JOURNAL OF MEMORY AND LANGUAGE 30, 725–745 (1991)

Recovery from Misanalyses of Garden-Path Sentences

FERNANDA FERREIRA AND JOHN M. HENDERSON

University of Alberta, Edmonton Alberta, Canada, T6G 2E9

In five experiments we examined the way in which readers reanalyze garden-path sentences, using grammaticality judgments as the dependent measure. The stimuli were two-clause sentences containing an ambiguous noun phrase which could function as either the object of the first clause or the subject of the second. Prior research has shown that the former analysis is generally preferred. In the first two experiments, we varied the number of words in the ambiguous phrase and found that reanalysis of garden-path sentences was more difficult with a longer ambiguous phrase. The third experiment established that this effect of phrase length is not attributable to the greater syntactic complexity of longer phrases. The fourth and fifth experiments demonstrated that the effect of phrase length is attributable to increasing the distance from the head of the ambiguous phrase to the disambiguating word of the garden-path sentence: Ambiguous phrases made long by the addition of prenominal adjectives were easy for the parser to reanalyze, but phrases made long by the addition of postnominal modifying prepositional phrases (Experiment 3) or relative clauses (Experiments 4 and 5) were hard for the parser to reanalyze. From these results, we argue that sentence comprehension requires the creation of phrase structure and the assignment of thematic roles to phrases, with the assignment taking place at the phrasal head. Reanalysis is affected by the ease with which thematic roles can be reassigned to misanalyzed phrases.

It is not unusual for readers to find that they have misunderstood a portion of a sentence. For example, with a sentence such as *Bill saw the girl with binoculars,* the reader may initially interpret the phrase *with binoculars* as a modifier of the action of seeing, only to discover later in the discourse that it is the girl who is in possession of the binoculars. In general, reanalysis has the following character: The language comprehension system encounters a portion of the input string that is ambiguous at some level of representation (in the example above, *with binoculars* is syntactically ambiguous). A strategy is adopted to resolve the ambiguity, and one of the possible interpretations is settled on. The strategy used may depend on the type of information available to the comprehension system at the point of the ambiguity. Then a later portion of the sentence or text disconfirms the initial analysis, and the other, less preferred interpretation must be selected. From this view of reanalysis, the following questions arise: First, what are the strategies the language comprehension system uses to resolve ambiguity? Second, what factors determine how easily a less preferred interpretation is selected?

The first question has been studied ex-

The research reported in this article was supported by Natural Science and Engineering Research Council of Canada operating grants to Fernanda Ferreira (OGP-37323) and John Henderson (OGP-41792). These experiments were reported at the CUNY sentence processing conference, New York, NY, March, 1990. We would like to thank Stephen Hayduk, Andrea Kross, Lucille Lavender, and Alan Petersen for their assistance with data collection and analysis. We also thank Chuck Clifton and Lyn Frazier for comments on an earlier version of this manuscript. Correspondence concerning this article and reprint requests may be addressed to either author at the Department of Psychology, P-220 Biological Sciences, University of Alberta, Edmonton, AB, Canada T6G 2E9. (Electronic mail address: usernand@ualtamts.bitnet or userjmhe@ualtamts.bitnet.)

0749-596X/91 $3.00

tensively in the area of syntactic parsing. Consider the following sentences:

(1) a. Because Bill drinks *wine* . . .
b. Because Bill drinks *wine* beer is never kept in the house
c. Because Bill drinks *wine* is never kept in the house.

The word *wine* in (1a) is syntactically ambiguous, because it could function as either the direct object of the verb in the first clause (as in 1b), or as the subject of a second clause (as in 1c). Frazier (1978, 1987) has proposed that the parser uses the principle of *late closure* to resolve the ambiguity. According to this principle, the parser prefers to attach new material into the phrase or clause currently open rather than to create new constituents. Thus, in (1a), the analysis on which *wine* is a direct object would initially be chosen, and sentence (1b) would be parsed quite easily. In contrast, a sentence such as (1c) would be more difficult. The disambiguating word *is* signals that *wine* cannot be the direct object of *drinks,* and so syntactic reanalysis of the ambiguous phrase is required. Because (1b) is consistent with late closure, it will be termed a *late closure* sentence. Sentence (1c), which requires that the first clause be closed early, will be termed an *early closure* sentence. Sentences such as (1c) that require reanalysis are termed *garden-path* sentences.

There is some debate in the area of sentence comprehension about the source of the garden-path in sentences such as (1c). For example, Frazier's model states that the garden-path occurs because an incorrect structural commitment is initially made based on the parser's principles of operation (Clifton & Ferreira, 1989; Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Frazier, 1978, 1987, 1989; Frazier & Rayner, 1982; Mitchell, 1987, 1989; Rayner, Carlson, & Frazier, 1983). Others have argued that the garden-path in (1c) arises because in (1a), all sources of information are suggestive of a direct object interpretation—the absence of a comma, the transitivity biases of the verb *drinks,* and the pragmatic appropriateness of wine as a drinkable object (Altmann, 1989; Holmes, 1987; Holmes, Stowe, & Cupples, 1989; Pritchett, 1988; Stowe, 1989; Tanenhaus & Carlson, 1989; Tanenhaus, Carlson, & Trueswell, 1989; Taraban & McClelland, 1988; Tyler, 1989). But whatever the source of the garden-path, an important question is: What determines the ease with which the parser can recover from it? This question is the focus of the present article.

To begin, compare sentence (1c) with sentence (2).

(2) Because Bill drinks *wine and other spirits* are never kept in the house.

The ambiguous region in (2) contains more words than in (1c), and the resulting sentence is more difficult to process (Frazier & Rayner, 1982; Warner & Glass, 1987). Using eye movement monitoring, Frazier and Rayner (1982) compared early and late closure sentences with short and long ambiguous regions. They found that fixation times were longer on the disambiguating word of early closure sentences, compared to late closure sentences, and times were longer for the early closure sentences when the ambiguous region was long rather than short. Warner and Glass (1987) also investigated early and late closure sentences with short and long ambiguous regions, using a grammaticality judgment task. They found that early closure sentences with a long ambiguous region were often considered ungrammatical. It thus appears that the longer the ambiguous portion of a garden-path sentence, the harder it is for subjects to reanalyze the sentence.

Why would length have this effect? Frazier and Rayner (1982) proposed that, with a longer ambiguous region, a syntactic misanalysis will have greater consequences for a sentence's semantic interpretation. They assume that the semantic analysis of a sentence lags slightly behind its syntactic analysis—for sake of argument, say one word

(Frazier and Rayner do not make this specific claim). Thus, with a two-word ambiguous region, only one word would have been semantically interpreted before reanalysis is initiated. But with a four-word region, three words would have been semantically interpreted before receipt of the disambiguating word. Thus, semantic interpretation would be further along with a long ambiguous region compared to a short one. Their hypothesis, then, is that the more a phrase has been semantically interpreted, the harder it is for the parser to reanalyze that phrase.

Warner and Glass (1987) suggest another possibility. They point out that as words are added to a phrase, syntactic nodes will have to be added as well. Their hypothesis is that reanalysis with a long ambiguous region is hard because more syntactic nodes need to be restructured. According to this syntactic hypothesis, it is not the extra words themselves that make reanalysis difficult. Region length affects ease of reanalysis indirectly, because more words necessitate extra syntactic nodes. To test this hypothesis, it is necessary to vary the number of words and the number of nodes independently. Unfortunately, it is impossible to add words to a phrase without adding nodes. However, it is possible to create two phrases with the same number of words but differing in the number of nodes. The Warner and Glass hypothesis predicts that reanalysis will be more difficult with a syntactically denser phrase, even with length held constant.

We conducted five experiments to examine the effects of varying the characteristics of the ambiguous region, using grammaticality judgments as our dependent measure. In the first two experiments, we extended the finding that the longer the ambiguous region of an early closure sentence, the more difficult its reanalysis, using a more homogeneous and carefully controlled set of materials than the ones used previously by Frazier and Rayner (1982) and Warner and Glass (1987) and eliminating some problems with the stimulus materials used by Warner and Glass. In the third, fourth, and fifth experiments, we varied the syntactic characteristics of the ambiguous region in order to determine whether it is the number of words in the ambiguous region that affects ease of reanalysis, the syntactic complexity of the ambiguous region, or yet another possibility, namely, the distance between the disambiguating word and the head of the phrase that has been syntactically misanalyzed.

We consider the grammaticality judgment task a suitable one for examining reanalysis. But to use such a task, it is important to be clear about the relationship between ease of reanalysis and a subject's tendency to label a sentence as ungrammatical. We assume that each subject sets a subjective criterion for labelling a sentence as ungrammatical and that a number of factors will affect where this criterion is set. If a sentence is "objectively" ungrammatical (e.g., a sentence like *Boy the danced those car*), subjects will have no trouble categorizing it as such; similarly, subjects will have little trouble with a simple grammatical sentence (e.g., *The boy danced all night*). However, not all sentences that are grammatical will be appropriately categorized. In particular, garden-path sentences will sometimes be labelled as ungrammatical even though a legitimate analysis is possible for the string, because subjects eventually terminate their search for a grammatical structure. To account for this time constraint, we also assume that subjects set a limit on how long they are willing to spend trying to find the right syntactic structure. If that time limit passes without success, the sentence will be labelled ungrammatical. The more difficult reanalysis is, the more likely the time limit will pass, and the more likely the sentence is to be labelled ungrammatical.

EXPERIMENT 1

In the first experiment, we examined the

effect of ambiguous region length on ease of reanalysis. This issue had been examined previously by Warner and Glass (1987), but their study had two problems that we sought to correct here. We will illustrate the problems with sentences (3) and (4), used in their experiment. (Sentences did not appear with commas in their study; they are included here for ease of exposition.)

(3) Early closure sentences:
   a. When the boys strike, *the dog* kills.
   b. Before the boy kills, *the man the dog bites* strikes.

(4) Late closure sentences:
   a. After the dog bites *the man*, the cat kills.
   b. When the horse kicks *the boy*, the dog bites the man.

Sentences (3) and (4) differ in closure: the (3) sentences are early closure, and the (4) sentences, late closure. The (a) versus (b) dimension concerns length: in (a), a critical part of the sentence is short, and in (b), it is long. Warner and Glass presented items such as (3) and (4) word-by-word using the Rapid Serial Visual Presentation (RSVP) technique and had subjects make grammaticality judgments at the end of each sentence. They found that the short early closure sentences, the short late closure sentences, and the long late closure sentences were almost always judged grammatical; but fewer than half of the long early closure sentences were considered grammatical. They concluded that length caused the parser difficulty only in the early closure condition.

The first problem with this experiment is that the late closure sentences in (4) are not comparable in terms of length of the ambiguous region to the early closure sentences in (3). To see this point, examine the trees in Fig. 1. For the early closure sentences (panels A and B in the figure), the ambiguous region is *the dog* for the (a) version and *the man the dog bites* for the (b) version. This pair is well constructed, because the ambiguous noun phrase (the NP capitalized in the figure) is short in A and long in B. But for the late closure sentences (C and D in the figure), the noun phrases constituting the ambiguous region are equally long in panels C and D. The extra length in D occurs at the end of the sentence (the NP *the man*), by which point the late versus early closure decision has been made. Thus, the

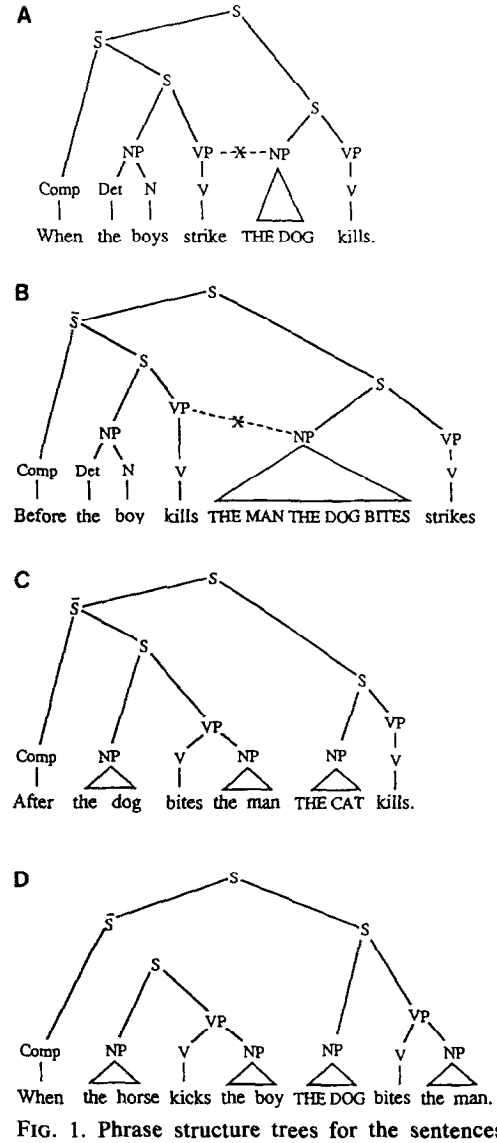

FIG. 1. Phrase structure trees for the sentences used by Warner and Glass (1987). The critical region is in capital letters. Tree A is early closure, short region; B is early closure, long region; C is late closure, short region; and D is late closure, "long" region.

two late closure sentences are no different in length, given the relevant decisions the parser has to make, and so the pattern of data Warner and Glass obtained with these two sentences does not speak to the issue of how the length of the ambiguous region of a sentence affects ease of reanalysis.

The second problem (noted also by Warner and Glass) is that the ambiguous region in panel B of Fig. 1 is not only long, it is also itself temporarily ambiguous. (To avoid confusion, we will refer to the ambiguous region following the main verb as the "critical" region in the first and second experiments, because both these experiments examine the issue of ambiguity within the ambiguous region.) Thus, the effect with this region obtained by Warner and Glass may not be due to its length, but rather to having two syntactic ambiguities in a sentence, one embedded within the other. To evaluate whether the effect is due to length or to ambiguity, it is necessary to compare (3b) to a sentence such as (5), where the critical region is long but unambiguous due to the presence of the complementizer *that*.[1]

(5) Before the boy kills, the man that the dog bites strikes.

Thus, our first step in investigating syntactic reanalysis was to examine the effect of increasing the length of the critical region, eliminating the problems we have outlined.

Method

Subjects. Eighteen University of Alberta undergraduates were paid $5 per hour for participating in the experiment. All subjects were native speakers of Canadian English and were unaware of the purposes of the experiment.

TABLE 1
SAMPLE ITEM, EXPERIMENTS 1 AND 2

Early closure versions
Short region
After the Martians invaded THE TOWN was evacuated.
Long, unambiguous region
After the Martians invaded THE TOWN THAT THE CITY BORDERED was evacuated.
Long, ambiguous region
After the Martians invaded THE TOWN THE CITY BORDERED was evacuated.

Late closure versions
Short region
After the Martians invaded THE TOWN the people were evacuated.
Long, unambiguous region
After the Martians invaded THE TOWN THAT THE CITY BORDERED the people were evacuated.
Long, ambiguous region
After the Martians invaded THE TOWN THE CITY BORDERED the people were evacuated.

Materials. Each experimental sentence appeared in one of the six versions illustrated in Table 1. The first three sentences in the table are early closure; the second three are late closure. The early versus late closure versions were constructed so that they differed only in the addition of the extra noun phrase for the late closure sentences (the phrase *the people* in the examples shown in Table 1).

Within each closure condition, the characteristics of the critical region were varied. The region was either short (e.g., *the town*), long and unambiguous (e.g., *the town that the city bordered*), or long and ambiguous (*the town the city bordered*). The latter two conditions differed in ambiguity because of the presence or absence of the complementizer *that*. With the complementizer, the phrase is unambiguously marked as a relative clause.

Six versions of 36 sentences were constructed. A list of all the experimental sentences can be obtained from either author. Each subject saw only one version of any

[1] It is perhaps more accurate to say that sentence (5) is *less* ambiguous than (3b). The sequence *Before the dog kills the man* could be interpreted as one sentence, followed by a new sentence containing a sentential subject, as in *that the dog bites becomes important.* However, given the rarity of the sentential subject construction, we will continue to refer to sentences such as (5) as "unambiguous."

one sentence, but saw all six conditions of the experiment. Across subjects, each version of each sentence was seen an equal number of times. The sentences were embedded in a list with 72 filler items. Half of these filler items were grammatical, and the other half were ungrammatical. The ungrammatical sentences were sentences such as *The player hit the ball the man who saw it* (i.e., typically missing obligatory constituents). The remaining grammatical sentences were of a variety of sentence types; none were syntactically ambiguous. The order of presentation of the items was randomized for each subject.

Procedure. Subjects were first shown a few examples of grammatical and ungrammatical sentences. The experimenter discussed this distinction with the subject, taking care to ensure that the subject understood the intended meaning of "grammatical." For instance, a sentence with a dangling preposition (e.g., *Who did the girl talk to?*) was given as an example of a grammatical sentence. Subjects were told to judge a sentence as ungrammatical if it seemed uninterpretable due to the absence of necessary words or the presence of too many words.

Subjects were then seated across from a CRT, with a button panel in front of them. Before each sentence, the message "Press any button for the next sentence" appeared on the screen. When the subject pressed a button, a fixation cross appeared in the center of the screen for 1000 ms. Thereafter, the first word of the sentence replaced the fixation cross at the center of the screen, the second word replaced the first word in the same location, and so on, until the end of the sentence. Each word was displayed for 250 ms with no interstimulus interval. After the last word, subjects made their grammaticality decision by pressing one response button with the left index finger if they judged the sentence to be ungrammatical and another response button with the right index finger if they judged the sentence to be grammatical. Decisions and decision times were automatically recorded. Subjects were encouraged to make their decisions as quickly and accurately as possible.

Design and analysis. The experiment was analyzed as a 2 × 3 factorial. The first variable was closure (early vs. late). The second variable concerns the critical region, which was either short (e.g., *the town*), long and unambiguous (e.g., *the town that the city bordered*), or long and ambiguous (*the town the city bordered*).

Results

The percentage of times that subjects judged the sentences to be grammatical is shown in Fig. 2. Grammaticality judgment times were not significant on any analysis (average judgment time was 1346 ms) and will not be discussed further (see Warner & Glass, 1987, for a similar lack of effects on judgment times). Grammaticality judgments were analyzed with both subjects (F_1) and items (F_2) as random effects.

As Fig. 2 illustrates, there was a main effect of closure such that early closure sentences were judged grammatical less often than late closure sentences (37% vs. 59%), $F_1(1,17) = 19.32, p < .01$; $F_2(1,35) = 55.90, p < .01$. The region variable was also significant, $F_1(2,34) = 70.68$, $p < .01$;

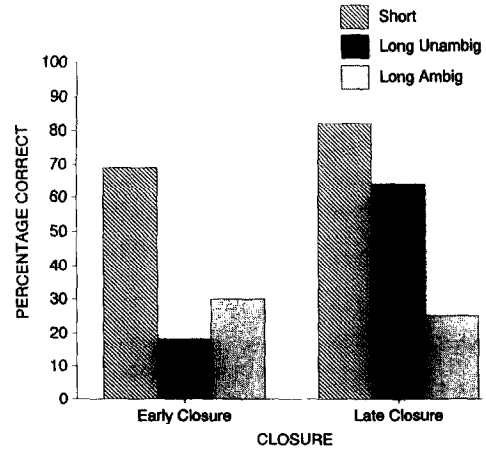

FIG. 2. Percentage of sentences correctly judged to be grammatical, Experiment 1.

$F_2(2,70) = 62.18$, $p < .01$. Sentences with short regions were judged grammatical more often than those with long unambiguous regions, which in turn were judged grammatical more often than sentences with long ambiguous regions (76%, 41%, and 27%, respectively).

In addition, there was a reliable interaction between these two variables, $F_1(2,34) = 12.41$, $p < .01$; $F_2(2,70) = 16.24$, $p < .01$. One way to understand this interaction is to consider just the following four points in Fig. 2: early closure short region, late closure short region, early closure long unambiguous region, and late closure long unambiguous region. Length uncontaminated by internal ambiguity affected both late and early closure sentences. For late closure sentences, the means for the short (82%) and the long unambiguous (64%) conditions differed significantly from each other, $F_1(1,17) = 9.49$, $p < .01$; $F_2(1,35) = 11.14$, $p < .01$. For the early closure sentences, these two conditions also differed reliably (69% vs. 18%), $F_1(1,17) = 93.32$, $p < .01$; $F_2(1,35) = 74.57$, $p < .01$. However, the effect of length was much greater for the early closure than for the late closure sentences, and this 2×2 interaction was reliable, $F_1(1,17) = 20.40$, $p < .01$; $F_2(1,35) = 21.71$, $p < .01$.

The early and late closure sentences containing a long, ambiguous critical region were judged grammatical equally often (30% vs. 25%), both Fs < 1. Finally, considering just the two short conditions, the early closure sentence was judged grammatical less often than the late closure sentence, $F_1(1,17) = 4.98$, $p < .05$; $F_2(1,35) = 5.52$, $p < .05$.

The grammatical filler sentences were judged grammatical 95% of the time; the ungrammatical fillers were judged grammatical 10% of the time.

Discussion

The results of this experiment differ from those obtained by Warner and Glass (1987). The first major difference is that here the length of the ambiguous region affected both early and late closure sentences. However, the effect of length was much greater for the early closure versions. One way to view this finding, which is consistent with the spirit of Warner and Glass's conclusions, is to claim that the length effect is qualitatively different for the two closure conditions. For the late closure sentences, it is possible that length did not affect syntactic processing. Instead, because the sentence is longer overall when the critical region is lengthened, there is a greater likelihood that subjects will occasionally miss a word of the sentence and, consequently, will more often misunderstand the sentence, compared to when the critical region is short. For the early closure sentences, length does affect syntactic processing. On encountering the disambiguating word of the sentence, the parser detects that it has misanalyzed the preceding string, and the longer the misanalyzed string, the more difficulty the parser has recovering from this misanalysis.

Another difference between our results and those obtained by Warner and Glass concerns the effect of making the critical region ambiguous. When a sentence is as difficult as it is in the two long ambiguous conditions, the early and late closure versions are successfully parsed on fewer than one-third of the trials, and the two closure types do not differ from each other. This finding is not surprising, however, because the existence of an ambiguity within the critical region invokes the minimal attachment principle (Frazier, 1978, 1987) as well as late closure. The minimal attachment principle states that words are syntactically organized into a phrase structure tree using as few syntactic nodes as possible, consistent with the well-formedness rules of the language. To see the operation of minimal attachment within the ambiguous critical region, consider what happens as the parser processes the late closure sentence *After the Martians invaded the town the city bordered the people were evacuated.* The

parser would first assume that the phrase *the town* is a simple noun phrase, due to the operation of minimal attachment. It would then attach the phrase as a direct object of *invaded,* by late closure. Next, the sequence *the city bordered the people* would be taken as a complete clause. Now the parser encounters the word *were,* signalling that the previous two clauses have to be somehow revised. Thus, even in the late closure version, the sentences containing a long, ambiguous critical region can be expected to cause the parser enormous trouble.

Finally, in contrast with Warner and Glass, we found that even with short critical regions, late closure sentences were easier than early closure sentences. However, the difference between our experiment and theirs is difficult to evaluate. One would expect that, during reanalysis, a wide variety of information sources would be used by the parser to attempt to discover the correct analysis (Frazier, 1987; Rayner, Carlson, & Frazier, 1983), including information such as verb transitivity and the overall pragmatic biases of the sentences. Because we did not use the same materials as Warner and Glass, it is possible that our materials differed enough on these dimensions that reanalysis, even in the short conditions, was harder in our experiments than in theirs.

The important finding from this experiment is that the length of the critical region affected the ease with which subjects could successfully parse garden-path sentences. Length was costly for both early and late closure sentences; however, it was far more costly for the former. In the second experiment, we replicated the first experiment with one change: Subjects read the sentences at their own pace rather than with RSVP. As in Experiment 1, subjects made a grammaticality decision after each sentence. Our purpose in conducting the second experiment was to attempt to increase the percentage of sentences judged to be grammatical. We were particularly concerned about the late closure sentences with a long, unambiguous region, because the grammaticality judgments were lower than we had expected. We reasoned that by allowing subjects to take as much time as they needed to read the experimental sentences, we would elevate grammaticality judgments, particularly in this one condition.

Experiment 2

Experiment 2 was similar to Experiment 1, except that subjects read the sentences at their own pace. At the end of each sentence, subjects were required to make a grammaticality judgment.

Method

Subjects. Eighteen subjects from the same pool as in Experiment 1 participated in the experiment.

Materials. The same materials were used here as in Experiment 1.

Procedure. The procedure was self-paced segment-by-segment reading with a cumulative display. Subjects read sentences one segment at a time. Segments either contained only a single content word, or a content word together with a preceding article. For example, the early closure sentence with a short region shown in Table 1 would be segmented as: *After / the Martians / invaded / the town / was / evacuated.*

On each trial, a single sentence appeared on the computer screen, but with an underline in place of every letter. When subjects were ready to read the sentence, they pushed a button, and the first segment became visible. After the subjects had read and understood the segment, they pushed the button for the next segment, and the previous one remained on the screen. Subjects proceeded in this fashion until the last segment of the sentence had been read. After the final segment, subjects were prompted to make their grammaticality judgments. Subjects were told to push one button if the sentence was grammatical, and to push a different button (using their

other hand) if the sentence was ungrammatical. As in Experiment 1, subjects were instructed before the experiment on the relevant sense of the term "grammatical."

Results

No significant effects were obtained for reading times on the final segment (average time spent on the final segment was 7002 ms). The lack of an effect for this final segment is probably due to our use of cumulative display, a presentation method that has been shown to be insensitive to online parsing processes (Ferreira & Henderson, 1990; Kennedy & Murray, 1984) as well as general comprehension processes (Just, Carpenter, & Woolley, 1982; Mitchell, 1984). We did not obtain significant effects for grammaticality judgment times either, and average judgment time was 505 ms. Given that this value is so much lower than that obtained in Experiment 1 (and in the remaining experiments) and that the time to read the final segment was so high, it is quite likely that subjects made their grammaticality decision while reading the final segment and then executed that decision rapidly once the cue was presented.

Figure 3 gives the mean percentage of sentences judged grammatical for the six conditions of the experiment. Early closure sentences were judged grammatical less often than late closure sentences (38% vs. 69%), $F_1(1,17) = 43.38, p < .01$; $F_2(1,35) = 59.89, P < .01$. In addition, there was a main effect of region type, $F_1(2,34) = 57.01, p < .01$; $F_2(2,70) = 85.96, p < .01$. Sentences with short regions were judged grammatical 80% of the time, with long unambiguous regions 48% of the time, and with long ambiguous regions 32% of the time.

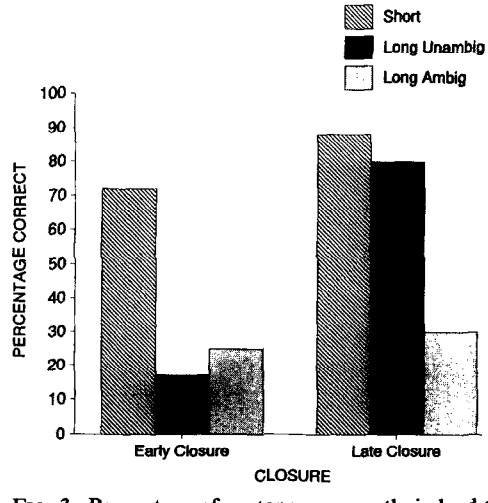

FIG. 3. Percentage of sentences correctly judged to be grammatical, Experiment 2.

The two variables interacted significantly, $F_1(2,34) = 27.05, p < .01$; $F_2(2,70) = 22.00, p < .01$. First, we will consider the pattern among the following four points in Fig. 3: early and late closure with a short region, and early and late closure with a long, unambiguous region. This 2 × 2 interaction was significant, indicating that the effect of length was smaller for the late closure than the early closure sentences, $F_1(1,17) = 39.33$; $p < .01$, $F_2(1,35) = 32.13$, $p < .01$. Specifically, for late closure, sentences with short critical regions were judged grammatical almost as often as those with long critical regions (88% vs. 80%), $F_1(1,17) = 3.40, p > .05$; $F_2(1,35) = 3.45, p > .05$. In contrast, for early closure, sentences with short critical regions were judged grammatical far more often than those with long critical regions (72% vs. 17%), $F_1(1,17) = 106.40, p < .01$; $F_2(1,35) = 76.12, p < .01$.

Early and late closure sentences containing a long, ambiguous critical region were judged grammatical equally often (25% vs. 30%), both Fs < 1. Early closure sentences with a short critical region were judged grammatical less often than their late closure counterparts (72% vs. 88%), $F_1(1,17) = 7.39, p < .05$; $F_2(1,35) = 5.98, p < .05$.

Grammatical filler sentences were judged grammatical 95% of the time; ungrammatical fillers were judged grammatical 4% of the time.

Discussion

The purpose of this experiment was to give subjects extra time to view the sen-

tences, in order to elevate the percentage of sentences judged to be grammatical. However, the extra time was of virtually no benefit for the early closure sentences; the results from the second experiment for the early closure items perfectly replicate those of the first. Allowing subjects to process the early closure sentences in a more leisurely fashion did not make it any more likely that subjects would successfully reanalyze those sentences. The additional time was of benefit, however, for the late closure sentences. In each of the three conditions, late closure sentences were judged grammatical more often in the second experiment than in the first. These results suggest that the region length manipulation has different consequences for late compared to early closure sentences. For late closure sentences, greater length simply increases the likelihood that subjects will fail to process a word, and thus that they will fail to comprehend the sentence; for early closure sentences, length exaggerates the garden-path, and even giving subjects more time in which to view the sentence has little effect.

In summary, the conclusion from the first two experiments is that increasing the length of the critical region in garden-path sentences increases the difficulty of reanalysis. The next question to be addressed is, why? One possibility is that the phenomenon is attributable directly to the increase in number of words. A second possibility is that the effect is only indirectly related to the increased word count and that the actual factor responsible for increasing the difficulty of reanalysis is the increase in the phrase's syntactic complexity, which is generally correlated with increased region length (Warner & Glass, 1987). The third experiment was designed to decide between these two possibilities.

Experiment 3

The purpose of the third experiment was to test the syntactic difficulty hypothesis proposed by Warner and Glass (1987). According to this hypothesis, it is not the addition of extra words that causes reanalysis to be difficult with a long region, but instead, the increase in syntactic complexity correlated with greater region length. We compared sentences in which the ambiguous regions were equated on number of words, but differed in syntactic complexity. The syntactic difficulty hypothesis predicts that reanalysis will be more difficult with the more complex region, even with region length held constant.

Consider, for example, the phrases *the birds that cheetahs eat* and *the bird with bright plumage*. In the former, length is created by adding a relative clause, and in the latter, by adding a prepositional phrase. The two phrases differ in number of syntactic nodes, as shown in Fig. 4, but have the same number of words (five). The phrase containing a relative clause requires seven nonterminal nodes; the one containing a prepositional phrase requires four. We assume that the more nodes in a phrase, the greater its syntactic complexity. The number of nodes has been shown to reflect underlying syntactic complexity: Ferreira

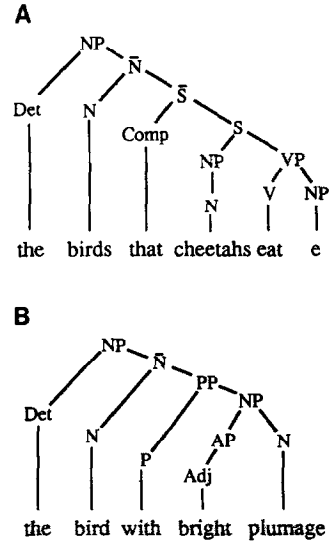

FIG. 4. Phrase structure trees for the ambiguous phrases used in Experiment 3. Tree A shows the structure for the phrase containing a modifying relative clause; B shows the structure for the phrase containing a modifying prepositional phrase.

(1991) found that the greater the number of nodes in a phrase a speaker was about to produce, the longer the initiation time for that phrase.

We compared early and late closure sentences containing three different types of ambiguous regions: a short region, a long region containing a relative clause, and a long region containing a prepositional phrase. If syntactic complexity causes reanalysis to be difficult, then the two long conditions should differ from one another, because they differ in complexity. If it is number of words in the ambiguous region that is critical, then the two long conditions should not differ, because they do not differ in number of words.

Method

Subjects. Twenty-four subjects from the University of Alberta subject pool participated in the experiment.

Materials. Early and late closure sentences were created with three different types of ambiguous regions. The six conditions of the experiment are shown in Table 2. The ambiguous region was either short (*the birds*), long due to the inclusion of a relative clause after the head noun (*the birds that cheetahs eat*), or long due to the inclusion of a prepositional phrase after the head noun (*the birds with bright plumage*). We constructed 36 sentences like the ones shown in Table 2. Each subject saw only one version of any one sentence, but each saw all six conditions of the experiment. Across subjects, each version of the sentences was seen an equal number of times. The same filler items were used as in Experiments 1 and 2. Presentation of the items was randomized for each subject.

TABLE 2
SAMPLE ITEM, EXPERIMENT 3

Early closure versions
Short region
When the men hunt THE BIRDS typically scatter.
Long region, relative clause
When the men hunt THE BIRDS THAT CHEETAHS EAT typically scatter.
Long region, prepositional phrase
When the men hunt THE BIRDS WITH BRIGHT PLUMAGE typically scatter.
Late closure versions
Short region
When the men hunt THE BIRDS the deer typically scatter.
Long region, relative clause
When the men hunt THE BIRDS THAT CHEETAHS EAT the deer typically scatter.
Long region, prepositional phrase
When the men hunt THE BIRDS WITH BRIGHT PLUMAGE the deer typically scatter.

Procedure. The sentences were presented in the same manner as in Experiment 1: Sentences were presented using RSVP, 250 ms per word, and subjects made a grammaticality judgment after each sentence.

Results

No significant effects were obtained with grammaticality judgment times (average time was 1185 ms). Figure 5 shows the percentage of sentences judged grammatical in each of the six conditions. Overall, early closure sentences were judged grammatical less often than late closure sentences (44% vs. 73%), $F_1(1,23) = 16.68$, $p < .01$; $F_2(1,24) = 17.55$, $p < .01$. The main effect

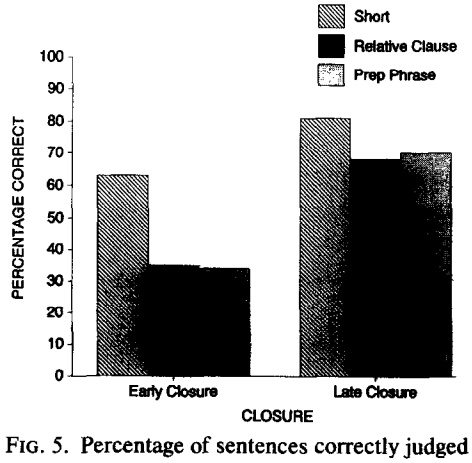

FIG. 5. Percentage of sentences correctly judged to be grammatical, Experiment 3.

of the region was also significant, $F_1(1,23) = 38.17, p < .01; F_2(1,23) = 31.75, p < .01$. Sentences with short regions were judged grammatical 72% of the time, those with long, relative clause regions were judged grammatical 52% of the time, and those with long, prepositional phrase regions were also judged grammatical 52% of the time.

In addition, the interaction between closure and region type was significant, $F_1(1,23) = 3.82, p < .05; F_2(1,23) = 4.48, p < .05$. First, consider just the late closure sentences. In the short region condition, sentences were judged grammatical 81% of the time; in the long, relative clause condition, sentences were judged grammatical 68% of the time; and in the long, prepositional phrase condition, sentences were judged grammatical 70% of the time. The difference among these three conditions was significant, $F_1(2,46) = 5.16, p < .01$, $F_2(2,70) = 3.89, p < .05$. The first two means differed reliably, $F_1(1,23) = 15.72, p < .01$, $F_2(1,35) = 7.40, p < .01$, but the relative clause and prepositional phrase conditions were equivalent, both Fs < 1.

A similar pattern held for the early closure sentences. The differences among the three region conditions (63%, 35%, and 34% for the short, long relative clause, and long prepositional phrase conditions, respectively) was significant, $F_1(1,23) = 14.42, p < .01$; $F_2(1,35) = 22.19, p < .01$. The first two conditions differed reliably, $F_1(1,23) = 22.12, p < .01$; $F_2(1,35) = 37.67, p < .01$, while the relative clause and prepositional phrase conditions did not, both Fs < 1. As in the previous two experiments, the interaction between the closure and region type variables is due to the fact that length was more costly for the early closure sentences than the late closure sentences.

Grammatical filler sentences were judged grammatical 94% of the time; ungrammatical fillers were judged grammatical 12% of the time.

Discussion

As in the previous experiments, early closure sentences were associated with worse performance than late closure sentences. In addition, we replicated the finding that the longer the ambiguous region of the early closure sentences, the harder reanalysis becomes—performance in the long, relative clause and the long, prepositional phrase conditions was worse than in the short condition. The purpose of this experiment was to investigate the reason for the effect of length. Is difficulty of reanalysis with a long ambiguous region due to the syntactic complexity of the constituent that must be reanalyzed, or the number of words in that constituent? The results were clear-cut: The relative clause and prepositional phrase conditions produced similar levels of performance even though they differed in syntactic complexity.

In summary, syntactic complexity does not affect difficulty of reanalysis. Instead, it appears that the more words in the ambiguous region of a garden-path sentence, the harder reanalysis is. (We use the word "appears" because, as Experiment 4 will show, this statement must be qualified.) Next, we turn to the question: Why does extending the length of the ambiguous region have this effect?

EXPERIMENT 4

Up to this point, we have demonstrated that garden-path sentences with long ambiguous regions are harder to reanalyze than ones with short regions. Furthermore, contrary to the hypothesis proposed by Warner and Glass (1987), this difficulty is not attributable to the greater syntactic complexity of the longer phrases. It appears that something about making the ambiguous region longer causes reanalysis to be difficult. In this experiment, we will investigate one possible reason that length may have this effect.

Consider the three early closure sen-

tences in Table 2. In the short condition, the ambiguous region consisted of a determiner and a noun (*the birds*). The word immediately following the head noun of the ambiguous phrase was the disambiguating word of the sentence, the word which signalled that the ongoing analysis of the sentence was not correct. Compare the short condition to the sentences in the two long conditions (which, recall from the third experiment, were equivalent in difficulty). With these sentences, the head noun was separated by three words from the disambiguating word. One possibility is that the greater the distance between the disambiguating word and the head of the phrase that has been misanalyzed, the harder reanalysis will be. According to this hypothesis, it is not length in and of itself that makes reanalysis difficult. Instead, it is the number of words between the head noun of the misanalyzed phrase and the disambiguating word that matters. To test this hypothesis, it is necessary to compare ambiguous phrases containing the same numbers of words, but differing in the location of their heads. Compare, for example, the following two sentences:

(6) a. While the boy scratched *the big and hairy dog* yawned loudly.
b. While the boy scratched *the dog that Sally hates* yawned loudly.

The two ambiguous regions have the same number of words, but differ in the location of the head of the phrase. In (a), the head of the noun phrase, the noun *dog,* is adjacent to the disambiguating word. By the head location hypothesis, the condition should be easy, in fact about as easy as a condition in which the ambiguous region were simply the phrase *the dog*. In (b), the head of the noun phrase is separated from the disambiguating word by three words, and so should be difficult.

Why might ease of reanalysis depend on the location of the head of the misanalyzed phrase with respect to the disambiguating word? One possibility is that once the head noun of a phrase has been encountered, the entire phrase can be assigned a thematic role—a semantic role such as agent, patient, recipient, and so on (Rayner et al., 1983). In a sentence such as (a), the phrase *the big and hairy dog* would be assigned the role of patient, because that phrase is assumed to be the direct object of the verb *scratched.* The parser next encounters the disambiguating word, which signals not only that the ambiguous phrase must be reanalyzed as the subject of an embedded clause, but also that the phrase must be assigned a different thematic role, one of agent. In a sentence such as (b), the thematic role of patient would be assigned at the head noun *dog,* and then that thematic role assignment would be maintained for three words before the parser encountered the disambiguating word. The longer the sentence comprehension system has been committed to one thematic role assignment, the harder it may be to give it up.

This hypothesis is a specific version of the hypothesis offered by Frazier and Rayner (1982). Recall their proposal that reanalysis with long regions is hard because semantic analysis of the sentence is further along. According to the present hypothesis, the semantic analysis that is at issue is the assignment of semantic/thematic roles to syntactic phrases. Our notion of "further along" is that a role is assigned to a phrase at its head, and the longer that role has been assigned to a phrase, the harder it is to assign the phrase a different role.

Yet another possibility is that ease of reanalysis depends on the location of the head of the misanalyzed phrase because phrases are attached to the ongoing phrase structure tree when their heads are encountered. Thus, on this hypothesis, the phrase *the big and hairy dog* would be internally analyzed up to the head noun *dog,* and only at the head would the phrase be attached (incorrectly) to the phrase marker as a direct object. Upon receipt of the disambigu-

ating word *yawned,* this attachment would be revised. Because of the recency of the attachment, revision would be quite easy. In contrast, a phrase such as *the dog that Sally hates* would be attached at the head noun, *dog*. Because the head occurs phrase-medially, the erroneous attachment would linger for three additional words before the parser encountered the disambiguating word. The longer an attachment has been made, the harder it may be to detach the phrase and move it to a new location in the phrase structure tree.

There may be other reasons that head location affects ease of reanalysis. We have simply outlined two hypotheses that strike us as plausible, but we will not attempt to distinguish between them. Our purpose in the present experiment is to establish whether the distance from the head noun of the misanalyzed phrase to the disambiguating word affects the ease with which subjects can recover from a garden-path.

Method

Subjects. Eighteen subjects from the University of Alberta subject pool participated in the experiment.

Materials. As in the previous three experiments, we contrasted early and late closure sentences. In addition, characteristics of the ambiguous region were varied so that the region was either (a) short, (b) long due to the inclusion of a relative clause, so that the head of the phrase occurred in the middle of that phrase, or (c) long due to the inclusion of prenominal adjectives, so that the head was phrase-final. A sample item is shown in Table 3.

Six versions of 36 sentences were constructed. Each subject saw only one version of any one sentence, but saw all six conditions of the experiment. Across subjects, each version of each sentence was seen an equal number of times. The same filler items were used as in the previous experiments. Presentation of the items was randomized for each subject.

TABLE 3
SAMPLE ITEM, EXPERIMENT 4

Early closure versions
Short region
While the boy scratched THE DOG yawned loudly.
Long region, relative clause
While the boy scratched THE DOG THAT SALLY HATES yawned loudly.
Long region, prenominal adjectives
While the boy scratched THE BIG AND HAIRY DOG yawned loudly.
Late closure versions
Short region
While the boy scratched THE DOG the girl yawned loudly.
Long region, relative clause
While the boy scratched THE DOG THAT SALLY HATES the girl yawned loudly.
Long region, prenominal adjectives
While the boy scratched THE BIG AND HAIRY DOG the girl yawned loudly.

Procedure. The procedure was the same as in the previous experiment.

Results

Again, no significant effects were obtained with grammaticality judgment times (average time was 1726 ms). Figure 6 shows the percentage of sentences judged grammatical in each of the six conditions. As the

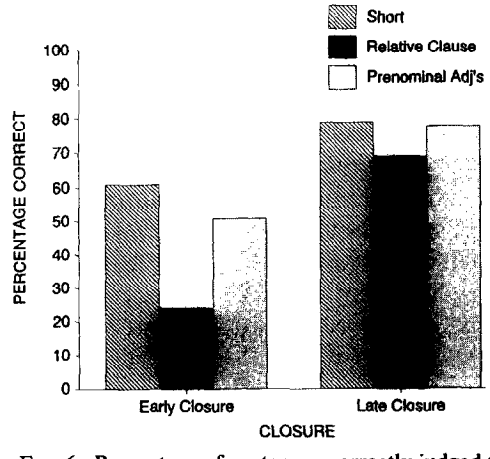

FIG. 6. Percentage of sentences correctly judged to be grammatical, Experiment 4.

figure reveals, there was a main effect of closure such that early closure sentences were judged grammatical less often than late closure sentences (45% vs. 75%), $F_1(1,17) = 58.30$, $p < .01$; $F_2(1,35) = 43.22$, $p < .01$. The main effect of the region was also significant, $F_1(2,34) = 16.40$, $p < .01$; $F_2(2,70) = 15.63$, $p < .01$. Sentences with short ambiguous regions were judged grammatical 70% of the time, those that were long due to the inclusion of a relative clause were judged grammatical 47% of the time, and those that were long due to the inclusion of adjectives were judged grammatical 64% of the time.

These two factors interacted significantly, $F_1(2,34) = 5.23$, $p < .05$; $F_2(2,70) = 3.64$, $p < .05$. For the late closure sentences, the region variable had no significant effect across all three conditions, $F_1(2,34) = 1.78$, $p > .15$; $F_2(2,70) = 1.14$, $p > .30$. The percentage judgments were 79% in the short condition, 69% in the long, relative clause condition, and 78% in the long, adjective condition.

For the early closure sentences, however, the region variable did have a significant effect, $F_1(2,34) = 16.36$, $p < .01$; $F_2(2,70) = 15.68$, $p < .01$. In the short condition, sentences were judged grammatical 61% of the time; in the long, relative clause condition, 24% of the time; and in the long, adjective condition, 51% of the time. The values of 61% and 51% did not differ significantly from one another, $F_1(1,17) = 1.77$, $p > .20$; $F_2(1,35) = 3.12$, $p > .05$, but the values of 51% vs. 24% did, $F_1(1,17) = 16.44$, $p < .01$; $F_2(1,35) = 14.18$, $p < .01$. Thus, the short and long, adjective conditions were equivalent, and performance was significantly better than in the long, relative clause condition.

Examining just the short region condition, early closure sentences were judged grammatical significantly less often than late closure sentences, $F_1(1,17) = 6.96$, $p < .05$; $F_2(1,35) = 8.16$, $p < .01$.

The grammatical fillers were judged to be grammatical 96% of the time; ungrammatical fillers were judged grammatical 11% of the time.

Discussion

The results of this experiment indicate that what exacerbates a garden-path is not simply adding more words to a constituent. The two long conditions (adjective and relative clause) were equivalent in length but differed in the location of the head of the misanalyzed phrase: In the adjective condition the head occurred early in the phrase, and in the relative clause condition, the head occurred late in the phrase. The finding that performance was worse in the relative clause condition suggests that the location of the head affects reanalysis processes. Specifically, the further the head of the misanalyzed phrase is from the disambiguating word, the harder it is to reanalyze that phrase. To strengthen this conclusion, we conducted one further experiment in which we minimized as much as possible the syntactic and semantic differences between the adjective and relative clause conditions.

Experiment 5

From the results of Experiment 4 we wish to conclude that the distance between the head of the ambiguous region and the disambiguating word affects the difficulty of reanalysis. This conclusion is based on the finding that performance with an ambiguous region such as *the big and hairy dog* was better than with an ambiguous region such as *the dog that Sally hates*. However, it is possible to counter that these phrases do not differ only in the location of the head; they also differ in the number of nouns they contain (one in the adjective version, two in the relative clause version) and in their semantic content (the adjective version consists of the propositions that the dog is big and the dog is hairy, while the relative clause version consists of the proposition that Sally hates a certain dog).

In the present experiment, we sought to minimize these differences by comparing a region such as *the big and hairy dog* with *the dog that is hairy*. These phrases have the same number of nouns: Each contains one, the word *dog*. The phrases are also more alike in semantic content: The first phrase expresses the two propositions given above, and the second expresses the proposition that the dog is hairy. It was not possible to compare *the big and hairy dog* with *the dog that is big and hairy*, because the latter contains two more words than the former, and it is critical for our arguments that the regions not differ in length. Notice, however, that whatever semantic difference remains between the two conditions goes against our hypothesis: The adjective condition, which we predict to be the easier of the two, is propositionally more complex than the relative clause condition. If we obtain this predicted difference in performance, we will have strengthened our conclusion that head location affects ease of reanalysis.

TABLE 4
SAMPLE ITEM, EXPERIMENT 5

Early closure versions
Short region
While the boy scratched THE DOG yawned loudly.
Long region, relative clause
While the boy scratched THE DOG THAT IS HAIRY yawned loudly.
Long region, prenominal adjectives
While the boy scratched THE BIG AND HAIRY DOG yawned loudly.
Late closure versions
Short region
While the boy scratched THE DOG the girl yawned loudly.
Long region, relative clause
While the boy scratched THE DOG THAT IS HAIRY the girl yawned loudly.
Long region, prenominal adjectives
While the boy scratched THE BIG AND HAIRY DOG the girl yawned loudly.

Method

Subjects. Thirty-six subjects from the University of Alberta subject pool participated in the experiment.

Materials. In this experiment we again contrasted early and late closure sentences and varied the type of ambiguous region. The ambiguous region was either (a) short, (b) long due to the inclusion of a relative clause, so that the head of the phrase occurred in the middle of that phrase, or (c) long due to the inclusion of prenominal adjectives, so that the head was phrase-final. Unlike Experiment 4, however, the adjective and relative clause conditions were made as similar as possible, as shown in Table 4. We created the relative clause condition from the adjective condition by taking the longer of the two adjectives and placing it in a postnominal adjectival phrase (e.g., *the big and hairy dog* → *the dog that is hairy*). We chose the longer adjective so that the two conditions would remain equal in number of syllables.

We constructed six versions of 36 sentences by making the appropriate changes to each of the experimental items used in Experiment 4. Each subject saw only one version of any item, but saw all six conditions of the experiment. Across subjects, each version of each sentence was seen an equal number of times. The same filler items were used in the previous experiments. Presentation of the items was randomized for each subject.

Procedure. The procedure was the same as in the previous experiment.

Results

Figure 7 shows the percentage of sentences judged grammatical in each of the six conditions. Early closure sentences were judged grammatical less often than late closure sentences (47% vs. 74%), $F_1(1,35) = 95.90$, $p < .01$; $F_2(1,35) = 45.11$, $p < .01$. The main effect of region was also significant, $F_1(2,70) = 34.43$, $p < .01$; $F_2(2,70) = 30.38$, $p < .01$. Sentences with short ambiguous regions were judged

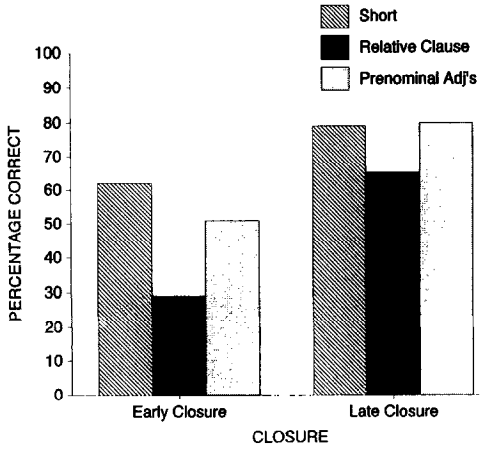

FIG. 7. Percentage of sentences correctly judged to be grammatical, Experiment 5.

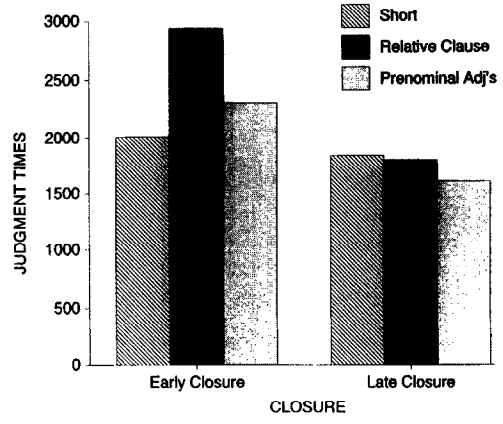

FIG. 8. Grammaticality judgment times in ms (correct responses only), Experiment 5.

grammatical 70% of the time, those that were long due to the inclusion of a relative clause were judged grammatical 47% of the time, and those that were long due to the inclusion of adjectives were judged grammatical 65% of the time.

As is evident in Fig. 7, the pattern of data replicates Experiment 4 (Fig. 6): The length manipulation was more costly for the early closure compared with the late closure sentences, $F_1(2,70) = 4.27$, $p < .05$; $F_2(2,70) = 4.51$, $p < .05$. Most importantly, a planned contrast showed that for the early closure sentences, the relative clause condition and the prenominal adjective condition differed significantly (29% vs. 51%), $F_1(1,35) = 17.94$, $p < .01$; $F_2(1,35) = 29.16$, $p < .01$.

Significant effects were also observed in the grammaticality judgment times, in contrast to the previous experiments. Figure 8 shows the mean correct judgment times in each of the six conditions. Early closure sentences were judged grammatical more slowly than late closure sentences (2416 vs. 1750 ms), $F_1(1,35) = 26.95$, $p < .01$; $F_2(1,35) = 48.03$, $p < .01$. There was also a marginal main effect of region, $F_1(2,70) = 8.30$, $p < .01$; $F_2(2,70) = 3.01$, $p = .05$. Grammaticality judgments for sentences with short ambiguous regions required 1879 ms, sentences with relative clause regions required 2252 ms, and those with adjectives required 2016 ms.

The closure and region factors interacted significantly, $F_1(2,70) = 7.49$, $p < .01$; $F_2(2,70) = 3.83$, $p < .05$. The pattern of this interaction was roughly the inverse of the percentage data, indicating that longer times were obtained in the same conditions as higher errors. Early closure sentences with a relative clause took subjects longer to judge grammatical than early closure sentences with prenominal adjectives (2939 vs. 2301 ms), though the effect was significant only by subjects, $F_1(1,35) = 8.56$, $p < .01$; $F_2(1,35) = 1.32$, $p > .05$.

The grammatical fillers were judged to be grammatical 92% of the time; ungrammatical fillers were judged grammatical 10% of the time.

Discussion

The purpose of this experiment was to reinforce our conclusion from Experiment 4 that head location affects ease of reanalysis. In the previous experiment, the ambiguous regions differed in more than just the placement of the head noun; they differed also in number of nouns and in their semantic content. In the present experiment, we equated the regions on number of nouns and made the regions as semantically similar as possible. Our finding that head position affects grammaticality judgments

(and judgment times), even with syllable length, region length, number of nouns, and overall semantic content equated, provides strong support for the hypothesis that the greater the distance from the head of the misanalyzed phrase to the syntactically disambiguating word, the harder the process of reanalysis becomes.

General Discussion

The purpose of this study was to investigate the factors that influence ease of syntactic reanalysis. That is, when the parser is garden-pathed, what determines how easily it can recover from its misanalysis? To investigate this question, we manipulated the form characteristics of the ambiguous phrase—its length, syntactic complexity, and head location. In the first and second experiments, we demonstrated that adding more words to the ambiguous region of garden-path sentences increased the difficulty of reanalysis. We then outlined three explanations of this effect: first, the effect could be due to adding the words themselves; second, it could be due to the increase in syntactic complexity that is typically correlated with increased word length; and third, it could be due to increasing the distance from the disambiguating word of the sentence to the head of the misanalyzed phrase. We found that the third explanation seems to be correct. Experiments 4 and 5 showed that it is not just adding words that affects reanalysis, because sentences equivalent in ambiguous region length differed in ease of reanalysis. Experiment 3 demonstrated that syntactic complexity does not affect ease of reanalysis, because sentences differing in syntactic complexity did not differ in ease of reanalysis. Experiments 4 and 5 showed that it is head location that matters. We compared two sentences with ambiguous phrases differing in head location, but equated on number of words. Our results show that reanalysis is harder when the head is further from the disambiguating word.

In the remainder of this section, we will discuss a model of reanalysis described in Ferreira and Henderson (1991), and show how it can account for the results of our five experiments. This model follows Rayner et al. (1983) in assuming the existence of (at least) two processors in sentence comprehension: a parser, whose task is to create a phrase structure tree for a sentence, and a thematic processor, whose job is to assign thematic roles to phrases. The thematic processor uses the lexical information stored with verbs to assign roles such as agent, patient, and theme to phrases. For example, the verb *put* assigns three thematic roles, as illustrated in the sentence below:

(7) Bill put the book on the shelf.

The phrase *Bill* is assigned the role of agent, *the book* is assigned the role of theme, and *on the shelf* is assigned the role of location.

Generally, *put* must assign these three roles to its arguments—that is, *put* has only one thematic structure. Other verbs, however, have more than one. For example, the verb *cheat* has at least two thematic structures: one where only the thematic role of agent is assigned to the subject of the sentence, as in *Mary cheated,* and one where agent is assigned to the subject and patient to the verb's direct object, as in *Mary cheated the other players.* Normally, the two thematic structures differ in frequency—in the example of *cheat,* the first thematic structure is more common than the second. Our model assumes that both thematic structures are made available upon access of the verb (Shapiro, Zurif, & Grimshaw, 1987, 1989; Stowe, 1989; Tanenhaus & Carlson, 1989), although the more frequently used structure will have a greater initial activation level.

The model assumes further that thematic roles are assigned to phrases after they have been syntactically categorized. More precisely, a thematic role is assigned to a phrase once a word is processed which can play the syntactic role of head. Thus, the sequence *the* would not be assigned a the-

matic role, nor would the sequence *the tall grey*. However, the sequence *the tall grey man* would be assigned a thematic role at the word *man*, and that thematic role assignment would be maintained if the phrase continued with a modifier after the head, such as *the tall grey man with the black hat*.

When a phrase has been syntactically misanalyzed, as in a garden-path sentence, the thematic role that has been assigned to the phrase must be revised as well. For example, consider a sentence from Experiment 3: *When the men hunt the birds typically scatter*. When the phrase *the birds* is initially syntactically analyzed as the direct object of *hunt* due to the operation of late closure (and/or possibly other factors such as the absence of a comma), it will correspondingly be assigned the thematic role of patient. When the disambiguating word is encountered, successful reanalysis requires not only that the ambiguous phrase be interpreted as the subject of a new clause, but also that it be assigned the role of agent of a new verb (*scatter*). Thus, reanalysis requires both that syntactic and thematic roles be revised.

Given that all thematic structures for a verb are available upon receipt of the verb, thematic reanalysis will require the parser to try to recover one of the thematic structures not initially chosen. For example, upon receipt of the verb *hunt*, two thematic structures would become available: one where *hunt* assigns only the role of agent, and one where *hunt* assigns both the roles of agent and patient. The initially chosen structure is the second one, where two roles are assigned. When the disambiguating word is encountered signalling that reanalysis is required, the first structure will have to be selected, the one where *hunt* assigns just the role of agent. The longer the thematic processor has been committed to the incorrect thematic structure, the harder it will be for it to recover the alternative, correct structure. This claim follows if we assume that both structures are available at the verb, but begin decaying rapidly. The longer the unselected structure has been unselected, the more it will have decayed, and the harder reanalysis will be.

How does this model account for the results of our experiments? First, the model accounts straightforwardly for the finding that the greater the distance between the head of the misanalyzed phrase and the disambiguating word, the harder reanalysis is. Consider what happens with a sentence such as *When the men hunt the birds (with bright plumage) scatter*. At the verb *hunt*, both thematic structures would become available. In the short condition, the parser then attaches the phrase *the birds* as the direct object of *hunt* (due to late closure, the absence of a comma, etc.). The existence of a direct object causes the thematic structure containing two thematic roles to be instantiated, and the other structure begins to decay. Next, the parser receives the disambiguating word *scatter*, which signals that the syntactic structure is incorrect. If the syntactic structure is wrong, the thematic structure will necessarily be wrong also. At this point, as argued by Rayner et al. (1983), the thematic processor proposes to the parser that it construct a new syntactic structure, one that can support the other legitimate thematic structure. That structure is one where only the role of agent is assigned, which corresponds to an intransitive syntactic structure. Because the parser receives the disambiguating word immediately after the thematic processor makes its erroneous thematic role assignment, the alternative thematic structure is still readily available. Now consider what happens in the condition where three words follow the head of the ambiguous phrase. Here, the incorrect thematic structure will be maintained for three words before the disambiguating word is encountered. Because more processing time has elapsed, the alternative, correct thematic structure has decayed substantially. As a result, reanalysis will be more difficult, because the correct thematic (and corresponding syntactic) structure will be less available. Thus, the

model predicts that the greater the distance between the head of a phrase (the point at which a thematic role is assigned) and the disambiguating word (the point at which syntactic and thematic reanalysis is initiated), the more difficult the process of reanalysis.

In summary, the thematic reanalysis model offers a coherent and plausible account of the results we obtained in our experiments. However, we must admit that the alternative hypothesis outlined in the introduction to Experiment 4 could also account for our results. Recall that according to that alternative model, ease of reanalysis is determined by the amount of time the parser has been committed to an incorrect attachment. The greater that amount of time, the more difficult reanalysis will be. If a phrase is only attached into an ongoing phrase structure tree when its head is encountered, then it follows that reanalysis will be easier when the head of the misanalyzed phrase is adjacent to the syntactically disambiguating word. Nevertheless, we believe there are reasons for preferring the thematic reanalysis model to this alternative. First, the former model has the virtue of invoking theoretical machinery that has been proposed independently elsewhere (Rayner et al., 1983; Tanenhaus & Carlson, 1989). Second, the model can easily accommodate effects of verb bias on reanalysis processes (Ferreira & Henderson, 1990, 1991), because the activation levels of the alternative thematic structures are influenced by their frequency of use. Ultimately, of course, the final choice of a model of reanalysis will be decided by further empirical testing.

It is also important to note that many factors besides the ones manipulated here could affect ease of parsing. For example, verb transitivity, semantic plausibility, and constraints from discourse could all potentially influence the ease with which the parser can recover from a garden-path. Nevertheless, we have identified one critical factor—the distance between the head of the misanalyzed phrase and the disambiguating word—and suggested a model to account for the powerful effect head location has on reanalysis processes.

References

Altmann, G. (1989). Parsing and interpretation: An introduction. *Language and Cognitive Processes*, **4**, 1–20.

Clifton, C., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, **4**, 77–103.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, **30**, 210–233.

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

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

Ferreira, F., & Henderson, J. M. (1991). The use of verb subcategorization information in syntactic parsing. In G. B. Simpson (Ed.), *Understanding Word and Sentence* (pp. 305–330). Amsterdam: North-Holland.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Sentence processing. In M. Coltheart (Ed.), *Attention and Performance, XII*. Hillsdale, NJ: Erlbaum.

Frazier, L. (1989). Against lexical generation of syntax. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 505–528). Cambridge, MA: MIT Press.

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

Holmes, V. M. (1987). Syntactic parsing: In search of the garden path. In M. Coltheart (Ed.), *Attention and Performance, XII*. Hillsdale, NJ: Erlbaum.

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

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

KENNEDY, A., & MURRAY, W. S. (1984). Inspection times for words in syntactically ambiguous sentences under three presentation conditions. *Journal of Experimental Psychology: Human Perception and Performance,* **10,** 833–849.

MITCHELL, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. Kieras and M. A. Just (Eds.), *New Methods in Reading Comprehension* (pp. 69–89). Hillsdale, NJ: Erlbaum.

MITCHELL, D. C. (1987). Lexical guidance in human parsing: Locus and processing. In M. Coltheart (Ed.), *Attention and Performance, XII.* Hillsdale, NJ: Erlbaum.

MITCHELL, D. C. (1989). Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes,* **4,** 123–154.

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

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.

SHAPIRO, L. P., ZURIF, E. B., & GRIMSHAW, J. (1987). Sentence processing and the mental representation of verbs. *Cognition,* **27,** 219–246.

SHAPIRO, L. P., ZURIF, E. B., & GRIMSHAW, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research,* **18,** 223–243.

STOWE, L. (1989). Thematic structures and sentence comprehension. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 319–357). Dordrecht: Kluwer Academic.

TANENHAUS, M. K., & CARLSON, G. N. (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 529–562). Cambridge, MA: MIT Press.

TANENHAUS, M. K., CARLSON, G. N., & TRUESWELL, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes,* **4,** 211–234.

TARABAN, R., & MCCLELLAND, J. R. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language,* **27,** 597–632.

TYLER, L. K. (1989). The role of lexical representations in language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 439–462). Cambridge, MA: MIT Press.

WARNER, J., & GLASS, A. L. (1987). Context and distance-to-disambiguation effects in ambiguity resolution: Evidence from grammaticality judgments of garden path sentences. *Journal of Memory and Language,* **26,** 714–738.

(Received October 25, 1990)
(Revision received February 20, 1991)