ARGUMENT STRUCTURE FRAMES: A LEXICAL COMPLEXITY METRIC?

A Thesis Presented

by

ANNA R. SCHMAUDER

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 1989

Psychology

ARGUMENT STRUCTURE FRAMES: A LEXICAL COMPLEXITY METRIC?

A Thesis Presented

by

ANNA R. SCHMAUDER

Approved as to style and content by:

Charles Clifton, Chairperson of Committee

Lyn Frazier, Member

Keith Rayner, Member

Seymour Berger, Department Head
Department of Psychology

ACKNOWLEDGEMENTS

I thank my committee members, Chuck Clifton, Lyn Frazier, and Keith Rayner, for their stimulating comments and discussion throughout the course of this research. Keith Rayner was a valuable source of information on eyetracking. Thanks to Lyn Frazier for bringing up some of the Big Issues in language processing and asking me to consider them. And I thank the chair of my committee, Chuck Clifton, for being so generous with his time, patience, suggestions, and snacks. I cannot imagine a better advisor.

Work on this thesis was supported in part by a pre-doctoral traineeship from NIMH and by a Class of 1905 Alumnae Fellowship from Mount Holyoke College.

Many of the ideas that led to this research arose during meetings with Chuck Clifton, Shari Speer, and John Huitema. The course I chose to pursue in this thesis is far less ambitious than those which we originally discussed; however, I have used many of the original objections brought up during these meetings.

Work in the eyetracking lab during the last two years may not have proceeded were it not for Robin Morris's expertise. I am long overdue in thanking her for teaching me to use the equipment and for consulting on the oddities of eyetracking.

For many wonderful conversations about cognitive psychology and for their support, I thank Shari Speer, Celia Klin, John Huitema, Ruth Litovsky, Sara Sereno, and Susan Boyce.

Finally, I thank my family for their support and Jim Wolper for his understanding and encouragement, and for being my best friend.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

According to one current view, the human sentence processing system consists of several functionally distinct modules (Fodor, 1983; Frazier, to appear). One module, the lexicon, contains lexical entries for all words known to the language user. Each lexical entry is thought to contain a minimal amount of information necessary for identifying and using the word represented in that lexical entry, but exactly how much and what kind of information is stored in a lexical entry is not clear. Specifying the amount and kind of information stored in lexical entries is relevant to research on sentence processing and to research aimed ultimately at describing the interaction between modules which make up the language processing system. However, discovering what kind of information is stored in lexical entries does not yield the answer to the further question of <u>how</u> information in the lexicon is used during sentence processing. And we need to know how lexical information is used during sentence processing before we can draw conclusions about modules within the language processor. The experiments reported here examine lexical information to determine which sorts of lexical information influence sentence processing and how the influence occurs.

Representational complexity studies, and, in particular, lexical complexity studies, provide one means for looking at the influence of representations on sentence processing. Previous work on representational complexity has investigated several complexity metrics. For instance, researchers have classified verbs according to the number of syntactic subcategorizations they call for or according to verbs' causativity, negativity, or factivity. However, few of these studies found any effect of lexical complexity on processing, and when effects of lexical complexity were found, often the studies were flawed (see below). Furthermore, few lexical complexity findings indicate that lexical complexity influences construction of the initial mental representation of a sentence.

Fodor, Garrett, and Bever (1968) tested for an influence of verbs' syntactic subcategorization frames on sentence processing. As part of their lexical entries, verbs have information about the syntactic class of the complements that can occur within the verb phrase of which they are heads. This syntactic information is encoded as a subcategoriztion frame. For instance, a verb like leave, which is optionally intransitive, has the subcategorizaton frame: leave [__(NP)], while a purely transitive verb like slap has a subcategorization frame that reflects the requirement that a NP follow the verb: slap [__NP]. Fodor, Garrett, and Bever (1968) compared pure transitive verbs (verbs which only subcategorize for a NP complement) to sentential complement verbs (verbs which subcategorize for S' in addition to NP). In one experiment, subjects heard a test sentence with two levels of self-embedding which contained either a sentential-complement verb or a transitive verb. Subjects paraphrased the test sentence as quickly and accurately as possible, and accuracy of paraphrase was recorded. Fodor, Garrett, and Bever found lower paraphrase accuracy for sentences containing verbs that took sentential complements than for sentences containing simple transitive verbs.

In a second experiment, Fodor, Garrett, and Bever compared transitives to sentential-complement verbs using an anagram task (subjects constructed a sentence from scrambled words). Subjects were more likely to create incorrect sentences and they failed to complete the task more often when the sentence to be constructed contained a sentential-complement verb than when the sentence contained a transitive verb. Fodor, Garrett, and Bever concluded that syntactic subcategorization frames associated with verbs' lexical entries affect sentence processing.

However, the sentence-paraphrase task and the anagram task themselves may have been the source of the processing complexity effect. That is, even if there was no effect of verb complexity on sentence comprehension, Fodor, Garrett, and Bever might have found an effect of verb complexity on paraphrasing or on anagram solution. These two tasks simply may not reflect immediate processing difficulty during sentence comprehension.

Hakes (1971) proposed that Fodor, Garrett, and Bever's tasks were insensitive to immediate processing difficulty. He undertook experiments to compare Fodor, Garrett, and Bever's paraphrase task results to results from a phoneme monitoring study. As in Swinney and Cutler's phoneme-monitoring study (reported in Cutler, 1983, and discussed on pp. 4-5 of this thesis), subjects in Hakes' phoneme-monitoring experiments listened for a word beginning with a prespecified phoneme while listening to a sentence for comprehension. The logic was that phoneme monitoring taps the same processes used for sentence comprehension, so sentences that are more difficult to comprehend should leave fewer resources available for use in the phoneme-monitoring task, leading to longer response latencies.

Hakes tested verbs that take simple NP objects (Transitives) against verbs that take both simple NP objects and NP complements (Complement verbs). Subjects performed either the phoneme monitoring task or the paraphrase task used by Fodor, Garrett, and Bever. Hakes found an effect of verb complexity with the paraphrase task, but he did not find evidence that verb complexity affects comprehension with the phoneme-monitoring task. Hakes concluded that verb complexity affects sentence processing, but that the effect is a structural complexity effect that only was detectable using the paraphrase task. On this account, phoneme monitoring reflected procedures that occur at a point earlier than that at which the verb complexity effect showed up.

These results suggest that syntactic subcategorization information does not guide initial parsing. Other researchers have asked whether semantic information influences lexical retrieval and integration of lexical items. Swinney and Cutler, reported in Cutler (1983), used a phoneme-monitoring task to determine whether factive words were more lexically complex than non-factive controls. Factives verbs imply that their sentence complements express true propositions, and Cutler argued that a word's factivity is a necessary part of the word's definition and thus of the word's lexical entry. Swinney and Cutler presented factive verbs (regret, know) and adjectives (important, crazy) in sentence contexts like:

> (1) The retired general deplored/declared a continued readiness for war on the part of the NATO partners.
> (deplored is factive, declared is nonfactive)

The target phoneme was /k/, and the critical word was "continued." No significant difference in reaction times to the target phoneme obtained when the word containing the target was preceded by a factive word as opposed to a non-factive control.

Using a classification task (subjects were asked: Is the sentence acceptable?), Cutler also found no effect of factivity when the same words as used by Swinney and Cutler were presented in sentences like:

> (2) The retired general deplored/declared the army's readiness for war.

From these results and the results of other experiments investigating lexical complexity measures, Cutler concluded that complex lexical entries do not adversely affect processing. Cutler's conclusion that factivity as a lexical complexity metric does not affect lexical access was based on differences reflected by phoneme monitoring response times. However, it is not clear what processes are tapped by the phoneme monitoring task (Foss and Gernsbacher, 1983; Cutler and Norris, 1979, Mehler, Segui, and Carey, 1978, Newman and Dell, 1978).

These problems with the phoneme monitoring task prompted Rayner and Duffy's (1986) investigation of verb complexity by monitoring readers' eye movements and examining fixation time on verbs during silent reading. They matched causative verbs (e.g., kill = cause to die) with noncausative verbs (in parentheses) in sentences like:

> The policeman frightened (encountered) the little girl.
> Paul never convinced (understood) the new president

Factives were compared to nonfactive verbs (in parentheses):

> The girl noticed (insisted) that the cake was moldy.
> The maid forgot (implied) that the sailor had left.

And they compared negative verbs (verbs whose lexical representations contain a negative element) to non-negatives (in parentheses):

> The teacher despised (rewarded) the unhappy child.
> The fireman ignored (advised) the town council.

Rayner and Duffy predicted that if the causative, factive, or negative complexity factors led to immediate processing difficulty, fixation times on the target verbs would be longer than fixation on matched controls.

Rayner and Duffy found that representational complexity of negative, causative, and factive verbs did not result in longer fixation time on the verbs (relative to their controls), a finding consistent with Cutler's (1983) conclusions. There were no effects of causativity, factivity, or negativity on either first fixation duration or gaze duration on the target verbs. However, Rayner and Duffy found that fixations on the word following negative verbs were longer than fixations on the word following positive verbs. They suggest that verbs with complex lexical representations may be more difficult to integrate into a sentence context once the lexical representation for the verb has been accessed.

Inhoff (1985) used eye movement records to investigate whether lexical presuppositions of factive verbs influence processing of their complements. He contrasted sentences with factive verbs and false complements with sentences identical but for substitution of nonfactive verbs for factive verbs. A complement was false because it was incompatible with the presumption of truth made by a factive verb; the complement was inconsistent either with a subject's world knowledge or with her definitional knowledge. In the example passage below, the complement "that two and two equals three" is incompatible with world knowledge; therefore the complement is false.

Subjects read 16 experimental passages containing a sentence with either a factive or a nonfactive verb and a false complement (passages contained no capitalization):

> tom and ann were the best
> first grade students in class.
> today was an arithmetic test.
> the teacher asked little tom. he knew/said
> that two and two equals three.
> does two and two equal three?

In addition, subjects saw 26 filler stories.

Inhoff (1985) looked for lexical complexity effects both on the verb and on the word in the complement that made the complement false ("three" in the above example). The results for both regions showed no significant differences in first fixation durations for factive versus nonfactive versions of the test passages either on the verb or on the critical word in the complement. The gaze duration measure also indicated no reliable difference in the verb region. However, analysis of gaze durations on the critical word in identical false complements of the verb indicated false complements incurred significantly longer gaze durations following factive verbs compared to when they followed a nonfactive verb.

On half of the trials, Inhoff included a condition in which a three-letter mask blocked the central characters of each fixation made by a subject. The mask moved with subjects' eyes as they fixated different regions. Inhoff presumed that the central mask would combine additively with reading times for critical regions if factivity influences lexical access only (Inhoff, 1984; Gordon, 1983). Inhoff's previous (1984) finding that the central mask interacts with contextual integration processes implied that the mask would interact with reading times for critical regions if factivity influences integration.

The reliable gaze duration effect in false complements following factive versus nonfactive verbs was accompanied by a mask-by-factivity interaction that approached significance. Inhoff concluded that the gaze duration effect in the complement shows that the presumed truth of the complement is included in the lexical entry for factive verbs, and, further, that this presuppostion influences text integration, as evidenced by the nearly significant mask-by-factivity interaction.

In summary, this research shows syntactic subcategorization frame complexity and causativity are not appropriate lexical complexity metrics. While a verb's negativity causes delayed processing difficulty (Rayner and Duffy, 1986), there is no evidence that negativity influences retrieval of the lexical representation of the verb. Nor does factivity influence retrieval of the verb's lexical entry (Inhoff's (1985) results only show influences of factivity on post-lexical-access integration).

In light of these results, Shapiro, Zurif, and Grimshaw's (1987) report of a lexical complexity effect (which, they claim, influences retrieval of a verb's lexical entry[1]) was surprising. Their finding merits further attention in light of current interest in modular models of the human sentence processing system.

Shapiro, Zurif, and Grimshaw (1987) investigated the effects on processing of two possible complexity metrics for verbs. Shapiro et al. pitted representational complexity due to different syntactic subcategorization possibilities for different verbs against representational complexity due to verbs' having more <u>argument structure</u> possibilities. Argument structures associated with verbs specify the number of semantic arguments each verb can take. Because thematic roles can be "realized differently in the syntax," (as seen in Shapiro et al.'s sentences (10) and (11):

(10) Joe [$_{VP}$sent [$_{NP}$the letter] [$_{VP}$to Sheldon]].

(11) Joe [$_{VP}$sent [$_{NP}$Sheldon] [$_{NP}$the letter]].),

Shapiro et al. represent possible arguments of verbs using variables whose values are the particular thematic roles which occur when the

---

[1]See general conclusions of Shapiro et al. (1989), where they say that their effects "[parallel] reports of the exhaustive retrieval of the multiple interpretations of polysemous nouns in contextually (referentially) biased sentences (Swinney, 1979, Tanenhaus, Leiman, and Seidenberg, 1979)." (Shapiro et al., 1989, p. 242)

verb is used. The verb categories used by Shapiro et al. (1987) are reported in Table 1, and their verbs can be found in Table 2.

Shapiro et al. found that only one type of structural information associated with verbs' lexical entries contributed to difficulties in on-line processing: they observed longer lexical decision times in a cross-modal lexical decision (CMLD) task as a function of argument structure complexity but not as a function of syntactic subcategorization structure complexity.[2] The mean reaction times (in msec) for verb types, collapsed over verbs and sentence types, were:

Transitives (626) < Nonalternating datives (672) = Alternating datives (679) = Two-complements (676) < Four-complements (731)

Shapiro et al. (1987) further showed that the number of different possible argument structure frames/arrangements (rather than the maximum number of arguments included in each argument structure frame) is what influenced competition for processing resources. They compared transitives (argument frame = (x,y)), datives that allow an optional third argument (in addition to (x,y), they have the argument frame (x,y,z)), and verbs which obligatorily take three semantic arguments (e.g., hand: Tom handed the money to the police) Obligatory three-place verbs have only the argument

---

[2] Shapiro et al's CMLD task differed from the cross-modal lexical decision priming task used by Swinney (1979) and Onifer and Swinney (1981) in that the lexical decision probe was unrelated to sentence context, so there was no priming relation between words in the aurally presented sentence and the visually presented LD probe.

frame (x,y,z). The mean reaction times obtained (in msec) for verb types, collapsed over verbs and sentence types, were:

Transitives (622) = Obligatory three-place (606) < Datives (647)

In another paper, Shapiro and Levine (1989) confirmed the argument structure complexity effect of Shapiro et al. (1987) and also showed that the complexity effect was no longer detected at a point about four syllables past the offset of the verb.[3]

From these results, Shapiro et al. (1987, 1989) and Shapiro and Levine (1988) drew the conclusion that several or all argument frames were activated initially and that this activation was reflected in processing cost. They found no evidence for syntactic subcategorization complexity in the experiments reported in Shapiro et al. (1987) and did not test for syntactic subcategorization complexity in later experiments. Also, Shapiro et al. (1987) claimed that if argument frames are the relevant processing complexity metric, the lexicon may be "organized primarily by representations referring to argument structure"(p. 244) and not by representations referring to syntactic subcategorization frames. They suggested that all semantic argument frames associated with a verb were activated immediately when the verb's lexical entry was activated; an effect similar to the immediate, multiple access of several meanings of polysemous nouns shown by Swinney (1979), Onifer and Swinney (1981), and Tanenhaus, Leiman, and Seidenberg (1979) using cross-modal, lexical-decision priming and cross-modal naming priming paradigms.

---

[3]Personal communication with L. Shapiro, 11 October, 1988.

It is not clear whether this means that the lexicon is ordered in a way that determines difficulty of access or whether argument structures guide post-access integration. In conclusion, they said: "early stages of sentence processing -- either access or sentence integration -- are claimed to act only on information that systematizes lexical entries -- that is, argument structure"(Shapiro et al., 1987, p. 244).

Shapiro, Zurif, and Grimshaw (1989) further explored the claim that all argument frames associated with a verb are immediately and exhaustively activated even when the sentence context in which the verb occurs biases the reader towards one possible argument structure. In the first experiment, transitives and datives (they mixed nonalternating and alternating datives in this category) were presented in two sentence frames, passivized clefts and questions:

(A) Passivized clefts:
Transitive:
It was [$_{PP}$for the boy] that [$_{NP}$the bike] was fixed yesterday.

Dative:
It was [$_{PP}$to the girl] that [$_{NP}$the letter] was sent last week.

(B) Questions:
Transitive:
[$_{PP}$For whom] was [$_{NP}$the car] fixed last week.

Dative:
[$_{PP}$To whom] was [$_{NP}$the box] sent yesterday.

Subjects listened to sentences and perfomed a CMLD task identical to the task in Shapiro et al. (1987). Results showed that transitive verbs (653 msec) led to quicker lexical decision responses that did datives (702 msec).

A second experiment using the same task compared two-complements to four-complements in passive and question sentence contexts:

(A) Passive:
Two-complement:
[$_{S'}$That the answer was wrong] was accepted by the boy.

Four-complement:
[$_{S'}$That the picture was missing] was discovered by the girl.

(B) Question:
Two-complement:
[$_{NP}$What plan] did the mayor accept for the empty lot?

Four-complement:
[$_{NP}$What secret] did the boy discover in the cave?.

Two-complement verbs led to faster responses on the lexical decision task (644 msec) than did four-complement verbs (718 msec).

Shapiro et al. (1989) conclude that the results provide further evidence for the argument frame complexity effect and new evidence supporting their claim that all of a verb's argument structure possibilities are activated immediately, even in instances in which sentence structure should bias towards activation of one frame.

Consider what Shapiro et al.'s (1987, 1989) and Shapiro and Levine's (1988) results tell us about the structure of the language processor. Shapiro et al.'s (1987, 1989) claim that the language processor acts only on argument structure information during sentence integration has a clear implication. If Shapiro et al. are right, they provide evidence that syntactic complexity of lexical entries is not a source of processing difficulty (the evidence being their lack of a syntactic subcategorization frame effect). They deny that major syntactic category information is used to guide initial parsing (that, for example, when the parser sees a verb, it expects some argument of a verb to follow), and also that semantic information like argument structure only influences reparsing. They support an interactive model, in which semantic information (argument structure and perhaps thematic information) guides the syntactic processor during construction of the first-pass parse of a sentence.

However, the results are compatible with at least two models of the language processing system. In addition to the interactive model described above, the data are compatible with a syntax-first model in which the thematic processor can influence revision of the parse tree which was constructed on the first-pass analysis according to phrase structure rules and Frazier's (1979) Minimal Attachment and Late Closure principles. On this syntax-first model, incoming lexical items are given the first available structural analysis determined by the syntactic module. After structure is assigned by the syntactic module, a thematic module checks that the suntactic structure is consistent with thematic information. The thematic module need not wait for the entire syntactic analysis of a sentence before performing its check; it is only necessary that the syntactic processor have assigned structure to part of the incoming string. When structural analysis of the syntactic module is incompatible with thematic information, the thematic parser suggests that the sentence be reanalyzed.[4] Shapiro's results do not decide between these two

---

[4]See Frazier (1989) for discussion of the role the thematic module plays in suggesting reanalyses to the syntactic module.

major theories about the structure and operation of the parser because Shapiro and colleagues did not show that the initial representation of a sentence is constructed in terms of argument structure representations.

In order to discover the structure of the language processing system and to select between these two models, I first must determine what level of processing the Shapiro results reflect. For instance, Shapiro and colleagues must provide evidence about construction of initial sentence representations in order to confirm or disprove a syntax-first model. Argument frame complexity may have influenced several levels of processing:

(i.) Lexical identification: A representationally complex lexical entry might be harder to match to sensory input if the lexicon is ordered in some way so it takes more resources to search for and/or activate a complex lexical entry.

(ii.) Memory load: Maintaining a complex representation after the lexical item has been identified may cause an increased memory load.

(iii.) Choice effect: At some point during sentence comprehension, there may be a choice between which of the possible argument frames associated with the lexical item to instantiate.

If the argument frame complexity effect influences lexical identification, the data will not determine which of the interactive and syntax-first models is correct. Lexical identification occurs before construction of the initial syntactic representation of the sentence, and, thus, also before operation of a thematic module would effect construction of this initial representation. If the complexity effect reflects a choice between argument frames associated with a lexical item, the data do not distinguish between the two proposed models either. Such a choice between argument frames might occur during reanalysis of the initial syntactic representation, an alternative that is compatible with an interactive or with a syntax-first model. A finding that complexity increases memory load after lexical identification would support the interactive model of the language processor. Thus, it is imperative that I determine what level of processing an argument frame complexity effect reflects.

While the cross-modal lexical decision task is sensitive to processing complexity (Clifton, Frazier, and Connine, 1984), the lexical decision task also can reflect postlexical processing (Seidenberg, Waters, Sanders, and Langer, 1984; Balota and Chumbley, 1985), so Shapiro and colleagues' results are ambiguous between an effect due to representational complexity at the prelexical access stage and an effect at a postlexical stage. Indeed, Shapiro et al. (1987) may have missed a syntactic subcategorization complexity effect if one occurred at a point earlier than that which the lexical decision task reflected. Before I try to determine at what point during processing argument structure complexity has an influence, I need to find a task that is more uniquely sensitive to early stages of lexical processing. While eye movements reflect postlexical processing and integration, first fixation durations are sensitive to very early stages of processing (Inhoff and Rayner, 1986; Rayner and Duffy, 1986; see Rayner and Pollatsek, 1987, for a review). Moreover, measurements of eye movements during silent reading do not reflect task-specific processes like those which may be reflected in the lexical-decision task. Both the eyetracking study and the cross-modal naming study reported in this thesis are thought to be more specifically sensitive to early stages of processing than lexical decision, and so the results may suggest answers to the three questions raised above.

The experiments reported in this thesis were designed to answer several questions. The first is that which Shapiro et al. (1987, 1989) asked: Is semantic argument frame complexity really the relevant complexity metric for verbs? The experiments were an attempt to verify whether argument structure complexity is the relevant lexical complexity measure. The lexical decision task has been criticized because it encourages processing that may not occur during sentence processing in non-experimental situations (see above). This raises the second question: Can either syntactic subcategorization frame complexity or semantic argument frame complexity be detected in a more natural on-line task like silent reading? An affirmative answer to this question would inform us about the operation of the language processor. The third question is whether a difference in 'lexical' complexity reflects lexical access differences or post-lexical integration differences. We must answer this question before we can decide whether lexical information affects processing in a way that makes lexical complexity effects relevant to discussion of the initial mental representation of a sentence.

CHAPTER 2

EXPERIMENT 1

During reading we fixate on words several times a second. It is during these fixations that we process what we are reading. Fixations vary in duration, and increased fixation durations on a word or between words result because, for some reason, the word or words in the perceptual span take more time and resources to process (Rayner, 1978). If Shapiro and colleagues are right, and argument frame complexity is a source of lexical complexity, then perhaps fixation durations on or near verbs reflect the verbs' argument frame complexity. The main goal of this experiment was to determine whether fixation durations are sensitive to verbs' argument frame or syntactic subcategorization frame complexity.

Past research has failed to find any lexical complexity metric other than word frequency reflected in increased fixation durations on the lexically complex word, so there is some question as to whether argument frame complexity should influence eye movements. For instance, Rayner and Duffy (1986) found no effects of negativity, causativity, or factivity on fixation durations on verbs, and Inhoff (1985) found no increase in fixation durations on factive verbs. Factives, negatives, and causatives are lexically complex in that they have as part of their mental representations either lexical presuppositions, a negative component, or component meanings (respectively). But this information is more semantic in nature than is the Shapiro et al. (1987) semantic argument frame. At the least, a verb's semantic argument frame is a structural representation where a lexical presupposition, for example, does not make structural predictions about the content of the complement. So despite the fact that Rayner and Duffy (1986) did not find longer fixations on causative, factive, or negative verbs, it is possible that semantic argument frames may influence fixation durations during reading.

Predictions

Verb categories were the same as those used by Shapiro et al. (1987) and Shapiro and Levine (1988): transitive, nonalternating dative, alternating dative, two-complement, and four-complement (see Table 1). Therefore, predictions for this first experiment were similar to those made by Shapiro et al. (1987). If syntactic subcategorization complexity contributes to processing difficulty, sentences with alternating datives (which have three different syntactic subcategorization frames) should yield longer fixation durations (on or after the verb) than sentences with all other verbs (because all other verbs only have one or two syntactic subcategorization frames).

However, if argument structure complexity contributes to processing difficulty, sentences with four-complement verbs (which have four argument structure frames) should yield longer fixation durations. Also, transitives should lead to the shortest fixation durations, since they have one argument frame. Nonalternating datives, alternating datives, and two-complements have two argument frame possibities. If neither the number of argument structure frames nor the number of syntactic subcategorization frames contributes to processing complexity, and if all other possible differences between verb classes can be eliminated, no verb class should cause significantly longer fixation durations. If both syntactic subcategorization complexity and argument complexity contribute to processing difficulty, the nonalternating datives should cause shorter fixation durations than the alternating datives (because alternating datives have an extra syntactic subcategorization frame although both kinds of dative have only two argument structure frames), and two-comps should result in shorter fixation durations than four-comps because four-comps have four argument structure frames compared to two-comps' two argument structure frames. Both kinds of complementizers have two syntactic subcategorization frames.

Verb frequency and length have been demonstrated to influence fixation durations during silent reading (Inhoff, 1984; Inhoff and Rayner, 1986; Rayner, 1977; Blanchard, 1985; Just and Carpenter, 1980). Frequency and length are confounded with verb category here, as in Shapiro et al. (1987), with the alternating datives being the most frequent, shortest verbs (see Table 4)[5]. As such, they were biased to incur shorter fixations. These predictions are summarized Table 3.

---

[5]Note that the frequency confound in Experiment 1 of this thesis is different than the frequency confound in Shapiro et al. (1987) because one verb has been added to each verb category here.

An argument frame complexity effect could reflect a variety of processes. Fixations on a word (here I consider only first-fixation or gaze durations and not regressive fixations) at least reflect the lexical access processes of matching perceptual input to a lexical entry plus retrieving information from the lexical entry once it has been located. In addition, fixation durations may reflect integration of the fixated word (and other words, if there are any, within the perceptual span) into preceding context, integration of the fixated word with following context that falls within the perceptual span (see Rayner and Pollatsek, 1987), and ambiguity resolution (Duffy, Morris, and Rayner, 1988).

In Experiment 1, the critical word was the main verb in each experimental sentence. I also examined fixation durations in post-verbal regions, expecting a delayed complexity effect (due to either syntactic subcategorization frames or to argument structure frames) to be possible. Following Rayner and Pollatsek (1987), I assume that fixation duration on a given verb at least reflects access of the verb's lexical entry. In general, when one finds no significant differences in fixation duration on a critical word but does find differences in fixation durations in regions after the critical word, the pattern of durations probably only reflects integration of the critical word into sentence context. This is because lexical access is thought to be completed before subjects look away from a word (Carpenter and Just, 1983, but cf. Kliegl, Olson, and Davidson, 1982; Rayner and Pollatsek, 1989). In Experiment 1, differences in the pattern of fixation durations in post-verbal regions would reflect integration of the verb (and subsequent material) with context (unless there is an unintentional confound of degree of difficulty of the post-verbal material). And a result of differences in fixation duration only on the verbs would suggest the possibility that differences across verb categories result from difficulty in identifying more complex lexical items.

Method

Subjects

Thirty-two members of the University of Massachusetts community were paid or given experimental credit for participating. Eight subjects' data were discarded due to bad calibrations that led to large numbers of track losses. All subjects had normal, uncorrected vision (determined by self-report), were native English speakers, and were naive with respect to the purpose of the study.

Materials

I used the same verbs used by Shapiro, Zurif, and Grimshaw (1987) in the first experiment they report (see Table 2). In addition, one new verb was added to each of the original five categories as follows:

Transitive: Adopt
N-A Dative: Release
Alt Dative: Give
2 Comp. : Expect
4 Comp. : Detect

These verbs were added to allow counterbalancing of sentence frame types across subjects.

Criteria used to classify verbs were identical to those reported by Shapiro et al. (1987) (see Table 1). Minimally, all verbs allowed a noun phrase complement. (However, some of the verbs were optionally intransitive, a fact which Shapiro et al. (1987) did not acknowledge: e.g., surrender, as in 'the enemy surrendered', and reserve, as in 'the family reserved ahead'.) The critical (argument frame complexity) contrast between two-complements and four-complements is based on Grimshaw's (1979) Q and E variables, ranging over the semantic types interrogative and exclamation respectively. Two-complement verbs take noun phrase complements and sentential complements that are propositional, while four-complement verbs allow in addition sentential complements of semantic types Q and E. Compare:

Two-complement:
NP: John expected the gift.
P: John expected that they would arrive soon.
Q: *John expected whether the child was old enough.
E: *John expected what a fool Bill was.

Four-complement:
NP: John discovered the gift.
P: John discovered that they would arrive soon.
Q: John discovered whether the child was old enough.
E: John discovered what a fool Bill was.

Some verbs selected by Shapiro et al. were classified questionably. For instance, they consider <u>regret</u> to be a two-complement verb, although it allows sentential complements of semantic type E: Bob regretted what a fool he had been the last time he saw Jill.

Since I used the same criteria, I inherited these incongruities. It is interesting to note that the current criteria Grimshaw used to classify verbs, as evidenced by a file of verbs and their classifications which David Swinney made available to me, does not classify verbs the same way Shapiro et al. (1987) did. In particular, the verbs recognize, indicate, and detect (added by me for experiment 1), were cited as two-complement verbs that allow noun phrase complements and propositional sentential complements only.[6]

As acknowledged by Shapiro et al. (1987), frequencies of the verbs they used were not controlled. In a footnote on pp. 241-242, Shapiro et al. claim that frequency effects in their first experiment do not account for the observed differences between verb category because the most frequent category, the four-comps, resulted in the longest lexical decision reaction times. Also, the transitives and non-alternating datives have the same mean frequency of occurrence, yet the transitives result in significantly faster reaction times. Shapiro et al. are correct in claiming that a frequency effect should have led to a reduction in reaction times for the four-complements.

With the materials used in Experiments 1 and 2, I also have a frequency problem (see Table 4), although mean frequencies for my verb categories differ from Shapiro et al.'s (1987) means since I added one verb to each category. The transitives and nonalternating datives are the least frequent verbs (Francis and Kucera, 1982: mean

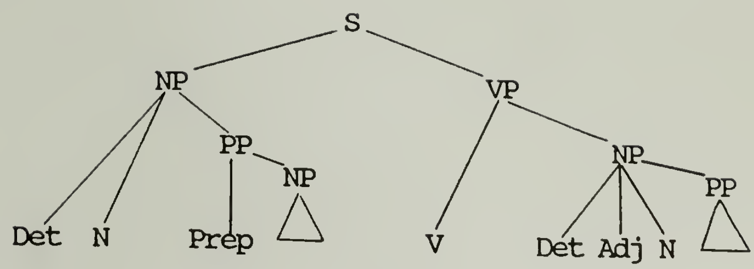
---

[6]Thanks to David Swinney for sending me the lexicon. Creation of the lexicon was supported by Jane Grimshaw and Ray Jackendoff's NSF grant number NSF IST-81-20403 awarded to Brandeis University.

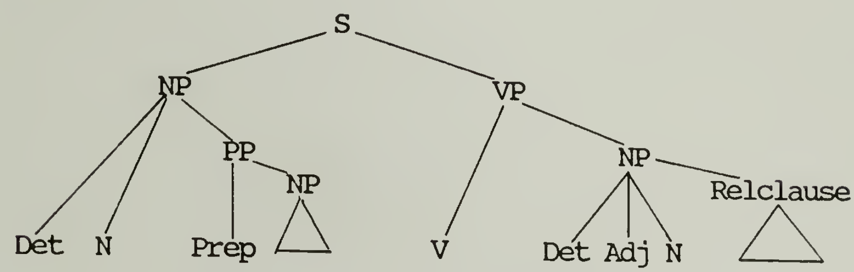
frequencies are 66 and 62 words per million, respectively). The two-complements and four-complements are of similar frequency (mean freqs. of 160 and 158 words per million respectively) but are less frequent than the alternating datives (297 words per million). If frequency alone influences fixation duration, and if there is no effect of syntactic subcategorization frame nor of semantic argument frame, the alternating datives should result in the shortest fixation durations, followed by the two- and four-complementss, with the nonalternating datives and transitives yielding the longest fixation durations.

Word length also differs across verb categories (see Table 4), and this will affect gaze durations on the verbs (Kliegl, Olson, and Davidson, 1982; Blanchard, 1985). Most verbs in the alternating dative category are shorter than verbs in any other category and should lead to shorter fixation durations for the alternating datives. A millisecond per character adjustment of gaze durations provides some compensation for length disparity across categories (Rayner, Sereno, Morris, Schmauder, and Clifton, 1989).

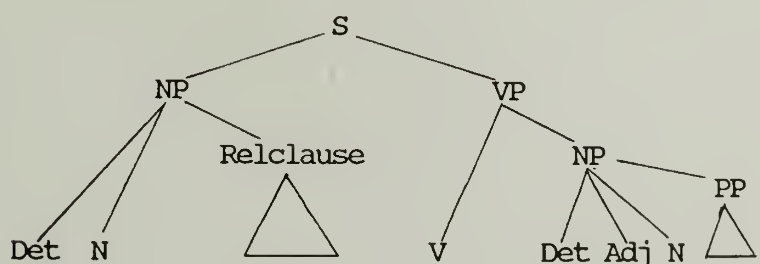
Each verb appeared in two sentence frames, but each subject only saw a verb once. In one version (the prepositional-phrase version), a simple noun-phrase subject was followed by a prepositional phrase, the verb, a modified object noun phrase, and either a relative clause or a prepositional phrase which is part of the object noun phrase:



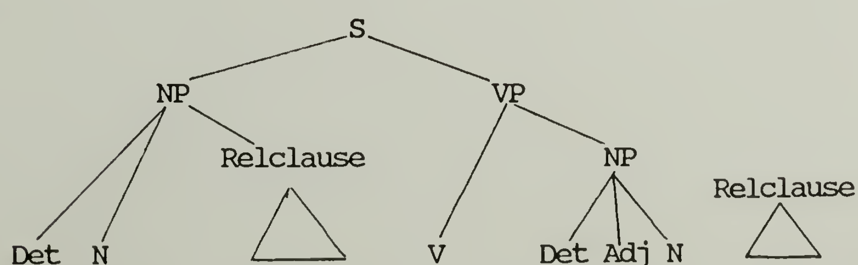
Or:



In the second version (the embedded-relative version), the same subject noun phrase, verb, and adjectival object noun phrase were used, but the pre-verbal prepositional phrase was replaced by an embedded relative clause of the form "who was ADJ":



Or:



In each of the two sentence frames, the number of syllables before (and, therefore, also through) the verb was the same, and for a given verb the post-verb context was identical in the prepositional-phrase version and the embedded-relative version.

In the Shapiro et al.(1987) sentences, it was possible to interpret the DET ADJ NOUN VERB as having a reduced relative reading in fourteen of the twenty-five DET ADJ NOUN VERB sentence versions. For example, in the sentence:

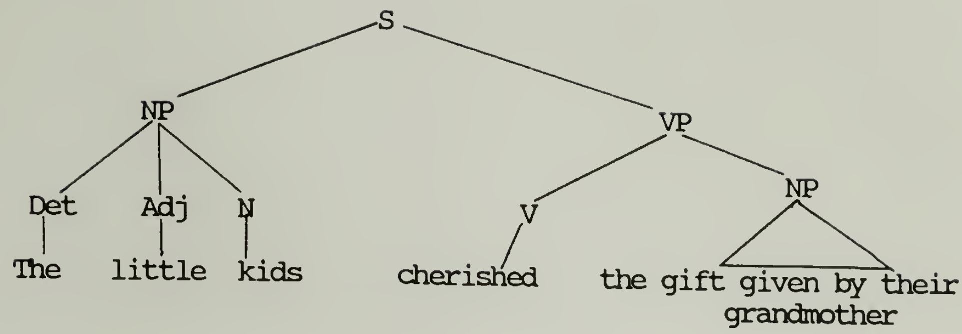
(1) The little kids <u>cherished</u> the gift given by their grandmother.
(see diagram A below)

until the reader reaches the object, nothing rules out the possiblity that the sentence will continue "by their parents were injured", as in:

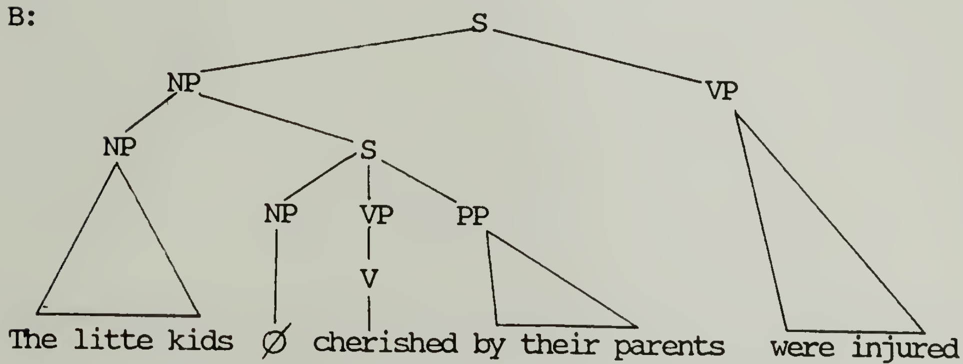
(2) The little kids <u>cherished</u> by their parents were injured.
(see diagram B)

Shapiro et al.'s results might reflect the fact that subjects had taken the reduced relative reading, realized that they had been garden-pathed, and then reanalyzed the sentence to get the non-reduced reading. On this alternate account, the complexity effect found by Shapiro et al. could be due to a garden-path effect which occurred with higher frequency in sentences containing four-complement verbs than in sentences containing other verb types.

A:



B:



I used the sentence frame with the prepostitional phrase preceding the verb to discourage readers from taking this reduced relative reading. It still was possible to take the subject noun phrase plus prepositional phrase as part of a relative clause, as in "The woman from New York cherished by her parents moved out anyway,"[7] but intuitively there seems to be no temptation to do so in a sentence like "The woman from New York cherished the pretty coat."

To insure that the test sentences did not differ in coherence or semantic complexity across verb categories, after constructing an initial set of materials, I randomly assigned preverbal context to the rest of the sentence (verb and postverbal context) and then edited a small number of sentences to eliminate semantic anomalies.

---

[7]Shari Speer pointed this reading out to me.

Only one sentence had to be modified significantly: after scrambling, an inanimate subject was matched to a predicate that required an animate subject, and so I replaced the subject with one that was animate. Also, in a few sentences, changes were made to maintain subject/verb and antecedent/consequent agreement after scrambling. This scrambling strategy reduced the possiblity that some verbs would appear more complex only because in some sentence frames, more semantically complicated material preceded the verb, causing spillover processing effects on the verb. Because of the nature of the materials, which consisted of a definite description that was (relatively) semantically unrelated to the action specified in the predicate, this random assignment of preverbal context to predicates did not lead to semantic disruption (see Appendix 1).

With six verbs in five categories there were thirty verbs to be tested. Each verb occured both in the prepositional-phrase sentence frame and in the embedded-relative sentence frame, so there were thirty pairs of sentences and a total of sixty test sentences (see Appendix 1 for materials). Each subject saw thirty sentences consisting of one pseudo-randomly selected sentence from each sentence pair, such that half of the sentences a subject saw were the prepositional-phrase sentence frame version and the other half were the embedded-relative version. Seventy-five distractor sentences were presented with the thirty test sentences. The distractor sentences provided variation in sentence structure within the experiment. Sixteen of the distractor sentences were of the form: NP

VP NP PP conjunction S', as in "John dismayed the woman with the bad news even though he tried to break it to her gently." These were part of an experiment testing prepositional phrase attachment (whether the prepositional phrase was an argument of the verb or of the object noun phrase) and Frazier's (1979) Minimal Attachment principle and served to test whether eye movements were sensitive to syntactic manipulations. Six sentences taken from Rayner and Frazier (1987) included temporarily ambiguous complements, as in "The pupils knew several solutions to the problem would be quite possible." The rest of the distractor sentences varied in form and semantic content (for example, some included fronted prepositional phrases) to provide further variation.

Apparatus

Subjects' eye movements were recorded by a Stanford Research Institute Dual Purkinje Eyetracker interfaced to an AT-class personal computer that controlled the experiment. The eyetracker has a resolution of 1 minute of arc. Horizontal and vertical position information was sampled every millisecond by the computer, and the existence of a fixation was determined by occurrence of five successive fixations in the same location. For each subject, the sequence of eye movements and the location and duration of fixations were stored on computer disk for later analysis.

Subjects were seated 23.5 inches away from a Sony Trinitron 1302 CRT on which the experimental sentences appeared. Four characters equalled one degree of visual angle. Letters were presented in lower case, except for the first letter of the sentence. Eye movements were recorded from the right eye, and viewing was binocular. The brightness of the screen was adjusted for each subject to a comfortable level.

Procedure

When a subject arrived, a bite bar was made to prevent head movement during the experiment. Subjects then were instructed that they would be reading sentences presented on a monitor and occasionally would press one of two response keys to answer questions about some of the sentences. The experimenter emphasized that subjects should read at a normal pace. Next subjects ran through a calibration procedure that took from 3 to 5 minutes. Subjects performed a practice block consisting of approximately 15 sentences, about one third of which were followed by a yes/no or true/false comprehension question about the sentence just read. Where there was a question, it appeared several lines lower on the screen than the experimental sentences did, and subjects pressed one response key to indicate no or false and another key to indicate yes or true. After the practice block, subjects were given an opportunity to ask questions before proceding with the experiment. Presentation of materials was divided into two blocks, and subjects were told that they could come off of the bite bar during a block if they needed to rest.

Design

There were three variables in the complete design of this experiment, all within subject in the subjects' analyses.

(A) Verb Category. There were five verb categories as in Shapiro et al. (1987): transitive, nonalternating dative, alternating dative, two-complement, and four-complement.

(B) Sentence Frame. Each verb occured in two sentence frames. In one frame, the prepositional-phrase frame, the subject noun phrase was modified by a prepositional phrase just prior to the verb. In the second frame, a relative clause modified the subject noun phrase prior to the verb.

(C) Region. For the initial set of analyses performed, there were three regions of interest in each sentence. The first region was all of the sentence material prior to the verb. The second region consisted of the verb, including the space before the first letter of the verb. Region 3 began with the space after the verb, and included the entire object noun phrase. A fourth, less interesting region included the remainder of the sentence. A second set of analyses was run after changing regions 2 and 3. The revised region 2 included the article following the verb (which was 'the' for 56 of the sentences and 'a' for 4 sentences). Region 3 then included the rest of the object noun phrase (the adjective and head noun).

Individual analyses were run for regions 1, 2, and 3, with verb category and sentence frame as within subjects variables.

Results

Each subject's data were processed to remove short fixations standing alone which are believed to be part of saccades and to merge short fixations (which probably reflect overshooting an intended landing position and subsequent adjustment following a saccade) with adjacent longer fixations. Fixations that were shorter than 80 milliseconds in duration and only one character away from the prior or next fixation were merged with that prior or next fixation. Fixations shorter than 40 milliseconds and less than three characters away from the prior or next fixation were deleted. Any fixation longer than 2000 msec was deleted.

Several measures were used to draw conclusions about what the results of experiment 1 reflect. I will report first fixation duration, first pass reading time, (gaze duration when the region consists of a single word) and total fixation time for the regions of interest, as well as accuracy on questions. First pass reading time is the sum of all left-to-right fixations made in a region and all within-region right to left movements (regressions). The first pass (gaze duration) measure stops when the subject either moves forward out of the region into a later region or regresses to a previous region, again crossing the boundary of the current region. Total fixation time in a region is the sum of all fixation durations in the region, including those made after leaving the region and then returning. In addition, eye movement patterns will be reported as the probability of regressions into and out of regions and the probability of fixating in a region. The regressions-out measure indicates the probability that a subject left the current region to regress to a previous region. The regressions-in measure indicates the probability that subjects made regressions from later regions into the current region.

Times reported as raw fixation times are not averaged to determine time per character, so one cannot accurately compare regions that differ significantly in length. Millisecond per character times are calculated by dividing raw times by the number of characters in the region. The msec/character measure is useful for comparison of regions that differ in length, but if lengths of the regions being compared are quite different, msec/character results may not be accurate (Rayner et al., 1989).

Subjects and items were treated as random factors in analyses of first fixation duration, first pass time (gaze duration where the region consists of a single word), and total fixation time. Unless otherwise indicated, fixation durations are raw fixation times, not msec/character measures.

Data will be presented in the following order. I will discuss analyses on the region containing only the main verbs first (referred to as region 2): first fixation durations (raw), gaze durations (raw and msec/character), regressions (into and out of the verb region) during gaze on a verb, and total fixation duration on the verbs (raw). These results are reported in Tables 5, 8, and 9. Next, I will present analyses on a region containing both the verb and the article after the verb (called revised region 2; results reported in Table 7). Analyses of first pass times (raw) on the entire object noun phrase (including article, adjective, and head noun) follow, along with correlation and regression analyses for the object noun phrase (see Tables 6, 8, and 9). Finally, I will discuss accuracy on comprehension questions and first pass (msec/character), total time (msec/character), and regression (in and out) analyses of the overall design, which includes verb categories (5 levels), sentence frames (2 levels), and analysis regions (4 levels) as independent variables (see Tables 8 and 9).

Fixation durations on the verb only

First I discuss analyses of region 2, which consisted only of the main verb and the space preceding it (see Tables 5, 8 and 9). Differences among the mean first fixation durations for the five verb categories neared, but did not reach significance in the subjects analysis ($F_1(4,23) = 2.32$, $p < .06$). The pattern of means was (means are listed from lowest to highest):

Alt. Dative,	Four Comp.,	Transitive,	Two Comp.,	Nonalt. Dative
(253)	(267)	(269)	(279)	(284)

Prepositional phrase sentence frame: 266
Relative clause sentence frame: 274

There was no sentence frame effect nor verb category by sentence frame interaction in the first fixation analysis, although the sentence frame effect neared significance in the items analysis ($F_2(1,25) = 3.67$, $p < .06$).

Gaze durations on the verbs reflected a significant main effect of verb category in the subjects analysis, ($F_1(4,23) = 3.84$, $p < .007$, $F_2(4,25) = 2.12$, $p > .1$), and a significant effect of sentence frame ($F_1(1,23) = 3.96$, $p < .06$, $F_2(1,25) = 6.9$, $p < .01$), with relative clause sentence frames leading to longer gaze durations on the verbs (mean gaze durations of 340 msec in region 2) than did the prepositional phrase frames (mean gaze durations of 317 msec in region 2). This marginal sentence frame effect suggests that the relative clauses were more difficult to process and that this processing difficulty carried over to the next word, in this case, the verb. The pattern of means was similar to that from the first fixation analysis (means are listed from lowest to highest):

Alt. Dative,	Four Comp.,	Transitive,	Two Comp.,	Nonalt. Dative
(294)	(320)	(332)	(337)	(360)

These gaze duration results, when considered together with the first fixation times on the verbs, suggested a true difference among verb types, although the pattern of means was inconsistent both with a syntactic subcategorization frame complexity effect and with a semantic argument frame effect. The alternating datives are much shorter (past tense forms of the alternating dative verbs average 5 letters long) and are higher frequency words (mean frequency is 297 words per million) than verbs in other categories (see Table 4). Nonalternating datives (mean frequency = 62 words per million) are lower frequency verbs than verbs in all other categories except the transitives, although they are not different in length.

The msec/character gaze duration analysis by subjects gave a different pattern of means (means are listed from lowest to highest):

Four Comp., Two Comp., Nonalt. Dative, Transitive, Alt. Dative
(33) (36) (38) (39) (48)

Verb category and sentence frame were both significant, $F_1(4,23) = 13.59$, $p = 0$, $F_2(4,25) = 3.3$, $p < .03$, and $F_1(1,23) = 5.91$, $p < .02$, $F_2(1,25) = 6.3$, $p < .02$, respectively. This pattern of means was probably due to problems with the msec/character measure, as mentioned above.

To investigate whether verb frequency and length in fact were responsible for the significant gaze duration verb category effect on the verbs, a correlational analysis was run. Frequency and number of letters correlated significantly with gaze durations on the verbs (frequency-gaze: $r = -.36$, $p < .026$; length-gaze: $r = .41$, $p < .012$). A subsequent multiple regression analysis run with raw gaze durations as the dependent variable and frequency and verb category as independent variables (frequency was entered first into the regression equation to test for remaining effect of verb category) yielded a nonsignificant verb category effect, two-tailed $t = .44$, $p < .66$.

Mean total fixation times on the verb (region 2) patterned the same as did mean first fixation durations and gaze durations (means are listed from lowest to highest):

Alt. Dative, Transitive, Four Comp., Two Comp., Nonalt. Dative
(414) (425) (440) (454) (466)

However, the verb category effect was not significant in the total time analysis, $F_1(4,23) = 1.54$, $p > .2$, $F_2(4,25) = .64$, $p > .6$.

The proportion of regressions into and out of the verb region (region 2) did not differ significantly across verb category in either the subjects or the items analysis (Regressions out of region 2: $\underline{F}_1(4,23) = 1.09$, $\underline{p} > .4$, $\underline{F}_2(4,25) = 1.9$, $\underline{p} > .1$; regressions into region 2: $\underline{F}_1(4,23) = 1.43$, $\underline{p} > .2$, $\underline{F}_2(4,25) = 1.3$, $\underline{p} > .3$).

Fixation durations on the verb plus the following article

To see whether there was any hint of a delayed verb complexity effect, regions 2 and three were redefined for a second set of analyses. The revised region 2 consisted of both the verb and the article following the verb, and region 3 became the remainder of the object noun phrase, that is, the adjective and head noun (see Table 7). Patterns of mean first fixation duration and first pass times were like those for the analysis where region 2 consisted only of the verb, and just as on the earlier analysis, verb category was significant only for the gaze duration measure, $\underline{F}_1(4,23) = 2.94$, $\underline{p} < .02$. The sentence frame effect in first pass times (gaze durations) was significant with the larger region 2, $\underline{F}_1(1,23) = 7.07$, $\underline{p} < .01$, whereas it was only marginally significant when region 2 consisted only of the verb, $\underline{F}_1(1,23) = 3.96$, $\underline{p} < .06$, indicating delayed processing difficulty in the sentence frames where relative clauses preceded the verb. Verb frequency and number of letters in the verb plus article region correlated with first pass time for this revised region, frequency-first pass, $\underline{r} = -.44$, $\underline{p} < .01$; length-first pass: $\underline{r} = .31$, $\underline{p} < .05$. A multiple regression analysis left no verb category effect after removing the contribution of frequency, two-tailed $\underline{t} = -1.13$, $\underline{p} > .3$.

Fixation durations on the object noun phrase

Looking at region 3 (see Tables 6, 8, and 9), which on the initial analysis consisted of the entire object noun phrase (DET, ADJ, NOUN), a significant verb category effect in the subjects analysis of first pass times, $F_1(4,23) = 4.74$, $p = 0$, did not occur in the items analysis, $F_2(4,25) = 2.05$, $p > .1$, indicating that the difference in the subjects analysis was not consistent across verb categories. These results are consistent with the patterns in region 2.

Accuracy on comprehension questions

Accuracy on comprehension questions, which followed one third of the test sentences, was high. Subjects responded correctly to questions 94% of the time overall (100% correct on true questions, 88% correct on false questions; subjects made false yes responses 12% of the time).

Analyses of the complete design

Using the complete design (with regions as a within subjects factor, see Tables 8 and 9), an Anova was run on first pass times with the millisecond per character correction. There was a significant verb category by region interaction, $F_1(12,92) = 5.42$, $p < .01$, due to long times again in the cells for the alternating dative verbs.

The analysis of variance run for the overall design on total times with the msec/character correction showed a large verb category effect, $F_1(1,23) = 14.21$, $p = 0$, and a large relative clause sentence frame effect, $F_1(1,23) = 23.37$, $p = 0$. This coincided with the fact that there were more regressive fixations in relative clause sentence frames.

Regressions-in and regressions-out analyses in the Anovas run for the complete design (included regions as a within-subjects variable) yielded no significant verb category effect (regressions out: $F_1(4,23) = .71$, $p > .6$, $F_2(4,25) = .58$, $p > .7$; regressions in: $F_1(4,23) = .8$, $p > .5$, $F_2(4,25) = .47$, $p > .8$). However, there was a highly significant sentence frame effect, suggesting more regressions into and out of the relative-clause region in the relative-clause sentence frame: regressions in: $F_1(1,23) = 28.47$, $p = 0$, $F_2(1,25) = 37.58$, $p = 0$; regressions out: $F_1(1,23) = 24.01$, $p = 0$, $F_2(1,25) = 45.15$, $p = 0$.

Discussion

The results of experiment 1 clearly indicate neither syntactic subcategorization frame complexity nor semantic argument frame complexity effects as reflected in eye movements. Not only were there no effects of these complexity metrics visible in first fixation durations or gaze durations on the verbs themselves, but also there was no delayed appearance of verb complexity (of any sort) on later regions of the sentence.

Experiment 1 confirms previous findings that first fixation duration and first pass times are sensitive to word frequency (Rayner and Duffy, 1986) and word length. The alternating dative verbs are quite a bit shorter than verbs in the other categories, and they have a higher frequency of occurrence.

Some of the filler sentences in experiment 1 contained a manipulation of prepositional phrase attachment sites as a test of argument preference (whether the prepositional phrase was an argument or an adjunct of the verb or of the object noun phrase) and Frazier's (1979) Minimal Attachment principle. Compare attachment of the 'with' prepositional phrase in:

> The baby disgusted the woman with his dirty diapers and she vowed never to have children of her own.
> The baby disgusted the woman with the high heels and she vowed never to have any children of her own.

An argument preference effect (preference for the prepositional phrase to be taken as an argument of the verb) obtained (gaze duration, $p < .04$). So syntactic complexity is reflected in eye movements, as has been demonstrated previously (Frazier and Rayner, 1987; Rayner, Carlson, and Frazier, 1983).

The lack of a delayed effect of verb argument structure frame complexity indicates that argument structure complexity does not influence ease of integration of the verb into the sentence in which it occurs, contrary to the suggestion made by Shapiro et al. (1987). One might object that the failure to present verbs from different categories in identical sentence frames unwittingly allowed semantic differences across sentence frames to wipe out whatever representational complexity effects there were. However, random assignment of context prior to the verbs with verb plus complement, the attempt to match object noun phrases across sentences on length and frequency, and the constancy of syntactic structure make it unlikely that differences in sentence frames obscured representational complexity effects.

Perhaps the eyetracking paradigm is not sensitive to argument frame complexity because eye movements reflect very early lexical and syntactic processing rather than the activity of the thematic or semantic processing modules. Shapiro et al's. (1987, 1989) results might only be detectable with a processing measure primarily sensitive to thematic and semantic complexity, like the lexical decision task.

To further investigate whether argument structure complexity influences processing, I performed a second experiment using the same sentences as Experiment 1. Experiment 2 used a dual task paradigm like the lexical decision paradigm used by Shapiro et al. (1987, 1989). Rather than make a lexical decision to the visually-presented probe, subjects named a visually-presented probe word, and naming latency was the dependent measure.

CHAPTER 3

EXPERIMENT 2

Results of Experiment 1 showed that neither verbs' argument frame complexity nor verbs' syntactic subcategorization frame complexity cause effects reflected in eye movements. Experiment 2 was undertaken to determine whether a dual-task paradigm like Shapiro et al. (1979) used would be a good measure of argument frame complexity or subcategorization frame complexity.

Shapiro et al.'s (1987) and Shapiro and Levine's (1988) results and conclusions suggest that several or all argument frames are activated initially and that this activation is reflected in processing cost. The experiments support either parallel access of all argument frames or rapid serial (and perhaps consistently ordered) activation of argument frames.

However, the experiments do not distinguish between effects that result from prelexical processes, those processes that occur during analysis of the speech input as subjects determine what word they are hearing, and effects that are the result of postlexical processes, the processes of selecting and integrating lexical information with context. The lexical decision task can reflect postlexical processing, as has been shown by Balota and Chumbley (1984), Seidenberg, Waters, Sanders, and Langer (1984) and West and Stanovich (1982). This tendency is particularly salient when lexical decision probes are related to context, and it is possible that some lexical decision probes were related to the critical verbs in the Shapiro et al. materials (or that the LD probes were reasonable continuations of the sentence in some instances). In any event, Seidenberg et al. (1984) argue convincingly that the lexical decision task does not distinguish between postlexical integration effects that are particular to the LD task and integration effects that occur during normal reading comprehension.

This suggests two ways in which Shapiro and colleagues' procedures can be improved upon. The first way is to improve cross-modal probes to insure that they are not related to sentence contexts if one continues to use a cross-modal task. But since the Shapiro et al. (1987) effects were quite large, and since the naming task seems to be less sensitive to postlexical effects than lexical decision (where postlexical effects are as defined above), I used a naming task like that used by Tanenhaus et al. (1979) in place of the CMLD task. The secondary naming task may serve as an index of processing load during sentence comprehension and so may reflect argument structure complexity.

In Experiment 2, the sentences used in Experiment 1 were presented binaurally over headphones. At the offset of the main verb of the sentence, a target word was presented on a video screen that sat in front of the subject. The task was to name the target word as quickly and accurately as possible while attending to the sentence which continued to play over the headphones.

Predictions

Predictions for Experiment 2 are the same as those for Experiment 1, with the exception that the dependent measure is not fixation duration but naming latency. So a main effect of verb category such that transitives < nonalternating datives = alternating datives = two-complements < four-complements will support the hypothesis that argument frame complexity influences processing. If syntactic subcategorization frame is the relevant complexity metric, the pattern of means should be: transitives < nonalternating datives = two-complements = four-complements < alternating datives.

Method

Subjects

Seventy members of the University of Massachusetts community were paid or given experimental credit for their participation. All had normal hearing and normal or corrected-to-normal vision (determined by self-report) and were naive with respect to the purpose of the study. Eight subjects' data were discarded because standard deviations were greater than 110 msec. Data from two subjects were lost due to a programming error.

Materials

The verbs and sentence contexts were the test materials described above for Experiment 1. There were a total of sixty test sentences, but as in Experiment 1, each subject heard only thirty sentences consisting of one pseudorandomly-selected sentence from the sentence-frame pair associated with each verb. Seventy-five filler sentences were presented with the thirty experimental sentences seen by each subject. The filler sentences represented various sentence structures, including garden-path sentences, to test whether the naming task was sensitive to structural complexity (see Appendix 1 for further detail).

Words used as naming targets for experimental sentences were selected according to the following criteria: nouns with an initial voiced or voiceless stop consonant (/p/,/t/,/k/,/b/,/d/,/g/), one syllable, 4-7 letters long, and frequency between 50 and 110 words per million (determined using Francis and Kucera (1982) word frequency norms). (Targets for filler sentences had an initial voiceless fricative or silibant, were one syllable and 4-7 letters long, and had frequencies less than 100 words per million.) After constructing an initial set of 30 naming target words, 5 different naming-target to verb pairings were created such that the six naming targets paired with transitive verbs in the first set of verb-target pairs were associated with nonalternating datives in the second set, with alternating datives in the third set, with 2-complements in the fourth set, and with 4-complements in the fifth set (see Appendix 2). Initial target-verb pairings for other categories were rotated similarly.[8] The goal was to reduce the chance that a particular verb-naming target pairing would contribute to across-verb-category differences in naming latencies.

---

[8]Shapiro et al. (1987) first ran 10 subjects with one lexical-decision-target/verb pairing then ran 10 subjects with a second target/verb pairing. Probes were assigned to verbs such that the first group saw a visual probe associated with a different verb class than the class the same probe occurred with for the second group.

Naming targets were unrelated to sentence context; in particular, the five target-to-verb pairings were constructed so that the naming target did not form a semantically plausible continuation of the sentence. These targets were presented at the offset of the verb during presentation of the thirty test sentences, and naming targets associated with filler sentences occurred at positions other than at the offset of the verb, so subjects would not develop expectations about when the target would appear. Targets affiliated with filler sentences differed in frequency and initial phoneme to provide further variation.

A male speaker recorded the sentences on one channel of a two-channel Teac X-10 tape recorder. Next, sentences were digitized, using a 4.5 kHz low-pass filter interfaced with an AT-class microcomputer. Using a waveform editing program written by Charles Clifton, the offset of the main verb was determined by listening for the end of the last phoneme in the main verb, and a tone was placed 1000 msec prior to the offset. The tone later signaled the experiment-running program to present the relevant naming target 1000 msec after the tone, coincident with the offset of the verb.

Probe positions for distractor sentences differed as follows. For the sixteen sentences testing the Minimal Attachment principle and argument status of prepositional phrases, probes were presented at the offset of the head noun of the prepositional phrase:

> The baby disgusted the woman with its dirty diapers (probe) and she vowed never to have any children of her own.

Probes for the six sentences with temporarily ambiguous arguments taken from Rayner and Frazier (1987) occurred at the offset of the copula:

> The pupils knew several solutions to the problem would be (probe) quite possible.

For the remaining fillers, tones triggering probe presentation for eight fillers were pseudo-randomly placed in the first quarter of the sentence, for another eight in the second quarter of the sentence, for a third eight sentences in the third quarter of the sentence, and for the final eight fillers, the probes occurred during the last quarter of the sentence. Thus, subjects could not predict when during a sentence the naming target would occur.

Procedure

When a subject arrived, she or he was seated in front of a CRT and instructed that she or he would be listening to sentences over headphones, and that at some point during a sentence, a word would appear on the CRT. She or he was told that her or his task would be to name the word as quickly and accurately as possible while still listening to the sentence. Subjects were instructed to take their time when answering comprehension questions. After a practice block of thirty trials, subjects proceeded with the experiment. Naming latency to visual probes was recorded by a voice key interfaced with a microcomputer on which the experimental program ran. Subjects were monitored to determine whether they correctly named the target word presented, and trials on which an error in naming occurred were thrown out. Naming errors were defined as instances in which the subject said a word other than the word she or he saw, began to say one word and changed words mid-utterance, or failed to complete naming the word she or he began to say. Aural presentation of the sentence continued during the naming task. At the offset of each sentence, a true/false or yes/no comprehension question was displayed on the CRT, and subjects pulled response triggers to indicate their answers. Subjects received feedback when they answered questions incorrectly (the message "WRONG ON QUESTION" was displayed on the screen); correct answers received no immediate feedback, but overall percent-correct-on-questions feedback was given twice during the experiment during breaks between blocks.

Design

The dependent measure was naming latency to the visually presented probe word. There was one between subjects variable, naming probe, and two within subjects variables, verb category and sentence frame, in the overall design.

(A) Naming probe. Each naming probe was paired with a verb in each of the five verb categories. There were five pairings of naming target to verb as described above in the materials section.

(B) Verb category. As in Experiment 1, there were five verb categories as in Shapiro et al. (1987): transitive, nonalternating dative, alternating dative, two-complement, and four-complement.

(C) Sentence Frame. Each verb occured in two sentence frames, and frames were identical to those in Experiment 1. In one frame, the prepositional-phrase frame, the subject noun phrase was modified by a prepositional phrase just prior to the verb. In the second frame, a relative clause modified the subject noun phrase prior to the verb.

Results

Data from eight subjects were replaced because the subjects had standard deviations greater than 110 msec. Data from two subjects were replaced due to a programming error that resulted in the loss of those subjects' data. This left data from 60 of the original 70 subjects run.

Subjects were quite accurate in their answers to comprehension questions (see Table 10), so it is reasonable to assume that they were listening to the auditorily presented sentences.

The pattern of mean naming latencies suggested that four-complements were more difficult than verbs in the other categories:

Transitive,	Nonalt. dative,	Alt. dative,	Two-comp.,	Four-comp.
(494)	(496)	(500)	(502)	(507)

This ordering of means was the same as that reported by Shapiro et al. (1987), except for their result of a slightly shorter mean for two-complements than for alternating datives. Analysis of variance collapsing over the verb-naming target pairing yielded a design with two within-subjects variables: verb category (five levels) and sentence frame (two levels) (see Table 11). No effects or interactions were significant: verb category, $\underline{F}_1$ (4,59) = 1.95, $\underline{p}$ > .1, $\underline{F}_2$ (4,25) = .67, $\underline{p}$ > .6; sentence frame, $\underline{F}_1$ (1,236) = 1.89, $\underline{p}$ > .17, $\underline{F}_2$ (1,25) = 1.17, $\underline{p}$ > .3.

No frequency or length effects were visible; the alternating datives led to naming latencies similar to those following verbs in other categories. Filler sentences that varied in syntactic complexity did not lead to differences in naming latencies.

Planned contrasts comparing two-complements with four-complements, nonalternating datives with alternating datives, and transitives with all datives were not significant (two-complements versus four-complements, $\underline{F}$ (1,59) = 1.09, $\underline{p}$ > .3; nonalternating versus alternating datives, $\underline{F}$ (1,59) = .53, $\underline{p}$ > .5; transitives versus both datives, $\underline{F}$ (1,59) = .79, $\underline{p}$ > .4).

Naming accuracy was 98% for both prepositional phrase versions and embedded relative clause versions of the test sentences.

Discussion

This failure to find evidence for any kind of complexity effect in the cross modal naming task is troubling. There is a slight possiblity that the naming probe position (which was immediately at the offset of the verb) was too early, and that complexity effects would occur at a later point during processing. However, Shapiro and Levine (1988) used two probe positions, one at the offset of the verb and one four syllables after the verb, and found that the semantic argument frame effect did not occur at the second probe position.

They concluded from this finding that the complexity effect was immediate rather than delayed.

The hint in the data that the four-complements might be slightly more difficult than verbs in other categories might suggest that Shapiro et al.'s (1987, 1989) results were accurate. However, the results may be due to the fact that the four-complements used in this experiment have uniqueness points which occur later in the word than uniqueness points in other verb categories:

Mean uniqueness points (number of phonemes from the start of the word):

Transitive	Nonalt. Dative	Alt. Dative	Two-Comp.	Four-Comp.
4.3	4.3	4.0	4.3	5.3

Since I used Shapiro et al's (1987) verbs as a subset, this might be the source of their effects too. Perhaps lexical identification processes have not reached the same point in the case of the four-complement verbs as have verbs in the other categories when the naming probe is presented. More processing resources therefore might be allocated to the word identification processes in the case of the four-complements, leaving fewer available resources for use in naming the secondary probe.

However, one also must consider the distance between uniqueness point and the offset of the verb. Newman and Dell (1978) found that length of the word preceding a target for phoneme detection significantly influenced mean response time in the phoneme detection task. Phoneme detection latency varied inversely with length of the preceding word. Mehler, Segui, and Carey (1978) reported similar results: reaction times in a phoneme monitoring task were longer when the target phoneme was presented following a short word, independent of the ambiguity or nonambiguity of the short word. These results suggest that longer words or words with uniqueness points further away from the naming target might lead to faster naming latencies, rather than leaving fewer processing resources available for the secondary naming task as I argue above.

Mean number of phonemes remaining in verbs after uniqueness points:

Transitive	Nonalt. Dative	Alt. Dative	Two-Comp.	Four-Comp.
3.2	3.7	1.7	3.5	4.0

These means suggest that the alternating datives should have been the only verbs to induce a difference in naming latencies due to the short distance between uniqueness points and presentation of the naming probe.

CHAPTER 4

GENERAL DISCUSSION

I have no evidence for a syntactic subcategorization complexity effect nor for a semantic argument frame complexity effect. Experiments 1 and 2 provided only negative evidence in answer to the question: is semantic argument frame complexity the relevant complexity metric to consider when constructing a theory of lexical representations?

In Experiment 1, all verb category effects were due to frequency and length, as shown by the fact that regression for frequency left no remaining verb complexity effect. This result and the demonstrated sensitivity of eye movements to syntactic complexity manipulations support the following conclusions.

Experiment 1 showed that not only was there no immediate argument complexity effect, but also there was no delayed complexity effect. This claim is strengthened by the analysis in which the verb and the following article were treated as one region and no verb complexity effect occurred other than one due to differences in word frequency across verb categories. If structure of the verb complement is used to rule out or support particular argument frames during early sentence parsing (for instance on the first-pass parse), I would have seen evidence of this effect in Experiment 1, since there are more frames to consider for four-complement verbs.

One might claim that I failed to get the complexity effect because sentence frames differed over verb categories in my materials.[9] But even though the object noun phrase (region 3) differed across sentences containing verbs from different categories, the syntactic structure of the complement was identical across sentences and the semantic complexity and frequency of lexical items, etc., was fairly homogeneous across categories as a result of the way materials were constructed. Therefore, if there was a delayed effect of subcategorization frame complexity or argument structure complexity, it probably was not obscured by disparities in verb complements across verb categories.

However, to be certain that differences in verb complements did not influence the results of Experiment 1, an eyetracking experiment comparing two-complements and four-complements that occur in identical sentence frames is now underway. The six two- and four-complements from the first two experiments are included, along with two additional verbs in each category, for a total of eight verbs in each category. Each two-complement verb is paired with a four-complement verb of similar frequency. Each of the two verbs in a verb pair then is inserted into two sentence frames such that a subject will see both verbs in a pair and the sentence frame/verb pair variable is between subjects:

> The friend who was sad <u>accepted</u> the terrific present that the attractive stranger sent.
> The friend who was sad <u>remembered</u> the terrific present that the attractive stranger sent.

---

[9]This objection to Shapiro et al. (1987) was raised by Roman Taraban, personal communication, November, 1988.

> The artist who was strange <u>accepted</u> the beautiful job that the young carpenter did.
> The artist who was strange <u>remembered</u> the beautiful job that the young carpenter did.

Comparison of gaze durations for the verbs and first pass reading times for the verb complements of verbs occurring in the same sentence frames will serve as an accurate test of argument frame complexity as reflected in eye movements during silent reading.

Experiment 2 also failed to demonstrate that argument frame complexity influences sentence processing. However, the task was insensitive to frequency and syntactic complexity manipulations too. It is possible that a delayed complexity effect may have occurred, but the second experiment was not designed to enable detection of a delayed effect.

These results suggest an answer to the question: do differences in lexical complexity reflect lexical access processes or post-lexical integration processes? Recall that Shapiro et al. (1987, 1989) and Shapiro and Levine (1988) could not show whether their complexity results were due to lexical identification, memory load, or choice effects. Argument structure information associated with a verb only becomes available once the verb is accessed, that is, once the phonological (or visual) code is successfully matched to the stored representation of the verb. (For the present discussion, it does not matter how the successful match is achieved.) At this point, argument frame complexity might influence reading in several ways. Perhaps it is harder for the reader to activate the information associated with a verb that has many argument frames merely because there is more information to dredge up. The added difficulty may take processing resources away from other tasks competing for the same limited resources. Note that this difficulty of activation occurs before any of the argument frame information associated with the verb can be used by the language processor, since the necessary information is not yet in working memory.

If this were the sense in which argument frame complexity caused processing difficulty, surely the difficulty should have appeared on or immediately after the verbs in Experiment 1. But neither Experiment 1 nor Experiment 2 provide data in support of this account.

Another reason argument frame or syntactic subcategorization frame complexity might influence processing is that it might take more processing resources to maintain various frames over time as the verb complement is being processed. On this story, I should have found evidence for argument frame complexity in the verb complement, even though the evidence may not have been apparent at the verb. Again, Experiment 1 and Shapiro and Levine's (1988) results provide no support for this possiblity.

Still, Shapiro and colleagues repeatedly found an argument frame complexity effect. There are two potential explanations for their findings. The first possible reason why I did not replicate the argument frame complexity effect with the cross-modal naming task in Experiment 2 hinges upon the filler sentences Shapiro et al.

(1987) used. To the 50 test sentences (five verbs from each of five verb classes in two sentence frames each) they added 100 "foils". Foils contained the same main verbs as the test sentences and differed from the test sentences in that foils could have complements other than noun phrase complements (e.g., PP, adverbial phrase, or sentential complements were allowed) and in that probe positions were at locations other than at verb offset in the foils. Each subject heard all 150 sentences and thus probably heard each of the thirty test verbs six times. Some aspect of this repetition, perhaps in combination with the lexical decision task, might have led subjects to develop a special strategy for dealing with the task, and the strategy may be what is tested by their experiment, rather than influence of argument frame complexity on language processing.

The second explanation for their finding of an argument frame complexity effect and my failure to replicate the effect is that the lexical decision task is specifically sensitive to post-lexical-access processes, and the complexity effect occurs at a post-lexical stage of processing. This explanation fits best with the evidence from Experiments 1 and 2, and with the work of Shapiro and his colleagues. If the argument frame complexity effect _is_ specific to the lexical decision task, one can interpret the contrast between their findings and my results from Experiments 1 and 2 as providing information about the structure of the language processing system, as follows.

Results of the two experiments reported in this thesis, taken together with other work that found an argument frame complexity effect, indicate that semantic argument frame complexity does not influence lexical access, nor does it result in a burden on memory as multiple argument frames are maintained until a particular frame is instantiated. Rather, the complexity effect Shapiro and colleagues observed is produced by operation of the thematic processor. When a verb initially is encountered by the language processor, the perceptual representation of the verb is matched to the verb's mental representation. This lexical access is influenced by factors like word frequency and length, and some lexical information, such as syntactic category, is automatically activated along with the word when lexical retrieval is successful. Contrary to Shapiro et al.'s (1987) claim, all argument frames are <u>not</u> activated during initial lexical access; the results of Experiments 1 and 2 reflect this fact. After the initial syntactic representation is constructed by the syntactic module of the language processor, the thematic module checks the parse proposed by the syntactic module for accuracy. To check the parse proposed by the syntactic processor, the thematic processor must access stored semantic argument frame information via the already-activated lexical entry for the verb. (The activated representation of the word acts as a pointer to stored argument frames.) My claim that argument frame complexity effects detected with the CMLD task reflect operation of the thematic processor is supported by Inhoff's (1985) finding of longer gaze durations on false complements following factive verbs. Those longer gaze durations reflected operation of the thematic processor as it found disagreement between the factive presuppositions of the verb and the assertion made in the verb complement.

There are two possible explanations for why argument frame complexity influences operation of the thematic processor:

(1) No argument frame information at all is recovered automatically when the verb's lexical entry is retrieved. Thus, the thematic processor must <u>always</u> confirm the incoming verb complement with all stored argument frames, and since there are more frames to check in the case of four-complement verbs, there is more work for the thematic processor to perform.

(2) Argument frames are ordered according to frequency of occurrence, and at the time of lexical retrieval, only one frame, the most frequent argument frame, is activated with the verb's lexical representation. Other, possible, argument frames are available only by reaccessing their represenations via the verb's activated lexical entry. The thematic processor will not have to reaccess the lexicon if the verb only has one argument frame nor if the most frequent argument frame is instantiated. However, when the incoming complement matches a non-preferred argument frame, the thematic processor searches the lexicon for other argument frames affiliated with the verb.

In either case, I propose that the lexical decision task is sensitive to the thematic processor's reaccessing the lexicon for one of three reasons. It may be because the procedures and mechanisms used to find and consider alternate frames are the same as those used to search the lexicon and access any word's lexical representation. The lexical decision task requires that subjects search the lexicon for an entry that matches the probe. The same processes used to accomplish this task are used by the thematic processor when it searches the lexicon for argument frames associated with a verb's lexical entry. The second explanation for sensitivity of lexical decision to argument frame complexity is that it taps a later stage of processing than does naming and a later stage than that which eye movements reflect. It is at this later stage of processing that the thematic processor does its work; thus, argument frame complexity only then has an influence. Finally, perhaps the lexical decision task artifactually induces the argument frame complexity effect because subjects attempt to integrate lexical decision probe words with the sentence context. They may do so to assist with the lexical decision, since a probe that fits with sentence context will be a word.

To determine which of these three causes led to a complexity effect detectable only with a lexical decision task, I will run an experiment using the cross-modal lexical decision paradigm used by Shapiro et al. (1987) with the same tapes used in Experiment 2 of this thesis. My materials differ from Shapiro's and are carefully controlled so probes are not compatible with sentence context. So, if I also find an argument frame complexity effect with my materials, I will have evidence against the possibility that lexical decision induces complexity effects because subjects attempt to integrate probes with sentence context.

Determining whether no argument frames are initially activated with a verb or whether only the most frequent argument frame is initially activated remains as a question to be addressed in future research. As an initial step towards discovering the time-course of activation of argument frames, I will give subjects the sentences from Experiments 1 and 2 with post-verbal material deleted and have them perform a sentence-completion task. Results will be scored to determine what the verbs' preferred argument frames and subcategorization frames are. These preferences will be used to determine whether results of Experiments 1 and 2 and Shapiro et al's (1987, 1989) results reflect violation of argument frame or subcategorization frame preferences. Evidence for violation of preferred frames will support my claim that argument frame complexity reflects operation of the thematic processor.

APPENDIX A

DATA TABLES

Table 1
VERB CATEGORIES FROM SHAPIRO ET AL. (1987)

Category name:	Subcategorization frames:	Argument frames:
Transitive:	[__NP]	(x,y)
Nonalternating Datives:	[__NP] [__NP PP]	(x,y) (x,y,z)
Alternating Datives:	[__NP] [__NP PP] [__NP NP]	(x,y) (x,y,z)
2-Complement:	[__NP] [__S′]	(x,y) (x,P)
4-Complement:	[__NP] [__S′]	(x,y) (x,P) (x,Q) (x,E)

Table 2
VERBS USED BY SHAPIRO ET AL. (1987)

Transitives:	Nonalternating-datives:	Alternating-datives:
secure	surrender	dig
fix	address	buy
measure	return	send
cherish	restore	lend
exhibit	donate	reserve
Two-complement:	Four-complement:	
regret	discover	
assume	recognize	
accept	remember	
claim	state	
maintain	indicate	

Table 3
PREDICTIONS FOR EXPERIMENT 1

Source of complexity	Order of fixation durations (short-->long)
Syntactic Subcat. Complexity	Transitive < NAD = 2-Comp = 4-Comp < AD (1 S.S.frame) (2 S.S.frames) (3 S.S.frames)
Argument Struct. Frame Complexity	Transitive < NAD = AD = 2-Comp < 4-Comp (1 A.S.frame) (2 A.S.frames) (4 A.S.frames)
Neither Synt. Subcat. nor Arg. Struct. Complexity	Transitive = NAD = AD = 2-Comp = 4-Comp
Both Synt. Subcat. and Arg. Struct. Complexity	Transitives will have the shortest fix. durs., NAD < AD and 2-Comps < 4-Comps
Verb frequency alone	AD < 2-Comp = 4-Comp < Transitives = NAD

Table 4
MEAN FREQUENCY AND WORD LENGTH FOR VERBS USED IN EXPERIMENT 1

Verb category	Mean frequency (words per million) (Francis and Kucera, 1982)	Mean length (letters)
Transitive	66	8
Nonalt. Dative	62	9
Alt. Dative	297	5
Two Complement	160	8
Four Complement	158	9

Table 5
EXPERIMENT 1 RESULTS:
Original Region 2

Mean fixation durations on the verb only
First Fixation and Gaze Duration measures

First fixation, raw times:

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
	Transitive	258	280	269
Verb	Non-alternating	270	298	284
Category	Alternating	258	248	253
	Two-complement	276	283	279
	Four-complement	271	262	267
	Mean	266	274	

Gaze durations, raw times (msec/char in parentheses):

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
	Transitive	323 (38)	341 (40)	332 (39)
Verb	Non-alternating	338 (35)	383 (41)	360 (38)
Category	Alternating	302 (48)	286 (47)	294 (48)
	Two-complement	312 (34)	361 (39)	337 (36)
	Four-complement	311 (32)	330 (34)	320 (33)
	Mean	317 (37)	340 (40)	

Table 6
EXPERIMENT 1 RESULTS:
Original Region 3

Mean fixation durations on the entire object noun phrase

First fixation, raw times:

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
Verb Category	Transitive	249	266	258
	Non-alternating	248	263	256
	Alternating	250	291	270
	Two-complement	277	276	277
	Four-complement	256	244	250
	Mean	256	268	

First pass, raw times (msec/char in parentheses):

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
Verb Category	Transitive	499 (30)	518 (31)	508 (31)
	Non-alternating	601 (35)	575 (34)	588 (34)
	Alternating	586 (34)	560 (32)	573 (33)
	Two-complement	662 (33)	588 (30)	625 (31)
	Four-complement	546 (29)	564 (29)	555 (29)
	Mean	579 (32)	561 (31)	

Table 7
EXPERIMENT 1 RESULTS:
Revised Region 2

Mean fixation durations on the verb plus article

First fixation, raw times:

		Sentence Frame Prep.Phrase	Embedded Relative	Mean
	Transitive	258	279	267
Verb	Non-alternating	269	298	284
Category	Alternating	251	259	255
	Two-complement	274	282	278
	Four-complement	273	262	268
	Mean	265	276	

First pass, raw times (msec/char in parentheses):

		Sentence Frame Prep.Phrase	Embedded Relative	Mean
	Transitive	364 (29)	385 (30)	374 (30)
Verb	Non-alternating	368 (27)	433 (29)	401 (29)
Category	Alternating	329 (34)	374 (37)	352 (37)
	Two-complement	340 (26)	391 (28)	366 (28)
	Four-complement	327 (23)	355 (24)	341 (24)
	Mean	346 (28)	388 (31)	

Table 8
EXPERIMENT 1 RESULTS:
Overall, prepositional phrase version

OVERALL DESIGN: sentence regions within subjects.
Mean fixation durations
ORIGINAL SEGMENTATION

PREPOSITIONAL PHRASE VERSION:

First fixation; Raw times: (msec):

	Region 1 |The neighbor from Main Street|	Region 2 cherished|	Region 3 the fancy gift|	Region 4 that her....|
Transitive:	210	258	249	278
Nonalternating:	211	270	248	246
Alternating:	205	258	250	262
Two-complement:	206	276	277	253
Four-complement:	220	271	256	249

First pass; Raw times (msec) and (msec/char, in parentheses):

Transitive:	842(29)	323(38)	499(30)	494(35)
Nonalternating:	999(36)	338(35)	601(35)	708(41)
Alternating:	866(34)	302(48)	586(34)	697(36)
Two-complement:	817(35)	312(34)	662(33)	648(35)
Four-complement:	930(35)	311(32)	546(29)	557(35)

Total time; Raw times (msec) (msec/char in parentheses):

Transitive:	927 (33)	376(45)	641(39)	593(40)
Nonalternating:	1171 (42)	424(44)	719(42)	741(43)
Alternating:	1012 (40)	380(64)	727(42)	784(40)
Two-complement:	944 (40)	412(44)	761(38)	750(40)
Four-complement:	1033 (39)	405(42)	659(35)	578(37)

Regressions In and (Regressions Out in Parentheses):

Transitive:	.17(0)	.16(.03)	.16(.14)	.11(.41)
Nonalternating:	.24(0)	.10(.03)	.11(.08)	.05(.41)
Alternating:	.26(0)	.10(.07)	.08(.08)	.07(.36)
Two-complement:	.22(0)	.14(.08)	.08(.08)	.06(.43)
Four-complement:	.19(0)	.12(.13)	.10(.11)	.08(.19)

Probability of fixation:

Transitive:	1.0	.94	.99	.92
Nonalternating:	1.0	.97	1.0	.96
Alternating:	1.0	.72	1.0	.94
Two-complement:	1.0	.99	1.0	.97
Four-complement:	1.0	.94	1.0	.92

Table 9
EXPERIMENT 1 RESULTS:
Overall, relative clause version

OVERALL DESIGN: sentence regions within subjects.
Mean fixation durations
ORIGINAL SEGMENTATION

RELATIVE CLAUSE VERSION:
First fixation; Raw times (msec):

	Region 1	Region 2	Region 3	Region 4
	|The neighbor who was lazy|	cherished|	the fancy gift|	that her....|
Transitive:	212	280	266	280
Nonalternating:	214	298	263	254
Alternating:	210	248	291	294
Two-complement:	224	283	276	253
Four-complement:	208	262	244	271

First pass; Raw times (msec) (msec/char in parentheses):

Transitive:	911(32)	341(40)	518(31)	665(43)
Nonalternating:	950(33)	383(41)	575(34)	673(36)
Alternating:	896(35)	286(47)	560(32)	747(40)
Two-complement:	947(37)	361(39)	588(30)	659(36)
Four-complement:	817(33)	330(34)	564(29)	621(37)

Total time;Raw times (msec) (msec/char in parentheses):

Transitive:	1175(42)	474(56)	742(45)	730(47)
Nonalternating:	1173(41)	508(54)	735(43)	746(40)
Alternating:	1142(44)	447(72)	828(48)	810(43)
Two-complement:	1205(48)	496(53)	796(40)	725(39)
Four-complement:	1049(42)	475(48)	743(39)	676(41)

Regressions In (Regressions Out in Parentheses):

Transitive:	.39(0)	.18(.14)	.14(.19)	.06(.52)
Nonalternating:	.41(0)	.12(.20)	.12(.12)	.00(.44)
Alternating:	.46(0)	.32(.26)	.07(.38)	.08(.35)
Two-complement:	.39(0)	.15(.16)	.12(.14)	.06(.46)
Four-complement:	.44(0)	.21(.19)	.10(.18)	.06(.45)

Probability of fixation:

Transitive:	1.0	.99	.99	.89
Nonalternating:	1.0	.99	.99	.87
Alternating:	1.0	.71	.99	.97
Two-complement:	1.0	.94	.97	.96
Four-complement:	1.0	1.0	1.0	.87

Table 10
EXPERIMENT 2 RESULTS:
Percentage correct, questions

Percentage correct responses to comprehension questions

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
	Transitive	.96	.91	.94
Verb	Non-alternating	.90	.97	.93
Category	Alternating	.92	.94	.93
	Two-complement	.88	.98	.93
	Four-complement	.95	.98	.96
	Mean	.92	.96	

Table 11
EXPERIMENT 2 RESULTS:
Naming latencies

Mean naming latencies, averaged over items (msec):

		Prep.Phrase	Sentence Frame Embedded Relative	Mean
	Transitive	491	497	494
Verb	Non-alternating	493	499	496
Category	Alternating	494	505	500
	Two-complement	502	502	502
	Four-complement	506	507	507
	Mean	497	502	

APPENDIX B

FIGURES

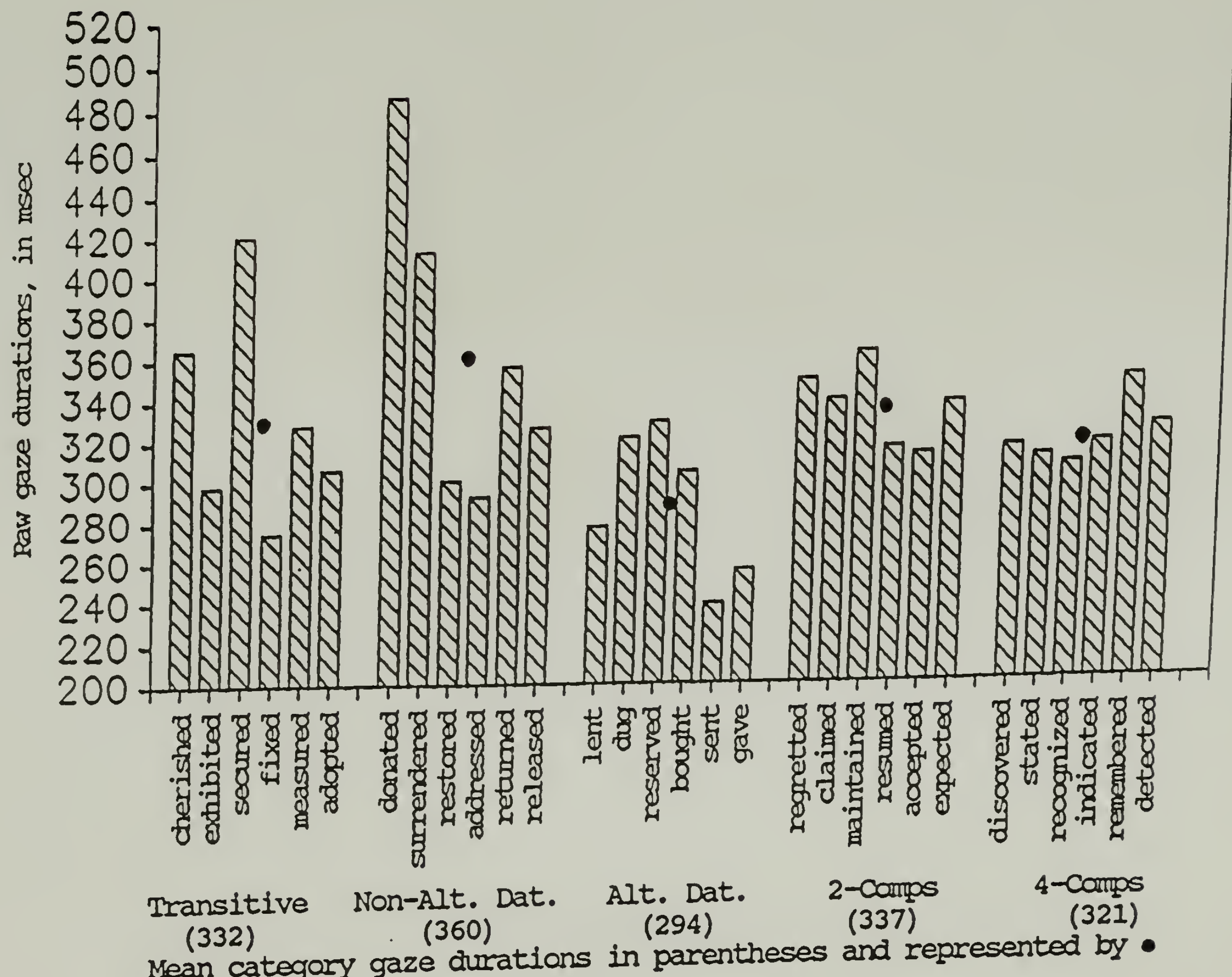


Figure 1. Experiment 1 Mean Gaze Durations on Verbs and Verb Category Means

APPENDIX C

MATERIALS FROM EXPERIMENT 1

(vertical bar indicates line break in original display)
TRANSITIVES:
The neighbor from Main Street cherished the fancy gift that her brother|brought.
The neighbor who was lazy cherished the fancy gift that her brother| brought.

The preacher from Utah exhibited the large project from the new| museum.
The preacher who was sick exhibited the large project from the new| museum.

The student at Harvard secured the creaky door that had swung open.
The student who was smart secured the creaky door that had swung open.

The doctor from Intensive Care fixed the rusty machine at the end of| the hall.
The doctor who was passionate fixed the rusty machine at the end of| the hall.

The president of Heritage Bank measured the big yard behind the house.
The president who was competent measured the big yard behind the house.

The writer from New England adopted the hungry baby that needed a home.
The writer who was angry adopted the hungry baby that needed a home.

NONALTERNATING DATIVES:
The firefighter from Amherst donated the famous painting with the| intricate frame.
The firefighter who was old donated the famous painting with the| intricate frame.

The clerk from Filenes surrendered the old location with the bad view.
The clerk who was short surrendered the old location with the bad view.

The teacher from New Orleans restored the lost wallet that he found| behind the barn.
The teacher who was happy restored the lost wallet that he found| behind the barn.

The father from North Dakota addressed the smug letter that he| intended the landlord to read.
The father who was successful addressed the smug letter that he| intended the landlord to read.

The tenant from the Plaza returned the torn magazine that she received|yesterday.
The tenant who was honest returned the torn magazine that she received|yesterday.

The grandmother from Mexico released the mean boy with the shaved| head.
The grandmother who was feeble released the mean boy with the shaved| head.

ALTERNATING DATIVES:
The salesman from out West lent the thick book about art history.
The salesman who was quick lent the thick book about art history.

The artist from Holland dug a shallow hole that went under a fence.
The artist who was strange dug a shallow hole that went under a fence.

The driver in Las Vegas reserved the hotel room in an expensive part| of town.
The driver who was dirty reserved the hotel room in an expensive part|of town.

The worker from New York bought the fake diamond that was on sale.
The worker who was new bought the fake diamond that was on sale.

The traveler from Africa sent the important package that contained| perishable goods.
The traveler who was hungry sent the important package that contained|perishable goods.

The child from Dallas gave the short speech in her first play.
The child who was cute gave the short speech in her first play.

TWO-COMPLEMENTS:
The lawyer from Miami regretted the hasty decision about the patient's|treatment.
The lawyer who was evil regretted the hasty decision about the patient's|treatment.

The juror from Maryland claimed the black purse which she had lost.
The juror who was wrinkled claimed the black purse which she had lost.

The farmer from England maintained the delicate balance which threatened|to elude him at any moment.
The farmer who was kind maintained the delicate balance which threatened|to elude him at any moment.

The nurse from Las Vegas assumed the unpleasant duties that no one else|would perform.
The nurse who was gentle assumed the unpleasant duties that no one else|would perform.

The man from Berlin accepted the beautiful trophy of a large fish.
The man who was young accepted the beautiful trophy of a large fish.

The builder from Chicago expected a special offer of a new job.
The builder who was lonely expected a special offer of a new job.

FOUR-COMPLEMENTS:
The officer from the First Precinct discovered the hidden prize with| moving parts.
The officer who was serious discovered the hidden prize with| moving parts.

The person with my mother stated the difficult problem that we faced.
The person who was ugly stated the difficult problem that we faced.

The friend from Iowa recognized the famous actor with the beautiful| wife.
The friend who was alone recognized the famous actor with the beautiful|wife.

The man from the Midwest indicated the guilty verdict which the jury| handed down.
The man who was handsome indicated the guilty verdict which the jury| handed down.

The mother from back East remembered the excellent cafe with the| romantic atmosphere.
The mother who was sad remembered the excellent cafe with the| romantic atmosphere.

The girl from London detected the thin smoke that rose from the house.
The girl who was nice detected the thin smoke that rose from the house.

APPENDIX D

NAMING TARGETS FROM EXPERIMENT 2

TRANSITIVES:

Verb:	Set 1:	Set 2:	Set 3:	Set 4:	Set 5:
cherish	bond	gate	cloud	curve	turn
exhibit	call	troop	guest	base	goal
secure	drive	gain	branch	grade	bird
adopt	pound	path	climb	band	guard
fix	pair	king	team	cause	press
measure	bank	page	tube	block	desk

NONALTERNATING DATIVES:

Verb:	Set 1:	Set 2:	Set 3:	Set 4:	Set 5:
donate	guard	bank	troop	branch	curve
surrender	press	pound	page	guest	grade
restore	turn	bond	path	climb	band
release	desk	drive	gate	cloud	block
address	bird	pair	king	team	base
return	goal	call	gain	tube	cause

ALTERNATING DATIVES:

Verb:	Set 1:	Set 2:	Set 3:	Set 4:	Set 5:
lend	band	guard	drive	gain	guest
dig	grade	press	bank	page	team
reserve	cause	bird	bond	king	climb
buy	curve	goal	pair	troop	branch
send	block	desk	call	gate	cloud
give	base	turn	pound	path	tube

TWO-COMPLEMENTS:

Verb:	Set 1:	Set 2:	Set 3:	Set 4:	Set 5:
regret	tube	grade	desk	pound	king
claim	team	curve	guard	call	gate
maintain	cloud	band	press	bank	gain
assume	guest	block	bird	bond	page
accept	branch	cause	turn	drive	troop
expect	climb	base	goal	pair	path

FOUR-COMPLEMENTS:

Verb:	Set 1:	Set 2:	Set 3:	Set 4:	Set 5:
detect	page	climb	band	goal	pair
discover	gates	cloud	grade	bird	bank
state	gain	branch	curve	press	pound
recognize	path	team	cause	guard	bond
indicate	troop	tube	base	desk	drive
remember	king	guest	block	turn	call

BIBLIOGRAPHY

Balota, D.A., and Chumbley, J.I. (1984). Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage. Journal of Experimental Psychology: Human Perception and Performance, 10, 340-357.

Blanchard, H.E. (1985). A comparison of some processing time measures based on eye movements. Acta Psychologica, 58, 1-15.

Carpenter, P.A., and Just, M.A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), Eye Movements in Reading (pp. 275-308). N.Y.: Academic Press.

Clifton, C., Frazier, L., and Connine, C. (1984). Lexical expectations in sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 23, 696-708.

Cutler, A., and Norris, D.G. (1979). Monitoring sentence comprehension. In W.E. Cooper and E.C.T. Walker (Eds.), Sentence Processing: Psycholinguistic Studies Presented to Merill Garrett (pp. 113-134). Hillsdale, NJ: Lawrence Erlbaum.

Cutler, A. (1983). Lexical Complexity and Sentence Processing. In G.B. Flores d'Arcais & R.J. Jarvella (Eds.), The processes of language understanding. N.Y.: Wiley.

Duffy, S.A., Morris, R.K., and Rayner, K. (1988). Lexical Ambiguity and Fixation Times in Reading. Journal of Memory and Language, 27, 429-446.

Fodor, J.A. (1983). Modularity of Mind. Cambridge, MA: MIT Press.

Fodor, J.A., Garrett, M., and Bever, T.G. (1968). Some syntactic determinants of sentential complexity, II: Verb structure. Perception & Psychophysics, 3, 453-461.

Foss, D.J., and Gernsbacher, M.A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. Journal of Verbal Learning and Verbal Behavior, 22, 609-632.

Francis, W.N., and Kucera, H. (1982). Frequency analysis of English usage. Boston, M.A.: Houghton-Mifflin.

Frazier, L. (To appear). Exploring the architecture of the language system. In J. Altman, Ed., Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives. Cambridge, MA: MIT Press.

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

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

Glucksberg, S., Kreuz, R.J., and Rho, S.H. (1986). Context Can Constrain Lexical Access: Implications for Models of Language Comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 323-335.

Grimshaw, J. (1979). Complement structure and the lexicon. Linguistic Inquiry, 10, 279-326.

Hakes, D.T. (1971). Does verb structure affect sentence comprehension? Perception & Psychophysics, 10, 229-232.

Inhoff, A.W. (1984). Two stages of word processing during eye fixations in the reading of prose. Journal of Verbal Learning and Verbal Behavior, 23, 612-624.

Inhoff, A.W. (1985). The effect of factivity on lexical retrieval and postlexical processes during eye fixations in reading. Journal of Psycholinguistic Research, 14, 45-56.

Inhoff, A.W., and Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. Perception & Psychophysics, 40, 431-439.

Just, M.A., and Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.

Kliegl, R., Olson, R.K., and Davidson, B.J. (1982). Regression analyses as a tool for studying reading processes: Comments on Just and Carpenter's eye fixation theory. Memory & Cognition, 10, 287-296.

Mehler, J., Segui, J., and Carey, P. (1978). Tails of words: Monitoring ambiguity. Journal of Verbal Learning and Verbal Behavior, 17, 29-36.

Newman, J.E., and Dell, G.S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. Journal of Verbal Learning and Verbal Behavior, 17, 359-374.

Onifer, W., and Swinney, D.A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. Memory & Cognition, 9, 225-236.

Rayner, K., Sereno, S., Morris, R., Schmauder, A., and Clifton, C. (In press). Eye movements and on-line language comprehension processes. Language and Cognitive Processes.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. Memory & Cognition, 4, 442-448.

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

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

Rayner, K., and Frazier, L. (1987). Parsing temporarily ambiguous complements. Quarterly Journal of Experimental Psychology, 39A, 657-673.

Rayner, K., and Pollatsek, A. (1989). The Psychology of Reading. Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., and Pollatsek, A. (1987). Eye Movements in Reading: A Tutorial Review. In M. Coltheart (Ed.), Attention and Performance XII: The Psychology of Reading. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.L., and Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 14, 489-537.

Seidenberg, M.S., Waters, G.S., Sanders, M., and Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. Memory & Cognition, 12, 315-328.

Shapiro, L.P., Zurif, E., and Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. Journal of Psycholinguistic Research, 18, 223-243.

Shapiro, L.P., Zurif, E., and Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. Cognition, 27, 219-246.

Shapiro, L.P., and Levine, B.A. (1988). Verb processing during sentence comprehension in aphasia. Manuscript submitted for publication.

Swinney, D.A. (1979). Lexical Access during Sentence Comprehension: (Re)Consideration of Context Effects. Journal of Verbal Learning and Verbal Behavior, 18, 645-659.

Swinney, D.A. (1982). The structure and time-course of information interaction during sentence comprehension: lexical segmentation, access, and interpretation. In J. Mehler, E. Walker, and M. Garrett (Eds.), Perspectives on mental representation. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Tanenhaus, M.K., Leiman, J.M., and Seidenberg, M.S. (1979). Evidence for Multiple Stages in the Processing of Ambiguous Words in Syntactic Contexts. Journal of Verbal Learning and Verbal Behavior, 18, 427-440.

West, R.F., and Stanovich, K.E. (1982). Sources of inhibition in experiments on the effect of sentence context on word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 385-399.