, Inc.

Understanding a sentence depends on assigning its constituents to their proper thematic roles. Thus in interpreting the phrase "with a stick" in "The boy hit the girl with a stick," we must determine whether the stick is an instrument of the action of hitting, or an object in the possession of the girl. This is typically analyzed as involving two separate decisions: one in which the place of the constituent in a syntactic structural description is determined and another in which the thematic role is assigned. We

refer to the first of these as attachment and to the second as role assignment. In many cases, as here, the two co-vary; a difference in attachment entails a difference in role. There are often several possible roles consistent with the same attachment.

In the two experiments reported in this paper we examine the kinds of information that people use to guide on-line attachment and role assignment decisions. We are particularly interested in examining a central issue concerning how people deal with such decisions. It has been proposed that people use only general syntactic principles to guide an initial syntactic parse of the sentence, in which attachment decisions are made. On this view it is only after constructing such a parse that people consider the possible role assignments of the syntactically organized constituents. Our question is, is this view correct? Or, as others have proposed, do people use the particular content of a sentence as one source of informa-

597

tion that influences initial attachment and role-assignment decisions? We refer to the former view as the syntax-first view and to the latter as content-guided processing.

Syntax-First Models

The suggestion that parsing decisions including attachment are guided by syntactic principles or strategies has been offered many times (Kimball, 1973; Fodor, Bever, & Garrett, 1974; Frazier & Fodor, 1978). Principles are said to be "syntactic" when the conditions for their execution refer to the grammatical categories of the words under consideration, the presence or absence of function words or other grammatical cues, and the configuration of the phrase tree for the sentence at the time the decision is made, but not to the actual lexical/semantic content of the sentence. An example of one of these strategies is the *canonical sentoid strategy* (Fodor et al., 1974) which states roughly that when the surface sequence Noun Phrase–Verb–Noun Phrase is encountered, assume that these items are syntactic Subject, Verb, and Object, respectively. This strategy often works, since many English sentences are consistent with the strategy—e.g., [1a]:

[1] John believed Mary
        (a) and Bill believed Sally.
        (b) stole the cookies.
        (c) to be a saint.

If we were to apply this strategy to [1b], the fact that "John," "believed," and "Mary" follow the canonical order of constituents, and the function word "that" is not present to signal the construction of a that-clause, we would initially analyze "Mary" as the Object Noun Phrase. However, upon reading the verb "stole," we would have to re-analyze the syntactic attachment of "Mary" and assign it the new function of Subject Noun Phrase. A model of human sentence processing based on syntactic principles would generally predict reprocessing in those instances in which later

material reveals that the syntax-based expectations were incorrect.

An illustration of this reasoning, and an important exemplar of the class of syntax-first models, comes from the work of Frazier (Frazier, 1978; Frazier & Fodor, 1978). Frazier has proposed a syntactic processing principle called *Minimal Attachment*. This principle states that initial syntactic decisions will favor the simplest attachment of a phrase into the phrasal representation of the sentence; that is, the parser favors the structure with the minimum number of nodes. In order to make this notion concrete, consider sentences [2] and [3]:

[2] The spy saw the cop with binoculars. (minimal)

[3] The spy saw the cop with a revolver. (nonminimal)

According to Minimal Attachment the initial attachment of the prepositional phrase (PP) in both sentences will be to the verb phrase (VP) node—i.e., it will function as a complement to the verb "saw," as shown in Fig. 1A for [2]. The reason is that such an attachment of the prepositional phrase does not, on Frazier's account, require the construction of new nodes in the syntactic tree that represents the structure of the sentence. Attachment as a constituent of a complex noun phrase (NP) (as in "the cop with a revolver"), on the other hand, does, according to Frazier, require the construction of a new node—the one that represents the complex noun phrase as a whole, under which the simple noun phrase "the cop" and the prepositional phrase are both attached.

Syntax guidance alone is not sufficient to account for all attachment decisions that are made in the course of processing, as [2] and [3] make clear. People interpret "with binoculars" in [2] as attaching to the verb phrase, filling the role of instrument of the verb, while they interpret "with a revolver" in [3] as attaching to the previous noun phrase, specifying a possession of the
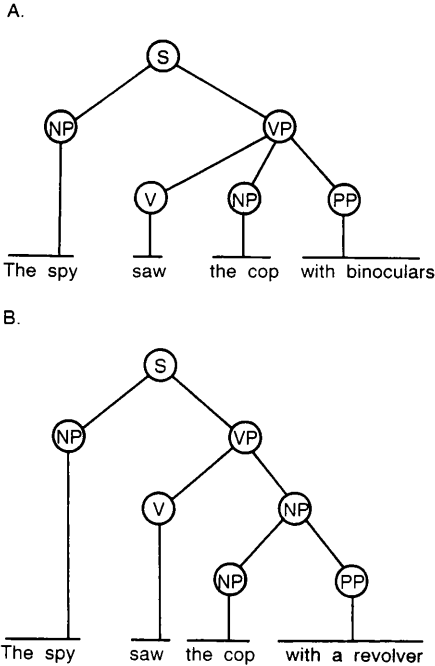
A.

B.

FIG. 1. Phrase trees. (A) Minimally attached prepositional phrase. (B) Nonminimally attached prepositional phrase.

policeman. Such examples illustrate that attachment is not ultimately immune to the influence of content. As already noted, though, the question is whether *initial* attachment decisions are affected by content, or whether content only enters in *after* an initial attachment is proposed by syntax for acceptance or rejection by an examination of content. According to Frazier (Rayner, Carlson, & Frazier, 1983; Frazier, 1986), content considerations in sentences like [2] and [3] can veto initial attachment considerations, but cannot guide them. Thus in [3], the initial attachment proposed by syntax is dispreferred by the "Thematic Processor," and so the prepositional phrase must be reattached nonminimally— as shown in Fig. 1B. Since syntactic reprocessing is required, on this account, it should take longer to process [3] than [2].

Content-Guided Processing

An alternative to the view that initial attachment decisions are governed only by general syntactic principles is the view that the content of the sentence that has been processed up to the point where an attachment decision must be made can contribute to initial attachment decisions. Several versions of this type of model have been proposed. It has often been suggested that the main verb in a sentence might specify an argument structure for the noun phrases and other phrases and complements that it appears with. On some accounts, the information associated with verbs has been quite general, indicating factors like the likelihood that a verb is transitive and will be followed by an Object Noun Phrase (Clifton, Frazier, & Connine, 1984; Mitchell & Holmes, 1985) or that a verb will take a clausal complement (Holmes, 1987; Mitchell & Holmes, 1985). On other accounts, the information that verbs carry is more specific, indicating likely predicate argument expansions for the verb, as in Lexical-Functional Grammar (Kaplan & Bresnan, 1982; Ford, Bresnan, & Kaplan, 1982). On yet other accounts, other constituents in a sentence besides the verb are candidates for influencing and guiding processing. For example, Crain and Steedman (1985) have proposed that a Subject Noun Phrase can influence whether a following verb is seen as the main verb or as a verb in a reduced relative clause. McClelland and Kawamoto (1986) describe a model in which all constituents in a sentence can influence attachment and role assignment decisions about all other constituents. Other models that are essentially content-guided are described by Schank (1972, 1975), Tyler and Marslen-Wilson (1977; Marslen-Wilson & Tyler, 1980), Oden (1978, 1983), Thibadeau, Just, and Carpenter (1982), and Kurtzman (1984). The content-guided view does not deny that syntactic principles may have some influence in processing. It simply claims that the content also plays a role. Thus content is one source of guidance; depending on the specific model there may be other sources as well.

In this paper we discuss two experiments

in which attachment and role assignment are examined in a word-by-word reading task using single sentences. Experiment 1 examines Frazier's claim that initial attachment decisions for prepositional phrases are governed by Minimal Attachment. The results indicate that it is not the ultimate attachment of the prepositional phrase to the Verb Phrase node or to the Object Noun Phrase node that determines reading times. Rather, reading times are predictable from the consistency of subjects' expectations for attachment with the attachment required for the prepositional phrase. These expectations vary from sentence to sentence, based on its content, and so the results demonstrate that the processing difficulty subjects have with sentences like [3] is not a result of a general syntactic principle, but of the violation of content-based expectations.

In Experiment 2 we carefully examine the circumstances under which reading times are slowed by a violation of content-based expectations. These expectations can specify a particular expected filler for the noun that completes the prepositional phrase, a thematic role for the phrase, and an attachment of the phrase consistent with the role. In Experiment 1, all three sorts of expectations tended either to be confirmed or disconfirmed together by the content of the prepositional phrase. It may be that the slowing of reading times reflects a process of reattachment, as we assumed for Experiment 1, following Rayner et al. (1983). But they might be slowed instead by a process of reassigning roles or of replacing the expected filler with an unexpected filler, even when it plays the same role and has the same attachment as the one subjects were expecting. In Experiment 2, we attempt to disentangle these issues.

EXPERIMENT 1

The striking claim of Minimal Attachment is that a single syntactic principle based on the complexity of alternative syntactic representations governs a wide range

of initial parsing decisions and determines which syntactic structures will be easy to process and which will require reprocessing and therefore extra time. These include, among others, the prepositional phrase attachment discussed above, analysis of reduced relative clauses (Rayner et al., 1983; Ferreira & Clifton, 1986), and sentential complements (Frazier & Rayner, 1982).

This experiment examines the evidence for the claim made for the minimal attachment of prepositional phrases in sentences like [2] and [3]. Prepositional phrase attachment represents only a small part of syntactic and thematic processing, and we stress that our evidence here cannot refute the claim that the Minimal Attachment principle might apply in other cases. The processing of sentences like [2] and [3] is an attractive place to begin our inquiry, though, since it was our preexperimental impression that the specific content of particular sentences appeared to govern whether a minimal or nonminimal attachment was expected in this case. Others, too, have conceded that prepositional phrase attachment is a likely candidate for guidance by content (Marcus, 1980).

The major source of evidence for minimal attachment of prepositional phrases is in a study by Rayner et al. (1983). They showed that the prepositional phrase noun-filler[1] (e.g., "revolver" in "with a revolver"), which disambiguated prepositional phrase attachment, and the words that immediately followed it, were read more quickly in sentences like [2], for which the correct attachment is the minimal attachment, than in sentences like [3], where the correct attachment is nonminimal. However, based on a review of the Rayner et al. materials, we suspected that the sentence content preceding the noun-

---

[1] This "noun-filler" is more properly referred to as part of the Object of the preposition heading the prepositional phrase, or as part of the Noun Phrase of the Prepositional Phrase. Using "noun-filler" allows us to refer clearly to the word of interest, and we will use it consistently throughout this paper.

filler in their materials produced expectations for verb phrase attachment, and that it was these expectations, rather than minimal attachment per se, that favored sentences like [2] over sentences like [3].[2] If this were true, we reasoned it would be possible to construct another set of sentences in which the content prior to the disambiguating noun-filler produced expectations for attachment to the Object Noun Phrase (nonminimal) and facilitated reading for nonminimal sentences. Thus, in Experiment 1 we used the Rayner et al. materials, which consisted of sentence pairs like [2] and [3], and we constructed an additional set of sentence pairs like [4] and [5]:

[4] The reporter exposed corruption in the article. (minimal)

[5] The reporter exposed corruption in the government. (nonminimal)

As we shall show, the content of the sentences leading up to the noun-filler in the Rayner et al. stimuli does indeed lead subjects to expect a prepositional phrase that attaches to the verb phrase (minimal) more often than one that attaches to the Object Noun Phrase (nonminimal), whereas the content in our stimuli produces the converse expectation, which is for nonminimal rather than minimal attachment. We then used these two sentence sets in an on-line reading task with the goal of determining whether the reading times in Rayner et al. were due to the Minimal Attachment principle or whether they could be accounted for with independent measures of expectations that depended on the specific content of the sentence. If expectations were used to guide processing, then we could replicate the Rayner et al. effects using their stimuli but not ours. If, however, the processing difficulty reported by Rayner et al. really reflects the operation of Minimal Attach-

ment per se, then we should expect to replicate their results with both sets of materials.

The method and measures that we have used here are not identical to those used by Rayner et al. (1983). They monitored eye movements in reading the sentences and measured fixation duration on the critical word and the next few words following. We have used a word-by-word reading task, in which subjects press a key to obtain each successive word, and we have measured the time spent on the critical word and the next few words following. While this method in which word reading times as measured by inter-word key press intervals may reflect slightly different processes than fixation dwell times and may introduce some distortions into the reading process, we reasoned that, if we were able to replicate Rayner et al.'s effects with their materials, such a replication would favor the view that our method does indeed reflect the factors that produced the reading time results in Rayner et al.'s experiment. An examination of performance on our new materials would indicate whether the operative factors were general syntactic ones or factors that depended on the particular content of the sentences.

Method

For the sake of simplicity, in our first study—Experiment 1A—we used sentence pairs like those in [2]–[5] that did not continue beyond the noun-filler for the prepositional phrase—the word that determines the ultimate prepositional phrase attachment decision. We then replicated our results in Experiment 1B using longer versions of the sentences, in which the noun-filler appeared within the sentence, followed by a word indicating the beginning of a new clause. The longer versions were used to confirm that the effects from 1A were not simply end-of-sentence effects and that they would hold up even when the ambiguous attachment occurred within a sentence. All the critical sentences that

---

[2] Kurtzman (1984) also suggests a content-based account of the Rayner et al. (1983) results. Oden (1978, 1983) considers the effects of differences in the plausibility of attachments like those in [2] and [3].

were used in these experiments are shown in Appendix I.

Subjects. The 28 subjects in Experiment 1A and 42 subjects in Experiment 1B were undergraduates at Carnegie Mellon University who participated for course credit or for $3. Twenty-four students from the same population rated the stimuli, and an additional 24 students at a local community college pretested the stimuli using a cloze task. All the participants were native speakers of English.

Stimuli. Thirty-two sentence pairs were used in this study. Each pair has one sentence with the prepositional phrase attached to the Verb Phrase node in a phrase tree representation,[3] and a matched sentence with the prepositional phrase attached to the Object Noun Phrase node. The former attachment will be referred to in the remainder of this paper as a *Verb Phrase Attachment* and corresponds to minimal attachment of the prepositional phrase; the latter will be referred to as a *Noun Phrase Attachment* and corresponds to nonminimal attachment. The two sentences in each pair are identical, except for the noun-filler. One set of 12 pairs included all the sentences that Rayner et al. (1983) used to test the minimal attachment of prepositional phrases.[4] These sentences were used with an interest in replicating the Rayner et al. finding. We constructed an

[3] Not everyone would agree that all the Verb Phrase Attachments are attached to the Verb Phrase node in the phrase markers for these sentences. According to a criterion that Marcus (1980) uses, "a prepositional phrase attaches to a verb phrase node only if it fills an 'intrinsic case' slot of the verb that is the head of the verb phrase; if it serves as a general modifier of the clause, then it attaches to the S node" (p. 84). Jackendoff (1972), on the other hand, tends to attach prepositional phrases to the S node only if they serve a "speaker-oriented" role (p. 95), as "in my opinion" would. The issue is not strictly relevant to minimal attachment, however, since neither S nor Verb Phrase node attachment requires the formation of a new node in the parse tree, and thus both are equally minimal. Nevertheless, we consider whether our results differ for more and less intrinsic prepositional phrases below.

additional set of 20 pairs in order to test the content-guided processing hypothesis. We will refer to the Rayner et al. (1983) stimuli as the Rayner, Carlson, and Frazier sentences and to our sentences as the Taraban and McClelland sentences. We will call that portion of either the Rayner, Carlson, and Frazier sentences or Taraban and McClelland sentences up to the noun-filler as the sentence *frame*. In the original Rayner, Carlson, and Frazier sentences, the noun-filler is followed by an additional phrase or clause, which was dropped for the sake of simplicity in Experiment 1A, as noted above. In Experiment 1B, we used the Rayner, Carlson, and Frazier stimuli verbatim, in which the prepositional phrase occurred within the sentence, and modified the Taraban and McClelland stimuli from Experiment 1A to include a clause that immediately followed the prepositional phrase. The first word of the clause in the Taraban and McClelland stimuli signalled the beginning of the clause, and the other two words were generally high-frequency words with a simple interpretation, like pronouns and auxillary verbs. Example clause continuations are *though she wasn't* . . . and *so that she.* . . . Continuations like these were used to give subjects a clear signal that the prepositional phrase was complete and that what followed was not attached to that prepositional phrase. For all pairs, including the Rayner, Carlson, and Frazier sentences, the same continuation was used for the two members of a sentence pair. The mean word frequency (Francis & Kucera, 1982) and mean word length for the noun-fillers by Source (Rayner, Carlson, & Frazier; Taraban & McClelland) and Attachment (Verb Phrase; Noun Phrase) are shown in Table 1.

[4] Rayner et al. actually had a longer and shorter version of the same basic sentence—"The spy saw the cop with binoculars but the cop didn't see him" and "The spy saw the cop with an extremely powerful pair of binoculars but the cop didn't see him"—for which they reported no significant differences in reading times. We used only the shorter version.

Stimulus pretesting. Subjects' expectations for prepositional phrase attachment for both sets of sentences were pretested in two ways. One group of pretest subjects was tested individually. These subjects came from the same population as those used in the experiment proper. Each of these subjects was shown Rayner, Carlson, and Frazier and Taraban and McClelland sentence frames one at a time and was asked to complete each of these frames out loud with the first word or phrase that came to mind. This part was like a cloze task. The subject was then shown one of the noun-fillers for the sentence frame and was asked to rate it using the scale shown here:

The completion is
5 the exact word or phrase I expected
4 not the exact word or phrase I expected, but close to it
3 easy to make sense of, but not the word or phrase I expected
2 hard to make sense of—I had to work at it
1 nonsensical or incomprehensible

This scale is worded in terms of "expectations" and we will refer to these ratings as measures of "fit-to-expectations," although other factors might influence the ratings, particularly at the lower end of the scale. After rating the noun-filler, the subject was asked to paraphrase the sentence as completed with that word. This allowed us to determine whether subjects interpreted the prepositional phrase as intended—i.e., either minimally or nonminimally, depending on the condition to which

the sentence belonged. One version of each of the 32 test sentence frames was mixed in with 40 distractor sentences and presented in random order to each subject on an IBM XT, under the constraint that each test sentence was separated from the next test sentence by at least one distractor sentence, and that both versions of a sentence frame were rated an equal number of times across subjects.

Paraphrase data collected for the test sentences indicated that three of these were interpreted differently than intended. The three pairs containing these sentences (one from Rayner, Carlson, & Frazier and two from Taraban & McClelland) were excluded from further consideration. This left 11 (out of 12) sentence frames in the Rayner, Carlson, and Frazier set and 18 in the Taraban and McCelland set. For these remaining 29 frames, the paraphrase data indicated that on 99% of the trials subjects interpreted the sentence with the noun-filler we provided as having the attachment that was intended for the experimental condition. The mean ratings for the two sets of stimuli, by condition, are shown in Table 2. These are based on the 99% of sentences for which the subject provided the condition-appropriate interpretation. The Verb Phrase Attachment noun-fillers from Rayner, Carlson, and Frazier were rated significantly higher than the Noun Phrase Attachment noun-fillers (3.65 vs. 2.93; $t(23) = 9.97, p < .001$). The results for Taraban and McClelland sentences were just the opposite, with the Noun Phrase Attachment noun-fillers rated significantly higher than

TABLE 1
EXPERIMENTS 1A AND 1B: MEAN WORD FREQUENCIES AND WORD LENGTHS FOR NOUN-FILLERS

Source	Attachment	Word frequency (tokens/million)	Length (in characters)
Rayner et al.	Verb phrase	71.7	7.3
	Noun phrase	47.3	7.3
Taraban & McClelland	Verb phrase	164.8	7.1
	Noun phrase	166.3	7.2

TABLE 2
MEAN RATINGS FOR NOUN-FILLERS IN EXPERIMENT 1

	Attachment	
Source	Verb phrase	Noun phrase
Rayner, Carlson & Frazier	3.65	2.93
Taraban & McClelland	2.99	3.97

the Verb Phrase Attachment noun-fillers (3.97 vs. 2.99; $t(23) = 14.20, p < .001$). (In this paper, t tests for ratings are all two-tailed.)

Another pretest was conducted in small groups using students at a local community college. These subjects were given a typed list of the Rayner, Carlson, and Frazier and Taraban and McClelland sentence frames, along with additional sentence fragments that were distractors, and were asked to complete each one with the first word or phrase that came to mind. The 29 test sentence frames and 70 distractors were presented to each subject in a different random order, under the constraint that each test item was separated from the next test item by at least one distractor item. After completing this cloze task, subjects provided a written paraphrase of the test sentences, which we used to clarify prepositional phrase attachment ambiguities in their completions. When presented with Rayner, Carlson, and Frazier sentence frames, subjects were inclined to complete the sentences with prepositional phrases that were minimally attached, with 59% completing the frame with a Verb Phrase Attachment prepositional phrase. When presented with Taraban and McClelland sentence frames, subjects showed a preference for nonminimal attachment, with 76% completing the frames with a Noun Phrase Attachment prepositional phrase.[5] Similar data were

available from the first group of pretest subjects. These subjects completed the Rayner, Carlson, and Frazier frames with Verb Phrase Attachments 70% of the time and completed the Taraban and McClelland frames with Noun Phrase Attachments 86% of the time.[6]

The results from both of the pretests demonstrated that by varying the content of the sentence frame, we could manipulate subjects' expectations for attachment of the prepositional phrase. There were clear differences in the expectations evoked by the two sets we pretested. The Rayner, Carlson, and Frazier set evoked more completions of the minimal attachment type (Verb Phrase Attachment) in the cloze task, and the Taraban and McClelland set evoked more attachments of the nonminimal type (Noun Phrase Attachment). When the test phrases conformed to the type of attachment subjects produced on their own, these phrases were rated as more consistent with

---

73% of the completions were consistent with their condition (Verb Phrase Attachment for Rayner, Carlson, & Frazier and Noun Phrase Attachment for Taraban & McClelland) and when they were absent 70% were, by subjects and by items. This difference was nonsignificant over subjects $F(1,22) = 0.81$, ns) and over items $F(1,19) = 1.24$, ns), indicating that these extra words did not produce any reliable differences in subjects' completions.

[6] We did not ask subjects who did the rating pretest to paraphrase the sentence that they generated with their continuation. However, based on our experience with the cloze pretest subjects, we determined that the intended reading rarely diverged from what appeared to us to be the most obvious reading. For example, in *The kids played all the records on the stereo* the prepositional phrase is most obviously the instrument of "played" and not the location of the records. An unexpected reading like the latter was found for only a very small percentage of the paraphrases.

---

[5] An examination of Appendix I shows that 20 of the prepositions in the prepositional phrase are followed by an article or adjective. In order to determine whether these words had an effect on subjects' completions, half of the cloze test sentence frames included these words and half did not, as a between-subjects variable. When these words were present

their expectations than if the phrases used the alternate attachment.

The 29 surviving pairs of critical sentences and 65 filler sentences were used in an on-line reading task described below. Seven of the distractor sentences were similar to the test sentences and evoked an expectation for Verb Phrase Attachment. These distractors were used to assure that each subject was exposed to an equal number of sentences that evoked Verb Phrase and Noun Phrase Attachments. The remaining distractors in Experiment 1A were simple sentences without prepositional phrases; in Experiment 1B, one quarter of the distractor sentences had an additional clause attached, using the conjunctions of the post-phrase clauses in that experiment. A simple comprehension question was constructed for each test and distractor sentence. For about half of the test sentences, the question probed for information related to the condition-appropriate attachment of the prepositional phrase. In the remaining cases, the question queried other information in the sentence.

Procedure. The sentences in this experiment were presented on an IBM XT using a *moving window* procedure (Just, Carpenter, & Woolley, 1982). The initial display for a sentence presented a dash in place of each letter of all words in the sentence and preserved spaces between the words. The subjects pressed a microswitch to bring up each word; when the switch was pressed for the next word, the previous word was replaced with dashes. Subjects read the sentences silently. A question was displayed following a button press after the final word. The question was answered orally. Each subject read 29 test sentences and 65 distractor sentences. Test sentences were selected randomly for each subject, under these conditions: each subject viewed one of the sentences from each sentence pair in the complete stimulus set; half of the sentences from our set and 5 or 6 out of 11 from Rayner, Carlson, and Frazier were Verb Phrase Attachment sentences; and both versions of a sentence frame appeared an equal number of times across subjects. Each test sentence was separated from the next test sentence by at least one distractor sentence. Eight distractor sentences were used for practice. Subjects were instructed to read the sentences as naturally as possible and were informed that the times spent reading each word in the sentence and initiating a response to the question would be recorded. Subjects' responses to the questions were recorded by the experimenter. The error rate for questions was in fact a low 6% for test sentences. If a subject answered a question incorrectly, the reading times for the associated sentence were excluded from the analyses below. This helped assure that our findings would characterize relatively error-free effective sentence processing.

Results for Experiment 1A

The most important finding in this experiment was a highly significant effect for subjects' expectations and a null effect for Minimal Attachment on reading times. When we examine the mean reading time for noun-fillers in the Rayner, Carlson, and Frazier set, we find a 112-ms facilitation for the Verb Phrase Attachment condition (719 ms vs. 831 ms). However, in the Taraban and McClelland set, this facilitative effect is reversed, with Noun Phrase Attachment sentences showing a 94-ms facilitation (Noun Phrase Attachment: 644 ms vs. Verb Phrase Attachment: 738 ms). These results, based on subjects, are illustrated in Fig. 2. Similar effects were found in the item analysis (Rayner, Carlson, & Frazier: Verb Phrase Attachment 727 ms vs. Noun Phrase Attachment 831 ms; Taraban & McClelland: Noun Phrase Attachment 648 ms vs. Verb Phrase Attachment 737 ms). When the reading times for the two conditions are combined across the Rayner, Carlson, & Frazier and Taraban & McClelland sentence sets, the mean reading time for noun-fillers in Verb Phrase Attachment sentences (729 ms) does not appear to be sig-

nificantly different from the mean reading time for noun-fillers in Noun Phrase Attachment sentences (737 ms).

These effects were examined more closely using an ANOVA with Source (Rayner, Carlson, & Frazier; Taraban & McClelland) and prepositional phrase Attachment (Verb Phrase; Noun Phrase). Given the above results, a highly significant interaction between Source and Attachment would be expected. Indeed this interaction was found in analyses over subjects ($F(1,27) = 10.89$, $p < .003$) and over items ($F(1,54) = 8.45$, $p < .005$). There was no effect whatsoever for Attachment, either by subjects ($F(1,27) = 0.21$, ns) or by items (Verb Phrase Attachment: 733 ms vs. Noun Phrase Attachment: 717 ms) ($F(1,54) = 0.25$, ns). Minimal Attachment predicted a significant processing advantage for Verb Phrase Attachment sentences over Noun Phrase Attachment sentences, which simply was not found.

FIG. 2. Experiment 1A. Mean reading times in ms for noun-fillers in the Rayner, Carlson, and Frazier sentences and Taraban and McClelland sentences.

When we plot the reaction time data against the rating data, as shown in Fig. 3, we notice the close similarity of the two. In general, processing times for syntactic attachments seem to grow smaller or larger as subjects' expectations are fulfilled or disappointed. Separate ANOVAs were used for a closer analysis of effects in the Rayner, Carlson, and Frazier and Taraban and McClelland sentences. The F values (df numerator $= 1$) for the main effects in these analyses, and in similar analyses for Experiment 1B, were used for one-tailed t tests. We justify a one-tailed test on the basis of the strong predictions we had for the direction of these effects. The 112-ms advantage in the Rayner, Carlson, and Frazier data for Verb Phrase Attachment ($t(27) = 2.63$, $p < .01$) and the 94-ms advantage in the Taraban and McClelland data for Noun Phrase Attachment ($t(27) = 3.25$, $p < .002$) were both significant over subjects, as were the associated effects over items (Rayner, Carlson, & Frazier: 104 ms; $t(20) = 1.99$, $p < .04$; Taraban & McClelland: 89 ms; $t(34) = 2.18$, $p < .02$). These are important findings, since they suggest that the Rayner et al. (1983) effects may have been due to expectations that were generated on-line by the subjects, based on the particular content of the sentences.

Figure 3 also shows an overall disadvantage for Rayner, Carlson, and Frazier sentences compared to Taraban and McClelland sentences (775 ms vs. 691 ms), which appeared as a significant effect for Source, by subjects ($F(1,27) = 13.33$, $p < .01$) and by items (Rayner, Carlson, & Frazier: 779 ms vs. Taraban & McClelland: 692 ms) ($F(1,54) = 6.81$, $p < .02$). Part of the offset between the two curves might be explained by lower word frequencies for the Rayner, Carlson, and Frazier test words (See Table 2)—lower frequency words generally take longer to process and could have raised the overall processing times for the Rayner, Carlson, and Frazier items. There is other evidence that the Rayner, Carlson, and Frazier sentences were more difficult over-
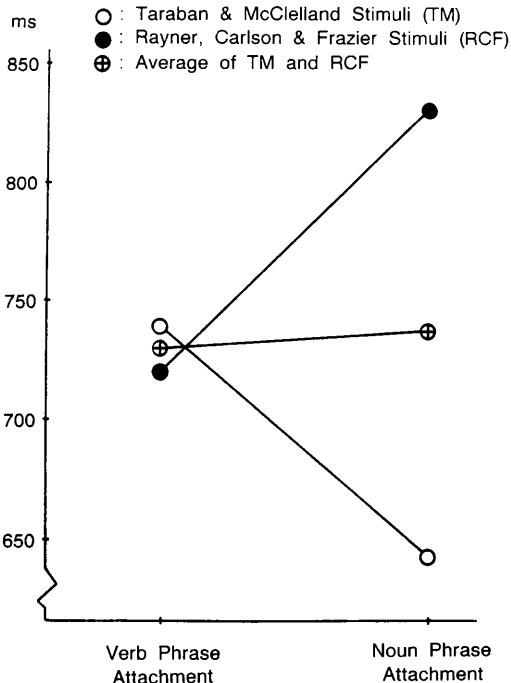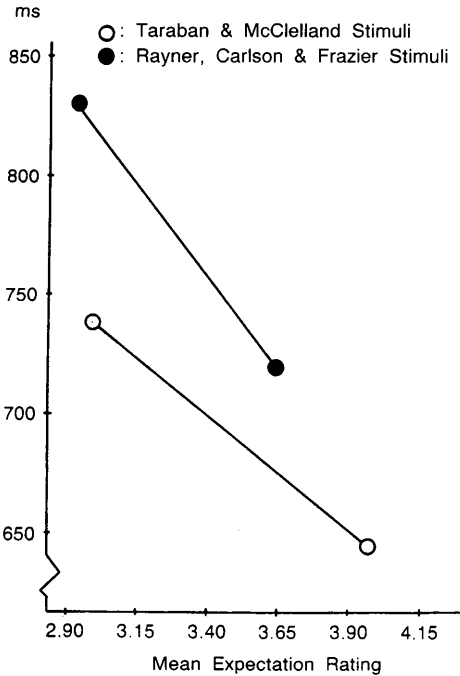
FIG. 3. Experiment 1A. Reading times in ms as a function of expectation ratings for the Rayner, Carlson, and Frazier sentences and the Taraban and McClelland sentences.

all. This is indicated by reading times in the region up to the preposition—for the Taraban and McClelland sentences, the average per character reading time was 70 ms, while for the Rayner, Carlson, and Frazier sentences, it was 78 ms.

Results for Experiment 1B

This experiment was a replication of Experiment 1A, using the longer versions of the test sentences shown in Appendix I. These longer versions allowed us to measure processing times for the noun-filler— the noun completing the prepositional phrase—when that word was not the last word in the sentence. This was done to allow us to separate attachment effects from end-of-sentence wrap-up effects (Just & Carpenter, 1980; Thibadeau et al., 1982; Daneman & Carpenter, 1983), which are evidenced by somewhat longer reading times for our noun-fillers relative to earlier

words in the test sentences in Experiment 1A. The primary data for Experiment 1B consisted of the time spent reading the noun-filler and the three words that followed it. The manner in which this experiment was conducted was otherwise identical to Experiment 1A.

Positioning the noun-filler in the middle of a sentence produced effects that were very consistent with the results in Experiment 1A for the Rayner, Carlson, and Frazier and Taraban and McClelland sentences. In order to examine the overall effects, we summed over the noun-filler and the three words that followed it, and then compared Verb Phrase Attachments to Noun Phrase Attachments using this formula: Verb Phrase Attachments *minus* Noun Phrase Attachments. The total effect, based on subjects, for the Taraban and McClelland sentences was 69 ms, and the total effect for the Rayner, Carlson, and Frazier sentences was −94 ms, the negative sign corresponding to a shift in the direction of expectations for the respective attachments. As shown in Fig. 4, which breaks these effects down word by word, there was no evident effect for disambiguation on the noun-filler itself. On both the first and second word following the noun-filler, there was about a 35-ms facilitation for the expected attachment (Noun Phrase Attachment for Taraban & McClelland and Verb Phrase Attachment for Rayner, Carlson, & Frazier). On the third word following the noun-filler the facilitation decreased, but was still present at a level of about 15 ms.[7]

An overall ANOVA using Source (Rayner, Carlson, & Frazier; Taraban & McClelland), Attachment (Verb Phrase; Noun Phrase), and word Position (the noun-filler and the three words following it) showed a significant effect for Source ($F(1,41)$ =

---

[7] Up to the third word after the noun-filler, all words were on the same line. Subsequent words, however, could occur on the next line on the screen. Thus reading times beyond the third word are confounded with line changes for some of the sentences. For this reason they are not reported.

O : Taraban & McClelland Stimuli (TM)
● : Rayner, Carlson & Frazier Stimuli (RCF)
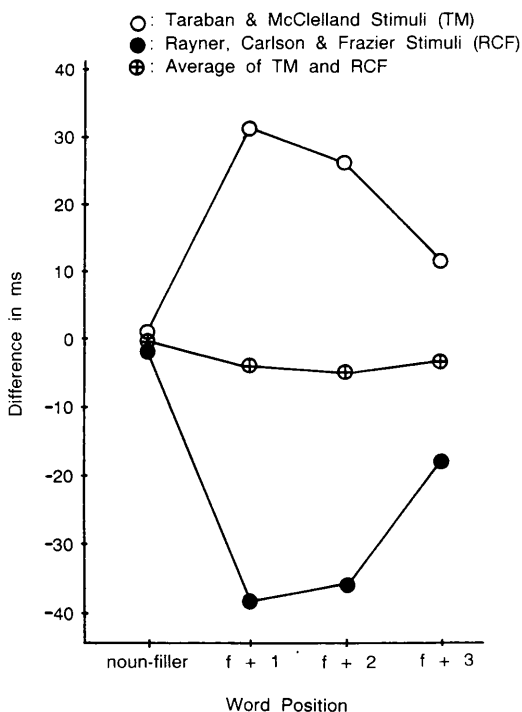⊕ : Average of TM and RCF

FIG. 4. Experiment 1B. The difference in reading times in ms between verb phrase attachment stimuli and noun phrase attachment stimuli at each position shown.

each word position. Word positions, corresponding to Fig. 4, will be referred to as *noun-filler,* and $f + 1$, $f + 2$, and $f + 3$, for the first, second, and third words, respectively, following the noun-filler. In the analysis over subjects, there was a significant effect for the Verb Phrase Attachment *minus* Noun Phrase Attachment difference in the Rayner, Carlson, and Frazier sentences at $f + 1$ (-38 ms: $t(41) = 2.50$, $p < .01$), $f + 2$ (-36 ms: $t(41) = 2.54$, $p < .01$), and $f + 3$ (-18 ms: $t(41) = 2.55$, $p < .01$), and corresponding effects in the Taraban and McClelland sentences: $f + 1$ (31 ms: $t(41) = 3.11$, $p < .002$) $f + 2$ (26 ms: $t(41) = 3.07$, $p < .003$), and $f + 3$ (11 ms: $t(41) = 1.66$, $p = .05$). Although the size of the effects was similar in the analysis over items, the effects were only marginally significant for the Rayner, Carlson, and Frazier sentences at $f + 2$ (-33 ms: $t(20) = 1.65$, $p < .06$). In the Taraban and McClelland sentences, there were significant effects at $f + 1$ (30 ms: $t(34) = 1.99$, $p < .03$), and $f + 2$ (25 ms: $t(34) = 1.69$, $p < .05$).

It is possible to make a distinction between prepositional phrases that fill an "intrinsic" role of the main verb and nonintrinsic prepositional phrases that provide clausal context (Marcus, 1980). Using the phrase tree notation of Fig. 1, "intrinsic" prepositional phrases attach to the VP-node, while "nonintrinsic" phrases attach to the S-node. Since no new nodes need to be created for prepositional phrase attachment to either of these points, both are potential sites for Minimal Attachment. It may be that intrinsic as well as nonintrinsic phrases would initially attach to the VP-node, though, due to the near proximity of that attachment point compared to the higher S-node (cf. Late Closure, Frazier & Fodor, 1978; Frazier & Rayner, 1982). Given these assumptions, one prediction is that nonintrinsic phrases would be more difficult to process than intrinsic phrases, since the nonintrinsic phrases would have to be reattached from their initial point of

5.19, $p < .03$) and Position ($F(3,123) = 9.57$, $p < .001$), as well as a significant Source × Attachment interaction ($F(1,41) = 14.71$, $p < .001$) and a significant Source × Attachment × Position interaction ($F(3,123) = 5.54$, $p < .001$), over subjects. In the analysis over items, there was a significant effect for Position ($F(3,162) = 14.75$, $p < .001$) and a marginal effect for the Source × Attachment interaction ($F(1,54) = 3.48$, $p < .07$) and for the Source × Attachment × Position interaction ($F(3,162) = 2.46$, $p < .07$). In general, these results suggested that reading times varied significantly among word positions and that the Rayner, Carlson, and Frazier and Taraban and McClelland stimuli were behaving differently.

One-way ANOVAs were again used to examine the significance of these effects separately for Rayner, Carlson, and Frazier and Taraban and McClelland sentences at

attachment. Depending on the order in which alternative sites of reattachment are considered, the nonintrinsic phrases could take as long or longer to reattach compared to Noun Phrase Attachments. We therefore wanted to know whether the reading time advantage that our Noun Phrase Attachments showed depended on whether they were compared to intrinsic or nonintrinsic phrases, which could have been mixed together in our Verb Phrase Attachment condition. The present authors therefore made a yes/no judgment on whether the Verb Phrase Attachments in the Taraban and McClelland stimuli were intrinsic; discrepancies in initial judgments were mutually resolved. Eleven of the 18 Taraban and McClelland Verb Phrase Attachment sentences were classified as "intrinsic" and the remaining 7 as "nonintrinsic," as shown in Appendix I. Table 3 shows the noun-filler word frequency and noun-filler word length for the more- and less-intrinsic Verb Phrase Attachment noun-fillers and for the Noun Phrase Attachment noun-fillers in the matched sentences; also shown are reading times for the noun-filler and the three words that followed in each case. An ANOVA using Attachment (Verb Phrase; Noun Phrase), Relation (Intrinsic, Nonintrinsic), and Position (the noun-filler and the three word positions following) showed nonsignificant effects for the interaction terms that would have indicated differences between "Intrinsic" and "Nonintrinsic"

phrases: Attachment × Relation ($F(1,41) = 1.57$, ns) by subjects and ($F(1,32) = 2.04$, ns) by items; and Attachment × Relation × Position ($F(3,123) = 0.57$, ns) by subjects and ($F(3,96) = 0.46$, ns) by items. The cell means in the Intrinsic relation condition demonstrate a large advantage for the Noun Phrase Attachment sentences. These results, therefore, do not suggest that our results depend critically on an intrinsic vs. nonintrinsic distinction for the 'less-expected' Taraban and McClelland Verb Phrase Attachment stimuli.

Discussion

This experiment showed that one set of sentences set subjects' expectations to favor verb phrase attachment of a post-verbal prepositional phrase, before they saw the noun-filler that disambiguated the attachment, while another set of sentences set expectations to favor Object-NP attachment. As is shown in Table 2, the difference between Taraban and McClelland sentences and Rayner, Carlson, and Frazier sentences is that subjects' expectations for the noun-filler that disambiguated the attachment of the prepositional phrase are high for Rayner, Carlson, and Frazier Verb Phrase Attachment noun-fillers and lower for the Noun Phrase Attachment noun-fillers. For Taraban and McClelland sentences, this is reversed. Differences in subjects' expectations for the Taraban and McClelland and Rayner, Carlson, and Frazier

TABLE 3
EXPERIMENT 1: MEAN WORD FREQUENCIES, WORD LENGTHS, AND EXPECTATION RATINGS FOR TARABAN & McCLELLAND NOUN-FILLERS ACCORDING TO THE ROLE OF THE PREPOSITIONAL PHRASE

	Role			
	Intrinsic		Nonintrinsic	
Attachment	Verb phrase	Noun phrase	Verb phrase	Noun phrase
Word frequency (tokens/million)	164	182	166	141
Length (in characters)	7.4	7.4	6.4	6.7
Expectation rating	3.09	3.79	2.83	4.24
Reading time	1506	1418	1374	1341

Note. The reading times in ms are mean total reading times for the noun-filler and the three words that follow it.

sentences were also evident in subject-generated completions for prepositional phrase fragments. The syntactic constituents in the Rayner, Carlson, and Frazier and Taraban and McClelland sets were of the same type, so the shift in expectations cannot be explained in terms of syntactic differences. The sentences up to the disambiguating word did differ in content, though, and it is the specific content that appears to have produced the expectations.

The major finding in this experiment is that these content-based expectations affect what happens in on-line reading. On the assumption that expectations guide processing, a slowdown in processing was predicted in those conditions in which subjects' expectations were inconsistent with the attachment required by the sentence. As shown in Fig. 3 expectation ratings were a good predictor of the time subjects would spend in processing the disambiguating region of the test sentences. Reading times followed these patterns of expectations, with higher conformity to expectations producing faster reading times. This supported the view that content-based expectations were not epiphenomenal with respect to reading, but rather captured critical information about the on-line reading of these sentences. Content does not appear to have the "eventual" effect often attributed to it in "syntax first" models, but rather has an immediate guiding influence on the course that processing follows; when subsequent information is inconsistent with content-based expectations, there is a slowdown in processing.

The disruption of processing occasioned by a mismatch with expectations reported in Experiment 1B appears primarily in the "reading time" associated with the two words after the noun-filler, and not on the noun-filler itself. It would be interesting to compare this pattern of results with that found by Rayner et al. (1983). Unfortunately, in their results they found no significant interaction of attachment with individual fixations before and after the noun-filler. They do report that the average duration of the three fixations before the noun-filler was not affected by attachment, while the average duration of the three fixations on or after the noun-filler was affected by attachment. It would appear that their effects are spread over the processing of the first few words in the disambiguating region; whether any effect was present on the first fixation on the noun-filler is simply not clear.

In assessing our own findings, it must be noted that it is very difficult to determine just where in the course of processing the observed effects are produced. Such questions are very difficult to answer, particularly if processing occurs concurrently at many levels, as we are inclined to believe. It is worth noting too that effects that show up while the word following the noun-filler is on the screen may well reflect processes that are contemporaneous with the processing of the noun-filler, at least in part; subjects may still be processing the noun-filler even after they have pressed a button calling the word that follows to the screen. It might also be noted that subjects may in some cases be holding back on reattachment until the word after the noun-filler makes it clear that the noun-filler really was the last word in the prepositional phrase.

What implications do these findings have for Minimal Attachment? Our results undermine the Rayner et al. evidence for Minimal Attachment in sentences like those tested here and suggest that the evidence they found for Minimal Attachment was due to the confounding of expectations with attachment in their stimuli. The data here do not entirely refute the theory, though. First, the violation of content-guided expectations appears to be the dominant source of reading time effects in these experiments, and Minimal Attachment may play a subtle role that is masked by these larger effects. Second, the prediction derived from Minimal Attachment for sentences like [2] and [3] depends on the ancillary assumption that no complex NP node

(see Fig. 1B) is formed between the verb phrase node and the node for the simple NP when a simple NP is processed during on-line processing. If the complex NP node were formed automatically in the course of processing the Object-NP ("the cop" in [3]), then there would be no difference in degree of minimality between Noun Phrase Attachments and Verb Phrase Attachments of the prepositional phrase, since no new nodes would need to be attached in either case. It could be, then, that Minimal Attachment is correct as a general principle, but simply does not apply in this case. Thus our findings leave open the possibility that content governs only in cases (such as prepositional phrase attachment) where syntax does not (cf. Marcus, 1980). However, there are findings which indicate a role for content in other cases as well (cf. Holmes, 1987), further weakening the appeal of the syntactic-guided processing view.

While our results do not strictly rule out the Minimal Attachment principle, they suggest that an important factor in on-line processing of sentences is the degree to which the ultimate attachment of constituents actually matches with subjects' expectations for these constituents. Predictions based on content-sensitive expectations do fit the interaction between the attachment required by a sentence and subjects' expectations for attachment, summarized in Table 4. Thus, all of our evidence is consistent with the view that subjects experience difficulty when their expectations are violated, and that these expectations are governed, at least in part, by content. Indeed it should be noted that the results we have reported cannot be accounted for by *any* syntactic principle of which we are aware— that is, by any principle that does not consider the content of the sentence—since our expectation effects occurred in sentences that differed in the content, and not in the syntactic constituents of the sentence frames. This does not mean, of course, that syntactic principles do not play some role in directing on-line processing decisions. For the moment, however, we will continue to focus on the role of content.

EXPERIMENT 2

Violation of expectations for the noun-filler appears to cause processing difficulty, as we found in Experiment 1. But several aspects of subjects' expectations were violated. In Experiment 1, we followed Rayner et al. (1983) in the view that the violation of the expected attachment was the source of processing difficulty for unexpected prepositional phrases in that experiment. This is not the only possibility, though, since violation of the expected attachment covaried with violation of the expected role of the prepositional phrase. Perhaps it was the violation of the expected thematic role of the prepositional phrase that determined processing difficulty, rather than the violation of the expected attachment per se.

There is yet another possibility. Perhaps it was neither the violation of the expected attachment of the prepositional phrase nor

TABLE 4
READING TIMES FOR NOUN-FILLER IN MS

| Source | Experiment | Attachment | | |
		Verb phrase	Noun phrase	Difference
Rayner, Carlson & Frazier	1A	719	831	−112
	1B	1420	1514	−94
Taraban & McClelland	1A	738	644	94
	1B	1455	1386	69

Note. Experiment 1B shows the mean total reading time for the noun-filler and the three words that follow it.

of the expected thematic role of the prepositional phrase that produced the effects we found, but simply a violation of the subject's expectations for a particular noun-filler. Subjects may have had a select pool of candidates, or perhaps a single candidate, in mind for the noun-filler, given the prior content of the sentence. For a Verb Phrase Attachment, for example, and a particular role, like Instrument, we can find better and worse instances of appropriate instruments in the context of a particular sentence. Processing difficulties may have been due to the consistency of the particular noun-filler in the sentence with the particular candidates they expected. The data from Experiment 1 do not help to separate out these various sources of influence. In fact, the two conditions for the Rayner, Carlson, and Frazier and Taraban and McClelland stimuli confound all three sources. For example, in [2] and [3] (repeated from above), while attachment changes from Verb Phrase Attachment to Noun Phrase Attachment, thematic roles go from Instrument of "see" to Modifier of "cop":

[2] The spy saw the cop with binoculars.
[3] The spy saw the cop with a revolver.

These changes in attachment and role are perfectly correlated with changes in the noun-fillers. Therefore, we do not know what part of the measured effect for the stimuli is actually due to attachment switching, relative to an expected attachment, what part is due to thematic role switching, relative to an expected role, and what part is due to the word used for the noun-filler relative to an expectation for a particular filler. Processing may in fact be subject to all three influences, but we do now know their relative effects.

We need to consider, then, how to measure the effects of violations of expectations for each of these sources of influence without confounding them with the others. One way is to identify a set of sentence frames such that each one evokes a consistent expectation for an attachment and role for the prepositional phrase. We can then find four different noun-fillers for each frame, creating four different prepositional phrases. These four prepositional phrases associated with the same frame differ according to the manner in which they violate subjects' expectations. We will represent these conditions as follows:

a. Fully Consistent
b. Less Expected Filler
c. Less Expected Role
d. Less Expected Attachment.

Phrases in the first condition are consistent with subjects' expectations for the attachment and role of the prepositional phrase, and the particular word used for the noun-filler is actually quite good for that role and attachment—the way in which we assess these various consistencies is discussed below. Relative to this "fully expected" condition, we sequentially violate expectations for filler; filler and role; and filler, role, and attachment, as follows. The second condition uses a less expected noun-filler, but one that is consistent with the expected attachment and role for the phrase. Each of the remaining conditions include the expectation violations of the previous conditions and add a new one: the thematic role of the prepositional phrase in condition (c) is a less expected role, in addition to the noun-filler itself being less expected; the attachment of the prepositional phrase in condition (d) is not the attachment subjects are generally expecting, in addition to the role and filler in the phrase being a less expected role and filler. By examining reading times for contiguous pairs of conditions, we can determine the effects of adding one type of expectation violation at a time, making it possible to move some distance toward determining which kinds of violations of expectations produce the strongest effects.

In order to better understand the design, it might be helpful to examine some sample stimuli, where a–d represent four experi-

mental conditions that use the same initial sentence frame and post-phrase continuation:

[6] The janitor cleaned the storage area with the
    (a) broom
    (b) solvent
    (c) manager
    (d) odor
because of many complaints.

[7] The hospital admitted the patient with
    (a) cancer
    (b) amnesia
    (c) bodyguards
    (d) apologies
while the other incoming patients calmly waited.

Noun-fillers (a) and (b) are consistent with the expected syntactic attachment and thematic role for the sentence, based on a cloze test like the one used in Experiment 1. Half of the sentence frames created an expectation for a Verb Phrase Attachment like the frame for [6], and the other half of the frames created an expectation for a Noun Phrase Attachment, like the frame for [7]. We will refer to the former as *Verb Phrase Bias* frames and to the latter as *Noun Phrase Bias* frames. Conditions (a) and (b) differ in the plausibility and expectedness of the noun-filler (explained below), with the (a) filler being highly plausible and expected and the (b) filler being less plausible and expected. The (c) noun-filler retains the same expected attachment as (b), but changes from the expected thematic role of (b) to a less expected role. Finally, the thematic role of the (d) noun-filler is unexpected, as it is for the (c) noun-filler, but the attachment of the (d) filler is additionally unexpected, relative to the (c) filler.

The preceding paragraph alludes to two potential sources of reading time effects related to different aspects of the congruity of the noun-filler with the sentence frame. One of these is the extent to which the noun-filler violates the expectations formed on the basis of the frame. The other is the plausibility of the event described by the sentence—the frame and noun-filler together—once the noun-filler has been encountered. An example should help to make this distinction clear. The sentence frame "The doctor examined the patient with a . . ." evokes expectations for a Verb Phrase Attachment and an Instrument role for the prepositional phrase. Although an instrument like a "magnifying glass" could be attached according to this expectation, it would be an unexpected and fairly implausible instrument. On the other hand, the attachment of the Object-NP of a prepositional phrase describing a disease, like "cancer," though it might be unexpected, would lead the sentence as a whole to seem quite plausible once its role and attachment had been established. The rating scale in Experiment 1 probably captured both expectedness and plausibility to some extent. Since we felt it was important to control both factors as much as possible, we made an effort to assess them separately, as described in the Method section. We have been concerned only with the plausibility of the frame and noun-filler, since the three words immediately following the noun-filler, for which reading times were examined, tended to be short function words that initiated a new clause and were deemed not to add specific content beyond indicating that a new clause was beginning.

A key feature of our design was its attempt to examine the effects of role-expectation violations and of attachment-expectation violations separately from any general effect of degree of expectedness or plausibility as defined above. Therefore, we wanted the plausibility and expectedness of the noun-filler to be as closely matched as possible in conditions (b), (c), and (d). In this way, the comparison of conditions (b) and (c) would be a relatively pure

indication of the processing cost of role switching, and the comparison of (c) and (d) would be a relatively pure indication of the processing cost of attachment switching, over and above the cost of role switching. A comparison of conditions (a) and (b), on the other hand, which differ in terms of both the plausibility and the expectedness of the noun-filler, would indicate the effects of manipulating these factors for the noun-filler itself, while holding everything else constant. Since our primary interest was not in separating effects of expectedness and plausibility, but only in controlling both factors, we have not attempted to decouple these factors explicitly in our materials, even though we have attempted to assess them both.

Method

Subjects. The subjects in this experiment were Carnegie Mellon University undergraduates and participated for course credit. These subjects were used for the same sequence of pretesting, rating, and on-line reading of the sentences that was used in Experiment 1. Sixty subjects did a cloze task for sentence frames, 36 subjects rated versions of the completed sentences using our noun-fillers, and another 44 subjects read these sentences for understanding in an on-line reading task. All the participants were native speakers of English.

Stimuli. Twenty-four sentence frames like [6] and [7] were used in this study. Each sentence frame was used to create 4 sentences that were identical except for the noun-filler, which disambiguated the attachment and thematic role of the prepositional phrase. Similar to Experiment 1, all four noun-fillers in a sentence set were followed by the same clause, which was clearly signaled by a conjunction. The two words following the conjunction were usually function words, auxiliaries, or other words without much potency, in themselves, to shift plausibility one way or another, which is important to mention since our reading time data include these words.

In order to construct these stimuli, we had to identify an acceptable set of thematic roles. The majority of thematic roles that we used came from Fillmore (1968, 1971). In those instances in which we had to identify a new role, we used a test found in Fillmore (1971), which is described in Appendix II. Construction of our stimuli required a broader use of thematic roles than is usually found in the literature, although it is consistent with Fillmore's view of the range of possible roles (e.g., Fillmore (in preparation)). Overall, the stimuli used 11 different thematic roles (Instrument, Goal, Source, Location, Manner, Time, Accompaniment, Attribute, Scope, Purpose, Content) and 5 different prepositions (with, in, for, on, and from).

The complete list of test sentences is provided in Appendix III. The mean word frequency (Francis & Kucera, 1982) and mean word length for the noun-fillers in the four conditions are shown in Table 5. In addition to the 96 test sentences that resulted from the 24 sentence frames, there were 66 distractor sentences that were either simple or compound and that did not contain prepositional phrases. A simple comprehension question was constructed for each test sentence and distractor. For about half of the test sentences, the question probed for information related to the condition-appropriate attachment and role. In the remaining cases, the question queried other information in the sentence.

Stimulus pretesting. Sentence frames were pretested using a cloze task, as in Experiment 1. There were 52 test sentence frames and 84 distractor sentence frames. A subject completed 26 test sentence frames and all the distractor frames. The test sentence frames were selected randomly for each subject, under the constraint that all the frames were completed an equal number of times across subjects. For each subject, the test frames were mixed randomly with the distractors, with at least 1 distractor appearing between one test sentence and the next. Subjects pro-

TABLE 5
EXPERIMENT 2: MEAN WORD FREQUENCIES AND
WORD LENGTHS FOR NOUN-FILLERS IN THE
4 CONDITIONS

Frame Type	Word frequency (tokens/ million)	Length (in characters)
Verb phrase bias		
Fully Consistent	68	6.8
Less Expected Filler	32	7.0
Less Expected Role	31	7.3
Less Expected Attachment	74	6.1
Noun phrase bias		
Fully Consistent	81	6.3
Less Expected Filler	129	6.3
Less Expected Role	84	5.8
Less Expected Attachment	62	6.6

vided a written paraphrase for the test sentences, as in Experiment 1. Twenty-four sentence frames were selected for the online reading task. Twelve of these evoked an expectation for a Verb Phrase Attachment and the other 12 evoked an expectation for a Noun Phrase Attachment. Specifically, for these 24 sentence frames, there was 95% consistency between cloze test results and the attachment that they were categorized as evoking; the consistency of thematic roles, given attachment, was also 95%.[8] (Individual percentages are listed

[8] As in Experiment 1, some of the prepositions in the critical prepositional phrase were followed by the articles "a" or "the," and in one case, the pronoun "his." In order to determine whether there was any additional influence on subjects' completions from these words, half of cloze trials appeared with the additional word and half appeared with only the preposition, as a between-subjects manipulation. There was no evident effect for the additional word on the likelihood that subjects would attach the prepositional phrase as a Noun Phrase Attachment or a Verb Phrase Attachment, by subjects ($F(1,14) = 0.38$, ns), or by items ($F(1,11) = 1.00$, ns).

with each sentence in Appendix III.) This assured that expectations for attachment and role just prior to the noun-filler would be highly consistent across subjects.

Noun-fillers for the 24 sentence frames were rated individually by subjects. The method was identical to Experiment 1, except that in this experiment we added a plausibility rating that always followed the expectation rating and paraphrase. This order was followed since it was not likely that subjects could judge the initial "expectedness" of a noun-filler if the judgment were delayed too long. The paraphrase served to focus the subject on the scenario described, in addition to providing the experimenter with information about the subject's interpretation of the sentence. For the sake of exactness, the instructions for both of the ratings are shown in Table 6. Each of the scales appeared on the screen as the subject made each judgment. As the table indicates, we used the word "realistic" in the instructions for the judgments we have been referring to as plausibility judgments because we felt this wording increased the extent to which the subject focused on the event described by the sentence, whose plausibility we were concerned that they consider, rather than on the sentence itself.

The mean expectation ratings, by subjects, for the 24 sentence sets, are shown in Fig. 5, and the mean plausibility ratings, by subjects, are shown in Fig. 6. These ratings are based on 94% of the total ratings, which were for the condition-appropriate interpretation—if a subject rated a sentence according to an alternative interpretation, the rating was not used. A one-way ANOVA for expectation ratings using the four conditions indicated a significant difference by subjects ($F(3,105) = 183.88, p < .001$) and by items ($F(3,92) = 71.84, p < .001$). A one-way ANOVA for the plausibility ratings also revealed highly significant differences between conditions, by subjects ($F(3,105) = 75.85, p < .001$) and by items ($F(3,92) = 20.91, p < .001$).

TABLE 6
RATING INSTRUCTIONS AND SCALES FOR EXPERIMENT 2

How expected was the final word?
Shortly after you view the completion for the sentence fragment, consider about where the completion
would fall on a list of possible completions that you might expect for the sentence. Use this scale:
  On a list of completions I might expect, the word would probably fall in the
    5   TOP 2 possible completions
    4   TOP 10 possible completions
    3   TOP 20 possible completions
    2   TOP 100 possible completions
    1   TOP 1000 possible completions
    0   Below TOP 1000 possible completions

How realistic is the sentence?
Focus on the scenario (situation) in the sentence that you just paraphrased. Your scenario should include
any information conveyed by the final word. Rate how realistic you think the scenario described is.
  The scenario described is
    5   EXTREMELY realistic
    4   VERY realistic
    3   MODERATELY realistic
    2   SOMEWHAT realistic
    1   BARELY realistic
    0   NOT AT ALL realistic

Not all possible paired comparisons be-
tween these ratings are theoretically impor-
tant to this study. The three that we are
interested in are these:

(1) Fully Consistent vs. Less Expected
    Filler

(2) Less Expected Filler vs. Less Ex-
    pected Role
(3) Less Expected Role vs. Less Ex-
    pected Attachment.

For the reasons presented above, which
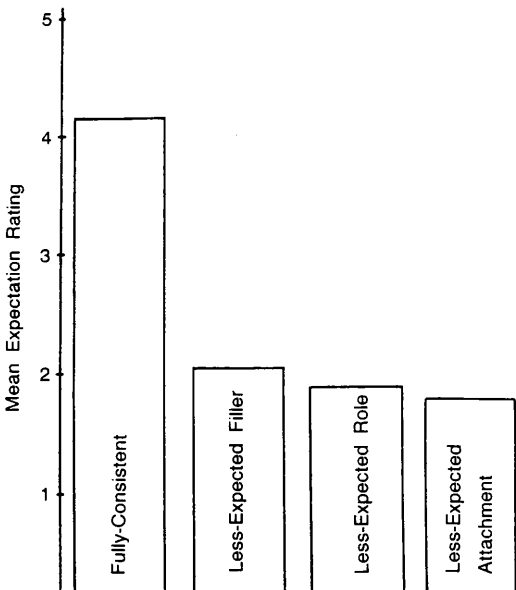have to do with measuring the effects of the

FIG. 5. Experiment 2. Mean *expectation* ratings by
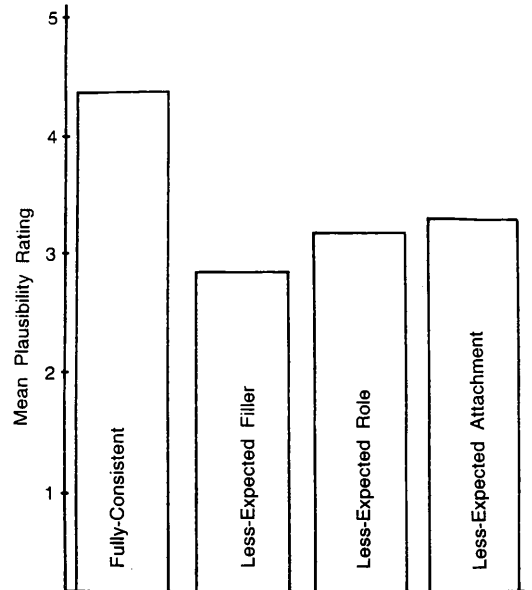condition.

FIG. 6. Experiment 2. Mean *plausibility* ratings by
condition.

noun-fillers, in particular, in one comparison and controlling for their influence elsewhere, we wanted to restrict differences in plausibility and expectedness to comparison (1). The results indicated that we were able to come quite close to this goal. *t* tests for the noun-filler expectation ratings indicated that a difference of 2.11 units between the Fully Consistent and Less Expected Filler conditions exceeded the critical difference of .29 units in the subject analysis; a 2.11 unit difference by items also exceeded its critical difference of .47 units. (Critical differences for ratings are based on two-tailed *t* tests, significant at the .05 level, and adjusted for the number of tests.) There were no significant differences between the Less Expected Filler and Less Expected Role conditions, nor between the Less Expected Role and Less Expected Attachment conditions. *t* tests showed a pattern of significant differences for the sentence plausibility ratings that was similar to the pattern for noun-filler expectation ratings. A difference of 1.53 units between the Fully Consistent and Less Expected Filler conditions exceeded the critical difference of .27 units in the subject analysis; a difference of 1.51 units by items exceeded a critical difference of .50 units for that comparison. A .32 unit difference between the Less Expected Filler and Less Expected Role conditions exceeded the critical difference in the subject analysis, but the .29 unit difference in the item analysis did not exceed the critical difference for that comparison. This slight advantage for the Less Expected Role condition favors the null hypothesis in the on-line reading task, since increased processing costs caused by role switching could be partially masked by the advantage in plausibility. There were no significant differences between the Less Expected Role and Less Expected Attachment conditions either by subjects or items.

Procedure. The procedure in this experiment was identical to the procedure in Experiment 1. Each subject read only one version of each of the 24 test sentence frames

shown in Appendix III and 66 filler sentences. This meant that half of the test sentence frames that any subject saw evoked an expectation for a Verb Phrase Attachment and the other half evoked an expectation for a Noun Phrase Attachment. Test sentences were selected randomly for each subject following a Latin Square design for the four experimental conditions that assured that a subject would see an equal number of noun-fillers from each of these conditions and that all of the completions would appear an equal number of times across subjects. Each test sentence was separated from the next by at least 1 distractor sentence. Eight distractor sentences were used for practice. As in the previous experiment, subjects were instructed to read the sentences as naturally as possible and were informed that the times spent reading each word in the sentence and initiating a response to the question would be recorded. Subjects' responses to the questions were recorded by the experimenter.

Results and Discussion

The experiment was designed to answer three questions:

• What is the effect of manipulating noun-filler expectedness and plausibility when the attachment and thematic role of the prepositional phrase are consistent with content-based expectations?
• What is the effect of switching the role of the prepositional phrase when this is not confounded with attachment switching, and when overall noun-filler expectedness and plausibility is controlled?
• What is the effect of attachment switching, over and above any effect of role switching, again controlling for noun-filler expectedness and plausibility?

We can provide preliminary answers to these by examining Fig. 7, which shows the mean reading times, by subjects, for the four experimental conditions—the times
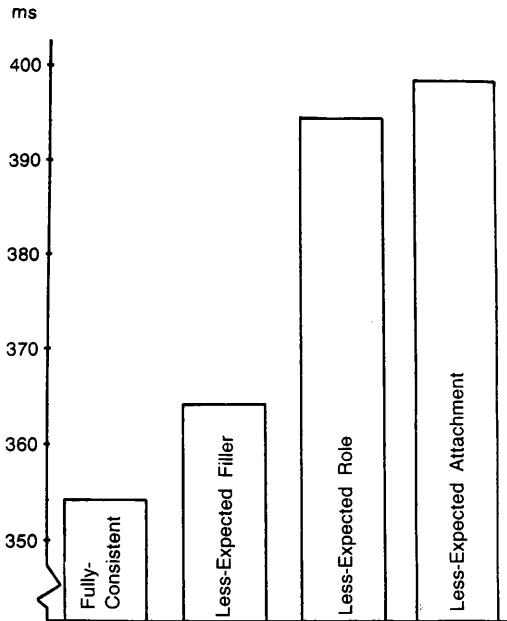
FIG. 7. Experiment 2. Mean reading times by condition. The times shown are average reading times in ms for the noun-filler and the three words that followed.

shown are average reading times for the noun-filler and each of the three words following it.[9]

In response to the first question, a visual examination of the means shows that "high" expectedness (approximately "top 10" on the scale) and "high" plausibility (between "Very" and "Extremely Realistic" on the scale) for noun-fillers in the Fully Consistent condition and "moderate" expectedness and plausibility (approximately "Top 100" and "Moderately Realistic") for noun-fillers in the Less Expected Filler condition produced a difference in reading times of 10 ms on average for the noun-filler and the three words following it. This produced a total advantage of about 40 ms over all four words when the noun-filler was more highly expected and plausible than when it was not. This effect was produced in sentences that

[9] Reading times beyond the third word were again confounded with line changes for some of the sentences, so they are not reported.

were consistent both with respect to subjects' expectations for attachment of the prepositional phrase and thematic role assignment of the prepositional phrase. In the remaining comparisons, noun-filler expectedness and plausibility are held very nearly constant and vary in a much narrower range than they vary between the first two conditions. Thus it seems very likely that these factors do not contribute substantially to any effects found in comparisons among the remaining conditions.

The Less Expected Filler condition shows reading times for sentences in which the noun-filler disambiguates the role of the prepositional phrase in accord with subjects' expectations for the thematic role of the phrase, while the Less Expected Role condition shows reading times when the role required by the sentence for the prepositional phrase is inconsistent with expectations. Attachment is held constant for sentence pairs in these conditions, and filler expectedness and plausibility are controlled. Phrases that could be assigned a role in accord with expectations (Less Expected Filler) produced a 30 ms advantage on average for the noun-filler and each of the three words following it compared to those phrases for which the role required by the sentence for the phrase was not in accord with role expectations (Less Expected Role). This produced a total advantage of about 120 ms over all four words for role-consistent phrases. The answer to the second question, then, is that violation of thematic role expectations appears to produce a substantial slowdown in processing.

In answer to the third question, it appears that violations of subjects' expectations for the syntactic attachment of the prepositional phrase do not produce any additional processing difficulty over and above those already produced by thematic role violations. The phrases that were neither in accord with attachment expectations nor thematic role expectations (Less Expected Attachment) showed only a 4 ms per word disadvantage for the noun-filler

and each of the three words that followed compared to the condition in which attachment was in accord with expectations but the thematic role required by the sentence for the prepositional phrase was not (Less Expected Role). A sentence with a prepositional phrase attachment that was not in accord with subjects' expectations resulted in a total additional disadvantage of about 16 ms over the critical four words if the expectation for a thematic role had also been violated.

An ANOVA using Frame Type (Verb Phrase Bias, Noun Phrase Bias), the four conditions corresponding to differences in Fit-to-Expectation (Fully Consistent, Less Expected Filler, Less Expected Role, Less Expected Attachment) and Word Position (the noun-filler and each of the three words that followed) was used in an initial analysis. There was a main effect for Fit-to-Expectation $(F(3,129) = 13.74, p < .001)$ by subjects and $(F(3,88) = 3.30, p < .03)$ by items. The effect for Word Position was also significant $(F(3,129) = 7.03, p < .001)$ by subjects and $(F(3,264) = 9.32, p < .001)$ by items. There was also a significant Fit-to-Expectation X Position interaction by subjects $(F(9,387) = 2.81, p < .003)$ and by items $(F(9,264) = 2.03, p < .04)$. The significant interaction indicated that the effects at some word positions differed between expectation conditions. This is evident in Fig. 8, which shows the mean reading times, by subjects, at the noun-filler and each of the three word positions that follow, for each expectation condition. The main effect for Frame Type was significant $(F(1,43) = 5.20, p < .03)$ by subjects but not by items $(F(1,88) = 0.59, ns)$; all the interaction terms with Frame Type were nonsignificant (p's > .20), by subjects and items. The main effect for Frame Type indicated that Noun Phrase Bias frames were somewhat harder overall (Noun Phrase Bias 384 ms, Verb Phrase Bias 372 ms, by subjects, for the noun-filler and each of the three words that followed; Noun Phrase Bias 386 ms, Verb Phrase Bias 376 ms, by
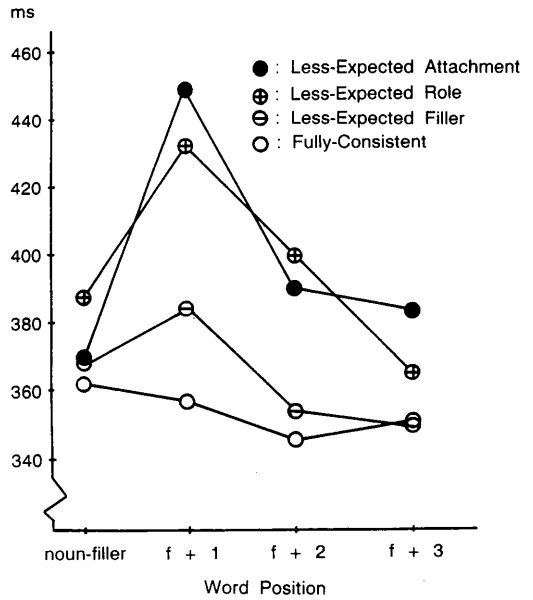

FIG. 8. Experiment 2. Mean reading times for the noun-filler and for each of the three words that followed, in ms.

items); the absence of interactions with Frame Type indicated that the remaining effects did not differ significantly among Noun Phrase Bias and Verb Phrase Bias frames. Note that the effect for Frame Type is not the same as an effect for attachment per se; such an effect is embedded with other factors in the interaction of Frame Type with Fit-to-Expectation. An analysis specifically focusing on this point is presented below.

Separate ANOVAs were used for further analysis of effects for the noun-filler and the three words that followed. The F values (df-numerator = 1) for the main effects in these analyses were used for one-tailed t tests, which are reported below. Each of the paired comparisons is based on the reading times for an equal number of Verb Phrase Bias and Noun Phrase Bias frames. This was important in order to control for any differences in the overall sentence frame on the processing of the prepositional phrase. The purpose of the t tests was to provide specific answers to the three ques-

tions above and to examine where in the temporal course of processing these effects emerged.

The analysis for reading times for the noun-filler itself revealed no significant effects of the fit-to-expectation, indicating that, as in Experiment 1, violations of expectations did not affect the screen duration of the noun-filler. Of course, subjects may well have continued to process the noun-filler after they pressed the key that caused this word to be removed and the next word to be presented.

A comparison of the Fully Consistent condition and the Less Expected Filler condition at f + 1 showed that a 28-ms advantage for the Fully Consistent condition, by subjects, was significant ($t(43) = 2.67$, $p < .01$), as was a 31-ms advantage by items ($t(46) = 1.83, p < .05$). This was the only position at which there was a significant effect. This effect suggests that the appearance of a less expected and less plausible filler in a phrase that otherwise conforms to a subject's expectations for attachment and role produces a small but detectable cost in processing. This effect appears immediately after the subject reads the word and the associated difficulty appears to be overcome quickly.

A comparison of the Less Expected Filler condition and the Less Expected Role condition at f + 1 showed a 46-ms advantage for the Less Expected Filler condition. This effect was significant by subjects ($t(43) = 2.66, p < .01$), as was a 46-ms advantage by items ($t(46) = 1.64, p = .05$). A similar comparison at f + 2 again showed that the 45-ms advantage for the Less Expected Filler condition was significant ($t(43) = 2.84, p < .005$), by subjects, as was the 51-ms advantage by items ($t(46) = 2.03, p < .03$). These effects suggest that thematic role information is important in processing a sentence. When these expectations are not fulfilled, the subject pays a considerable processing cost that tends to persist.

An examination of Fig. 8 shows that the curve for the Less Expected Attachment condition, in which both attachment and thematic role are inconsistent with subjects' expectations, is nearly coincident with the curve for the Less Expected Role condition, in which the thematic role, but not the attachment, is inconsistent with subjects' expectations. The 19-ms advantage for Less Expected Role vs. Less Expected Attachment at f + 3 was significant by subjects ($t(43) = 1.65, p = .05$), but a corresponding 17-ms advantage, by items, was not ($t < 1.0$). These results provide only slim evidence that violations of attachment expectations add anything to observed processing effects beyond the effects that role expectations already contribute.

Experiment 2 was not designed as a test of the Minimal Attachment hypothesis. However, the design did permit us to conduct a post-hoc ANOVA for Minimal Attachment in this experiment, using the attachment expectation evoked by the sentence Frame (Verb Phrase Bias, Noun Phrase Bias), the ultimate Attachment required by the prepositional phrase (Verb Phrase Attachment, Noun Phrase Attachment), and Word Position (the noun-filler and each of the three words that followed). In order to control for any possible overall differences in sentence frames, we used the Less Expected Role and Less Expected Attachment conditions, thereby using all the same frames for both the Verb Phrase Attachment and Noun Phrase Attachment effects. Examining first the rating results for these materials, there was a slight effect in the expectation ratings favoring the Verb Phrase Attachment condition (Verb Phrase Attachment: 1.96 vs. Noun Phrase Attachment: 1.71), which was significant by subjects ($t(35) = 2.24, p < .03$), but not items, ($t(44) = 1.16$, ns). While this effect tends to favor Minimal Attachment, there was a slight offsetting disadvantage in the plausibility ratings (Verb Phrase Attachment:

3.14 vs. Noun Phrase Attachment: 3.27). This was neither significant by subjects ($t(35) = 0.85$, ns) nor by items ($t(44) = 1.37$, ns). In the analysis of reading times, there was only a single significant effect, which was for Word Position ($F(3,129) = 7.04, p < .001$) by subjects and ($F(3,132) = 7.48, p < .001$) by items. There was an overall advantage for minimal attachment (Verb Phrase Attachment 387 ms, Noun Phrase Attachment 405 ms, by subjects; Verb Phrase Attachment 391 ms, Noun Phrase Attachment 409 ms, by items— average reading times for the noun-filler and each of the three words that followed) but this effect was not significant nor were any of the remaining main effects or interactions (all $ps > .10$). While these findings are not inconsistent with some partial role for minimal attachment, they do not provide much evidence of the importance of this principle for predicting reading times in this task.

GENERAL DISCUSSION

The results from Experiment 2 corroborate a conclusion drawn from Experiment 1, which is that the particular content of a sentence evokes expectations for the on-line processing of constituents in advance of the input that fully specifies the constituent, and that violations of these expectations result in a slowing of processing. Experiment 2 also indicates which aspects of subjects' expectations were the ones whose violation gave rise to the observed processing effects. Apparently, mild violations of subjects' expectations for actual content of a prepositional phrase, together with a mild reduction in the plausibility of the particular filler for the role to which it is assigned, does have a significant impact on processing, but the effect is rather small and short-lived within the range of violation of expectation and plausibility considered here, as long as the expected role and attachment are not changed. Violations of subjects' expectations for the thematic role of the prep-

ositional phrase appear to exert a stronger influence on reading times. When the thematic role required for the prepositional phrase by the particular sentence is not in accord with the role expectations of the subject, there is no additional slowdown of processing when the required syntactic attachment of the prepositional phrase is also in violation of a subject's expectations.

Are subjects really using content to form expectations in anticipation of upcoming constituents? Throughout this paper we have argued that our findings suggest that subjects *anticipate* the thematic role and attachment of a prepositional phrase before it is complete and that processing is slowed when the complete prepositional phrase fails to satisfy these expectations. Is this argument justified? In answering this question, it is worth considering an alternative possibility that might be offered, namely that we are not so much observing a violation of expected role and attachment, but a difficulty in integrating a completed prepositional phrase with prior material. To make this argument concrete, consider sentences [2] and [3] again:

[2] The spy saw the cop with binoculars.
[3] The spy saw the cop with a revolver.

Subjects process [2] more quickly than [3], it might be argued, not so much because the prepositional phrase in [3] actually violates their expectations but because they find it more difficult to integrate with the prior context of the sentence, in the absence of any prior expectations.

This interpretation seems quite consistent with the results of Experiment 1, since in that experiment, it is quite likely that (on the average, anyway) the prepositional phrases that violated the subjects' role assignment and attachment expectations fit together less well with the prior content of the sentence. This being the case, it could well be that the difficulty subjects had with the unexpected versions of the prepositional phrases in that experiment could

have reflected a difficulty in finding a co-
herent interpretation of the entire sentence.
Some of the results of Experiment 2, how-
ever, are more difficult to reconcile with
this kind of view. For in Experiment 2, we
found that prepositional phrases whose
noun-fillers were matched for expectedness
and plausibility could still differ in the
amount of slowing they produced, as a
function of whether the role that the prep-
ositional phrase was interpreted as filling
matched the role that the subject expected
based on the prior content of the sentence.
Such a difference seems to indicate that the
prior content of the sentence leads the sub-
ject to set up some mental structure that is
more ready to accommodate a preposi-
tional phrase filling one role than a prepo-
sitional phrase filling another. When a com-
pleted prepositional phrase fits this ex-
pected role—even if it does not fit it
particularly well—processing goes through
much more quickly than it does when the
prepositional phrase requires a different
role (and possibly a different attachment as
well).

In summary, the results of Experiment 2
reinforce the general conclusion that sub-
jects generate expectations for the role as-
signments of upcoming constituents and
use these to guide the processing of these
constituents, and with the assumption that
these expectations are based in part on the
content of the sentence, rather than simply
its syntactic form. The findings further sug-
gest that the assignment of constituents to
roles, or at least, the reassignment of con-
stituents when role assignment expecta-
tions are violated, may be a rate-limiting
aspect of on-line reading processes.

Implications for Models of
   Sentence Processing

A basic task that our subjects face in
reading materials like the ones we consider
here is to determine the thematic role of the
various constituents of the sentences. Our
results have implications for several kinds

of models of the processes that occur in
arriving at these role assignment decisions.

Syntax-first models. Our results provide
little support for models in which an auton-
omous parsing mechanism proposes a sin-
gle parse based deterministically on general
syntactic principles alone, subject to accep-
tance or rejection by a "Thematic Proces-
sor" (Rayner et al., 1983; Frazier, 1986;
Ferreira & Clifton, 1986). On the particular
assumption that the principle that governs
syntactic processing decisions is Minimal
Attachment, we would expect subjects to
have difficulty with all prepositional
phrases that do not require a Verb Phrase
Attachment reading. This is not what we
found for those sentences labeled Taraban
and McClelland in Experiment 1—i.e., the
set of sentences that set expectations for
Object-NP attachment of a prepositional
phrase prior to the availability of the disam-
biguating noun-filler. The finding that some
sentences with comparable surface constit-
uent arrangements favored prepositional
phrases that require minimal attachment
while others favored prepositional phrases
that require nonminimal attachment makes
it very difficult to see how any syntactic
principle—be it Minimal Attachment or
something else—could be determining the
constituent structure made available for
thematic processing in a consistent way.
Our results are not inconsistent with ver-
sions of syntax-first models in which the
syntactic mechanism makes several alter-
native parses available for simultaneous
consideration by a Thematic Processor.
Some versions of such an account might
give syntax a deterministic role in some
cases but would leave it to the thematic pro-
cessor to select among alternatives that the
syntactic processor could not itself reject.
Others might allow the syntactic processor
to provide a range of possibilities with
weightings indicating the "goodness" of
each possibility (see Marcus, 1980; or
Kurtzman, 1984, for a discussion of both of
these possibilities). In any event, it would

appear that the factors governing the effects that were obtained in the present experiments would have to be attributed to the inner workings of the Thematic Processor, rather than to decisions made in the syntactic processor itself.

Lexically guided models. A second class of models are those that use the specific lexical content of a sentence, particularly the verb, in order to guide syntactic and semantic processing. One of the most highly articulated models of this sort is the LFG model of Kaplan and Bresnan (1982). In this model, lexical entries or "forms" for verbs appear with predicate-argument specifications that define a mapping between the thematic arguments of a verb and the grammatical functions (Subject, Object, etc.) associated with these arguments. Ford et al. (1982) suggested that the various forms of a given verb have different "strengths" or "saliences," perhaps based on general frequency of usage, and that the strongest form determines the initial syntactic analysis. Thus, to use an example from Ford et al., if the strongest form for "want" specifies only a Subject and an Object, while the strongest form for "position" specifies a Subject, an Object, and a Prepositional Phrase Complement, the parser will interpret "The woman positioned the dress on the rack" in accordance with Minimal Attachment (treating the prepositional phrase as a constituent of the verb phrase, but will interpret "The woman wanted the dress on the rack" in opposition to Minimal Attachment, treating the prepositional phrase as a constituent of the complex Noun Phrase "the dress on the rack." Generally, as in this example, discussions of sentence processing under LFG have focused on the way in which the lexicon can govern attachment decisions. However, it seems clear that the lexical forms can also be used to govern role assignment decisions as well. Indeed, our data might be taken as suggesting that the slowdown in processing that occurs when a constituent that does not fit an expected role results from the processing activity associated with replacing the strongest lexical form with another that is more appropriate to the particular content of the sentence.

The LFG model, then, suggests a mechanism that allows information associated with verbs to govern expectations for attachment and role assignment. However, there are difficulties that must be faced before the LFG model could be said to provide a full account of the outcome of attachment and role assignment processes. First of all, the model would need to be augmented to capture the fact that information associated with noun phrases also appears to set up expectations whose violation causes processing delay. An unexpected role for a prepositional phrase that complements a Noun Phrase produces a disruption in processing, just as an unexpected role associated with a prepositional phrase that complements the main verb causes disruption. Perhaps this could be accommodated by proposing lexical forms for nouns as well as verbs. A fundamental problem, though, lies in the fact that remote information, outside the clause containing a prepositional phrase whose attachment is at issue, appears to influence attachment decisions. An example of this can be found in Ford et al. (1982), who report that subjects' attachment preferences, measured off-line, were different for [8] and [9]. While subjects' interpretations favored Noun Phrase Attachment of "for Susan" in [8] ("the package that was for Susan"), they favored Verb Phrase Attachment in [9] ("carried for Susan"):

[8] When he arrived at our doorstep, I could see that Joe carried a package for Susan.

[9] Whenever she got tired, Joe carried a package for Susan.

A final problem arises in specifying the conditions under which a reanalysis is required. The LFG theorists have been clear

624   TARABAN AND MCCLELLAND

about how content that is incompatible with grammatical category requirements could trigger reanalysis of constituent relations (Kaplan & Bresnan, 1982; Ford et al., 1982). However, there are no clear mechanisms for accommodating the influence of content when it does not produce outright ungrammaticality. An illustration of this problem comes from our own materials. In a sentence frame like the following from Experiment 1, *I read the article in the* _____, most subjects expected a completion consistent with a Noun Phrase modifier interpretation of the prepositional phrase (e.g., magazine or newspaper). Yet when confronted with the noun-filler "bathtub," most subjects interpreted the prepositional phrase as the location in which the "reading" took place. It appears that the expected thematic role and attachment for the prepositional phrase were rejected because "reading in the bathtub" is more plausible than finding reading material in the bathtub, even though this is not a strictly unacceptable interpretation and even though it was consistent with the initial preference exhibited by most subjects. In general, it is not clear how the appropriate attachment and role assignment decisions can be reached without in some sense weighing the relative merits of the alternatives.

One possible way of dealing with these problems would be to suppose that lexical guidance, as postulated by LFG, influences initial "bottom-up" attachment and role assignment decisions, but that ultimately a thematic processor of the kind envisioned by Frazier (1986; Rayner et al., 1983), in which different alternatives are weighed for plausibility and fit to the entire context, would be required to select the most plausible reading. Appealing to a thematic postprocessor may allow the LFG account to be saved in the face of evidence that suggests that the ultimate interpretations readers get appear to depend, rather generally, on all of the constituents of a sentence and prior

pragmatic context (Crain & Steedman, 1985). However, such a move undermines much of the appeal of the LFG account, since it once again places much of the action in the underspecified "Thematic Processor," and robs the initial processing mechanisms of access to any but the simplest content information. Experimentally, though, the materials we used in the present experiments do not rule out the possibility that initial processing decisions are based only on expectations associated with the heads of constituents to which an item might be attached (both Verb Phrases and Noun Phrases). Further research is planned to examine this issue.

Multicue models. A third class of models consists of models in which syntactic cues and information derived from the content of the sentence are used together to guide on-line attachment and role assignment processes. Such models include interactive models of the type proposed by Rumelhart (1977), Just and Carpenter (1980), Marslen-Wilson and Tyler (1980), Tyler and Marslen-Wilson (1977), Thibadeau et al. (1982), and McClelland (1987), in which construction of syntactic and conceptual representations occur in parallel, exerting mutual influences on each other, as well as models in which both kinds of cues are used in constructing a conceptual representation of the content of the sentence, and syntactic attachment is implicit in the relations among constituents that are captured in the conceptual structure (McClelland & Kawamoto, 1986; MacWhinney, 1987). In these models it is generally assumed that graded activation and competition allows multiple syntactic and semantic cues to work together or to compete with each other, with the interpretation that is most consistent with the cues and most internally consistent winning out over other alternatives.

Models of this type have a number of attractive properties. (1) They are naturally consistent with the possibility that syntac-

tic cues could dominate in some cases, but that content cues could dominate in other cases, depending on the relative strength of cues in each case. (2) They provide a natural mechanism for each constituent of a sentence to contribute to the interpretation of all other constituents. (3) They seem completely consistent with the results of the present experiments and with the attractive possibility that all aspects of the content of a sentence can influence initial processing decisions.

The drawback of these models is that most are quite incomplete or underspecified. It is a matter of ongoing research to develop computationally sufficient models of this type. Meanwhile, models of this class make a number of testable predictions that can be assessed in further experiments. First, they predict that expectations that guide processing can be generated in response to characteristics of other constituents of the sentence besides the verb, and even by pre- or extra-sentential context. Second, they predict that processing will be influenced, not only by the plausibility and expectedness of the target reading of the sentence, but also by the existence of other, competing readings. We are currently developing experimental tests of both of these predictions.

CONCLUSION

Taken together, the results from Experiments 1 and 2 help to establish the potency of content-based expectations on sentence processing and to indicate the nature of the difficulties that violations of these expectations produce. We have found that noun-fillers that violate subjects' expectations lead to increased processing times on the noun-filler if it is the last word of a sentence, or on the one or two words after it, if it is not the final word of a sentence. What matters most is violation of the subject's expectations for the thematic role of the prepositional phrase. Whether the expected syntactic attachment is violated causes lit-

tle, if any, additional difficulty over and above that associated with violation of role expectations.

The expectations that govern these effects depend on the specific content of the sentence, rather than on any syntactic principle such as Minimal Attachment. The fact that the effects depend on sentence content is demonstrated by the fact that their direction—whether they be for one kind of role/ attachment combination or another—varies from sentence to sentence as the content of the sentence varies. This is not to say that there might not be some syntactic bias affecting processing to some slight degree that falls below the threshold of reliable detection in these studies; it is only to say that these syntactic biases if they exist are not the predominant factor influencing the time it takes to attach and assign prepositional phrases, as indicated in our word-by-word reading time measures.

At this point we do not have the data that would allow us to say what aspects of the content of a sentence underly these expectations. It remains for further research to establish whether all constituents of a sentence, and even extra-sentential context, can potentially influence initial expectations for attachment and role assignment, or whether the verb has some special privileged status in this regard.

APPENDIX I

Note. Test sentences from Experiments 1A and 1B. The sentences in Experiment 1A appeared without the post-prepositional phrase continuation; the sentences in Experiment 1B included the continuation. NPA: Noun Phrase Attachment; VPA: Verb Phrase Attachment.

Taraban and McClelland Sentences

The thieves stole all the paintings in the
    (i) museum (NPA) (ii) night (VPA) (non-intrinsic) while the guard slept.
The couple admired the house with a

(i) garden (NPA) (ii) friend (VPA) (non-intrinsic) but knew that it was over-priced.

The tourist learned the route through the
(i) mountains (NPA) (ii) interpreter (VPA) (intrinsic) while traveling on vacation.

The administrator announced the cuts in the
(i) budget (NPA) (ii) meeting (VPA) (non-intrinsic) even though he knew it would create hard feelings.

The engineers designed the bridge over the
(i) river (NPA) (ii) summer (VPA) (non-intrinsic) but a year passed before it was built.

The report described the government's programs in
(i) education (NPA) (ii) detail (VPA) (intrinsic) but most people ignored it.

The police arrested the mastermind behind the
(i) crimes (NPA) (ii) hideout (VPA) (non-intrinsic) but they forgot to read him his rights.

The spy had the plans for a
(i) weapon (NPA) (ii) price (VPA) (intrinsic) but he was caught before he could sell them.

I read the article in the
(i) magazine (NPA) (ii) bathtub (VPA) (nonintrinsic) while I was waiting for the cookies to bake.

The President suggested a solution to the
(i) problem (NPA) (ii) people (VPA) (intrinsic) although he knew it would be rejected.

The corporate executive considered the issues under
(i) discussion (NPA) (ii) pressure (VPA) (intrinsic) because his career depended on the outcome.

The woman married the man with
(i) money (NPA) (ii) delight (VPA) (intrinsic) while her friends looked on with envy.

The doctor cured the woman with
(i) tuberculosis (NPA) (ii) penicillin

(VPA) (intrinsic) even though his colleagues had thought it unlikely.

The hospital admitted the patient with
(i) cancer (NPA) (ii) urgency (VPA) (intrinsic) because she required intensive care.

The reporter exposed corruption in the
(i) government (NPA) (ii) article (VPA) (nonintrinsic) even though he risked making many enemies.

The woman flaunted the expensive ring around her
(i) finger (NPA) (ii) friends (VPA) (intrinsic) even though it wasn't paid for yet.

John ordered a pizza with
(i) pepperoni (NPA) (ii) enthusiasm (VPA) (intrinsic) when he was finished studying for his calculus exam.

The Vietnam veteran identified his old buddy from the
(i) war (NPA) (ii) photo (VPA) (intrinsic) even though many years had passed since he had seen him.

Rayner, Carlson, and Frazier Sentences

The spy saw the cop with
(i) binoculars (VPA) (ii) a revolver (NPA) but the cop didn't see him.

The little girl tried to cut the apple with plastic
(i) knives (VPA) (ii) coating (NPA) though she wasn't very successful.

The landlord painted all the walls with
(i) enamel (VPA) (ii) cracks (NPA) though it didn't help the appearance of the place.

John played the records with
(i) Jim's needle (VPA) (ii) deep scratches (NPA) last night to see how bad they sounded.

Jane finally decided to read the books on the
(i) train (VPA) (ii) list (NPA) so that she wouldn't fail her history test.

The overworked scientist only read the news reports on
(i) Sundays (VPA) (ii) tornados (NPA) because he had so little time.

The executive only called people on the
(i) intercom (VPA) (ii) payroll (NPA) because he was paranoid.

The kids played all the albums on the
(i) stereo (VPA) (ii) shelf (NPA) before they went to bed.

Grandfather could only read the numbers in bright
(i) rooms (VPA) (ii) colors (NPA) because of his poor eyesight.

That kid hit the girl with a
(i) whip (VPA) (ii) wart (NPA) before he got off the subway.

The doctor examined the patient with a
(i) stethoscope (VPA) (ii) toothache (NPA) but he couldn't determine what the problem was.

APPENDIX II: IDENTIFYING AND TESTING DISTINCT THEMATIC ROLES

Identifying a reasonable set of thematic roles for our stimuli was particularly challenging, since the issue of how many thematic roles are actually required for a grammar has not been settled. The difficulty of resolving this issue has been acknowledged by Fillmore (1968, 1971) in a number of places, as well as by others (Winograd, 1983). Fillmore's work provides a starting point for generating a list of possible thematic roles, particularly those associated with the verb.

In addition to roles found in Fillmore, we used Jacobson's (1964) definitions for adverbs and the primary definitional senses found in Wood (1967) for the prepositions used in our stimuli in order to generate additional candidate thematic roles. It was then necessary to distinguish distinct roles from those that were similar. One case is that between Instrument and something like Effective Substance, as in the two sentences, "The landlord painted the walls with the brush," and "The landlord painted the walls with the enamel." Further, there has been no discussion, that we are aware of, of the assignment of thematic roles to

postverbal prepositional phrases that modify the Object-NP. A distinction between these roles is necessary for the Noun Phrase Attachment stimuli. Thankfully, variations of some of the roles already proposed in Fillmore seem to apply here as well. For example, the prepositional phrase in "The boy spent the money in his pocket" could probably be classified as a Location. There are other instances in which the specific role is less evident, as in "The hospital admitted the patient with cancer."

In order to distinguish between roles, both for Verb Phrase and Noun Phrase Attachments, we adopted a test from Fillmore (1971). This test tests for the acceptability of the juxtaposition of two constituents in a coordinate conjunction construction. The basic idea is that if two noun phrases can be brought together in the same coordinate structure, then they must have the same thematic role. Thus, in this test, two noun phrases are conjoined and the felicity of their attachment as a unit to the preposition is considered. The prepositional phrase itself is either attached as a Noun Phrase or Verb Phrase Attachment. For example, consider "The doctor examined the patient with a stethoscope and rubber mallet." The conjunctive construction should seem acceptable, and the nouns seem to both be instruments. On the other hand, in "The doctor examined the patient with care and a rubber hammer," the conjunctive phrase seems awkward, and indeed, the first noun appears to express the manner of examination, while the second appears to express the instrument. Thus, the test appears to yield correct results in clear cases. When we apply this test to the Instrument and Effective Substance distinction considered above, we find that the two do not contrast in the same way as manner and instrument roles. For example, "The landlord painted the walls with a roller and latex paint" seems fine. Since these two arguments can be conjoined, we would not propose two distinct roles in this case.

Appendix III

V: verb phrase attachment
N: noun phrase attachment
F+: higher plausibility noun-filler
F−: lower plausibility noun-filler
AC: accompaniment role
AT: attribute role
C: content role
G: goal role

I: instrument role
L: location role
M: manner role
PR: purpose role
SC: scope role
S: source role
T: time role

Pretest Results ($n = 30$)

		Filler-frequency	Percentage attachment	Percentage role
V,M,F+ F− I N,AC	We watched the magic acts with while we were at the circus.	(amazement) 10 (disinterest) 3 (binoculars) 2 (hawks) 7	100	90
V,M,F+ F− I N,AT	The nurse undid the bandage with because of her expertise.	(care) 89 (pride) 45 (tweezers) (adhesive) 6	100	70

V,M,F+
 	
  F−
AC
N,L
	The conscientious researcher pursued
	  the problem with
	
	while all of his colleagues simply
	  ignored it.
		(enthusiasm) 29
		(dignity) 35
		(volunteers) 34
		(welfare) 53
			93	97

V,I,F+ F− M N,AT	The little girl cut the apple with a while she was playing in the kitchen.	(knife) 86 (fork) 20 (smirk) 2 (blemish) 3	100	100
V,I,F+ F− M N,AT	The bully hit the girl with a while the other children watched.	(stone) 66 (pillow) 11 (shriek) 4 (handicap) 7	93	100
V,I,F+ F− M N,AT	The policeman beat the suspect with a even though the suspect was armed.	(club) 178 (newspaper) 104 (vengeance) 10 (scar) 20	100	100

V,I,F+
 	
  F−
AC
N,AT
	The janitor cleaned the storage area
	  with the
	
	because of many complaints.
		(broom) 2
		(solvent) 8
		(manager) 114
		(odor) 22
			100	97

V,I,F+ F− AC N,AT	The landlord painted the wall with a	(brush) 36 (ladder) 6 (tenant) 14 (scratch) 17	100	80

APPENDIX III—*Continued*

		Filler-frequency	Percentage attachment	Percentage role
	but all his efforts didn't improve the looks of the wall.			
V,M,F+ F− I N,AT	The rescue crew reached the victim in	(minutes) 242 (seconds) 57 (canoes) 8 (shock) 33	97	90
	while the hospital personnel prepared for the victim's arrival.			
V,T,F+ F− M N,AT	The choir sang the carol on	(Christmas) 27 (Hanukkah) 1 (demand) 123 (parchment) 1	100	93
	because most people enjoyed hearing it.			
V,G,F+ F− L N,AT	The entrepreneur used the money on	(investments) 49 (stationery) 2 (vacation) 55 (hand) 717	97	97
	because he figured it would be the best thing to do.			
V,PR,F+ F− L N,S	The police protected the prime witness from the	(suspect) 7 (rain) 73 (balcony) 7 (rape) 3	93	96
	because of their concern for the witness.			
N,AT,F+ F− AC V,M	The hospital admitted the patient with	(cancer) 24 (amnesia) (bodyguards) 1 (apologies) 8	73	100
	while the other incoming patients calmly waited.			
N,AT,F+ F− AC V, I	The woman married the man with the	(money) 275 (trophies) 10 (groom) 5 (vow) 6	93	100
	while her friends looked on in disbelief.			
N,L,F+ F− AT V,T	The thieves stole all the paintings in the	(museum) 42 (library) 90 (frames) 96 (afternoon) 122	100	100
	while the security officer was on break.			
N,C,F+ F− SC V,M	The administrator discussed the cuts in	(spending) 12 (lighting) 7 (sight) 97 (secret) 52	93	100
	because a lot of employees would be affected.			

Appendix III—*Continued*

		Filler-frequency	Percentage attachment	Percentage role
N,C,F+ F− T V,T	The executive announced the reductions in the even though he knew they would create hard feelings.	(budget) 62 (research) 172 (future) 108 (evening) 149	100	100
N,L,F+ F− AT V,M	The document described the Ku Klux Klan's activities in because of the threat that the Klan posed.	(Alabama) 29 (Alaska) 22 (robes) 10 (code) 55	80	100
N,CO,F+ F− T V,G	The mayor explained his master plan for the even though there wasn't really much public interest.	(town) 281 (river) 183 (month) 327 (audience) 131	87	96
N,CO,F+ F− L V,M	The reporter destroyed the story on because he felt it might ruin his credibility.	(fraud) 13 (God) 335 (tape) 39 (impulse) 32	87	96
N,CO,F+ F− L V,L	The philanthropist appreciated the story on his because it was so sentimental.	(generosity) 7 (mother) 280 (lap) 19 (deathbed) 2	100	100
N,CO,F+ F− L V,M	The manager accepted the report on but he expected some verification to follow.	(profits) 46 (workers) 123 (record) 214 (faith) 110	80	100
N,CO,F+ F− L V,L	The rock star co-authored the bestseller on but he still wasn't satisfied with his accomplishments.	(drugs) 69 (surfing) (newsstands) 1 (tour) 48	100	97
N,T,F+ F− C V,G	The high-school senior stated his goals for the but he didn't believe anyone really cared.	(future) 108 (hour) 325 (dance) 94 (principal) 26	100	80

REFERENCES

CLIFTON, C., FRAZIER, L., & CONNINE, C. (1984). Lexical expectations in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* 23, 696–708.

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

DANEMAN, M., & CARPENTER, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 9, 561–584.

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

FILLMORE, C. (1968). The case for case. In E. Bach and R. T. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart, & Winston.

FILLMORE, C. (1971). Some problems for case grammer. In O'Brien (Ed.), *Report of the 22nd Annual Round-Table Meeting on Linguistics and Language Studies.* Washington, DC: Georgetown University Press.

FILLMORE, C. (in preparation). *On grammatical constructions.*

FORD, M., BRESNAN, J., & KAPLAN, R. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

FODOR, J. A., BEVER, T., & GARRETT, M. F. (1974). *The psychology of language.* New York: McGraw-Hill.

FRANCIS, W., & KUCERA, H. (1982). *Frequency analysis of English usage.* Boston: Houghton Mifflin.

FRAZIER, L. (1978). *On comprehending sentences: Syntactic parsing strategies.* Ph.D. dissertation, University of Connecticut. Bloomington: Indiana University Linguistics Club, 1979.

FRAZIER, L. (1986). Theories of sentence processing. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing.* Cambridge, MA: MIT Press.

FRAZIER, L., & FODOR, J. (1978). The sausage machine: A new two-stage parsing model. *Cognition,* 6, 291–325.

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.

HOLMES, V. M. (1987). Syntactic parsing: In search of the garden path. In M. Coltheart (Ed.), *Attention and performance XII.* London: Erlbaum.

JACKENDOFF, R. (1972). *Semantic interpretation in generative grammar.* Cambridge, MA: MIT Press.

JACOBSON, S. (1964). *Adverbial positions in English.* Stockholm: Studentbok.

JUST, M. A., & CARPENTER, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review,* 87, 329–354.

JUST, M. A., CARPENTER, P. A., & WOOLLEY, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General,* 111.

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

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

KURTZMAN, H. (1984). *Studies in syntactic ambiguity resolution.* Ph.D. dissertation, MIT. Bloomington: Indiana University Linguistics Club.

MACWHINNEY, B. (1987). Competition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

MARCUS, M. (1980). *Theory of syntactic recognition for natural language.* Cambridge, MA: MIT Press.

MARSLEN-WILSON, W. D., & TYLER, L. (1980). The temporal structure of spoken language understanding. *Cognition,* 8, 1–71.

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

MCCLELLAND, J., & KAWAMOTO, A. (1986). Mechanisms of sentence processing: Assigning roles to constituents. In J. L. McClelland and D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Volume 2. Cambridge, MA: MIT Press.

MITCHELL, D. C., & HOLMES, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language,* 24, 542–559.

ODEN, G. (1978). Semantic constraints and judged preference for interpretations of ambiguous sentences. *Memory and Cognition,* 6, 26–37.

ODEN, G. (1983). On the use of semantic constraints in guiding syntactic analysis. *International Journal of Man–Machine Studies,* 19, 335–357.

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.

SCHANK, R. C. (1972). Conceptual dependency: A

theory of natural language understanding. *Cognitive Psychology,* 3, 552–631.

SCHANK, R. C. (1975). *Conceptual information processing.* Amsterdam: North-Holland.

THIBADEAU, R., JUST, M. A., & CARPENTER, P. A. (1982). A model of the time course and content of reading. *Cognitive Science,* 6, 157–203.

TYLER, L, & MARSLEN-WILSON, W. D. (1977). The on-line effects of semantic content on syntactic processing. *Journal of Verbal Learning and Verbal Behavior,* 16, 683–692.

WINOGRAD, T. (1983). *Language as a cognitive process.* Menlo Park, CA: Addison–Wesley.

WOOD, F. T. (1967). *English prepositional idioms.* New York: MacMillan Co.