JOURNAL OF MEMORY AND LANGUAGE 29, 413–432 (1990)

Evidence for the Immediate Use of Verb Control Information in Sentence Processing

JULIE E. BOLAND, MICHAEL K. TANENHAUS, AND SUSAN M. GARNSEY

University of Rochester

When a verb is followed by an infinitival complement, the particular verb determines whether its subject or object is the understood subject of the infinitive. Thus, the verb "controls" the interpretation of the infinitive (e.g., *John promised/persuaded Mary to wash*). Frazier and colleagues have argued that verb control information is not immediately accessed and used in sentence processing based on whole-sentence comprehension times. The studies reported here examined the use of verb control using an on-line plausibility monitoring task. Subjects immediately detected incongruities that depended upon their having correctly used control information, indicating that verb control information is rapidly accessed and used. It is argued that the results support an approach to language comprehension that emphasizes the importance of lexical representations in rapidly integrating many of the different sources of linguistic and nonlinguistic knowledge that need to be coordinated during language comprehension. © 1990 Academic Press, Inc.

The lexical representation of a word, especially a verb, provides the reader or listener with considerable information about how that word combines syntactically and semantically with other words in the sentence and discourse (Marslen-Wilson, Brown, & Tyler, 1988; Tanenhaus & Carlson, 1989; Tanenhaus, Garnsey, & Boland, in press; Tanenhaus, Carlson, & Trueswell, 1989; Tyler, 1989). However, there is little consensus about how and when lexical representations are used in sentence processing. Some researchers have argued that the use of lexical information is delayed such that initial syntactic decisions are made without reference to relevant combinatory lexical information (Frazier, 1987, 1989; Mitchell, 1987, in press), while others have suggested that this information provides for rapid and relatively seamless communication among parsing, discourse context, and structurally relevant real-world knowledge (Marslen-Wilson et al., 1988; McClelland, 1987; Tanenhaus et al., 1989; Tanenhaus & Carlson, 1989; Tyler, 1989). Thus, the issue of when combinatory lexical information is accessed and used is directly relevant to the ongoing debate about the modularity of the language processing system (Fodor, 1983; Marslen-Wilson & Tyler, 1987; Tanenhaus, Carlson, & Seidenberg, 1985).

This article examines the time course of the use of verb control information, which is a particular type of combinatory lexical knowledge, in sentence processing. Control information determines how particular verbs influence interpretation of the understood subjects of infinitival complements. So, for example, the **subject** of the verb *promise* is also the understood subject of the infinitive in *John$_i$ promised Mary* __$_i$ *to wash himself,* while the **object** of the verb *persuade* controls the interpretation of its infinitive in *John persuaded Mary$_i$* __$_i$ *to*

This research was supported in part by NIH grant HD-22271 and NSF grant BNS-8617738. We would like to thank Greg Carlson and Gary Dell for their helpful insight and suggestions, John McCrary for help with the statistical analyses, and Janet Nicol for several helpful discussions as well as for providing us with her materials. We would also like to thank Gail Mauner and an anonymous reviewer for many helpful comments. Susan M. Garnsey is now at the Department of Psychology, University of Illinois at Champaign-Urbana. Requests for reprints should be sent to Michael K. Tanenhaus, Department of Psychology, Meliora Hall, University of Rochester, Rochester, NY 14627.

0749-596X/90 $3.00

wash herself. For ease of exposition all phonologically unrealized linguistic elements will be called gaps. Each gap will be marked with a "$\_\_$" and have a common subscript with its "filler," which is the noun phrase that is associated with the gap when the sentence is correctly interpreted. When a gap has been associated with its filler, they are said to be "coindexed."

In a sentence with a "controlled" infinitive complement, either the subject or the object of the matrix verb is understood as , i.e., controls the interpretation of, the implicit subject of the infinitive. Some examples of verb control are shown below. Sentence (1a) illustrates the use of a "subject control" verb (*promise*), and sentence (1b), an "object control" verb (*tell*).

(1) a. The girl$_i$ promised her brother $\_\__i$ to sing.
b. The girl told her brother$_j$ $\_\__j$ to sing.

In sentence (1a), the subject of *promise* is understood as the subject of the infinitive, i.e., the girl is promising her brother that she will sing, whereas in sentence (1b), the object of the matrix verb is understood as the subject of the infinitive, i.e., the girl is telling her brother that he should sing. We are following standard linguistic convention and distinguishing true infinitive complements from purpose or rationale clauses such as in *Mary told her brother the joke to make him laugh*. These clauses appear freely after most verbs regardless of control properties and can be paraphrased as "for the purpose of . . ." or "in order to. . . ."

Most control verbs have unambiguous control properties. Verbs that require a direct object before an infinitive phrase, such as *force* in (2b), are typically object control verbs. Verbs that do not allow an object before the infinitive, like *start* in (2a), are always subject control verbs. Only a few verbs like *promise* are transitive, yet subject control.

(2) a. The girl$_i$ started $\_\__i$ to sing.
b. The girl forced the woman$_j$ $\_\__j$ to sing.

Some verbs have ambiguous control properties. For example, *beg* is used as a subject control verb in sentence (3a) and as an object control verb in sentence (3b).

(3) a. The girl$_i$ begged $\_\__i$ to sing for the woman.
b. The girl begged the woman$_j$ $\_\__j$ to sing.

Although *beg* has ambiguous control properties, only the subject control interpretation is possible in (3a) because beg is used intransitively. In (3b) both the subject and object control interpretations are possible, but the object control interpretation is preferred. There is a strong preference for the object control interpretation whenever a control verb is used transitively. In (3b) it is clear that *beg* is being used transitively at the first word following the verb, but in Wh-questions such as those in (4a), both the transitivity and the control relationship can remain ambiguous for several words.

(4) a. Subject control: Which woman$_i$ did the girl$_j$ beg $\_\__j$ to sing for $\_\__i$?
b. Object control: Which woman$_i$ did the girl beg $\_\__i$ $\_\__i$ to sing?[1]

The ambiguity arises because *woman* is a filler that must be associated with some ob-

[1] In (4b), the first gap is an object gap ("trace" in the terms of the Government and Binding linguistic theory) and the second is a subject gap ("PRO" in Government and Binding terminology). (4a) also contains an object gap, but it is at the end of the sentence and is a prepositional object. Government and Binding Theory differentiates between trace and PRO since it analyzes object gaps as the result of movement and subject gaps as generated in deep structure. In addition, PRO has its own thematic role while trace shares a thematic role with an overt noun phrase; the two kinds of gaps occur in different environments, and they are subject to different constraints. We should note, though, that current linguistic theories disagree about whether or not there is actually an empty syntactic category associated with the understood subject of an infinitive phrase (see Sells, 1985).

ject gap in the sentence. In (4a), the object gap follows the preposition, *for*, and *the woman* is understood as the object of the preposition. Since *beg* is being used intransitively, only the subject control interpretation is possible. In (4b) however, the object gap follows the verb, *beg*, and *the woman* is interpreted as the direct object of the verb. Since *beg* is being used transitively, the object control interpretation prevails.

Frazier, Clifton, and Randall (1983) used sentences like those in (5), which are a longer and somewhat more complex version of those in (4).

(5) a. Subject control: Everyone liked the woman$_i$ (who$_i$) the little child$_j$ begged $\_\__j$ to sing those stupid French songs for $\_\__i$.
b. Object control: Every one liked the woman$_i$ (who$_i$) the little child begged $\_\__i$ $\_\__i$ to sing those stupid French songs.

Sentences were presented one word at a time, with each word centrally displayed on a CRT for 300 ms. As quickly as possible after the last word was presented, subjects responded "got it" or "missed it" to indicate whether or not they had understood the sentence. The important result was that subjects understood sentences like (5a) faster than sentences like (5b). Frazier et al. reasoned that if subjects committed themselves to a single interpretation (either subject of object control), reading times should be long when the disambiguating phrase (in the above example, this is the presence or absence of *for*) did not match the chosen interpretation. They accounted for the subject control advantage by proposing that comprehenders adopt a "most recent filler" (MRF) strategy in which the most recent potential filler is initially associated with a gap. In both (5a) and (5b) above, the MRF strategy will assign *the little child* to the subject gap in the infinitive phrase, i.e., the child as the singer. This is the correct assignment in (5a), the subject control sentence, but in (5b) the woman, not the child, is singing. The temporary misassignment of *child* to the subject gap causes a "garden path," which is corrected when verb control information becomes available.

When confronted with ambiguous control information (as with *beg*), comprehenders are forced to either resort to a heuristic like the MRF strategy for the initial filler-gap assignment, or delay filler-gap assignment until unambiguous information is available. But even when Frazier et al. (1983) used verbs with unambiguous control properties, as shown below in (6), subjects understood the sentences with subject control verbs faster than those with object control verbs.

(6) a. Subject control: Everyone liked the woman$_i$ (who$_i$) the little child$_j$ started $\_\__j$ to sing those stupid French songs for $\_\__i$.
b. Object control: Everyone liked the woman$_i$ (who$_i$) the little child forced $\_\__i$ $\_\__i$ to sing those stupid French songs.

In fact, having unambiguous control information did not reduce the recent filler advantage, a result replicated by Clifton and Frazier (1986). Frazier et al. argued that subjects continued to used the MRF strategy, even when the strategy led to a filler-gap assignment that was incompatible with the control properties of the verb. They concluded that the parser does not have access to verb control information when it initially associates the filler with a gap, providing support for modularity within the sentence processing system. Further, they argued that the delayed use of verb control information is consistent with the representational framework adopted in the Government and Binding Theory (Chomsky, 1981), but not with alternative linguistic theories, a point that has been vigorously debated (Crain & Fodor, 1985; Fodor, 1988; Ford & Dalrymple, 1988).

The ability of the MRF strategy to explain the Frazier et al. (1983) data has been questioned (Fodor, 1988; Ford & Dalrymple, 1988). Rather than detailing the various arguments, we will simply note that the Frazier et al. evidence for delayed use of verb control information is indirect. Longer end-of-sentence response latencies for distant filler sentences than for recent filler sentences are taken as evidence that readers misassigned the most recent filler to the subject gap. What is lacking is direct evidence about when subject gaps in infinitive complements are interpreted and when control information is used.

In order to determine when the subject gap is interpreted, we used an "embedded anomaly" logic which has proved useful in a series of studies investigating the processing of sentences with filler-gap dependencies (Boland, Tanenhaus, Carlson, & Garnsey, in press; Garnsey, Tanenhaus, & Chapman, 1989; Tanenhaus et al., 1985; Tanenhaus, Boland, Garnsey, & Carlson, 1989; Tanenhaus et al., in press). In general terms, the embedded anomaly approach manipulates the plausibility of a fronted Wh-phrase (a salient filler) for a particular gap or potential gap in the sentence. Since sentences with implausible fillers are only implausible if the filler is associated with the subject gap, the point at which plausibility effects occur indicates when the filler has been associated with the gap. In the present experiments, we manipulated plausibility of a fronted direct object as the implicit subject of an infinitive phrase.

Compare sentences (7a) and (7b).

(7) a. The cowboy signalled the outlaw __to surrender his weapons quietly.
b. The cowboy signalled the horse __to surrender his weapons quietly.

It is plausible for cowboys to signal both horses and outlaws, but sentence (7b) is less plausible than (7a) because horses cannot surrender. The subject gap must be correctly interpreted in order to notice this. If *cowboy* is incorrectly assigned to the gap, the sentence appears plausible since cowboys can surrender. We assume that real world knowledge can be used to assess plausibility as interpretations are built. The MRF strategy predicts that *horse* will be correctly associated with the subject gap in (7b). However, if *horse* is fronted, as in the Wh-question, *Which horse did the cowboy signal to surrender his weapons quietly?*, the MRF strategy predicts that *cowboy* will initially be associated with the subject gap since it will be closer to the gap. Thus, the reader would be unaware of the oddity until the mistaken analysis was corrected.

Our experiments address two questions. The first is a test of the MRF strategy: will plausibility effects in the recent filler versions precede plausibility effects in the distant filler versions? The second question concerns when verb control information is used: how early will plausibility effects occur? The second issue is independent of the first since plausibility effects could be equally late in both types of sentences, indicating that verb control information guides initial coindexing but that it becomes available relatively late in processing.

Experiment 1

In Experiment 1, we used a word-by-word sense-monitoring task to compare sentences in which the MRF strategy predicts initial misassignment with sentences in which the MRF strategy predicts correct initial assignment. Frazier et al. (1983) used object control verbs for their distant filler sentences and subject control verbs for their recent filler sentences. We manipulated filler distance by comparing question (distant filler) and declarative (recent filler) versions of object control sentences.

The four experimental conditions and the critical word positions are described for a sample sentence set in Fig. 1. In the first two conditions in Fig. 1, the sentences become implausible at *surrender*, because *horse* is an implausible subject. In the two declarative conditions the correct fillers, *horse* and *outlaw*, are the most recent fillers for the subject gap (subject of *to surrender*). Therefore, correct initial assignment is predicted by both the MRF strategy and the position that verb control information is

CONDITION		Critical Word Positions					
		V1 or N	to	V2	V2+1*	V2+2*	V2+3*
Implaus-Quest	Which horse did the cowboy	signal	to	surrender	to	the	authorities
Implaus-Declar	The cowboy signalled the	horse					
Plaus-Quest	Which outlaw did the cowboy	signal					
Plaus-Declar	The cowboy signalled the	outlaw					

FIG. 1. A sample sentence set from Experiment 1 is shown with critical word positions specified. The four condition names indicate the sentence type (declarative or question) and the plausibility of the filler as the subject of V2, which is the same as the overall plausibility of the sentence. **The last three words were either a noun phrase, a prepositional phrase, an adverbial phrase, or a noun phrase followed by an adverbial.*

used immediately. However, in the two question conditions *horse* and *outlaw*, respectively, are the correct fillers but not the most recent. (*Cowboy* is more recent.) The MRF strategy predicts correct initial assignment of *horse/outlaw* to the object gap (Clifton & Frazier, 1986), but predicts that *cowboy* will be incorrectly assigned to the subject gap. Corrective reassignment occurs when verb-control information becomes available. Because *cowboy* is a plausible subject of the infinitive, delayed detection of the implausibility of the implausible questions compared to the implausible declaratives would support the MRF hypothesis. If verb control information is used to select the filler for the subject gap, there should be no delay in implausibility detection in the Wh-question compared to the simple declaratives. If verb control information is used immediately, the plausibility effect should be at, or shortly after, the embedded verb in both sentences.

Method

Materials. Experiment 1 used 28 sentence sets of the form illustrated in Fig. 1. The full set appears in Appendix A. The sentences were constructed using 28 object control verbs. None of them could be used as subject-control verbs. Sentence completion norms were collected for a set of object control verbs. Verbs were selected from this set using the following criteria: they were never used as a subject control verb, they were virtually always used transitively, and in most of responses they were used with an overt or implied infinitive complement. (Details are available from the authors). All of the critical sentences had a transitive matrix verb, followed by either an object NP (declarative conditions) or a gap (question conditions), an infinitive VP with a subject gap, and either a prepositional phrase, an adverbial phrase, or a noun phrase three words in length. Both the subject gap and the object gap were coindexed with the object NP. All sentences were plausible up to the embedded verb. The subject of the matrix verb was always plausible as the subject of the embedded infinitive. In the implausible conditions, the object noun phrase was not a good Agent for the infinitive verb, so the implausible sentences became implausible when the embedded verb was read.

The four conditions were counterbalanced across four lists such that one version of each sentence appeared on each list and equal numbers of sentences in each condition occurred on all lists. A variety of distractor sentences were constructed using non-control verbs and ordered semirandomly with the critical sentences. Each list consisted of 28 critical sentences and 52 distractor sentences for a total of 80 sentences. Half of the critical sentences and six of the distractor sentences were anomalous; the remaining 60 sentences on each list were plausible.

Subjects. Thirty-two male and female undergraduates completed the experiment either in partial fulfillment of course requirements of for a minimal sum. All subjects were native English speakers.

Procedure. Sentences appeared on an IBM PC monitor. Subjects controlled the word-by-word presentation rate by pressing a key. Each key-press caused a word to appear, with the words accumulating across the screen. All of the critical sentences fit on a single line, though a few of the distractor sentences continued onto the next line. Subjects were told to read rapidly and carefully. They were instructed to continue pressing the same key as long as the sentence continued to make sense. If the sentence stopped making sense, they were to press a different "no" key as soon as possible. When a "no" response was given, presentation of the sentence was halted and a new trial began. At the end of a sentence, the last key press caused either a period or a question mark to appear. (The punctuation did not appear at the same time as the final word in the sentence). Before the experiment began, subjects were shown sample sentences that did and did not make sense. An explanation of when and why the sentence stopped making sense was given for each example that did not make sense. None of the examples contained control verbs or infinitive phrases. Subjects completed 20 practice trials before beginning the actual experiment.

Results

A record of which button was pressed was kept for six word positions in each critical sentence, beginning one word prior to the beginning of the infinitive phrase (see Fig. 1) and continuing until the end of the sentence or until the subject pressed a button indicating that the sentence had stopped making sense. We also collected reaction times for each button press in the critical region. We refer to these as the reading time data.

The reading time data provide a general idea of reading rates, and functions as a secondary measure of plausibility. However, a cautionary note is in order. The reading time data represent only those trials on which subjects continued to judge a sentence plausible. Therefore, responses of subjects who judged a sentence to be implausible at a particular word position are not included in the reaction time analysis at that word position, or any following word positions for that sentence. For the plausible conditions, where there are very few "no" responses, this does not present a problem, but for the implausible conditions there will be a large amount of missing reading time data after the point at which the sentences become implausible. Thus, the reading times represent a biased sample of the responses. Trials on which the anomaly is especially salient are likely to elicit a "no" response, so they will not be included in the reading time data. Trials on which subjects are not sure if a sentence is implausible are more likely to elicit a positive response and therefore be reflected in the reading time data and not in the judgment data. Thus, the most plausible of the implausible items are over-represented in the reading times for implausible conditions. Moreover, since there is missing data, the reading time means are more variable then they would otherwise be. As a consequence of these problems, the judgment data are more reliable and we will focus our attention on them.

Nonetheless, it is important to look at the reading times because they allow us to evaluate the possibility that control information is only used some proportion of the time, or by only a proportion of the subjects. For example, suppose that control information was used to interpret the sentence correctly just in those experimental sentences in which subjects responded "no" at V2 (see Fig. 1), but that control information was unavailable or unused at that point in the remainder of the experimental sentences. Plausibility effect would show up in the judgment data at V2, reflecting that portion of the Sentences in which control information was used. However, there would not be a plausibility effect in the reading times if the only time subjects used control information was when they responded "no" at V2. Under such circumstances, the reading times should exhibit the pattern of data which Frazier et al. (1983) predict: an early plausibility effect in the declarative sentences and a late plausibility effect in the questions.

Judgments. At each word position, a count was kept of the number of trials in each condition on which a subject responded "no," meaning that the sentence had stopped making sense. A summary of the cumulative percentages of "no" responses is shown in Fig. 2a. The percentage of "no" responses sharply increases at position V2 (the infinitive verb) and continues rising until the end of the sentence for both of the implausible conditions. There is no delay of plausibility effect for the question condition relative to the declarative conditions. By the end of the sentence, subjects had responded "no" for 77% of the implausible questions, 77% of the implausible declaratives, 8% of the plausible questions, and 7% of the plausible declaratives. A 4(list) × 2(sentence type) × 2(plausibility) ANOVA on this measure revealed an effect of plausibility by both subjects and items [$F1(1,28) = 581.11$, $p < .01$, $MSe = 267$; $F2(1,24) = 219.00$, $p < .01$, $MSe = 622$] and no other effects. Thus, both implausi-

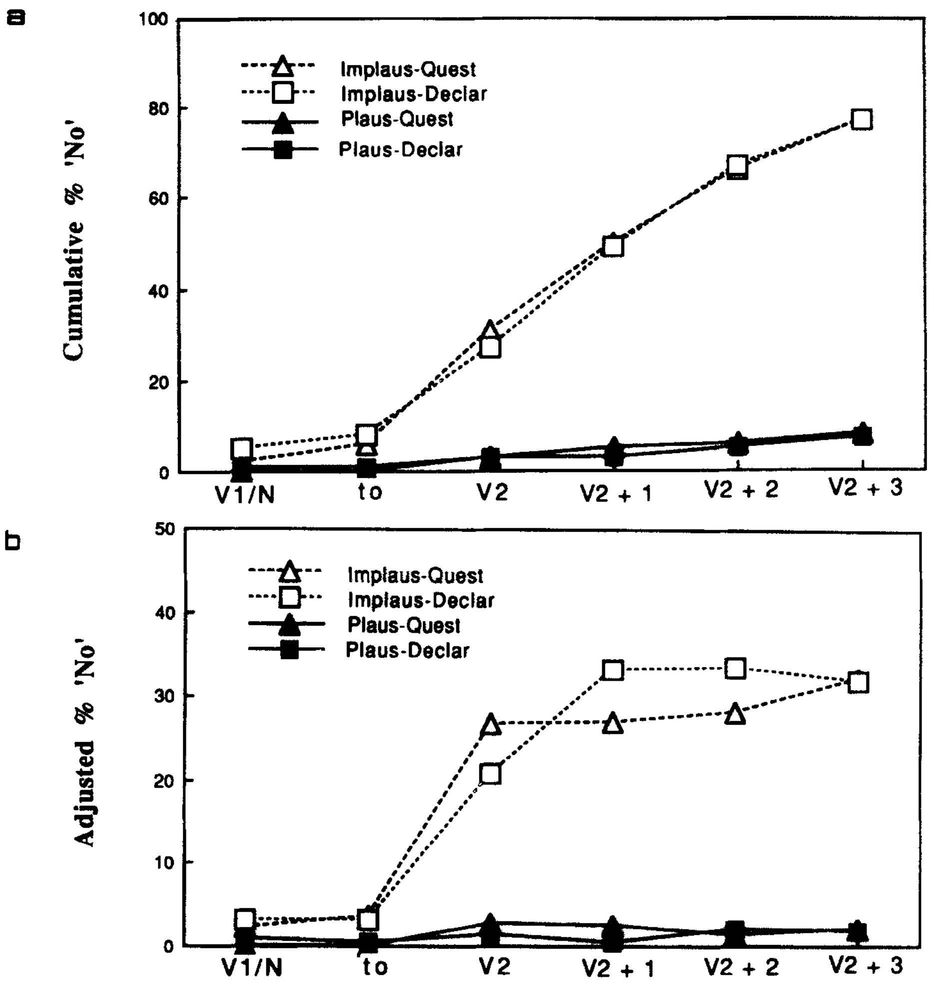

FIG. 2. (a) Cumulative percentage of "no" responses at each word position in Experiment 1. (b) Adjusted percentage of "no" responses at each word position in Experiment 1.

ble conditions were judged to be implausible more often than the plausible conditions, and questions and declaratives did not differ from one another.

While the cumulative percent of "no" responses at the end of the sentences provides a valid index of overall plausibility, it was also important for our purposes to know how plausibility effects changed across the sentence. However, the cumulative percentage of "no" responses at each word position is a problematic measure because its value at each position is strongly correlated with the value at the preceding position. The simple frequency of "no" responses at each word position is also dependent on the number of "no" responses at the preceding positions, because the number of opportunities to respond "no" depends on the number of trials on which a "no" response has already been made. This is because a trial ends whenever a subject responds "no." To minimize the dependence of later values on earlier ones, the data were transformed before being analyzed. The number of "no" responses at each word position was converted to a percentage of the "remaining possible no"s in the following way. There were seven trials per condition, so at the beginning of the sentence the number of trials on which a subject could possibly respond "no" was seven. At later word positions the number of remaining possible no's was equal to seven minus the number of trials on which the subject had responded "no" earlier in the sentence. For example, if a subject responded "yes" on all of the trials in a condition up through the second critical word position, then responded "no" on two out of the seven trials at the third critical word position, that subject's percentage of possible "no's" for that position would be 2/7 or about 29%. If the subject then responded "no" at the fourth critical word position on two more trials, the percentage of remaining possible "no"s at the fourth position would be 2/5 or 40%, because only five trials remain on which it is possible to respond "no."

The percentages of "remaining possible no"s are presented in Fig. 2b. Percentages for both subjects and items were submitted to a 4(list) × 2(sentence type) × 2(plausibility) × 6(word position) ANOVA. There were significant main effects by subjects and by items for plausibility [$F1(1,28) = 166.20$, $p < .01$, $MSe = 498$; $F2(1,24) = 89.94$, $p < .01$, $MSe = 969$] and word position [$F1(5,140) = 37.02$, $p < .01$, $MSe = 225$; $F2(5,120) = 33.01$, $p < .01$, $MSe = 281$]. (We report the Huynh-Feldt (Huynh & Feldt, 1976) adjusted probability values for all analyses involving the word position factor, which has more than two levels, since the results at different positions are not independent of one another. We present the original, unadjusted degrees of freedom). There was also a significant interaction for plausibility × word position in both the subject and item analyses [$F1(5,140) = 30.51$, $p < .01$, $MSe = 227$; $F2(5,120) = 27.42$, $p < .01$, $MSe = 289$]. Crucially, there was no plausibility × sentence type interaction [$F1 < 1$; $F2 < 1$].

Differences between the conditions at particular target positions were evaluated by making post hoc pairwise comparisons of both the subject and item means using the Ryan-Einot-Gabriel-Welsch multiple F (REGWF), which controls the experimentwise error for multiple comparisons, with $\alpha = .05$ (Einot & Gabriel, 1975; Ryan, 1959; Welsch, 1977). Starting at V2 and continuing until the end of the sentence, there were more "no" responses to both implausible questions and implausible declaratives compared to plausible questions and plausible declaratives at each word position. The proportion of "no" responses to implausible questions did not differ from implausible declaratives, nor did plausible questions differ from plausible declaratives.

Reading times. Mean reading times are shown in Fig. 3. It is apparent in the figure that plausibility effects were not delayed for the question conditions relative to the declarative conditions: word-by-word reading times became slower for implausible

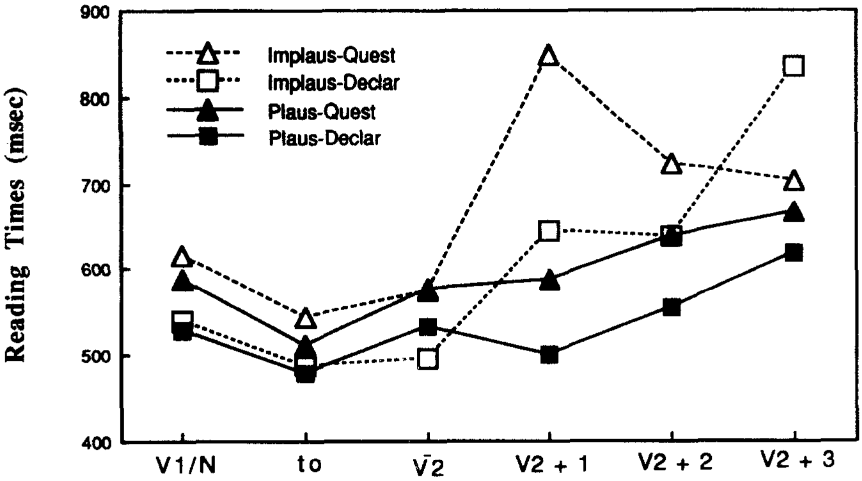

FIG. 3. Mean reading times for positive responses at each word position in Experiment 1.

conditions compared to plausible conditions beginning at V2 + 1 for both the questions and the declaratives. The percentage of trials for which reading time data is missing is essentially the inverse of the cumulative percent of "no" responses (shown in Fig. 2a). The actual percentage of trials used in the reading time analysis is as follows: 98% at V1/NP(S), 96% at "to," 82% at V2, 71% at V2 + 1, 62% at V2 + 2, 56% at V2 + 3. Occasional time-outs (responses longer than 3 s) account for the small discrepancies between the data presented here and the data in Fig. 2a.

Reading times for each condition at each word position were averaged by subjects and by items for use in the ANOVAs. At each word position those values more than 2.5 standard deviations above or below the subject mean for that word position were truncated and replaced with the cutoff value. In the implausible conditions, there were occasionally cells of subjects or items for which there were no positive responses, and therefore no reading times, at the last word positions. These missing cells were replaced using the following formula: subject or item mean plus condition mean minus grand mean (Winer, 1962). Replaced cells accounted for 3% of the subject means and 3% of the item means. (The subject or item mean was computed by averaging all the condition means for that subject or item at the relevant word position. The condition mean was figured across subjects or items at the relevant word position. The grand mean was the mean of all the subject or item means at the relevant word position).

Subject and item means were each submitted to 4(list) × 2(sentence type) × 2(plausibility) × 6(word position) ANOVAs identical to those used for the judgment data. There were significant main effects by subjects and by items for sentence type [$F1(1,28) = 9.04$, $p < .01$, $MSe = 79$; $F2(1,24) = 30.75$, $p < .01$, $MSe = 30$] and word position [$F1(5,140) = 13.77$, $p < .01$, $MSe = 52$; $F2(5,120) = 5.71$, $p < .01$, $MSe = 53$]. The main effect of plausibility was significant in the subject analysis [$F1(1,28) = 20.93$, $p < .01$, $MSe = 49$], but not in the item analysis [$F2(1,24) = .75$, $p > .10$, $MSe = 42$]. There was also an interaction of plausibility × word position which was significant in the subject analysis [$F1(5,140) = 5.42$, $p < .05$, $MSe = 40$], and marginally significant in the items analysis [$F2(5,120) = 1.92$, $p < .10$, $MSe = 26$].

Differences between conditions at particular word positions were again evaluated with the post hoc REGWF at $\alpha = .05$. In the subject analysis at V2 + 1, responses to implausible questions were slower than in the other three conditions. However, in a less conservative planned comparison at this position, implausible declaratives were also slower than plausible declaratives in the subject analysis [$F1(1,28) = 5.45$, $p < .05$, $MSe = 62$]. At V2 + 3, implausible declaratives were slower than plausible declaratives. No other comparisons were different at any of the word positions in the subject analysis. No comparisons were different in the items analysis.

Readers may be puzzled as to why effects that are significant or nearly significant in the subject analysis do not approach significance in the item analysis. Conditions in which reaction times are slow correspond to conditions in which there are a high proportion of "no" responses. Assume that a subject responds "no" in six out of the seven trials at the word following the embedded verb in the question-implausible condition. The subject's mean for that condition would then be based on the single item that he or she responded to. This will increase the variability of the data, but the mean should not be systematically distorted because the remaining score should be an unbiased estimate of the subject's true mean for that condition. That same subject will contribute a score to only one of the seven item means in that condition. This would not introduce a bias in the item means if fast and slow subjects were equally likely to respond "no." However, in these experiments fast subjects are somewhat more likely to respond "yes" in the implausible conditions. (Note that this is a form of speed-accuracy trade-off). As a result, fast subjects contribute more scores to the item means in the implausible conditions than slow subjects, thus reducing the size of the effects. In order to confirm that this was the case, we replaced all missing scores for each subject with the subject's mean score for the other items in the same condition as the item (trial) with the missing score. When this was done, the item analyses showed the same statistical pattern as did the subject analyses. (Details are available from the authors).

Discussion

In contrast to the predictions made by the MRF strategy, plausibility effects were seen at the same point for the Wh-questions and the declarative sentences. Thus, we found no evidence that it took more time to fill a gap with a distant filler than with a recent filler. This result indicates that control information, rather than a distance heuristic, is used to interpret infinitive complements. Secondly, the data indicate that verb control is available early since we saw plausibility effects at the earliest possible point (V2) in the judgment task. This is the first point at which the sentences with implausible fillers could become implausible. Lastly, this effect cannot be due to some small proportion of the subjects using control information, or many of the subjects using control information on a small proportion of the trials. Although the plausibility effects are weaker and one word later in the reading time data than the judgment data, there is no evidence that the plausibility effect is delayed in the Wh-questions compared to the declaratives. In fact, in the reading time data, the plausibility effects are larger for the Wh-questions compared to the declaratives.

In Experiment 1, the crucial effect, i.e., no difference in the timing of plausibility effects between declaratives and questions, was a null effect. Experiment 2 provides an opportunity for the immediate use of verb control to predict a difference in plausibility effects. The experiment was designed such that the same Wh-phrase filler produces an anomaly when the sentence contains an object control verb but not when it contains a subject control verb. This also controls for any differences in lexical relatedness between the fillers and the verb in the infinitive complement. Some of the plausible filler words in Experiment 1 were more semantically related to the embedded verb (the point of implausibility) than were the implausible filler words. Thus, it might be argued that some of the effects could be attributed to the relatedness of the lexical items. For example, *frog* and *hop* are more semantically related than *snake* and *hop*. It is possible, then, that this relationship between *hop* and *frog* in *The boy forced the frog to hop over the rock.* and *Which frog did the boy force to hop over the rock?* and the absence of a similar relationship between *hop* and *snake* in *The boy forced the snake to hop over the rock.* and *Which snake did the boy force to hop over the rock?* might have contributed to the differences in reading times (and perhaps even judgments) after *hop*. In Experiment 2, the same fillers and embedded verbs were used, thus controlling for lexical relatedness.

EXPERIMENT 2

In order to directly manipulate verb control, the control properties of the matrix verb (subject control or object control) were crossed with the plausibility of the (Wh-phrase) filler as the subject of the verb in the infinitive clause. The conditions with object control verbs are similar to the question conditions in Experiment 1. In those sentences the Wh-phrase is always understood to be the subject of the infinitive complement. However in the subject control sentences, the Wh-phrase will never be understood as the subject of the infinitive complement if control information is correctly used. Rather, it will be associated with a later gap. Therefore, plausibility effects are predicted only for the sentences with object control verbs.

Consider the sentences in Fig. 4. *Sub-j(ect)* and *Obj(ect)* refer to the control properties of the matrix verb. *Plaus(ible)* and *Implaus(ible)* refer to the plausibility of the Wh-phrase filler **as the subject of the infinitive clause**, not the overall plausibility of the sentence. The object control sentence with the implausible filler is implausible because horses cannot surrender; but the subject control version with the implausible filler is plausible because cowboys can surrender. We predicted more "no" responses at *surrender* and the immediately following words in *Obj-Implaus* compared to *Obj-Plaus*. This is simply a replication of the plausibility effect we found in Experiment 1. In contrast, we expected no plausibility effects for the subject control sentences: *Subj-Implaus* and *Subj-Plaus* should be equally plausible because the "implausible" filler is not associated with the subject gap, but rather a later gap where it is plausible. Such results would support our hypothesis because the control properties of the matrix verbs cause the difference in plausibility between *Obj-Implaus* and *Subj-Implaus*. Crucially, plausibility effects are no longer confounded with lexical relatedness.

Method

Materials. We constructed 16 sets of sentences like the set in Fig. 4. All of the materials are in Appendix B. In each set, the object control pair is identical to the subject control pair except for the matrix verb and its control properties (and often the last few

<table>
<tr><th rowspan="2">CONDITION</th><th rowspan="2"></th><th colspan="7">Critical Word Positions</th></tr>
<tr><th>N</th><th>V1</th><th>to</th><th>V2</th><th>V2+1*</th><th>V2+2*</th><th>V2+3*</th></tr>
<tr><td>Obj-Implaus</td><td>Which horse did the</td><td>cowboy</td><td>signal</td><td rowspan="4">to</td><td rowspan="4">surrender</td><td rowspan="4">to</td><td rowspan="4">the</td><td rowspan="4">authorities</td></tr>
<tr><td>Subj-Implaus</td><td>Which horse did the</td><td>cowboy</td><td>refuse</td></tr>
<tr><td>Obj-Plaus</td><td>Which outlaw did the</td><td>cowboy</td><td>signal</td></tr>
<tr><td>Subj-Plaus</td><td>Which outlaw did the</td><td>cowboy</td><td>refuse</td></tr>
</table>

FIG. 4. A sample sentence set from Experiment 2 is shown with critical word positions specified. The four condition names indicate the verb type (object control or subject control) and the plausibility of the filler *as the subject of V2,* which is not the same as overall plausibility of the sentence. The Obj-Implaus condition is the only one in which the overall sentence is implausible. **The last three words were either a noun phrase or a prepositional phrase.*

words in the sentence). The Wh-phrase is the object of the matrix verb in the object control sentences, and also the subject of the infinitive (i.e., it is the "surrenderer"). In the subject control versions, the Wh-phrase is either the object of the infinitive or the object of a preposition later in the sentence, and the subject of the matrix verb, *the cowboy,* is the subject of the infinitive (i.e., it does the surrendering). Of the three nouns (*horse, outlaw,* and *cowboy*), *horse* is the only implausible surrenderer. The *Obj-Plaus* and *Subj-Plaus* versions are plausible by any account. Three-word endings immediately followed the embedded verb. Endings were either noun phrases or prepositional phrases. The reason for the different endings was that it was often not possible to use the same ending across conditions within a set and still have coherent sentences. This was necessary because in order to construct strong plausibility differences we often used two-place transitive verbs as the infinitive verb. In the subject control versions, the Wh-phrase filler was the object of these infinitive verbs, so the requirement of direct object was met. However, in the object control versions, the Wh-phrase filler was the subject of the infinitive verb so another (explicit) noun phrase was needed to fill the infinitive's object position. Thus, the object control sentences typically end with noun phrases while the subject control sentences typically end in prepositional phrases.

Eighty distractor sentences were constructed using non-control verbs. Distractors included Wh-sentences and declarative sentences with various kinds of sentential complements, with the bulk of them being simple and straightforward. The distractors were primarily sensible, but included some sentences that were nonsensible for semantic or syntactic reasons. Critical and distractor sentences were ordered semirandomly and the four versions of the critical sentences were rotated to form four lists, each with 96 sentences. The critical sentences were counterbalanced across lists so that one condition of every sentence set appeared on each list and every list had four sentences from each condition.

Subjects. Thirty-two undergraduates completed the experiment either in partial fulfillment of course requirements or for a minimal sum. All were native English speakers.

Procedure. The procedure used in Experiment 1 was used again here. To summarize, subjects controlled sentence presentation rate by pressing a button to get each word to appear on the computer screen. If the sentence stopped making sense, they pressed a "no" button and the trial ended. Subjects were told to read rapidly and carefully.

Results

A record of which button was pressed was kept for seven word positions in each critical sentence, beginning with the subject of the matrix verb and continuing until the end of the sentence or until the subject pressed a key to indicate that the sentence had stopped making sense. Reading times were also collected for button presses in the critical region. As in Experiment 1, we will refer to the percentage of "no" responses as the judgments and the reading times for the positive responses as the reading times.

Judgments. At each word position, a count was kept of the number of trials in each condition on which a subject responded "no," meaning that the sentence had stopped making sense. A summary of the cumulative percentage of "no" responses at each word position is shown in Fig. 5a. The percentage of "no" responses for the implausible object control sentences sharply increases at the infinitive verb and continues rising until the end of the sentence. There was no similar increase in any of the other conditions. Subjects responded "no" by the end of the sentence for 70% of the implausible object control trials, and 20, 14, and 16% of the plausible object control, implausible subject control, and plausible subject control trials, respectively. A 4(list)

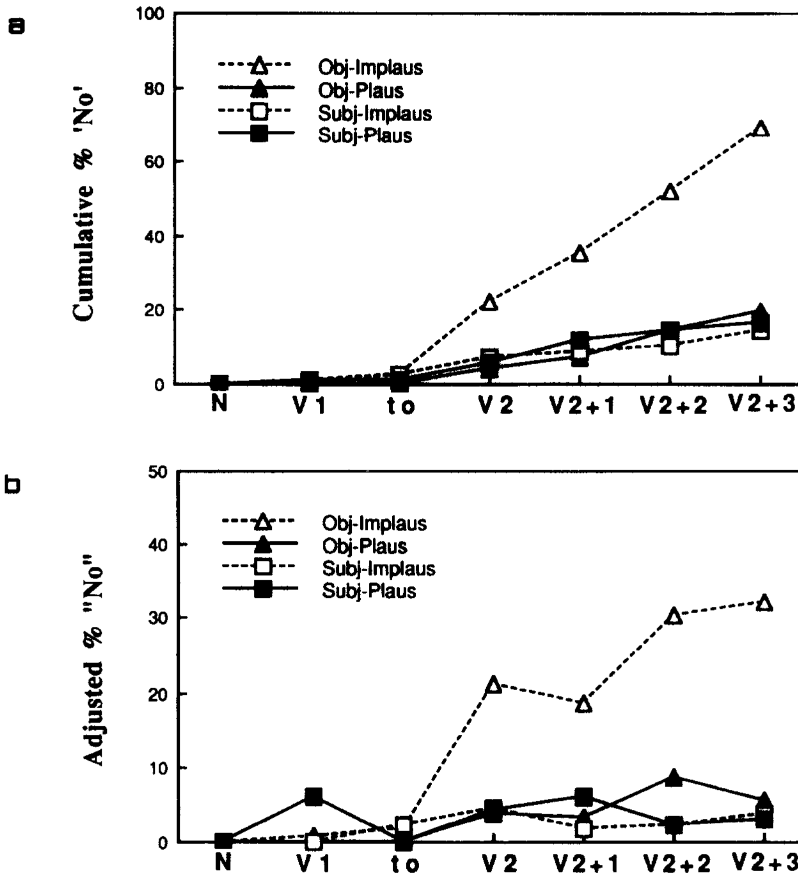

FIG. 5. (a) Cumulative percentage of "no" responses at each word position in Experiment 2. (b) Adjusted percentage of "no" responses at each word position in Experiment 2.

× 2(verb type) × 2(plausibility) ANOVA on this measure revealed main effects of plausibility [$F1(1,28) = 79.65$, $p < .01$, $MSe = 228$; $F2(1,12) = 16.29$, $p < .01$, $MSe = 557$] and verb type [$F1(1,28) = 96.74$, $p < .01$, $MSe = 284$; $F2(1,12) = 39.21$, $p < .01$, $MSe = 351$], which were due to a plausibility by verb type interaction [$F1(1,28 = 92.69$, $p < .01$, $MSe = 236$; $F2(1,12) = 58.61$, $p < .01$, $MSe = 187$].

As in Experiment 1, analyses across word position were done on the percentage of remaining possible "no"s rather than the simple percentage of "no"s in order to minimize the dependence of later values on earlier ones. The percentages of remaining possible "no"s are presented in Fig. 5b. Mean percentages by subjects and by items were submitted to a 4(list) × 2(verb type) × 2(plausibility) × 7(word position) ANOVA. There were significant main effects by subjects and by items for plausibility [$F1(1,28) = 29.03$, $p < .01$, $MSe = 414$, $F2(1,12) = 11.88$, $p < .01$, $MSe = 287$], verb type [$F1(1,28) = 36.59$, $p < .01$, $MSe = 418$; $F2(1,12) = 42.40$, $p < .01$, $MSe = 99$], and word position [$F1(6,168) = 23.79$, $p < .01$, $MSe = 187$; $F2(6,72) = 8.98$, $p < .01$, $MSe = 184$]. In both the subject and item analyses there were significant interactions for plausibility × verb type [$F1(1,28) = 35.34$, $p < .01$, $MSe = 376$; $F2(1,12) = 44.40$, $p < .01$, $MSe = 99$], verb type × word position [$F1(6,168) = 15.16$, $p < .01$, $MSe = 183$; $F2(6,72) = 6.32$, $p < .01$, $MSe = 130$], and plausibility × word position [$F1(6,168) = 11.00$, $p < .01$, $MSe = 162$; $F2(6,72) = 3.97$, $p < .01$, $MSe = 140$]. In addition, the three-way interaction of plausibility × verb type × word position was significant in the subject analysis [$F1(6,168) = 11.27$, $p < .01$, $MSe = 164$] and in the item analysis [$F2(6,72) = 5.09$, $p < .01$, $MSe = 109$].

Differences between the conditions at particular target positions were evaluated by making post hoc pairwise comparisons of both the subject and item means using the REGWF with $\alpha = .05$. Beginning at position V2 and continuing to the end of the sentence, there were more "no" responses in the implausible object control condition compared with all other conditions in both the subject analysis and the item analysis.

Reading times. Mean reading times are shown in Fig. 6. As in the judgment data, a plausibility effect was seen only for the object control conditions. In addition, the reading times for plausible object control sentences diverge from those of subject control sentences as position V2 + 2. Only positive responses were included in the reading time analysis. Mean reading times for each condition at each word position were computed by subjects and by items for use in the ANOVAs. The reading time data are subject to the same problems described in Experiment 1. At each word position, those values more than 2.5 standard deviations above or below the subject mean were truncated and replaced with the cutoff value. Cells with no positive responses were replaced in the same way as in Experiment 1. Missing values accounted for 2% of the subject means and .5% of the item means.

The means were submitted to a 4(list) × 2(verb type) × 2(plausibility) × 7(word position) ANOVA identical to that used for the judgment data. There was a significant main effect by subjects and by items for word position [$F1(6,168) = 26.62$, $p < .01$, $MSe = 56$; $F2(6,72) = 18.43$, $p < .01$, $MSe = 37$]. The main effect of verb type was significant in the subject analysis [$F1(1,28) = 12.58$, $p < .01$, $MSe = 97$] and marginal in the item analysis [$F2(1,12) = 4.73$, $p < .10$, $MSe = 77$]. There was also an interaction of verb type × word position in the subject analysis [$F1(6,168) = 3.64$, $p < .01$, $MSe = 42$], but not in the item analysis [$F2(6,72) = 1$].

Differences between conditions at particular word positions were again evaluated with the post hoc REGWF at $\alpha = .05$. In the subject analysis at V2 + 1, responses in the implausible object control condition were slower than those for the implausible subject control condition. However, a less conservative planned comparison test at V2 + 1 also found times for the implausible object control condition to be slower than those in the plausible object control condition in the subject analysis [$F1(1,28) = 6.91$, $p < .05$, $MSe = 80$]. The difference between plausible subject and object control conditions at V2 + 2 and V2 + 3 was

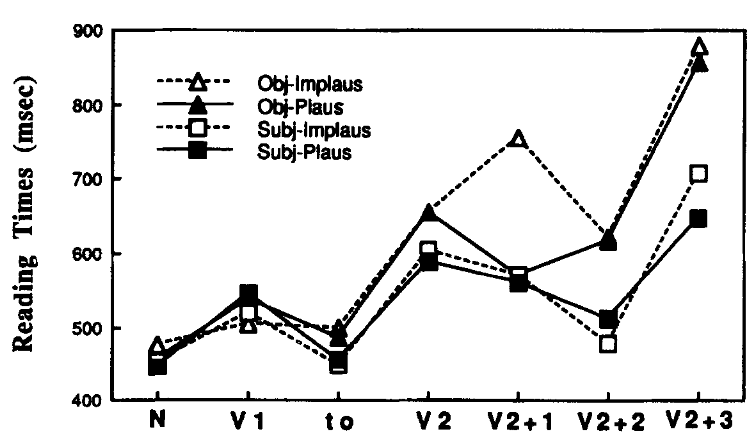

FIG. 6. Mean reading times for positive responses at each word position in Experiment 2.

nonsignificant in the post hocs. The difference was significant, though, in a less conservative planned comparison test; V2 + 2: [$F1(1,28) = 10.01, p < .01, MSe = 18$]; V2 + 3: [$F1(1,28) = 4.77, p < .05, MSe = 14$]. No comparisons were different in the item analysis. The item means are subject to the same problems discussed in the results section of Experiment 1. As in Experiment 1, when scores were replaced for each subject, the item analyses showed the same statistical pattern as the subject analyses.

Discussion

As predicted, a plausibility effect was seen for the object control versions, but not the subject control versions. This cannot be due to lexical level effects since the same noun phrases and infinitive verbs were used in the subject control sentences as the object control sentences. Thus, verb control was clearly used to make the correct coindexing assignments. The timing of the plausibility effects was the same as that obtained in Experiment 1, indicating again that verb control information is used immediately.

Note also the slower reading times at the last two word positions in the plausible object control condition compared to the plausible subject control condition. If there is nothing intrinsically more difficult about object control sentences compared to subject control sentences, then the plausible object control sentences should be read just as quickly as the subject control conditions. The difference, therefore, suggests that object control sentences may be more difficult than subject control sentences, although the difficulty arises late, after the subject gap has been correctly associated with a filler. Similar results were reported by Crain and Fodor (1985). Such an effect is important because it could explain the subject control advantage found by Frazier et al. (1983) and Clifton and Frazier (1986) with end-of-sentence response times. A reason to be cautious about such an interpretation, though, is that for most of the sentence sets, the words in the last three positions were different for the subject and object control versions (see Appendix B). The reason for these differences was explained in the description of the materials. We cannot rule out the possibility that the wording difference could account for the difference in reading time between plausible subject and object control conditions at these word positions. However, given the results of Crain and Fodor, it seems likely that our results reflect a genuine complexity difference between subject and object control sentences.

General Discussion

The two experiments reported here examined how verb control information is used to interpret the empty subject of infinitive complements. Our experiments were motivated by the Frazier et al. (1983) claim that initial coindexing decisions are made without reference to control information, but instead use the MRF strategy in which the most recent possible filler is interpreted as the understood subject of an infinitive clause. This strategy predicts that people will initially misinterpret the subject gap in Wh-questions such as (8) below. The mistake is predicted because the subject of the matrix verb *signal* (i.e., *the cowboy*), is the most recent filler, but the control properties of *signal* specify that its object (i.e., the fronted Wh-phrase, *which outlaw*) must be the subject of the infinitive. In contrast, following the MRF strategy would result in the correct interpretation of the subject gap in declarative sentences such as (9).

(8) Which outlaw$_i$ did the cowboy signal $\_\__i$ $\_\__i$ to surrender to the authorities?

(9) The cowboy signalled the outlaw$_i$ $\_\__i$ to surrender to the authorities.

We tested the MRF strategy by comparing Wh-questions like (8) with declarative sentences like (9). The sentences become implausible when the subject gap in the infinitive complement is correctly interpreted. These materials allowed us to address two related but independent questions. The first question is whether the correct (object control) interpretation of the subject gap is delayed in Wh-questions relative to declarative sentences. We found no evidence that this is so, which suggests that control information guides the interpretation of the subject gap. The second question concerns how early in processing control information is used. The reliable plausibility effects in the judgment times of Experiments 1 and 2 at the embedded verb are evidence that verb control information is used very early, if not immediately. Given our results, it seems unlikely that the difficulty of sentences with object control verbs reported by Frazier et al. (1983) and replicated by Clifton and Frazier (1986) was due to a temporary garden path caused by associating the wrong filler with the subject gap.

As with any on-line task, though, it is important to consider the extent to which the results might be due to the task demands placed on the subject. The sense-monitoring task raises two concerns. The first is that the reading times reported are slower than those obtained in other self-paced reading tasks and, crucially, slower than the experimenter-paced presentation times used by Frazier et al. (1983). While it is possible that the slower reading times allowed time for verb control information to be made available this is unlikely given the findings of Clifton and Frazier (1986). They compared presentation rates of 350 ms per word (as was used by Frazier et al.) and 550 ms per word (which is more comparable to the reading times reported here) and found that both presentation rates elicited the subject control advantage reported in Frazier et al. Thus, reading time per se is unlikely to be the relevant variable.

Potentially more problematic is the sense-monitoring aspect of our task. It is important to emphasize that subjects were not explicitly instructed to make plausibility judgments at each word in the sentence. Rather they were told to read through the sentences, pressing a button for each word, and to press a different button if the sentence stopped making sense. While we believe that our task reflects processing typical of careful reading, subjects may have developed interpretations for sentences more rapidly in our experiments than in experiments using other tasks. This could conceivably have created a situation in which it would be possible for control information to over-ride the MRF strategy. However, this is unlikely because other studies in our lab which used this task to examine the effects of context on the processing of other types of sentences have found evidence for garden-pathing, even when context might have been expected to prevent the garden paths. Thus, it is not the case that this task is simply insensitive to local, structurally motivated garden paths. We can only conclude that there was no garden path because verb control information was correctly used. And if verb control information can be correctly used, why should the system resort to the MRF strategy? Such a heuristic could be motivated by a need to make early coindexing commitments combined with a delay in availability of verb control information. However, there is no reason to believe that the processing system would be forced into coindexing commitments prior to the availability of verb control information since no structural commitments rest on the coindexing of the subject gap (Nicol & Osterhout, 1989). What we cannot rule out is the possibility that our task hastened the availability of verb control information. But this does not compromise the evidence reported here against the MRF strategy.

However, one would still like to explain why Frazier et al. (1983) obtained the results they did. The reading time data from Experiment 2 may provide an answer which is unrelated to any processing strategies which the different tasks might have induced. The explanation is suggested by the longer reading times for the plausible object control sentences than for the plausible subject control sentences late in the sentence after the region where the plausibility results clearly demonstrated that control information had been correctly used. Thus, Experiment 2 provides some evidence for a complexity difference between object and subject control sentences that is unrelated to any misuse of control information. Frazier et al.'s (1983) end-of-sentence measure probably reflects this complexity difference rather than a failure to use verb control information.

There still remains the question of why one would find this complexity difference. Fodor (1988) has suggested that object control questions are more complex than subject control questions for semantic reasons. One possible semantic explanation is that the event structure denoted by an object control verb with an infinitive complement is more complex than the event structure of a subject control verb. Constructing a mental model or event structure might thus be more costly in terms of processing resources for an object control sentence than for a subject control sentence. For a subject control verb, the same entity functions as the "doer" in two related events, the event denoted by the matrix verb and the embedded event denoted by the verb in the infinitive complement. In contrast, for an object control verb, the two events have different "doers" and there is an additional relationship because the subject of the matrix verb is in some sense responsible for the embedded event (following Farkas, 1988). If an explanation along these lines is correct, the complexity difference that has been used to argue against the early use of control actually provides evidence that people use the semantic information provided by control to rapidly construct a representation of the events described in the sentence.

Evidence against the MRF strategy has also been obtained in a recent series of experiments reported by Nicol and Osterhout (1989) (see also Nicol & Swinney, 1989). They tested for the reactivation of antecedents for subject and object gaps using a cross-modal lexical decision paradigm. For object gaps, they consistently found priming only for the correct filler regardless of distance. In contrast, they did not find immediate priming for the correct filler for the implied subject of the infinitive complement. In fact, they did not find any pattern of priming for only the correct filler of the subject gap even by the end of their sentences. Instead they found some evidence that both correct and incorrect fillers were primed. Nicol and Osterhout concluded that control information is used correctly but that the interpretation of subject gaps is relatively slow. While this conclusion differs from ours, it is difficult to evaluate the Nicol and Osterhout claims in the absence of clear priming patterns for the infinitive subject gap.

What are the implications for the modularity debate of our demonstration that control information is used correctly and early in sentence processing? We would argue that in fact there are none. If Frazier et al. (1983) had been correct in their claim that coindexing decisions are made without reference to control information, that would have been clear evidence that initial syntactic processing does not make use of at least one aspect of combinatory lexical information. Moreover, it would have been the strongest type of evidence that can be offered in support of a modular system, namely, evidence that relevant information from one module is not used by another module. However, evidence that control information is used correctly does not provide evidence for or against the general hypothesis that the comprehension system is modular. It is only relevant to the particular hypothesis explored in Frazier et al. We do not mean to say that verb control phenomena no longer suggest interesting questions about language processing. On the contrary, we believe the study of the use of verb control will continue to be an important area, since it provides a domain in which to examine how the syntactic and semantic structures associated with words, and verbs in particular, are used during sentence processing. The evidence that we have reported here about the immediate and correct use of verb control information complements other recent studies (Boland et al., in press; Marslen-Wilson et al., 1988; Tanenhaus et al., 1989; Trueswell, Tanenhaus, & Garnsey, 1989) demonstrating that the language processing system fully exploits the combinatory lexical information associated with verbs to rapidly coordinate a range of different information types during language comprehension.

APPENDIX A

Materials from Experiment 1

The first line in each set contains the plausible and implausible question conditions. The second line contains the plausible and implausible declarative conditions. In each case, the implausible filler is given first, followed by the plausible filler.

1 Which planaria/parrot did the professor train to tell really stupid jokes?
The professor trained the planaria/parrot to tell really stupid jokes.

2 Which executive/assets did the shifting market cause to devaluate early this year?
The shifting market caused the executive/assets to devaluate early this year.

3 Which snake/frog did the boy force to hop over the rock?
The boy forced the snake/frog to hop over the rock.

4 Which pet/girl did the boy help to organize the club meeting?
The boy helped the pet/girl to organize the club meeting.

5 Which dog/private did the sergeant command to mop the washroom floor?
The sergeant commanded the dog/private to mop the washroom floor.

6 Which janitor/president did the legislator convince to veto the suffrage amendment?
The legislator convinced the janitor/president to veto the suffrage amendment.

7 Which employer/workers did the union order to strike during the dispute?
The union ordered the employer/workers to strike during the dispute.

8 Which cow/son did the farmer urge to lock the barn door?
The farmer urged the cow/son to lock the barn door.

9 Which cat/puppy did the girl teach to bark on her command?
The girl taught the cat/puppy to bark on her command.

10 Which horse/outlaw did the cowboy signal to surrender his weapons quietly?
The cowboy signalled the horse/outlaw to surrender his weapons quietly.

11 Which child/mobster did the psychopath persuade to assassinate the new president?
The psychopath persuaded the child/mobster to assassinate the new president.

12 Which calculator/robot did the genius program to dance like a ballerina?
The genius programmed the calculator/robot to dance like a ballerina.

13 Which submarine/airplane did the pilot radio to fly south of Moscow?
The pilot radioed the submarine/airplane to fly south of Moscow.

14 Which customer/doctor did the pharmacist warn to prescribe fewer strong narcotics?
The pharmacist warned the customer/doctor to prescribe fewer strong narcotics.

15 Which mechanic/secretary did the author hire to type his new book?
The author hired the mechanic/secretary to type his new book.

16 Which dean/student did the professor remind to study for the exam?
The professor reminded the dean/student to study for the exam.

17 Which dog/man did the neighborhood rely on to organize its block parties?
The neighborhood relied on the dog/man to organize its block parties.

18 Which theatre/actress did the playwright pressure to sing in his play?
The playwright pressured the theatre/actress to sing in his play.

19 Which snake/man did the performer coax to sing in the nightclub?
The performer coaxed the snake/man to sing in the nightclub.

20 Which elephant/recruit did the trainer require to say his name clearly?
The trainer required the elephant/recruit to say his name clearly.

21 Which patient/doctor did the intern assist to transplant the man's heart?
The intern assisted the patient/doctor to transplant the man's heart.

22 Which concert/flutist did the conductor direct to play the featured solo?
The conductor directed the concert/flutist to play the featured solo.

23 Which toddler/teenager did the parents permit to babysit their only child?
The parents permitted the toddler/teenager to babysit their only child.

24 Which sister/brother did the invalid pay to father his firstborn child?
The invalid paid his sister/brother to father his firstborn child.
25 Which bachelor/client did the therapist advise to divorce his cheating wife?
The therapist advised the bachelor/client to divorce his cheating wife.
26 Which prisoner/jury did the lawyer inspire to convict the innocent defendant?
The lawyer inspired the prisoner/jury to convict the innocent defendant.
27 Which baby/grandson did the grandma encourage to discuss current political events?
The grandma encouraged the baby/grandson to discuss current political events.
28 Which reptile/visitor did the zookeeper allow to photograph the baby seal?
The zookeeper allowed the reptile/visitor to photograph the baby seal.

APPENDIX B

Materials for Experiment 2

The first line of each set contains the object control version with implausible and plausible fillers. The second line contains the subject control versions with the same two fillers.

1 Which baby/grandson did the grandma encourage to guide the pony away?
Which baby/grandson did the grandma attempt to guide toward the pony?
2 Which horse/outlaw did the cowboy signal to surrender to the authorities?
Which horse/outlaw did the cowboy refuse to surrender to the authorities?
3 Which infant/volunteer did the nurses teach to protect the other workers?
Which infant/volunteer did the nurses continue to protect from getting ill?
4 Which pet/twin did the girl urge to draw in the book?
Which pet/twin did the girl pretend to draw from the book?
5 Which mechanic/assistant did the supervisor hire to phone the potential clients?
Which mechanic/assistant did the supervisor plan to phone about the payments?
6 Which janitor/foreman did the boss pressure to fire his associate workers?
Which janitor/foreman did the boss agree to fire with the others?
7 Which criminal/policeman did the sheriff convince to arrest the new suspect?
Which criminal/policeman did the sheriff start to arrest at the scene?
8 Which caterpillar/aunt did the child coax to imitate a movie star?
Which caterpillar/aunt did the child learn to imitate in this way?
9 Which kitten/suitor did the lady permit to hug the big teddy-bear?
Which kitten/suitor does the lady like to hug in the evening?
10 Which monkey/veterinarian did the zookeeper allow to send the new medicine?
Which monkey/veterinarian did the zookeeper forget to send the new cage?
11 Which victim/worker did the firechief order to rescue the burning ladder?
Which victim/worker did the firechief hope to rescue with the ladder?
12 Which prisoner/jury did the lawyer direct to convict the defending attorney?
Which prisoner/jury did the lawyer swear to convict after the defense?
13 Which beggar/executive did the banker persuade to loan his son money?
Which beggar/executive did the banker decide to loan a small sum?
14 Which toddler/girl did the neighbor remind to babysit on Saturday night?
Which toddler/girl did the neighbor remember to babysit on Saturday night?
15 Which steer/wrangler did the cowhand force to wrestle in the corral?
Which steer/wrangler did the cowhand refuse to wrestle without a rope?
16 Which country/group did the diplomat warn to leave the others alone?
Which country/group did the diplomat intend to leave in early January?

REFERENCES

BOLAND, J. E., TANENHAUS, M. K., CARLSON, G., & GARNSEY, S. M. (in press). Lexical projection and the interaction of syntax and semantics in parsing. *Journal of Psycholinguistic Research.*

CHOMSKY, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

CLIFTON, C., JR., & FRAZIER, L. (1986). The use of syntactic information in filling gaps. *Journal of Psycholinguistic Research,* **15**, 209–224.

CRAIN, S., & FODOR, J. D. (1985). How can grammars help parsers? In D. R. Dowty, L. Kartunnen, & A. M. Zwicky (Eds.), *Natural language processing: Psychological, computational, and theoretical perspectives.* New York: Cambridge University Press.

EINOT, I., & GABRIEL, K. R. (1975). A study of the powers of several methods of multiple comparison. *Journal of the American Statistical Association,* **70**, 574–583.

FARKAS, D. (1988). On obligatory control. *Linguistics and Philosophy,* **11**, 27–58.

FODOR, J. A. (1983). *Modularity of mind.* Cambridge, MA: The MIT Press.

FODOR, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9, 427–474.

FODOR, J. D. (1988). On modularity in syntactic processing. *Journal of Psycholinguistic Research*, 17, 125–163.

FORD, M. & DALRYMPLE, M. (1988). A note on some psychological evidence and alternative grammars. *Cognition*, 29, 63–71.

FRAZIER, L. (1987). Theories of syntactic processing. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language processing*. Cambridge, MA: MIT Press.

FRAZIER, L. (1989). Against lexical generation of syntax. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

FRAZIER, L., CLIFTON, C., JR., & RANDALL, J. (1983). Filling gaps: Decision principles and structure in sentence comprehension. *Cognition*, 13, 187–222.

GARNSEY, S. M., TANENHAUS, M. K., & CHAPMAN, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, 18, 51–60.

HUYNH, H., & FELDT, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1, 69–82.

MARSLEN-WILSON, W. D., BROWN, C., & TYLER, L. K. (1988). Lexical representations in spoken language comprehension. *Language and Cognitive Processes*, 3, 1–16.

MARSLEN-WILSON, W. D., & TYLER, L. K. (1987). Against modularity. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language processing*. Cambridge, MA: MIT Press.

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

MITCHELL, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.

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

NICOL, J., & OSTERHOUT, L. (1989). *Re-activating antecedents of empty categories during parsing*. Manuscript submitted for publication.

NICOL, J., & SWINNEY, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18, 5–20.

RYAN, T. A. (1959). Multiple comparisons in psychological research. *Psychological Bulletin*, 56, 26.

SELLS, P. (1985). *Lectures on contemporary syntactic theories*. Stanford, CA: CSLI.

STOWE, L. A., TANENHAUS, M. K., & CARLSON, G. N. (1989). *Filling gaps on-line: Use of lexical and semantic information in sentence processing*. Manuscript submitted for publication.

TANENHAUS, M. K., BOLAND, J. E., GARNSEY, S. M., & CARLSON, G. N. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research*, 18, 37–50.

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

TANENHAUS, M. K., CARLSON, G. N., & SEIDENBERG, M. S. (1985). Do listeners compute linguistic representations? In D. R. Dowty, L. Kartunnen, & A. M. Zwicky (Eds.), *Natural language processing: Psychological, computational, and theoretical perspectives*. New York: Cambridge University 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.

TANENHAUS, M. K., GARNSEY, S. M., & BOLAND, J. E. (in press). Combinatory lexical information and language comprehension. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

TRUESWELL, J., TANENHAUS, M. K., & GARNSEY, S. M. (1989). *Semantic influences on parsing: Use of thematic role information in syntactic disambiguation*. Manuscript submitted for publication.

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

WELSCH, R. E. (1977). Stepwise multiple comparison procedures. *Journal of the American Statistical Association*, 72, 359.

WINER, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.

(Received July 5, 1989)

(Revision received November 15, 1989)