Syntactic Effects from Lexical Decision in Sentences: Implications for Human Parsing

by

Barton Day Wright

B.A. Swarthmore College
(1976)
M.A. University of Pennsylvania
(1978)

SUBMITTED TO THE DEPARTMENT OF
PSYCHOLOGY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS OF THE
DEGREE OF

DOCTOR OF PHILOSOPHY
IN PSYCHOLOGY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1982

Signature of Author Barton Day Wright
Department of Psychology
October 8, 1981

Certified by
Merrill F. Garrett
Thesis Supervisor

Accepted by
Richard Held
Chairman, Departmental Graduate Committee

Syntactic Effects From Lexical Decision in Sentences:
Implications for Human Parsing

Abstract

A new experimental paradigm was explored as a tool for understanding human parsing. The overall pattern of results suggests that the paradigm does indeed reflect the operation of the human parser, and that the parser predicts the heads of predictable phrases.

An English sentence was presented across a computer CRT one word at a time (at reading speeds), from left to right. The sentence was terminated prematurely by the appearance of a lexical decision target on the same line. The grammatical structure of the complete sentence affected lexical decision RT, depending on the part of speech of the target. Semantic and pragmatic effects were controlled by using targets which were unpredictable in the sentence context. In some experiments, target word pairs differed only in derivational morphology (e.g. TRANSLATES/TRANSLATION), making nonsyntactic explanations especially unlikely.

In Experiment 1, modal verb contexts with main verb targets and preposition contexts with noun targets produced lower RTs than the opposite pairings. In Experiment 2, transitive verb contexts with noun targets and subject noun phrase contexts with verb targets yielded lower RTs than the opposite pairings. The use of adjective and adverb targets with the same sentence contexts did not yield an expected RT elevation of the noun phrase - adjective combination. In Experiment 3a, where contexts ended with a double-object verb followed by an indirect object noun phrase, RTs for noun and verb targets did not differ, but error rates were high. In Experiment 3b, the indirect object noun phrase was replaced by a prepositional phrase, and the noun targets yielded lower RTs. Experiment 3 results suggest difficulty in processing indirect objects. In Experiment 4, adjective targets produced lower RTs when heads of phrases (after stative verb - adverb sequences) than when legal but not phrase heads (after object-taking verbs). Experiments 5 and 6 did not yield expected effects, and explanations are discussed. For Experiments 7 to 9, the paradigm was modified slightly by providing continuation of the sentence beyond the target word. Results of Experiments 1 and 2 were replicated in the error results, though some RT effects were still found. Experiment 10 explores the behavior of closed class targets in the paradigm.

Thesis Supervisor: Dr. Merrill F. Garrett

Title: Professor of Psychology

To oppressed people everywhere,

but especially to those of the third world

who suffer unduly under capitalism

Contents

Acknowledgments

I am grateful to my thesis advisor, Merrill Garrett, for his advice and ideas throughout this research. I thank the other members of my thesis committee for serving on it: Jerry Fodor, Richard Held, and Joan Bresnan, and especially Joan for her attentive comments on an earlier draft. Sean True was of great assistance with the computer. He not only made available (and wrote) the DAMON language under which the experiments were controlled, but also answered endless questions on problems digital. Many people have been of assistance in diverse ways, for instance by checking my sents and ences, discussing the literature, pondering recalcitrant data, and resolving statistical dilemmas: Carmen Egido, Peter Gordon, Jim Hodgson, Howard Kurtzman, Janet Lang, Nina Maragioglio, Connie Parkes, Molly Potter, and Janell Schweickert (but especially Nina and Connie).

My most profound debt is to Edward Walker, who has been an unfailing source of encouragement and moral support throughout my years at MIT. He has also been of great help with every aspect of the research process, from formulating hypotheses to wiring switches to revising write-ups.

Biographical Note

Bart Wright was born in Ann Arbor, Michigan in 1954, and raised in Durham, New Hampshire. He attended Swarthmore College (B.A., 1976), where he majored in Psychology, toward the end dabbling in pigeons (with Barry Schwartz) and human perception (with Hans Wallach). In 1977 he began the Ph.D. program in psychology at the University of Pennsylvania. There he came under the influence of Lila Gleitman and Daniel Osherson, and acquired an interest focused on syntax. He left Penn for M.I.T. in 1978.

The struggle for social justice has, in various forms, occupied much of Bart's attention. He combatted nuclear power as an active member of the Clamshell Alliance in 1976-77, and was from 1978 until its demise in 1980 a member of Cooperative College Community, a group attempting to form a socialist tuition-free college. His present focus for action is on preventing the U.S. arms buildup.

List of Examples

List of Tables

* * * INTRODUCTION * * *

The experiments reported in this thesis are directed toward an understanding of human parsing. Human parsing is the transformation of a sentence as a string of words into a structured representation, which is in turn used to determine the meaning of the sentence. It is the segregation of words into phrases, phrases into clauses, and the determination of the grammatical roles the words, phrases and clauses play with respect to one another. It is widely believed that these processes must be carried out to understand the innumerable utterances and written sentences which confront us every day (Fodor, Bever, and Garrett, 1974; Clark and Clark, 1977).

Many good ideas and a few interesting theories have been presented in recent years on the subject of human parsing (e.g. Bever, 1970; Kaplan, 1972; Kimball, 1973; Wanner, Kaplan, and Shiner, 1975; Limber, 1976; Frazier and Fodor, 1978; Marcus, 1980; Ford, Bresnan, and Kaplan, in press). In contrast to most other areas of cognitive psychology, this work is characterized more by theories and speculations than by experimental investigations. The scarcity of experimental work is due at least in part to the difficulty of conceiving of paradigms of research which will measure parsing processes.

One method that has been used is the collection of intuitions (Frazier and Fodor, 1978; Marcus, 1980). Someone reads or listens to a sentence and introspects about the processing he or she is carrying out. The judgment called for is reasonably precise, e.g. whether conscious reprocessing of the sentence is necessary, or which of two similar sentences is easier to process. This is an obvious first choice for a method, since it is very easy and direct. It also has a venerable history as the chief method of linguists in developing and refining theories of transformational grammar.

However, the judgments requested by the linguist are different in kind from those requested by the psycholinguist. The linguist asks whether a sentence, once fully processed, is a grammatical sentence of the language. The psycholinguist asks whether the processing of the sentence, as it is heard or read, is difficult. Intuitions about such rapidly occurring mental processes are often unreliable and unhelpful (Posner, 1973). Certainly one wants experimental corroboration of intuitions.

One requirement of experimental tasks for testing parsing theories is that they tap sentence processing as it is occurring. If one does not take a measure of the reader or listener's performance until the processing of the sentence has been completed, it is most likely that no trace of the course of processing will remain. Consider the sentence, "John lifted a hundred pound bags." Interpreted correctly, it means that John lifted a hundred bags, each containing one pound of something. Marcus (1980) claims, on the basis of intuitions, that this is a garden path sentence. People start to interpret it on the model of "John lifted a hundred pound bag" and must subsequently revise their analysis due to incompatible information. If one presented these sentences in a list, and asked for recall later, it would be perfectly reasonable if their recall were uninfluenced by the momentary processing difficulty; the correct representation of the sentence is likely to be stored with no indication of the difficulty encountered in achieving it.

Psycholinguistics has accumulated a substantial inventory of on-line processing measures, sharing the basic feature that a subject is required to perform some task irrelevant to comprehension while reading or listening to a sentence. The irrelevant task may be phoneme monitoring, mispronunciation detection, click detection, or speech shadowing (for a review see Cutler and Norris, 1979). All these tasks attempt to provide a measure of processing

load for some well-localized point in the sentence. They assume that a fixed processing capacity must be divided between the two competing tasks. Longer latency to perform the irrelevant task (or poor performance) indicates an increased processing load. In principle, any prediction of differential processing difficulty made by a parsing theory could be tested by these techniques. In practice, many of the techniques have limitations (such as sensitivity) which restrict their usefulness.

The experimental paradigm used in the present research is lexical decision within a sentence context. Lexical decision is a task that has been used primarily to study lexical access, or how words are recognized for linguistic use (e.g. Taft and Forster, 1976). In the lexical decision task, a subject is presented with a letter string which is either a word or an orthographically legal nonword (such as "bonch"). The subject, under time pressure, makes one of two responses to indicate whether or not the letter string is a word. A faster response for one type of word is usually taken to indicate that that type of word is accessed more quickly. Some investigators have placed their lexical decision targets after sentence contexts (Fischler and Bloom, 1979, 1980; Kleiman, 1980; Schuberth and Eimas, 1977; Schuberth, Spoehr, and Lane, 1981; Simpson, 1981). In such cases, lexical decision becomes an on-line measure, though not a measure of overall processing difficulty. Instead, the lexical decision time is taken to reflect directly the effect of the sentence context on recognition of the word itself.

Swinney (e.g. 1979) has used a lexical decision task within a sentence context, but with an entirely different emphasis. In the Swinney paradigm, a visual lexical decision target appears during the auditory presentation of a sentence. Ambiguous words in the sentence prime lexical decision targets which are related to either reading of the ambiguous word, even in sentences

which strongly bias for one reading of the ambiguous word. The important point is that Swinney's focus is on the ambiguous words that are the source of the priming, rather than on the priming itself.

Like Swinney's work, the present experiments use a lexical decision task within a sentence context, and focus on the source of an influence on lexical decision. Word recognition latency serves merely as a convenient dependent variable. In contrast to the focus on meaning relationships between sentence contexts and lexical decision targets that has characterized previous studies, the present work is concerned entirely with syntactic variables.

The Paradigm

In the experiments in this paper, a sentence of English is presented in normal orthography from left to right across a computer CRT, one word at a time, with words remaining on the screen after they appear. At some point in this presentation, a lexical decision target appears; the word (or nonword) target is in capital letters and it is presented to the right of the incomplete sentence, just where the next word of the sentence would have appeared. After the subject responds, the screen is cleared and the next trial begins. (This paradigm is modified somewhat in Experiments 7, 8, and 9.)

A constant concern with such experimental tasks is whether they tap the same processes as everyday language use. While this paradigm differs in many ways from normal reading, it also approximates many aspects well. The sentence contexts are meaningful and cover a broad range of topics; the text is presented in lower case type, except where capitalization is normally called for; and the presentation rate approximates normal reading speed (either 300 or 200 words per minute).

A basic assumption of the paradigm is that the various components of the language comprehension system which are normally engaged during sentence perception are active when the target appears and will influence performance on the lexical decision task. The component(s) concerned with semantic and pragmatic interpretation should produce relatively faster response times to targets which would be expected next in the sentence based on real world knowledge and logical constraints. Similarly, the component concerned with syntactic interpretation -- the parser --should produce relatively faster response times to targets that are in accord with the parser's expectations.

To observe the syntactic effects, the influence of the semantic and pragmatic components must be neutralized. This is accomplished by comparing RTs for two target words which differ in their syntactic properties, but are both of extremely low predictability for the sentence fragment they occur after. With universally weak pragmatic relations between sentence context and target word, we can make a fair claim to be observing truly syntactic effects. In Experiments 2 and 8, we take the additional step of equating semantic and pragmatic relationships by contrasting, after the same sentence context, targets such as TRANSLATES and TRANSLATION.

Note that if subjects were presented with an entire set of materials in which the lexical decision target was always semantically unrelated to the sentence context, they might adopt a strategy which insulated the lexical decision task from the effects of sentence processing. To discourage such strategies, roughly half of the word items in the experiments reported here are fillers in which the target word is a satisfactory and meaningful completion of the sentence. Also, the number of words before the point of interruption varies, so that subjects have no way of knowing beforehand where the interruption will occur.

Assuming that the paradigm can in fact capture syntactic effects, it

will be used to measure relative "preferences" of the parser for words of different parts of speech after a variety of sentence context types. The hope in initiating this line of inquiry is that it will provide some useful constraint on the nature of the human parser and its operation.

There are a number of basic questions about human parsing and about its relation to the experimental paradigm which this research finesses, in order to concentrate on the questions of relative "preference" for given syntactic types at specific loci in a sentence. For instance, it is not known whether the parser actively predicts linguistic structures, or whether it merely accepts words passively. It is not known whether the parser makes very fine discriminations of the acceptability of following words, or a very gross binary discrimination. It is not known whether the parser's effect on lexical decision is to facilitate preferred words, or to inhibit nonpreferred words. And it is not known whether this effect acts on the recognition of the word itself (access) or rather on post-access processes within the lexical decision task. The usefulness of the paradigm does not depend on particular positions on any of these questions, but it is possible that results from the paradigm will provide evidence on some of them.

The fact that these questions are unanswered does not jeopardize the validity of the intended comparisons, but it does pose problems of terminology. In the exposition, I will frequently write, "the parser prefers A to B" or "the parser predicts a noun phrase" or "the parser produces an elevated time to word B", even though the parser is not animate and does not prefer, the parser might not predict anything, and the parser might be facilitating other words instead of inhibiting word B. Cumbersome and lengthy expressions would be required to avoid these imprecise usages.

* * * PRELIMINARY EXPERIMENTS * * *

EXPERIMENT 1

The purpose of this experiment was to determine the basic effectiveness of the paradigm for cases where we have strong grounds for expecting parsing contrasts. First some terminology. Each item consists of the initial portion of a sentence of English followed by a word or nonword lexical decision target. The incomplete sentence preceding the target in an item will be called a sent. Example 1 shows a quadruple of items, consisting of a pair of sents crossed orthogonally with a pair of target words.

Example 1.

MV	If your bicycle is stolen, you must FORMULATE
MN	*If your bicycle is stolen, you must BATTERIES
PV	*For now, the happy family lives with FORMULATE
PN	For now, the happy family lives with BATTERIES

The first sent (M) ends with a subject noun phrase followed by a modal verb. It is easy to continue this sent with a main verb (V) such as FORMULATE, but virtually impossible to continue it with a plural noun (N) such as BATTERIES. (The asterisk before versions MN and PV signifies that the sent cannot be grammatically continued with the target word). If the experimental assumptions are minimally correct, we expect the reaction time for version MV to be lower than that for version MN. For the sent ending in a preposition (P), the prediction is reversed. It is very hard to continue this sent with an uninflected verb such as FORMULATE, but very easy to continue it with a plural noun such as BATTERIES. Correspondingly, we expect the reaction time for version PN to be lower than that for version PV.

The link between the ease with which a sentence can be continued and the expected operation of the parser should be clear. When the target does

not itself complete a sentence, and when no sentence can be formed by adding additional words, it follows that the parser has no structural role to assign to the target word. All adequate parsers would reflect this fact in some way; it might well be visible in this task as a "preference" for targets allowing legal continuations over those that don't.

The legal versions (MV and PN) are chosen so that the target word is not only a legal continuation, but also predictable because of the requirements of the grammar. In particular, the target word is the correct part of speech to be an obligatory part of a predictable phrase. In the PN case, the preposition sent allows the prediction of a noun phrase object of the preposition, which the noun target can fulfill. In the MV case, the modal sent allows the prediction of a verb phrase, which the verb target can fulfill. By making the "bad" cases as bad as possible, and the "good" cases as good as possible, we increase the chances that the distinctions that separate good from bad include at least one which the human parser is sensitive to, and will reveal to us in this task.

METHOD

The entire experiment described below was performed using 10 subjects (1a), and the preliminary results were sufficiently interesting to warrant a replication in which the particular items were changed (1b). Because the two experiments are conceptually identical, they will be discussed together.

Subjects

MIT undergraduate volunteers were paid for their participation. All were native speakers of English.

Materials

The complete set of Experiment 1 materials with word targets is shown in Appendix A. The quadruple of items in Example 1 is prototypical. 12 such quadruples were constructed, each consisting of a pair of sents crossed with a pair of target words. One member of each sent pair ends in a modal verb, the other in a preposition. Sents range in length from 5 to 12 words. The two sents in each pair are the same length in words, and matched approximately for typed length (i.e. the total number of characters of the two was matched.) No attempt was made to match the sents of each pair closely for grammatical structure.

The first of each pair of target words is an uninflected verb; the second a plural noun. All verbs and nouns are unambiguous as to grammatical category (Note 1). The noun and verb in each pair were matched for length in letters and syllables, and approximately for frequency of occurrence in written English according to the Kucera-Francis (1967) count. Words range in length from 5 to 9 letters.

Nonword pairs were constructed with the "noun" nonwords ending in -s, and the "verb" nonwords not ending in -s. These nonword pairs were assigned to 12 different pairs of sents that were designed by the same criteria. In Experiment 1b (the replication) the word targets were assigned to this group of sents, and the nonwords were assigned to the sents that took word targets in Experiment 1a.

30 filler items were constructed in which the target words were a natural continuation of the sent. In 24 of these cases, the target word also completed a sentence of English. Example filler items are shown in Example 2.

Example 2.

There are six churches in this small TOWN
As the man fell down the stairs, he broke his GLASSES
Good raisins are dried in the sun PATIENTLY
At noon the whistle in the factory went OFF
At the same time we must not forget OUR

30 additional filler items with nonword targets were included to maintain a balance of word and nonword items (Note 2).

The materials were grouped into two sets to be seen by two different groups of subjects. For each quadruple of experimental items, versions MV and PN were presented to one subject group, and versions PV and MN were presented to the other. This was counterbalanced so that each subject saw the 4 item versions equally. The items were arranged in a single pseudorandom order (by sent) for both subject groups. The 132 items of the experiment were divided into two equal blocks, and preceded by 12 representative practice items.

Procedure

The subject was seated before the CRT of a TERAK computer in a dimly lit room, holding a small box with two side-by-side microswitches. The subject operated one switch with each hand, the right for word responses and the left for nonword.

A fixation cross appeared at the left margin of the CRT before each trial. At the start of the trial the fixation mark disappeared, and the sent appeared from left to right across the screen, a new word appearing every 200 msec (300 words per minute). 100 msec after the appearance of the last word of the sent, a brief tone sounded. 300 msec after the tone onset, the target word appeared in capital letters to the right of the last word. The sent and target remained on the screen until the subject made a lexical decision. There were thus 3 cues indicating the target word: a tone was sounded 300 msec before its appearance, it appeared 400 msec rather than 200

msec after the previous word, and it appeared in all caps. The reaction time and response (yes or no) were recorded with millisecond accuracy and stored automatically by the computer. The presentation of items was self-paced by the subject via a foot pedal, though a minimum of one second separated trials.

Subjects were instructed as to the procedure and the nature of the lexical decision task. They were informed that some target words would fit with the preceding sentence and some would not, but it was emphasized that they should not concern themselves with such differences. Subjects received no feedback on their performance.

RESULTS AND DISCUSSION

For both the initial experiment (Experiment 1a) and the replication experiment (Experiment 1b), 10 subjects were run, 5 in each condition. Four subjects in the initial experiment and 2 in the replication were replaced due to average RTs over 850 msec or RT standard deviations over 300 msec. None exceeded the 10% error criterion. RTs more than two standard deviations above the subject's mean were replaced with that subject's cutoff value.

The error rate (including filler items) in the initial experiment was 3.1%, and in the replication 4.5%.

Previous robust findings with the lexical decision task were replicated in this paradigm. For instance, for the Experiment 1a fillers, words yielded decision times 88 msec faster than nonwords (min $F'(1,40) = 9.6$, $p < .01$). The semantically predictable word targets of the filler items yielded RTs 72 msec faster than the semantically unexpected word targets of the experimental items (min $F'(1,48) = 9.7$, $p < .01$). And, for the experimental materials of Experiments 1a and 1b combined, RT correlates with

the log of Kucera-Francis frequency ($r(23) = -.46$, $p = .024$). The other experiments in this thesis generally show comparable findings.

Table 1 shows the RT results for Experiment 1 itself.

Table 1. Experiment 1.
Mean Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word.
Error percentages are in parentheses.

	Initial (1a): Target Verb	Initial (1a): Target Noun		Replication (1b): Target Verb	Replication (1b): Target Noun	
Sent: Modal	652 (10)	707 (3)	679	668 (3)	718 (3)	693
Sent: Prep.	696 (7)	669 (2)	681	738 (7)	668 (12)	703
	674	688		703	693	

The predicted result was that versions MV and PN would produce faster RTs than versions MN and PV. This prediction is supported. The interaction effect in a 2 x 2 analysis of variance (Note 3) is 82 msec in the initial experiment, 120 msec in the replication. This is significant in the initial experiment (min $F'(1,18) = 4.9$, $p < .05$) and approaches significance in the replication experiment (min $F'(1,17) = 3.5$, $p < .10$). Neither main effect approaches significance in either experiment. Note that the replication is not only a replication over subjects. Though the particular target words are the same, the sents and the pairing of sent with target word are completely different in the initial and replication experiments. If the results of the initial and replication experiments are combined and treated as a single experiment, the interaction effect is stronger (min $F'(1,34) = 7.6$, $p < .01$), while main effects are still far from significance.

The experimental paradigm has provided RT differences distinguishing items where the target is an illegal continuation from those where the

target is an obligatory part of a predictable phrase. Modal verbs engender faster decision times for verbs; prepositions engender faster times for nouns. This holds even when the sent context bears no useful relation to the meaning of the particular verb or noun target. More general implications are best discussed following Experiment 2.

EXPERIMENT 2

Two differences from Experiment 1 mark this experiment: 1) different syntactic contrasts are tested, and 2) an additional control for semantic effects is used.

The target word pairs of Experiment 1 (such as FORMULATE/BATTERIES) were matched for length, frequency, and number of syllables. However, they also differ in a host of ways, including their meaning and use in the language. To make a stronger case that the effects observed with this paradigm are truly syntactic, target word pairs such as TRANSLATES/TRANSLATION were used in Experiment 2. These are in an important sense the same word, differing only in the suffixes which determine their part of speech (Note 4).

Instead of using modal verbs and prepositions as the final words of the sentence context, nouns and verbs were used. In particular, each noun was the last word in a singular subject noun phrase, and each verb was the main verb of a tensed clause. Verbs that usually take a direct object were used. When crossed with the target word pairs, quadruples such as that in Example 3 result.

Example 3.
NV The crowd near the church indicates that an important funeral TRANSLATES
NN +The crowd near the church indicates that an important funeral TRANSLATION
VV *The towers on the horizon indicate that the barriers isolate TRANSLATES
VN The towers on the horizon indicate that the barriers isolate TRANSLATION

The targets in the two "good" cases (NV and VN) have the same status as in the good cases of Experiment 1: they are the correct part of speech to be an obligatory part of a predictable phrase. One of the "bad" cases (VV) is also parallel to Experiment 1: it cannot be grammatically completed.

The other "bad" case (NN) is of great interest. A noun can often be followed by another noun, as can be seen in another possible continuation of

the sent in version NN, "The crowd near the church indicates that an important funeral PROCESSION (is about to leave.)" The first noun functions as an adjective, though this construction is so productive that one would not want to say that FUNERAL is an adjective as well as a noun. The NN version is thus not syntactically illegal. But we do expect that the parser would prefer a verb to a noun after receiving FUNERAL, i.e. that it would prefer version NV over version NN. (The " + " in Example 3 signifies the expected processing difficulty). This preference follows from the attachment of FUNERAL as the head of the noun phrase, and this attachment decision is consistent with both the Minimal Attachment Principle of Kimball (1973) and the Syntactic Preference Principle of Ford, Bresnan, and Kaplan (in press). This NV-NN comparison will determine whether the paradigm differentiates between two legal cases, one of which we believe is "preferred" by human parsers. In Experiment 1, the "bad" cases (PV and MN) could not be completed at all; the item had become ungrammatical.

In sum, we predict the decision times for versions NV and VN to be lower than those for versions VV and NN. Version NN might be somewhat lower than version VV because version NN can be continued grammatically.

Another section of Experiment 2 broadened the repertoire of target word categories. The very same sents ending in nouns and verbs were paired with adjective and adverb targets, to yield quadruples such as that shown in Example 4.

Example 4.

NA ++The crowd near the church indicates that an important funeral HAPPY
NR The crowd near the church indicates that an important funeral HAPPILY
VA The towers on the horizon indicate that the barriers isolate HAPPY
VR The towers on the horizon indicate that the barriers isolate HAPPILY

Version NA can be completed by taking the adjective to be introducing a reduced relative, as in "The tallest building near the hotel [which is] EASY -- to spot is the Empire State Building." This is a construction which

we expect to be troublesome to the parser, on the basis of intuitive data as well as such principles as Minimal Attachment. The adjective in the VA version can easily serve as part of the object noun phrase. Versions NR and VR are easily completed, since adverbs occur freely in these positions. Thus, on the basis of intuition, we expect version NA to yield an elevated RT, compared to the other three versions.

Both the adjective- and adverb-target materials just discussed and the verb- and noun-target materials presented earlier require balancing nonword items. Since the word targets are marked with suffixes, the nonword targets are as well. This allows a systematic look at whether sentence context can influence the rejection of nonwords marked clearly as to their grammatical category. The same kind of sents were followed by matched nonword pairs such as FABLORATES/FABLORATION and SERICAL/SERICALLY, each of which is marked quite clearly with suffixes indicating its part of speech.

METHOD

Subjects

MIT undergraduate volunteers were paid for their participation. All were native speakers of English. None had participated in Experiment 1.

Materials

The entire set of Experiment 2 materials appears in Appendix B; the two quadruples shown in Examples 3 and 4 are prototypical. Consider first Example 3. Twelve such quadruples were constructed, each consisting of a sent pair matched with a target word pair. The first member of each sent pair ended in a noun, the second in a verb. In particular, the noun would naturally be construed as ending a singular noun phrase serving as the

subject of its clause. The verb was always the main verb of a tensed clause, and verbs were chosen that require noun objects. The two sents in a pair were matched rather closely for length, overall grammatical structure, and characteristics of the final words.

12 verb/noun target word pairs were constructed, the noun derived from the verb by addition of a derivational suffix. The verb occurred in its third person singular form, inflected with an "s". Both members of all pairs were unambiguous as to grammatical category. Pairs were also carefully selected so that they shared all of their meanings. Matching for frequency proved impossible, since derived nouns are nearly always less frequent than their base verbs. The ratio of verb to noun frequencies (summed over inflections) was approximately 2:1 for all pairs. The nouns were generally longer than their corresponding verbs, in number of letters and number of syllables.

The very same 12 sent pairs were combined with 12 adjective/adverb pairs, creating items such as are shown in Example 4. The adverb in each pair was formed from the adjective by adding the -ly morpheme. Both members of each pair were unambiguous as to grammatical category. As with the verb/noun pairs, the adjective and adverb shared all their meanings. Frequency matching was again impossible. The adjective to adverb frequency ratio was roughly 3:1.

For each of the 12 sent pairs there were a total of 8 versions, four with verb/noun targets as in Example 3, and four with adjective/adverb targets as in Example 4. These items were shown to four subject groups in a counterbalanced design. Each subject saw a particular sent only once, and only one member of a target word pair. No subject saw both TRANSLATES and TRANSLATION, for instance. Each of the four subject groups received an even distribution of the 8 different types of items.

Nonword targets were constructed, and similarly assigned to 12 different sent pairs which were constructed according to the same criteria as the word-item sent pairs. 12 "verb"/"noun" nonword pairs such as FABLORATES/FABLORATION were constructed, superficially comparable to the verb/noun word pairs. All the "verb" nonwords had one of the suffixes -ate, -ize, or -ify, followed by the -s morpheme. All the "noun" nonwords were formed by adding the suffix -ion to the verb. 12 "adjective"/"adverb" pairs such as SERICAL/SERICALLY were constructed on the model of the word pairs. All the "adjective" nonwords ended in -able, -ous, -ful, or -al. The "adverb" nonwords were formed by adding -ly to the "adjective" nonwords.

As in Experiment 1, half the items seen by the subject were word and half were nonword. Of the word items, half were fillers constructed so that the target word was a natural continuation of the sent (and usually a natural completion as well) (Note 5).

The items were presented in different pseudo-random orders to different subjects. The entire randomized set of 208 items was divided into two equal blocks, each preceded by 10 warm-up items. There was also an initial block of 30 representative practice items.

Procedure

The procedure was identical to that of Experiment 1, with one modification. An initial execution of this experiment using the Experiment 1 presentation rate of 200 msec per word revealed only hints of expected effects. A careful look at the data suggested that many subjects were falling behind in their reading of the sent, particularly for longer sents. Since they had not reached the end of the sent when the target word appeared, their reaction times did not show an effect of syntactic context. The somewhat greater length and complexity of Experiment 2 materials might

account for difficulties in Experiment 2 not encountered in Experiment 1.

The presentation rate for Experiment 2 was slowed to 300 msec per word (200 words per minute), the target word still appearing 400 msec after the last word of the sent. The delay preceding the target word was much less noticeable than at the faster presentation rate, but subjects still had no difficulty identifying which word was the target of their lexical decision.

The entire experiment took a half hour.

RESULTS AND DISCUSSION

40 subjects were run, 10 in each condition. 8 subjects were replaced due to an error rate over 10%; 10 were replaced due to a mean RT over 850 msec, or an RT standard deviation over 250 msec. The overall error rate (including filler items) was 3.9%. RTs more than two standard deviations above the subject's mean were replaced with that subject's cutoff value.

Word Items With Verb and Noun Targets

Table 2 shows the results for the materials exemplified in Example 3.

Table 2. Experiment 2, Word Items, Verb/Noun Targets. Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word. Error percentages are in parentheses.

		Target: Verb	Target: Noun	
Sent-Final Word	Noun	660 (9)	714 (7)	685
	Verb	706 (10)	673 (7)	690
		683	694	

The results are as predicted; the syntactically illegal VV version and the syntactically nonpreferred NN version yield higher RTs than the syntactically preferred VN and NV versions. The 87 msec interaction effect in a 2 x 2 analysis of variance is significant (min $F'(1,43) = 8.7$, $p < .01$). Neither main effect approaches significance.

The failure to find a main effect of target type is somewhat surprising. Recall that the noun and verb targets have systematic differences in length and frequency, which would lead us to suspect a faster time for the verb targets. Perhaps the absence of a main effect of target type is due to some unsuspected effect of sentence context canceling out

length and frequency effects. To address this possibility, two additional groups of 10 subjects performed a simple lexical decision task on the targets of Experiment 2, without sentence contexts. Once again, no subject saw both TRANSLATES and TRANSLATION. The verb and noun targets did not differ in this baseline condition either, indicating that there is no basic difference between these verbs and nouns in recognition time, despite their frequency and length differences.

The important result of this experiment is the interaction effect. Taken together with the results of Experiment 1, it provides good evidence that the experimental paradigm works. The structure of an incomplete sentence interacts with the grammatical category of the next word, as measured by reaction time to accept the target word in a lexical decision task. This has been observed with two completely different pairs of sent syntactic types in the two experiments.

Though this work is aimed at an understanding of parsing, these results take on great interest when viewed in another light. We have observed two very clear instances of syntactic priming in a lexical decision task. Semantic priming using the lexical decision task is a common finding, but syntactic priming in a lexical decision task has not previously been directly displayed. These results show syntactic influence on lexical recognition processes, though we cannot at this point say what aspect of the process is being affected.

The very close semantic relation of the members of target word pairs (such as TRANSLATES/TRANSLATION) in this experiment makes a semantic explanation of the results highly unlikely. Any influence of the sentence context on the meaning of TRANSLATES should be equally an influence on the meaning of TRANSLATION. The fact that the two words behave differently can plausibly be attributed only to their syntactic category. In turn, the

influence on the decision times must derive from a syntactic manifestation of sentence processing.

The results for the NN case bode well for the usefulness of the technique in studying human parsing. Consider that the VV case ("The towers on the horizon indicate that the barriers isolate TRANSLATES") cannot be grammatically completed. It is no surprise that the parser balks at the word TRANSLATES in this case. However, syntactically well-formed completions are available for all 3 of the other cases, including the crucial NN case. The finding of an elevated lexical decision time for the NN case is good evidence that we are tapping a processing phenomenon; the NN case is more difficult to process than the VN and NV cases, even though the syntactic structures that result from successful processing are equally acceptable.

Word Items With Adjective and Adverb Targets

Table 3 shows the results for the materials exemplified in Example 4.

Table 3. Experiment 2, Word Items, Adj/Adv Targets. Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word. Error percentages are in parentheses.

		Target: Adj	Target: Adverb	
Sent-Final Word	Noun	672 (8)	698 (13)	685
	Verb	648 (13)	676 (6)	662
		660	687	

The prediction was that the adjective-after-noun cell (version NA) would show an elevated RT relative to the other 3 cells. No hint of this effect is visible, and the 2 msec interaction effect in a 2 x 2 analysis of

variance is far from significance (Note 6). Neither main effect approaches significance. As with the verb and noun targets, the adjectives and adverbs did not differ in the lexical decision task when presented without sentence contexts. Versions VA and NR yielded more errors than the other two versions, but the effect is not significant.

The contrast in the intuitive acceptability of the adjective in the NA and VA versions seems as great as the contrast for verb and noun targets. Finding the predicted effect in the one case but not the other introduces a range of interesting possibilities. These are laid out in the next section, Hypotheses of the Parser/Paradigm Interaction.

Items With Nonword Targets

The left side of Table 4 shows the results for the items with "verb" and "noun" nonword targets. These materials can be pictured by replacing the target words TRANSLATES and TRANSLATION in Example 3 with nonwords such as FABLORATES and FABLORATION. The right side of Table 4 shows the results for the corresponding materials with "adjective" and "adverb" nonword targets. These can be pictured by replacing HAPPY and HAPPILY with SERICAL and SERICALLY in Example 4.

Table 4. Experiment 2, Nonword Items.
Reaction Time (Milliseconds) as a Function of
Syntactic Category of Target Nonword and Sent-Final Word.
Error percentages are in parentheses.

	Target "Verb"		Target "Noun"			Target "Adj"		Target "Adverb"		
Sent: Noun	724	(2)	756	(3)	740	744	(5)	757	(8)	750
Sent: Verb	702	(4)	740	(5)	721	712	(5)	762	(7)	737
	713		748			728		760		

In neither case does the statistical interaction in the 2 x 2 analysis of variance show any hint that a (conceptual) interaction between sentence context and nonword category information has occurred.

The main effect of sent type is far from significance, but the main effect of target type is significant for the "verb"/"noun" targets (min $F'(1,44) = 5.6$, $p < .05$), and approaches significance for the "adj"/"adv" targets (min $F'(1,48) = 3.7$, $p < .10$). In both sets of results the longer nonword target requires more time to reject than the shorter nonword target. Thus the rejection time for "nouns" is longer than that for "verbs", and the time to reject "adverbs" is longer than for "adjectives". The effect might well be due to the differences in morphology, rather than the length difference. If so, the finding may have a bearing on the role of word morphology in lexical access. It has no obvious bearing on parsing.

The failure to find an interaction between syntactic context and suffixes on nonwords indicating their part of speech does argue against one kind of explanation for the results obtained with verb and noun word targets in this experiment. If that result were due to an interaction of sentence context with the surface forms of category-determining suffixation (such as "ate", "ion", or "ify"), then one would expect to observe such an interaction with nonwords as well. For the syntactic force of such suffixes to be felt in this task, they apparently must occur on real words.

* * * HYPOTHESES OF THE PARSER/PARADIGM INTERACTION * * *

Lexical decision within sentence contexts can show purely syntactic effects of the context on the recognition times for noun and verb targets. This has been observed in two different experiments. In Experiment 1, the determining contexts ended in modal verbs and prepositions, while in Experiment 2 they ended in main verbs and singular subject noun phrases. A semantic or pragmatic account of the results is unlikely because the lexical decision targets were extremely unlikely continuations of the sentence contexts, and especially so because minimal pairs such as TRANSLATES/TRANSLATION were used in Experiment 2. The failure to find any effect of context on rejection times for nonwords with similar endings (such as FABLORATES/FABLORATION) argues that the relevant property of the target words is not simply the surface suffixes.

A purely syntactic effect of sentence context on lexical decision is a novel result in and of itself, and it establishes the necessary precondition for pursuit of specific questions about syntactic processes. The goal of this research is to determine whether the effects of syntactic context are a by-product of the operation of the human parser, and then to go on to use the paradigm to study the parser itself. Progress towards this goal requires a reexamination of the results that have accrued.

Experiments 1 and 2 were motivated largely by educated guesses about the operation of the human parser and its relation to the experimental paradigm. The range of explanations for parser/paradigm interactions was too large to be considered systematically, and the results of Experiments 1 and 2 have so far been discussed only in terms of the initial expectations which motivated them. Now, with the results of these experiments in hand, the space of hypotheses has been cut down to manageable size. In this

section I will spell out five hypotheses which seem plausible. Those which account for all the data will be tested in Experiments 3 through 6.

All the hypotheses concern the linguistic description of the differences between low RT and high RT items. The mechanisms that might lie behind these descriptions are discussed, but are not themselves part of the hypotheses.

The first three hypotheses all assume that a reasonably complete syntactic analysis of the sent is engaged within the experimental task. In contrast, the fourth hypothesis (Local) holds that the effects derive only from a local syntactic relationship between the sent-final word and the target. The fifth hypothesis (Adjective) is a specific and incomplete hypothesis, proposing that adjectives behave differently in this task from nouns and verbs.

Of the three sentence-level hypotheses, the first (Heads-of-Phrases) proposes a general description of the effect with respect to constituent structure, quite plausibly reflecting the parser and revealing something of interest as to its operation. The second (Bad Continuation) suggests that raised RTs result when the target word is a bad continuation of a sentence, while the third (Single Word Completion) suggests that lowered RTs result when the target word is a completion of a sentence. These latter two hypotheses are unlikely to reflect the operation of the parser in an interesting way.

A. The Heads-of-Phrases Hypothesis (HOPH)

The syntactic head of a phrase is roughly the single word of which all other words in the phrase are arguments or modifiers. The head of a noun phrase is a noun (or pronoun); the head of a verb phrase is the main verb. The head is the only constituent of a phrase which is obligatory in the

unmarked case (i.e. in the absence of specific information to the contrary). The syntactic features of the head determine the syntactic features of the entire phrase. For instance, the head noun of the singular noun phrase "a bag of tomatoes" is the singular "bag" rather than the plural "tomatoes". The concept of the head of a phrase also has more precise definitions, for instance within X-bar theory (Jackendoff, 1977).

HOPH states that RT will be lowered if the target can function as the head of a phrase which can be predicted with certainty by the grammar as of the last word of the sent.

For instance, refer back to Example 1, depicting the materials of Experiment 1. In the sent "If your bicycle is stolen, you must", the grammar requires that a verb phrase follow. A subject noun phrase (you) and a modal verb (must) have been received, and the grammar requires a verb phrase. This need not be the very next element (an adverb or prepositional phrase could intervene), but it must occur. HOPH states that since the verb target FORMULATE can be the head of this predictable verb phrase, its RT will be lowered. Similarly, after the sent "For now, the happy family lives with", the grammar requires a noun phrase, since "with" has introduced a prepositional phrase which needs a noun object. Since the noun BATTERIES can function as the head of this noun phrase, its RT is lowered.

The materials of Experiment 2 are represented in Example 5, with all four targets (verb, noun, adjective, and adverb) paired with each sent.

Example 5.
VV The towers on the horizon indicate that the barriers isolate TRANSLATES
VN The towers on the horizon indicate that the barriers isolate TRANSLATION
VA The towers on the horizon indicate that the barriers isolate HAPPY
VR The towers on the horizon indicate that the barriers isolate HAPPILY

NV The crowd near the church indicates that an important funeral TRANSLATES
NN The crowd near the church indicates that an important funeral TRANSLATION
NA The crowd near the church indicates that an important funeral HAPPY
NR The crowd near the church indicates that an important funeral HAPPILY

After "isolate" in the V sent, the grammar allows the prediction of a noun phrase, since "isolate" requires a noun object. HOPH predicts correctly that the RT to TRANSLATION will be lowered, since it can function as the head of the noun phrase. HAPPY can function as the head of an adjective phrase modifying the noun object of "isolate". But this adjective phrase is optional (not a certain prediction), and therefore HOPH predicts no effect upon the RT for HAPPY. This is a correct prediction.

After "funeral" in the N sent, a subject noun phrase has been received. The grammar allows prediction of a verb phrase. As expected, the verb TRANSLATES has a lowered RT. Neither the adjective HAPPY nor the adverb HAPPILY can function as the head of a predicted phrase. The status of the noun TRANSLATION is tricky. It can function as the head of the subject noun phrase itself, so one might expect it to show a lowered RT. But for TRANSLATION to function as a predicted head, the noun "funeral" must have been initially analyzed as a modifier. If a commitment is made by the parser at all, there is good reason (e.g. Minimal Attachment) to think "funeral" is analyzed as the head, not a modifier.

HOPH accounts for all the data in hand, and is tested in Experiments 3 through 6. The Heads-of-Phrases Hypothesis has important implications for the structure and operation of the human parser. These are discussed later.

B. The Bad Continuation Hypothesis

Perhaps slow RTs result whenever the target word precludes a grammatical completion of the sentence. This could well result from the operation of the parser, in the following way: As long as the unfolding sentence has one grammatical interpretation, the parser is a black box and reveals nothing of its inner workings in this task. But if the sentence has no grammatical interpretation, the parser fails. This failure is reflected in a high RT to the target word, which cannot be integrated into the sentence. This hypothesis has a high a priori plausibility because it utilizes a distinction which we know to be characteristic of the parser: some strings can be parsed and some cannot. Correspondingly, the truth of this hypothesis would be disappointing, because the experimental technique would merely be revealing what can be ascertained by consulting intuition.

This hypothesis accounts for the results of Experiment 1. The cases yielding slow RTs there consisted of a modal verb followed by a noun, and a preposition followed by a main verb. In both cases, the item cannot be grammatically completed. It also accounts for the case in Experiment 2 in which a main verb is followed by another main verb (inflected with an -s). However, it does not account for the second case in Experiment 2, where a noun is followed by another noun. This construction is perfectly grammatical, yet it yields an RT elevated to the same degree as the others.

Since this explanation is not consistent with the data, it will not be discussed further.

C. The Complete Sentence Hypothesis (CSH)

The Complete Sentence Hypothesis states that lower RTs result when the entire item constitutes a complete sentence of English.

It is reasonable to suggest that there are processing consequences related to the perception of a complete sentence, independent of any considerations specific to this experimental task. But the nature of the filler items in this task provides additional reason to suspect a key role for complete sentences. In all the filler items with word targets, the target word not only continues a sensible English sentence, but completes it as well. (The filler items were chosen to be complete sentences so they would maintain subject interest, and encourage the comprehension processes usually applied to sentences). Overall, half the time that the target is a word, the item is a complete sentence, while when the target is a nonword, the item is of course never a complete sentence. This imperfect correlation might (or might not) accentuate a natural tendency for "yes" responses to come more rapidly when the entire item is a complete English sentence.

Of course, the experimental items virtually never constitute perfectly acceptable and natural English sentences, but it turns out that those yielding low RTs usually constitute syntactically well-formed complete sentences. Referring to Example 1, the MN and PV versions are in no way complete sentences, while versions MV and PN are. Referring to Example 3, showing materials from Experiment 2 with verb and noun targets, the VV and NN versions are not complete sentences (notwithstanding that version NN can be completed by adding more words). Versions VN and NV are complete as they stand. None of the Experiment 2 materials with adjective and adverb targets (exemplified in Example 4) constitute complete sentences, and there are no RT differences.

One refinement of the hypothesis derives from an examination of the two complete versions with verb targets, exemplified by MV in Example 1 and NV in Example 3. Both types of item are complete at the category level, in that a verb completes the sentence. But for the many verbs used in these

experiments which require complements, these versions are incomplete at the level of strict subcategorization. As a result, CSH must pertain to completion at the basic category level, at least as regards the last (target) word of the item.

A mechanism which would plausibly drive CSH is a task-specific process that uses grammatical information to improve performance on the lexical decision. This task-specific process would utilize the parsing analysis of the sent, and predict that the lexical decision target will be the part of speech which is needed to complete the sentence. How often will such a process result in a correct prediction? Considering filler items with word targets, it turns out that in roughly half the cases the target is the predictable part of speech. In the other half of the cases there is either no word required to complete the sentence (the target is grammatically optional) or else the prediction is not unique and covers a variety of parts of speech. Depending on the experiment, the experimental items that would benefit from this strategy comprise between one fourth and one half of the total. Overall, then, roughly half of the word items would be helped by this strategy. The exact proportions do not matter, since we do not know the relative benefits of being right versus the costs of being wrong.

CSH is consistent with all the data of Experiments 1 and 2. It is tested further in Experiments 3 through 6.

D. The Local Hypothesis

The Local Hypothesis states that the syntactic effect of sentence context is very local, determined entirely by the last word of the sent. Perhaps this last word "primes" the recognition of words of grammatical categories that often follow it.

In Experiment 1, the combinations of sent-final words and targets that

yielded low RTs were modal verb - main verb and preposition - plural noun. These are good candidates for priming pairs. In Experiment 2, the pairings yielding low RTs were noun - verb and verb - noun. Why these pairs should produce priming is perhaps less obvious than for Experiment 1, but the idea would gain support from findings and speculations in psycholinguistics. For instance, in a parsing proposal presented by Bever (1970), the English parser is looking for noun-verb-noun sequences to identify as clauses. The alternation of nouns and verbs in this scheme might explain nouns engendering a verb prediction and verbs engendering a noun prediction.

The section of Experiment 2 with adjective and adverb targets did not yield any RT differences. In accord with this finding of no effect, no pairing of verbs or nouns with adjectives or adverbs would be expected to yield priming.

An even more primitive local effect might be at work in Experiment 2. The two cases yielding high RTs there involve a noun-noun sequence and a verb-verb sequence. Gleidman (1970) has suggested that the repetition of words of identical category might cause processing difficulty, which could in turn yield an elevated RT.

The Local Hypothesis is consistent with all the data of Experiments 1 and 2. It is tested further in Experiments 3a and 3b.

E. The Adjective Hypothesis

In the above hypotheses, considerable weight has been attached to the failure in Experiment 2 to find an effect of context on adjective targets. Along with the explanations considered above which attribute this failure to the grammatical relation of sent and target, one must consider the simple possibility that adjectives as a lexical category behave differently in this task from the noun and verb targets which have yielded expected effects.

Perhaps their recognition times cannot be influenced by sent contexts. This could just be an idiosyncratic fact about adjectives.

It is also true that adjectives typically serve a different role in sentences than nouns and verbs. Adjectives are usually less central to the grammatical structure, since they are usually modifiers while nouns and verbs are the heads of subject, object, and predicate phrases. There is little reason to think that this typical use pattern would translate into a different processing status for adjectives, but the possibility cannot be dismissed.

The Adjective Hypothesis is consistent with all the data of Experiments 1 and 2, and is tested within Experiment 4. Note that the Adjective Hypothesis says nothing about the effects obtained with noun and verb targets, so it is an incomplete explanation.

* * * TESTING THE HYPOTHESES: EXPERIMENTS 3 TO 6 * * *

The materials for Experiments 3 through 6 were mixed together and presented in one experimental session. 20 subjects in two subject groups participated. After examining these results, the materials of Experiments 3, 5, and 6 were modified. Together with the unchanged materials of Experiment 4, these were presented to 20 additional subjects, again in two subject groups. The first execution of Experiments 3 and 6 are referred to as 3a and 6a; the second as 3b and 6b. Three of the 40 subjects had participated in Experiment 2, over six months earlier.

EXPERIMENT 3a

One of the hypotheses which explains the results of Experiments 1 and 2 is the Local Hypothesis. According to this hypothesis, the effects we have observed are a local effect of the last word of the sent upon the lexical decision target. In Experiment 3 we pit this hypothesis against the three hypotheses which assume an integration of the sent as a linguistic unit to be one component of the observed effects.

Sample materials for Experiment 3a are shown in Example 6.

Example 6.

NV *The students decided that they would give their teacher ROTATES
NN The students decided that they would give their teacher ROTATION

The local relationship between the sent-final word and the target is the same as that in part of Experiment 2; a noun is followed by either a verb (ROTATES) or another noun (ROTATION). In Experiment 2, the predicted (and obtained) effect was a lower RT for the verb. If the Local Hypothesis is correct, we should observe the same effect here.

However, if an integration over the entire sent is involved, we predict a lower RT for the noun. The item ending in ROTATES cannot be grammatically completed, while the item ending in ROTATION is a syntactically complete sentence in itself, with ROTATION functioning as the direct object of the verb "give", and "their teacher" as the indirect object. All 12 of the sent contexts end with a verb such as "give" which can take indirect objects, followed by a noun phrase such as "their teacher" which is most naturally construed as an indirect object.

A lower RT for version NN will strongly suggest that the Local Hypothesis is false. A lower RT for version NV will strongly suggest that local effects are responsible. A finding of no difference could have many interpretations. One is that both local effects and global syntactic

effects are at work, and they are canceling each other. A second is that the Local Hypothesis is false, but the sentence contexts are insufficient for the parser to predict a noun phrase. Consider that the prediction of a lowered RT for version NN requires: 1) the identification of the verb "give" as one that can take an indirect object, 2) an understanding that the noun phrase "their teacher" is unlikely in this context to be the direct object (i.e. what was given), and 3) the convergence of this syntactic, semantic, and pragmatic information onto a prediction of a noun object. If not all of these processes occur, or even if they do occur but have not taken place in the 400 msec following the presentation of "teacher", there is no reason to expect any RT difference.

METHOD

Since Experiments 3 through 6 were mixed together in the same session, several of the following sections apply to the other experiments as well: Subjects, Filler Items and Randomization, Procedure, and Overall Results.

Subjects

MIT undergraduate volunteers were paid for their participation. All were native speakers of American English.

Materials

The entire set of Experiment 3a materials appears in Appendix C. Twelve item pairs such as shown in Example 6 were the materials for this experiment. All sents end with a verb (such as "give") which can support the indirect object construction, followed by a singular noun phrase two or three words long, always ending in a singular noun. The sents were

constructed to induce the indirect object reading of the sent-final noun phrase. The target words were the very same verb/noun pairs that were used in Experiment 2, consisting of a verb and a noun related by derivational morphology. As always, targets were unexpected continuations of their sents. Each subject saw only one version of an item (either NV or NN). Morphologically similar nonword targets (not used in any previous experiment) were assigned to similar sents to form 12 balancing nonword items.

Filler Items and Randomization

Experiments 3 through 6 contributed a total of 144 items to be seen by each subject group. Half of those contributed by each experiment had nonword targets, so of course half of the total had nonword targets. In addition, 144 filler items were constructed and shown to both subject groups. Of these, 72 had nonword targets. The other 72 had word targets which completed a natural and sensible English sentence. As before, from the subject's point of view half of the items have nonword targets. Of the word items, half are complete and natural English sentences, while the other half are not.

The items were presented in different pseudo-random orders to different subjects. The entire randomized set of 288 items was divided into three equal blocks, each preceded by 5 warm-up items. There was also an initial block of 30 representative practice items.

Procedure

The procedure was the same as for Experiment 2, which differs from Experiment 1 only in the presentation rate. Experiments 3 to 6 thus had the Experiment 2 presentation rate of 300 msec per word, rather than the faster

200 msec rate of Experiment 1. The experiment took approximately 40 minutes.

OVERALL RESULTS

Twenty subjects were run, 10 in each condition. Three subjects were replaced because of obvious inattention to the task; 3 were replaced due to an error rate over 10%; 1 was replaced due to an RT standard deviation over 250 msec. None exceeded the mean RT cutoff of 850 msec. The overall error rate (including filler items) was 3.3%. RTs more than two standard deviations above the subject's mean were replaced with that subject's cutoff value.

RESULTS AND DISCUSSION

The top of Table 5 shows the results for Experiment 3a; Experiment 3b is described in the next section.

Table 5. Experiments 3a (top row) and 3b (bottom row). Reaction Time (Milliseconds) as a Function of Syntactic Category of Target. Error percentages are in parentheses.

	Target	
	Verb	Noun
Experiment 3a	709 (18)	732 (12)
Experiment 3b	728 (7)	665 (3)

There is a 23 msec difference favoring the Local Hypothesis, but this is far from significance ($p > .3$ for both subject and item analysis). The more striking aspect of these results is the error rate, averaging 14.6%. This is far above the 3.3% error rate for these subjects over all the items they

saw. It is also higher than the 9.1% error rate shown by these very same target words in Experiment 2. The difference in error rate between the NN and NV versions is not significant.

One of the predicted explanations for such a finding of no effect was insufficient time to process the sent and predict a noun phrase. The high error rate is likely to be an indication of difficulty in processing the sent, which is interfering with the accuracy of the lexical decision task. On this assumption, the sent contexts in Experiment 3b were modified to provide a slightly different basis for the prediction of a noun phrase, and to be easier to process.

EXPERIMENT 3b

Sample materials for Experiment 3b are shown in Example 7.

Example 7.

NV *The students gave to their teacher ROTATES
NN The students gave to their teacher ROTATION

As in Experiment 3a, the sent ends in a noun, so that the Local Hypothesis predicts lowered RT for the verb target. In contrast to Experiment 3a, the sent ends in an object-taking verb (not necessarily one which can take an indirect object) followed by a prepositional phrase. The only steps that must be performed to obtain a noun prediction are 1) the recognition of "give" as an object-taking verb, and 2) the recognition that the prepositional phrase "to their teacher" cannot function as that object. No semantic or pragmatic computations are required. The Experiment 3b sents were also shorter than those used in Experiment 3a, in an attempt to keep processing load as low as possible and maximize the chance of a noun phrase prediction.

METHOD

The Subjects, Filler Items and Randomization, and Procedure sections are the same as for Experiment 3a. Experiment 3b was part of a separate experimental session from 3a, but these sections of 3b were unchanged from 3a. None of the subjects in Experiment 3b had participated in Experiment 3a.

Materials

The entire set of Experiment 3b materials appears in Appendix C. 12 item pairs such as shown in Example 7 were the materials for this experiment. All sents end with a verb (such as "obtain") which requires a

noun object, followed by a prepositional phrase that is 3 or 4 words long, ending in a singular noun. The target words were the very same verb/noun pairs that were used in Experiment 3a (and also Experiment 2). In 9 of 12 cases, the sent-final word was also identical to that of Experiment 3a, so that the pairing of sent-final word and target was unchanged. The prepositional phrase was always introduced by "to" or "for", and always denoted the benefactor or recipient. Thus the case relations between verb and following material were unchanged from 3a to 3b. The 3b sents were on the whole shorter than the 3a sents.

OVERALL RESULTS

Twenty subjects were run, 10 in each condition. Six were replaced due to a mean RT over 850 msec, or an RT standard deviation over 250 msec; none exceeded the 10% error criterion. The overall error rate (including filler items) was 2.7%. RTs more than two standard deviations above the subject's mean were replaced with that subject's cutoff value.

RESULTS AND DISCUSSION

The bottom of Table 5 shows the results for Experiment 3b. The noun targets were recognized 63 msec faster than the verb targets, and the difference is significant (min $F'(1,30) = 4.2$, $p < .05$). The Experiment 3b error rate (5.0%) is dramatically lower than the 3a error rate (14.6%), and the difference is significant (min $F'(1,25) = 4.8$, $p < .05$).

The finding of faster RTs to the noun targets in this experiment is inconsistent with the Local Hypothesis. The local relations between the sent-final word and the target are unchanged from Experiment 2, yet the RT effect is reversed. We can safely conclude that the influence on RT derives from an integration of the entire sent.

The contrast in the Experiment 3a and 3b findings is striking, given the small differences in the materials. Experiment 3b is characterized by a low error rate and the reliable finding of an expected RT difference, while 3a is characterized by a high error rate and no reliable RT difference. This strongly suggests some processing disruption in the Experiment 3a materials. The paradigm is not designed to measure processing load or disruption, but this of course says nothing against the examination of such data when it appears. The following consideration of the Experiment 3 results is a digression from the plan of Experiments 3 through 6, but it is directly relevant to parsing theories.

There are two systematic differences between the sents of Experiment 3a and 3b. One is their length; the sents in Experiment 3a are longer than in 3b. This is an obvious candidate for the source of the processing disruption. But other experiments do not show comparable length-error relationships, so length is unlikely to be the crucial variable.

The other systematic difference between the sents of Experiment 3a and 3b is whether the constituent after the verb is a noun phrase (3a) or a prepositional phrase (3b). In 3a, the noun phrase (NP) is best interpreted as the indirect object of the verb, while in 3b the prepositional phrase serves the same semantic function (benefactor or recipient). In both cases, the sent can be completed to form a full and natural sentence of English by the addition of an object NP. The naturalness of the resulting sentences is comparable between Experiment 3a and 3b; some people find the sentences required by the 3b sents to be less natural because of the prepositional phrase intervening between verb and object. Yet it is the Experiment 3a sents which produce the high error rates and show no RT effect. One is led to look for the difference in the details of parsing, rather than some overall evaluation of acceptability.

In the Experiment 3b sents, the verb must take a noun object, but is followed by a prepositional phrase. In Experiment 3a sents, the verb must take one noun object (its direct object) and may also take another noun object (its indirect object) which precedes the required object. Having received the verb and the NP which follows it, the parser faces a local ambiguity. The NP it has received could be the direct object, in which case no further NP is necessary or even acceptable. Or the NP it has received could be the indirect object, in which case a further NP (the direct object) is necessary. The ultimate resolution of ambiguities such as this could well be due to a consideration of the meaning of the two possible readings, for instance how likely the entity denoted by the NP is to be the direct as opposed to the indirect object of the verb. Assume for the moment that this information is not available within the time frame of this task. In this case, the parser has essentially three options: to reserve judgment until more words have been seen, to decide the NP is the direct object, or to decide the NP is the indirect object.

If the parser decides the NP is the indirect object, one would expect it to predict (Note 7) the direct object NP, and in this case we should observe in Experiment 3a the effect which we observe in 3b. We don't.

If the parser reserves judgment, then it would not predict an NP, and the failure to find an RT difference is explained. However, the high error rate is still in need of explanation. One could argue that the need of the parser to reserve judgment and hold unassembled fragments of the sentence in memory results in a high processing load. This in turn interferes with the accurate execution of the lexical decision task. This possibility accounts for the data well, but I find the remaining possibility even more compelling.

If the parser decides the NP is the direct object, one would expect no

prediction of another NP to follow. This accords with our finding of no RT effect. To account for the high error rate, one must reintroduce the consideration of meaning relations. After the parser decides that the NP it has received is the direct object, semantic and pragmatic components determine the meaning of the sent. In all 12 of the Experiment 3a sents, the direct object interpretation of the NP is pragmatically inappropriate. This causes a processing disruption which interferes with the accuracy of the lexical decision task, thus accounting for the high error rate. The pragmatic inappropriateness also sets in motion a syntactic reanalysis of the role of the sent-final NP. This is then an example of a minor "garden path" situation: the parser is led down the garden path by its incorrect treatment of the NP, and must retract its assignment of grammatical role to the NP in order to recover. Some writers on parsing have suggested that garden paths exact a heavy processing "cost" (e.g. Marcus, 1980; Frazier and Fodor, 1978). It is quite reasonable to suppose that this "cost" would be such as to interfere with the lexical decision task.

A parsing proposal by Ford, Bresnan, and Kaplan (in press) would maintain that the action of the parser at the end of an Experiment 3a sent depends crucially on the particular verb in the sent. On this view, the 3a verbs have different "lexical forms" for the direct object and indirect object complement structures. According to this proposal, whichever use of the verb has greater salience (which depends on such variables as frequency of use) will determine which lexical form is entertained first. With a larger sample of systematically chosen items and frequency counts of the lexical forms of the various verbs, one could test the Ford, Bresnan, and Kaplan proposal by a variant of Experiment 3a.

In the last several paragraphs we have been assuming that semantic and pragmatic information is available to the parser only after a substantial

delay. In fact, the results of Experiments 3a and 3b argue against a very rapid assimilation of pragmatic and semantic information into the parsing process. If the pragmatic information were available when the parser had to choose a role for the sent-final NP, there is no reason to think the parser would either be led astray or need to postpone the decision. The pragmatic information would allow selection of the indirect object role for the NP. Another NP would be predicted by the parser, and we would observe a low error rate and an RT difference, as in Experiment 3b.

EXPERIMENT 4

The Adjective Hypothesis holds that the recognition time for adjective targets cannot be influenced by the operation of the parser. The Heads-of-Phrases Hypothesis holds that no effects were found with adjective targets in Experiment 2 because the adjective could not function as the head of a phrase predictable as of the last word of the sent. The Complete Sentence Hypothesis holds that no effects were found because the adjective never completed a sentence. Experiment 4 opposes the Adjective Hypothesis to HOPH and CSH by providing an environment in which an adjective target will function as the head of a predictable adjective phrase, and will also complete the sentence. Example 8 shows sample materials used in Experiment 4.

Example 8.

RA	The interesting clock seems very TOLERABLE
R-	The interesting clock seems very RALORALET
VA	Your visiting friend should enjoy TOLERABLE
V-	Your visiting friend should enjoy RALORALET

The R sent (R stands for adverb) ends with a verb which frequently takes adjective complements followed by a degree adverb. Though degree adverbs can introduce adverb as well as adjective phrases, it seems reasonable to assume that an adjective phrase can be predicted by the end of the sent. The V sent ends after an object-taking main verb. As in Experiment 2, the grammar allows prediction of a noun phrase, of which the adjective TOLERABLE can be a legal part but not the head. In version RA TOLERABLE is the head of the predictable adjective phrase. Versions R- and V- consist of the same sents matched with a nonword target. These will provide a baseline measure of the processing difficulty of the two sents.

If version RA yields faster RTs than version VA, the Adjective Hypothesis can be rejected. A finding of no difference (though inherently

more difficult to interpret confidently) would support the Adjective Hypothesis.

METHOD

For the Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 3a.

Materials

The entire set of Experiment 4 materials is shown in Appendix D. Sixteen item quadruples like that shown in Example 8 served as the materials. The first sent in each sent pair ends with a verb which takes adjective complements freely, followed by a degree adverb which usually occurs with an adjective phrase. The verbs used were: be, remain, appear, seem, and become. The adverbs were: completely, absolutely, barely, highly, extremely, utterly, nearly, and very. The second sent in each pair ends in a verb which usually takes a noun phrase object and never takes adjective complements. The sents were matched for length.

The target words were category-unambiguous adjectives. Each adjective was paired with a nonword target of the same length.

RESULTS AND DISCUSSION

The effect we wish to measure for this experiment is the difference of the RA and VA versions, each first corrected for an effect of the sent itself by subtracting its nonword control. As a formula, this "effect score" is (RA - R-) - (VA - V-), which is equal to the interaction effect in a 2-way analysis of variance, with sent type as one factor and target type (word or nonword) as the other. After the first execution of the experiment (with Experiment 3a) the interaction effect was 33 msec in the right

direction, and it was near significance for both the subject analysis ($F(1,19) = 4.2$, $p = .054$) and the item analysis ($F(1,15) = 3.1$, $p = .098$). When Experiments 3 through 6 were run again, Experiment 4 materials were left unchanged, with the expectation that more subjects would bring the effect within the bounds of significance. The combined results of the two executions of the experiment are shown in Table 6.

Table 6. Experiment 4.
Reaction Time (Milliseconds) as a Function of Target (word or nonword) and Sent-Final Word. Error percentages are in parentheses.

		Target: Adjective		Target: Nonword		
Sent-Final Word	Adverb	618	(3)	668	(2)	643
	Verb	659	(5)	675	(2)	668
		639		672		

The interaction effect remained virtually unchanged at 34 msec, and was now significant for the subject analysis ($F(1,38) = 6.4$, $p = .016$) but little better than before for the item analysis ($F(1,15) = 3.3$, $p = .089$). An examination of the "effect score" for each item revealed a single item with a score of +196, 2.6 standard deviations from the mean. Such an item can be legitimately excluded from the analysis on the grounds that it is grossly unrepresentative of the pool of items. Just to be somewhat conservative, the item with the lowest effect score (-155) was eliminated along with this offending item. With the two extreme items eliminated, the effect becomes significant (min $F'(1,33) = 4.7$, $p < .05$). The adjective target yields a lower RT when it is the head of a predicted adjective phrase.

The main effects in this analysis of variance are of no particular theoretical interest, but both are near significance. The effect of target

type (word vs nonword) is significant for subjects ($F(1,38) = 8.8$, $p = .0052$), but not for items ($F(1,13) = 2.2$, $p = .16$). The effect of sent type is significant for both subjects ($F(1,38) = 15.6$, $p = .0003$) and items ($F(1,13) = 5.6$, $p = .034$), but not quite for min F' (min $F'(1,24) = 4.1$, $p < .10$).

The faster RT to the adjective target when it is the head of a predicted phrase lays to rest the Adjective Hypothesis. With the elimination of the Local Hypothesis in Experiment 3, the only remaining viable hypotheses are the Complete Sentence Hypothesis and the Heads-of-Phrases Hypothesis. Before turning to Experiment 6, where these two remaining hypotheses are pitted against each other, we consider Experiment 5, which was designed to address a different question. The results have some bearing on the interpretation of Experiment 6.

EXPERIMENT 5

The implicit assumption in all the hypotheses laid out in the Hypotheses section is that the influence of sentence context on lexical decision RT is a binary effect, having only two levels, such as "on" and "off". The hypotheses differ in the conditions under which the effect is "on" or "off". There is no a priori reason why the effect should operate in this way. My own intuition before collecting results with this paradigm was that predictable parts of speech (heads of phrases) would produce fast RTs, parts of speech in unpredictable but common constructions would produce slower RTs, uncommon constructions would produce yet slower RTs, and illegal constructions would produce the slowest RTs of all. All RT differences for previous experiments have separated predictable heads of phrases from other cases, and there has been no need to invoke anything other than a binary effect to explain the results. The one case in which we contrasted two cases, neither of which was a predictable head of a phrase, was in Experiment 2 with adjective targets. The adjective after the verb context (unpredictable but common) did not differ from the adjective after the noun context (unpredictable and very uncommon). This is the single piece of evidence that effects with the paradigm are only binary. It deserves replication.

A sample quadruple of Experiment 5 materials appears in Example 9.

Example 9.

DA New fertilizers allow farmers to cultivate their CRUCIAL
DV *New fertilizers allow farmers to cultivate their EXCLUDE
NA ++Farm experts believe that the hungry beetles CRUCIAL
NV Farm experts believe that the hungry beetles EXCLUDE

The first sent ends in a noun phrase determiner, while the second ends with a plural noun phrase in subject position. The target words are adjectives and verbs, both uninflected. The DA-NA comparison is a replication of the

adjective result of Experiment 2, though the sent contexts are different. The environment there in which an adjective is likely ended with a verb; here it ends with a determiner. The environment there in which an adjective is extremely unlikely was after a singular subject NP, whereas here it is after a plural subject NP. The versions with verb targets (DV and NV) were included to ensure that these contexts can produce the by-now-familiar RT difference between a predicted head of a phrase and an illegal continuation; the verb target is a predicted head in version NV and illegal in version DV.

Assuming that the NV and DV versions produce an RT difference, we can see whether the adjective result of Experiment 2 is replicated. If versions DA and NA do not differ, the result is replicated. If version DA yields faster RTs than version NA, we will have evidence for more than a binary effect of sentence context deriving from predictable heads of phrases.

METHOD

For Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 3a.

Materials

The entire set of Experiment 5 materials with word targets is shown in Appendix E. Twelve item quadruples like that shown in Example 9 were constructed. The first sent of each pair ends with an object-taking verb, followed immediately by a noun phrase determiner. The determiners used were articles (a and the), possessive pronouns (our, your, their, and my), and one quantifier (every). These were selected because they cannot by themselves constitute a noun phrase. The second sent of each pair ends with a plural noun which is the final noun of a plural NP in subject position. The targets are adjectives and verbs, unambiguous as to their category, and

matched for length, frequency, and number of syllables. Twelve additional sent pairs of the same kind were assigned nonword targets to serve as balancing nonword items.

The above description applies to Experiment 5 as it was executed the first time, with Experiment 3a. In its second execution (with Experiment 3b) the sents that had taken word targets were matched with the nonword targets, and the sents that had taken nonword targets before were matched with the word targets. When the two executions of Experiment 5 are considered together, there are 24 different item quadruples, though the 12 target word pairs each appear in two item quadruples. Each item was seen by 10 subjects. Each sent in the 24 quadruples also appeared with two different nonwords, allowing a baseline estimate of the processing difficulty of the sents.

RESULTS AND DISCUSSION

Following the first execution of Experiment 5, there were not even any suggestive RT differences among the 4 versions (the lowest and highest RT cell means differed by 9 msec). For the second execution, the particular items were changed (by switching word and nonword targets) but the kinds of items were unchanged. The results were little more impressive. Table 7 shows the combined results of both executions of Experiment 5.

Table 7. Experiment 5, Word Items.
Reaction Time (Milliseconds) as a Function of
Syntactic Category of Target and Sent-Final Word.
Error percentages are in parentheses.

		Target: Adjective		Target: Verb		
Sent-Final Word	Determiner	643	(4)	671	(6)	657
	Noun	651	(3)	656	(2)	653
		647		663		

A direct comparison of the DV and NV cells shows a difference of 15 msec in the expected direction, but far from significance. The 8 msec difference between the DA and NA cells is likewise far from significance. Consideration of the nonword baselines does not change the picture (nonwords were rejected 10 msec faster after the determiner than the noun phrase sents); they will not be considered further. As in Experiment 2, there is no difference between the unpredictable but common continuation (DA) and the extremely uncommon continuation (NA). But interpretation of this result is seriously undermined by the failure to find an effect with verb targets (DV vs NV). The verb target is grammatically illegal after the determiner, but after the plural subject NP, it is a predicted head of a phrase and completion of the sentence. All of the hypotheses considered in the Hypotheses section would predict a faster RT for version NV than DV.

In searching for an explanation, one's attention is drawn to the aspects of the sent contexts which are new to this experiment. The noun phrase sents end in plural nouns; noun phrase sents in previous experiments have ended in singular nouns. One suspicious feature of plural nouns is that they sound the same as possessive nouns. Thus if an auditory recoding is taking place (or if apostrophes are not used as a syntactic cue), plural nouns might be confused with possessive nouns. They might be treated as

ambiguous, resulting in no clear syntactic predictions. In particular, perhaps no prediction of the verb phrase emerges, so that the verb target is not in fact the head of a predicted phrase in version NV. This is possible, but it is overshadowed by the behavior of determiners, the new and different aspect of the other sent type.

Recall that three kinds of determiners were used, articles (a, the), possessive pronouns (my, your, their, our) and one quantifier (every). The quantifier "every" occurred too seldom (in only 2 of the 24 determiner sents) to allow any confident analysis. However, a post hoc analysis of the 13 item quadruples with articles as determiners and the 9 with possessive pronouns was illuminating. Table 8 shows the mean results when these two kinds of items are separated.

Table 8. Experiment 5, Post Hoc Analysis. Items are divided according to the kind of determiner in the determiner sent.
Mean Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word.

	Article Items		Possessive Pronoun Items	
	Target		Target	
	Adjective	Verb	Adjective	Verb
Determiner Sent:	664	708	639	631
Modal	682	667	619	659

For the article items, there is now a 41 msec tendency in the expected direction for verb targets (NV-DV), which approaches significance ($F(1,12) = 4.2$, $p = .063$) (Note 8). For adjective targets, there is an 18 msec tendency in the direction we would predict, though far from significance. For the possessive pronoun items, on the other hand, these effects are in the opposite direction from what we would predict! The significance of this overall different pattern of results for article and

possessive pronoun determiners is revealed by a significant 3-way interaction ($F(1,20) = 9.9$, $p = .0051$) in an analysis of variance with target type, sent type (determiner or noun phrase) and determiner type (article or possessive pronoun) as the factors (Note 9).

What properties of articles and possessive pronouns might lead to their creating such different effects in this experiment? Syntactically, they are specifiers with virtually identical properties. Their primary difference is semantic. In contrast to articles, possessive pronouns indicate possession, and have an anaphoric relation to some previous noun phrase. The articles are also much more frequent in the language. I have no suggestions for how these properties would lead to the differences shown in this experiment.

The post hoc analysis furnishes a more hopeful outcome than a simple finding of "no effect," which would be contrary to all hypotheses put forth to explain other results with the paradigm. The verb effect we expect is obtained with the articles "a" and "the". It is not obtained with possessive pronouns. We do not know why, but we do know that it has something to do with possessive pronouns. The experimental paradigm is yielding clear results which distinguish linguistic types, even if in some cases we do not have a theory which explains such differences.

This discussion can be summarized briefly. In Experiment 5, the predicted verb effect was not found, and therefore the failure to find an adjective effect carries no weight. Suspicion is cast on the ability of plural noun phrases and determiners to create the syntactic environments we expect. More speculatively, it appears that articles create the syntactic environments we expect, while possessive pronouns do not.

EXPERIMENT 6a

The Heads-of-Phrases Hypothesis (HOPH) and the Complete Sentence Hypothesis (CSH) are the only remaining hypotheses which account for all the data in hand. Their predictions have converged for every previous experiment. Their predictions diverge where the grammar allows the prediction of the head of a phrase, but where this phrase cannot complete the sentence.

A sample quadruple of Experiment 6a materials is shown in Example 10.

Example 10.

DN	Besides the pieces of fruit which the ENTRY
D-	Besides the pieces of fruit which the ENDOT
MN	*Besides the straw which the horse could ENTRY
M-	Besides the straw which the horse could ENDOT

Both sents are marked as subordinate constructions (by "besides", in this case), and both end within an incomplete relative clause. The first sent of each pair ends after a single specifier of the subject noun phrase. The second sent ends after a modal verb within the relative clause. The target word is a noun. As in Experiment 4, the nonword targets paired with the same sents serve as a baseline measure of the processing difficulty of the sent.

In the determiner sent, the grammar allows the prediction of a noun phrase, of which the noun target ENTRY can be the head. In the modal sent, the noun ENTRY is an illegal continuation. HOPH predicts a lower RT for the noun in DN than MN. CSH predicts no RT difference, because in neither case does the item constitute a complete sentence.

Both kinds of sent leave one not only within an introductory clause or prepositional phrase, but also within an object relative clause. This serves to contrast HOPH not only with CSH itself, but also with more inclusive relatives of CSH. For instance, a variant of CSH might hold that

a lowered RT would result whenever the entire item is a complete clause.

METHOD

For Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 3a.

Materials

The entire set of Experiment 6a materials is shown in Appendix F. Twenty item quadruples such as that shown in Example 10 were used. Both sents of each pair begin with 1) an introductory preposition or conjunction assuring that what follows is not a main clause, 2) the beginning of a noun phrase, after which there is 3) a relative pronoun introducing a relative clause of the noun phrase. From this point the sents continue differently. In the determiner sent, all that follows the relative pronoun is a determiner (a, the, every, my, or our). In the modal sent, the relative pronoun is followed by the relative's subject NP and a modal verb. The sents are matched for length in words, and approximately for typed length. The target words are singular nouns. Twelve of the 20 target nouns are those used in Experiments 1a and 1b. Each target word is matched with a nonword of the same length.

RESULTS AND DISCUSSION

The results of Experiment 6a are shown in the left half of Table 9.

Table 9. Experiments 6a and 6b.
Mean Reaction Time (Milliseconds) as a Function of Target Type (Word or Nonword) and Sent-Final Word. Error percentages are in parentheses.

	Experiment 6a					Experiment 6b				
	Target					Target				
	Noun		Nonword			Noun		Nonword		
Det.	626	(6)	651	(4)	639	636	(5)	651	(4)	643
Sent:										
Modal	632	(5)	662	(1)	647	646	(5)	642	(2)	649
	629		657			641		646		

The effect to be measured for this experiment is like that in Experiment 4: the difference of the DN and MN versions, each first corrected for effects of the sent itself by subtracting its nonword control. The formula for the "effect score" is (DN - D-) - (MN - M-). And as in Experiment 4, this score is equal to the interaction effect in a 2-way analysis of variance, with sent type as one factor and target type (word or nonword) as the other. The interaction effect for Experiment 6a is 5 msec (in the wrong direction) and far from significance. The main effect of sent type is far from significance, but that of target type (word vs nonword) is significant for the subject analysis ($F(1,19) = 5.7$, $p = .027$), but not for the item analysis ($F(1,19) = 2.3$, $p = .14$).

The finding of no effect in this experiment favors CSH over HOPH.

EXPERIMENT 6b

One possibility for the finding of no effect in Experiment 6a was that there were simply too many layers of embedding, creating too great a processing load for clear syntactic predictions to emerge. Perhaps the parser falls behind in its analysis of the sent. One simple step toward remedying this possible problem is to remove the introductory preposition or conjunction, thus removing one layer of embedding. This was the strategy in Experiment 6b, the second execution of Experiment 6, run with Experiment 3b.

A sample quadruple of 6b materials appears in Example 11.

Example 11.

DN The pieces of fruit which the ENTRY
D- The pieces of fruit which the ENDOT
MN *The straw which the horse could ENTRY
M- The straw which the horse could ENDOT

METHOD

For Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 3b.

Materials

The complete set of Experiment 6b materials is shown in Appendix F. The materials were those of Experiment 6a, modified slightly. The introductory preposition or conjunction was removed. In four of the sents minor word substitutions were made, to restore naturalness after elimination of the initial word.

RESULTS AND DISCUSSION

The results of Experiment 6b are shown in the right half of Table 9. The 19 msec interaction effect is in the direction predicted by HOPH but far

from significance. Neither main effect approaches significance.

The failure to find a significant effect once again favors CSH over HOPH. The 19 msec tendency in the direction predicted by HOPH must dampen more than usual our natural tendency to be wary of arguing from a finding of no difference. There are also more specific reasons to be suspicious of the results of Experiments 6a and 6b.

The elimination of the introductory preposition or conjunction that distinguishes Experiment 6b from 6a was intended to lessen the processing difficulty of the sents. With this change, the interaction effect moved 24 msec in the direction predicted by HOPH. There is reason to believe that the processing load is still formidable, particularly in the determiner sent. According to HOPH, it is this sent (of the two) which must yield the prediction of an NP of which the noun target serves as head. But the last two words of the sent indicate not only the beginning of a relative clause, but the beginning of an object relative. Object relatives are notoriously hard to process. If the parser is struggling with this construction, the prediction of a noun phrase may not emerge in time to influence the lexical decision task.

The other reason for caution in interpreting Experiment 6 results is that it relies on determiners to create the syntactic environment for a noun prediction. Determiners were implicated in the failure of Experiment 5. They just might be responsible for the failure of Experiment 6. A post hoc analysis on Experiment 6 parallel to the Experiment 5 analysis did not show different effects for articles and possessive pronouns. But the lexical decision targets in Experiment 5 were adjectives and verbs, whereas here they are nouns.

On balance, the results of Experiment 6 are inconclusive. CSH and HOPH remain as contending explanations of results with this paradigm, but they are the sole contenders.

* * * A MODIFIED PARADIGM * * *

PRELIMINARIES

The Task-Specific Incentives for Completion Strategies

In the Hypotheses section where the Heads-of-Phrases Hypothesis (HOPH) and the Complete Sentence Hypothesis (CSH) were first discussed, the supposition was that HOPH, if true, derived from the human parser itself. CSH, on the other hand, was presumed to derive from a task-specific process, not the parser. These assumptions may not be true. One can imagine the parser being constructed so as to predict words that complete entire sentences (a parser-driven CSH). But this seems unlikely because it is too specific a situation to be very important to the parser's operation. One can also imagine a task-specific process that lowers the RT to the head of any predictable phrase (a task-specific HOPH). This could conceivably be independent of the parser, but more plausibly would represent a task-specific inference extracted from the parser by some "general problem solver". For instance, consider the situation after the sent "For now, the happy family lives with". The representation of this sent constructed by the parser might allow a task-specific process to predict that a noun phrase will follow, and consequently to predict a noun for the lexical decision target. The parser itself, in its own task of determining the structure of sentences, might not make any such predictions. Or if the parser is taken to be a more flexible device, it may make such predictions in this task that it would not make outside of the experimental situation.

Most experimental tasks in cognitive psychology are open to the possible criticism that results may be task-specific. In this case there is a particular reason to worry about task-specificity. The target word is

always the last word of the item that the subject sees. This point was made in our original discussion of CSH. To apply this observation to HOPH, we note that the relevant unit of completion might not be the sentence, but the syntactic phrase.

In this light, the more interesting question is not whether HOPH is true, but whether the experimental results are due to the operation of the parser or to some task-specific process. In the next series of experiments, we change the paradigm slightly, by providing a continuous presentation of a sentence across the computer screen, the appearance of words uninterrupted by the lexical decision task. In so doing, we eliminate the elements of the task that might be encouraging an unnatural attention to the completion of linguistic units. If we can replicate our previous findings with this modified paradigm, it will undermine the view that the effects we have observed are due to a task-specific process. With its most plausible mechanism discredited, the position of CSH in turn will be weakened.

The Modified Paradigm

Many aspects of the experimental technique are unchanged in the new series of experiments. Sents are still presented from left to right across a computer CRT, one word every 300 msec, with words remaining on the screen after they appear. The target is still set off from the sent by its appearance in capital letters. There are minor changes concerning the presentation of the target. Previously the target was announced by a tone 300 msec before its appearance, but now the tone sounds concurrently with the target. The slightly longer 400 msec delay before the appearance of the target word has been eliminated in favor of a uniform 300 msec for all words.

The primary modification is that the target word is not the end of the

item; further words continue to appear while the subject is performing the lexical decision task. These words appear in lower case type at the same 300 msec rate as the sent and target, unaffected by the subject's response. These following words almost always constitute the final portion of a sentence of English; accordingly, the post-target words in an item will be called the ence. Every item the subject sees has an ence which is at least two words long.

In previous experiments it was possible to design the stimulus materials so that each item would fit on a single 78-character line. This is not always possible with the modified paradigm. Since it is important to use the sents and targets of previous experiments, while adding ences to each, the 78-character limit is often exceeded. When necessary, part or all of the ence appears after a double-spaced carriage return, i.e. two lines down and starting at the left margin. The division of sentences in this way is of course typical of everyday reading.

Whether or not the ence is carried onto the next line, the end of the ence is marked by a slash " / " which signals to the subject that the next item may be obtained by a press of the foot pedal.

The reason for presenting ences is to encourage subjects to view their decision as occurring internal to a sentence, thus discouraging any task-specific concern with the ends of phrases or sentences. However, this change also introduces the possibility of an effect of the ence upon the lexical decision task. This possibility is explored in a preliminary way within Experiment 7. However, it is a possibility which must be considered in all experiments, to avoid contaminating the effects we are primarily interested in. Lexical decision times in the previous experiments have generally been between 400 and 1000 msec, with occasional RTs which are much higher. The RT is of course measured from the onset of the target. In this

modified paradigm, the first word of the ence appears after 300 msec, the second after 600 msec, the third after 900 msec, and so forth. It seems unlikely that words of the ence beyond the second could affect RT with any consistency. The third appears not only 900 msec after the target, but at some considerable distance from the target (either to the right, or to the left on the next line).

We thus assume that the first two words of the ence might influence RT, while the remainder cannot. The design of item quadruples is influenced by this assumption. The first two words of the ence are carefully matched within a quadruple, while the remainder often vary freely in the service of creating sensible and natural sentences. Over the course of the experiment, a high density of true sentences is desirable to keep subjects processing all items as if they were sentences.

Overview of Experiments 7 to 9

As with Experiments 3 to 6, the materials of Experiments 7 to 9 were mixed together and presented in a single session. There were 4 conditions, with 5 subjects in each. None of the 20 subjects had participated in any of the earlier experiments.

Five subjects per cell is not many; Experiments 2 through 6 were all based on 10 subjects per cell. Because the number of subjects is limited, I will be more liberal in arguing from results which do not meet ordinary significance levels. By the same token, one must view the conclusions drawn from this series of experiments as tentative.

Experiment 7 is adapted from the materials of Experiment 1. It is included in part to determine whether the robust results obtained with the original paradigm can be replicated with the modified paradigm. It is also designed to investigate the influence of right context (the ence) upon the

lexical decision in this task.

Experiment 8 follows closely from Experiment 2. The section of Experiment 2 with verb and noun targets yielded the clearest results of any with the original paradigm; it is included here to see whether such clear and basic effects are found after the paradigm modification. The section of Experiment 2 with adjective and adverb targets contrasted a likely continuation and an extremely unlikely continuation, neither of which was a predictable phrase head. The inclusion of these materials allows us to ask whether likely continuations yield faster RTs than unlikely continuations, now that the task-specific salience of complete sentences has been removed. The sections of Experiment 2 with nonwords are included in Experiment 8 to see whether morphology can make nonwords assume a grammatical category in the modified paradigm.

Experiment 9 is adapted from Experiment 6a, with substantial changes. It pits HOPH and CSH against each other in the modified paradigm, where the task-specific versions of each have lost their principal logical support.

EXPERIMENT 7

Experiment 7 is adapted from the materials of Experiment 1. One of its purposes is to see whether the result of Experiment 1 will be replicated with the modified paradigm. There is no guarantee that it will.

The other purpose is to see whether right context can affect the lexical decision task. This question is of considerable interest because of its relation to the problem of lookahead in parsing. When the parser has received some word in the middle of a sentence, the number of possible structural roles the word can assume is often large. This problem of local ambiguity can be reduced considerably if the parser withholds its assignment of a structural role for the word until one or more following words have been received and subjected to a preliminary analysis. This is known as lookahead, and many (Kimball, 1973; Frazier and Fodor, 1978; Marcus, 1980) but not all (Kaplan, 1972) writers on parsing assume that it occurs to some degree.

Within this experimental task, it is possible that the parser will continue to operate after the lexical decision target appears, and that the ease with which it assigns a structural role to the target word will influence the lexical decision. If lookahead is used by the parser, the identity of the word following the target might in turn influence the process by which the parser assigns a structural role to the target word. Alternative explanations of a right-context effect do exist, but this is the line of inquiry which will lie open if we can find an effect of the first word of the ence upon the lexical decision. In Experiment 7 we contrast ence-initial words which are completely natural continuations with those which are bad (ungrammatical) continuations.

Example 12 shows a sample quadruple of Experiment 7 items.

Example 12.
MVD
If your bicycle is stolen, you must FORMULATE a complaint with the police
MVV
If your bicycle is stolen, you must FORMULATE have complaint with the police
PVD
For now, the happy family lives with FORMULATE a complaint with the police
PVV
For now, the happy family lives with FORMULATE have complaint with the police

(The word "police" on the line following versions MVV and PVV reflects the fact that this word was presented on the next line.) The sents are modal and preposition sents (from Experiment 1), followed by a verb target (also from Experiment 1), which here appears in all four versions. One of the ence types (D) begins with a determiner, and more specifically one of the two articles a or the. It is followed by the remainder of an object noun phrase (and sometimes more material). The other ence type (V) is identical to its matched D type, except that one of the two verbs be or have is substituted for the determiner. At least two words of each ence are presented on the same line as the target.

If the modified paradigm behaves like the original one, we should observe a main effect of sent type, with modal sents yielding faster RTs than the preposition sents. This would constitute a replication of the half of the original Experiment 1 result which we can test here. (There are no noun targets in Experiment 7.)

An effect of right context might emerge in one of two ways. A simple main effect of ence type might be found, with the D ences yielding faster RTs than the V ences. However, it seems likely that no beneficial effect of the acceptable ence will be observed in version PVD, where the mismatch between sent and target has already made the item ungrammatical. If so, the effect of right context might emerge as an interaction effect, with versions PVD and PVV yielding equally slow RTs, version MVV yielding a faster RT

because the sent and target are in accord, and version MVD yielding a still faster RT because sent, target, and ence are all in accord.

METHOD

Since Experiments 7 through 9 were mixed together in the same session, several of the following sections apply to Experiments 8 and 9 as well: Subjects, Filler Items and Randomization, Procedure, and Overall Results.

Subjects

Young adult volunteers from the Boston area were paid for their participation. All were native speakers of American English. None had participated in earlier experiments.

Materials

The entire set of Experiment 7 materials appears in Appendix G. Twelve item quadruples such as that shown in Example 12 were the materials for this experiment. The sents and targets were taken from Experiment 1a. The sents were in most cases identical, though some of the longer ones were shortened so that two words of the ence would fit on the same line. The Experiment 1 verb targets were used in the present experiment, with one exception. The verb COMPETE does not take noun objects, and was replaced with the verb FORGET. Since our legal continuation ences begin with an NP, it is important that all target verbs can (and often do) occur with noun objects.

The determiner ences were constructed to be natural completions of the modal sents, if an appropriate verb is substituted for the target. The first word of each was one of the articles a or the, introducing the object noun phrase. The verb ences were formed from the determiner ences by

replacing the articles with one of the verbs have or be.

Corresponding nonword materials were formed by a virtually identical procedure from the nonword materials of Experiment 1a. There was one difference in the ences. Internally ungrammatical ences were necessary in the word materials of this experiment, but were avoided in the nonword materials. Instead of forming the verb ence from the determiner ence by replacing the determiner, the verb was inserted in front of the determiner. This produces ences such as "have the view of all of us" rather than "have view of all of us".

No subject saw more than one item from a quadruple.

Filler Items and Randomization

Experiments 7, 8 and 9 contributed a total of 112 items to be seen by each subject group. Half of those contributed by each experiment had nonword targets, so of course half of the total had nonword targets. In addition, 112 filler items were constructed and shown to all four subject groups. Of these, 56 had nonword targets, and 56 word targets. All filler items were complete and natural English sentences, with some word within them either designated as the target word, or else replaced with a nonword target. These fillers were modified from fillers used with the original paradigm. Modification usually involved adding additional words, and often involved designating a different word as target.

Half of the items seen by each subject had word targets, the other half nonword targets. Half of the word-target items were complete and natural sentences; in many of the others a natural sentence could be formed by replacing the target word. Over half of the nonword-target items could be transformed into complete and natural sentences by replacing the target nonword.

The items were presented in different pseudo-random orders to different subjects. The entire randomized set of 224 items was divided into two equal blocks, each preceded by 5 warm-up items. There was also an initial block of 30 representative practice items.

Procedure

The subject was seated before the CRT of a TERAK computer in a dimly lit room, holding a small box with two side-by-side microswitches. The subject operated one switch with each hand, the right for word responses and the left for nonword.

Each trial started with the clearing of the CRT and the appearance of a fixation cross at the left margin. 300 msec later the fixation cross disappeared, and the item began appearing from left to right across the screen, a new word appearing every 300 msec. This uniform rate applied equally to sent, target, and ence, including words of the ence appearing on the next line. In contrast to the original paradigm, there was no alerting signal before the appearance of the lexical decision target. The target appeared in capital letters and was accompanied by a tone. 300 msec after the final word of the ence, a slash " / " appeared immediately to the right as a signal that the item had ended. At this point the subject was free to initiate the next trial by a push of the foot pedal, assuming that the lexical decision had been completed.

Subjects were given a 5-item demonstration of the task by the experimenter, and then instructed as to the procedure and the nature of the lexical decision task. They were informed that some sentences would make sense and some would not. They were requested to attempt to understand the sentences as they appeared, and were told that they would take a sentence recognition test at the end of the experiment. On the other hand, they were

asked not to reread the items after they had reached the end, and to make no special effort to remember the items.

The recognition test contained 20 items that had appeared, and 20 which had not. All of the nonoccurring items were modified from occurring items, 10 by replacement of the sent, and 10 by replacement of the ence. All subjects scored above chance on this test.

The entire experiment took approximately 40 minutes.

OVERALL RESULTS

Twenty subjects were run, 5 in each of the 4 conditions. Two subjects were replaced due to error rates over 10%, and six because of mean RTs over 850 msec. Three of these six (but no other subjects) exceeded the 300 msec cutoff for RT standard deviation. The overall error rate (including filler items) was 4.3%. RTs more than two standard deviations above the subject's mean were replaced with that subject's cutoff value.

RTs were higher than those observed in earlier experiments. In particular, each experiment (and the filler items) yielded overall mean RTs approximately 70 msec above the previous experiment (or set of filler items) from which their materials were drawn. This change in overall RTs is not surprising, given the substantial changes in the experimental technique. The elimination of the alerting stimulus in the modified paradigm is a likely cause of this RT elevation; in the original paradigm the tone provided a 300 msec warning of the impending lexical decision.

RESULTS AND DISCUSSION

The left side of Table 10 shows the results of Experiment 7 with word targets.

Table 10. Experiment 7.
Mean Reaction Time (Milliseconds) as a Function of Syntactic Category of Sent-Final and Ence-Initial Word. Error percentages are in parentheses.

	Word (Verb) Target:					Nonword Target:				
	Ence-Initial Word					Ence-Initial Word				
	Determiner		Verb			Determiner		Verb		
Modal Sent:	645	(2)	690	(2)	668	762	(2)	761	(7)	762
Prep.	703	(8)	720	(3)	712	708	(8)	729	(5)	719
	674		705			735		745		

The effect of Experiment 1 which we are trying to replicate using this modified paradigm is a main effect of sent type (for RTs). An effect of 44 msec was found, in the predicted direction. It is significant for both subjects ($F(1,19) = 4.6$, $p = .045$) and items ($F(1,11) = 5.8$, $p = .034$) but not for min F' (min $F'(1,30) = 2.6$, $p > .05$). There is also a main effect of sent type for errors, near significance for both subjects ($F(1,19) = 3.1$, $p = .096$) and items ($F(1,11) = 4.7$, $p = .054$). This effect is not in the direction to suggest a speed-accuracy tradeoff, but rather indicates that difficult items yield both slow RTs and high error rates. This concordance of RT and error results was not found in any experiment with the original paradigm. There, the theoretically interesting effects were found for RTs, while errors assumed no interpretable pattern.

The effect of right context (the ence) was expected to show up either as a main effect of ence type, or as an interaction of the sent and ence variables. For errors, these two effects are far from significance. The pattern in the RTs is in line with both possibilities. There is a 31 msec main effect of ence type in the predicted direction, which is significant for subjects ($F(1,19) = 5.9$, $p = .025$) but not for items ($F(1,11) = 1.9$, $p = .20$). There is also a 28 msec interaction effect of the expected type,

which is far from significance for both subjects and items. These effects certainly do not demonstrate that right context can affect the lexical decision. But they are reasonably encouraging, given the small numbers of items and subjects.

The right side of Table 10 shows the results of Experiment 7 for the nonword targets. There was a 43 msec main effect of sent type for RTs, significant for both subjects ($F(1,19) = 6.8$, $p = .018$) and items ($F(1,11) = 9.7$, $p = .0099$) but not quite for min F' (min $F'(1,30) = 4.0$, $p < .10$). The significant effect of sent type is in the opposite direction from that found with the word targets, and was not found in Experiment 1. None of the other effects for the nonwords (errors or RTs) were near significance.

Overall, the results of Experiment 7 provide evidence that the effects found with the original paradigm did not arise from a task-specific concern with complete sentences. The target in the modified paradigm is never the last word of the item, yet the effect of sent type (with word targets) replicates one of the RT effects found with the original paradigm. The finding of a complementary error effect strengthens this result.

EXPERIMENT 8

The result from the original paradigm which most convincingly demonstrated a syntactic effect of sentence context on lexical decision was the effect found in Experiment 2 with verb and noun targets. These materials were used in Experiment 8 to see whether such effects could be found with the modified paradigm. A sample item quadruple is shown in Example 13.

Example 13.

NV
The crowd near the church indicates that an important funeral TRANSLATES traffic patterns seriously

NN
+The crowd near the church indicates that an important funeral TRANSLATION from something they want

VV
*The towers on the horizon indicate that the barriers isolate TRANSLATES traffic patterns seriously

VN
The towers on the horizon indicate that the barriers isolate TRANSLATION from something they want

The sents and targets of Experiment 8 are identical to those of Experiment 2, without exception. As before, we expect versions NN and VV to yield higher RTs than NV and VN.

The ences do not constitute an independent variable in Experiment 8, but instead vary with the target word. The ence for versions NV and VV begins with a noun phrase introduced by an adjective. This provides a grammatical completion of the NV version, but not of the VV version; note that the VV version has already become ungrammatical by virtue of the target word.

The ence for versions NN and VN begins with a prepositional phrase and is designed to complete the VN version grammatically. Note that the first two words of this ence are a grammatical continuation of version NN. In Experiment 2, the target in version NN was not a bad continuation, yet it

yielded a higher RT. This was important in ruling out the Bad Continuation Hypothesis. We would like to have this same line open to us in Experiment 8, and thus it is important that the right context of version NN not produce a bad continuation which could influence the lexical decision. Though the remainder of the ence does not complete version NN, this is virtually certain not to affect performance in the task. The NN ence past the second word sometimes differs from the VN ence in order to have a more plausible semantic relation to the NN sent.

In Experiment 2, the materials with adjective and adverb targets did not produce a slower RT for the adjective after the noun sent, where it is a highly unlikely continuation, than after the verb sent, where it is a likely continuation. If the task-specific factors drawing attention to complete sentences are removed, perhaps RT differences will emerge that distinguish the highly unlikely continuation from the likely one.

Example 14 shows sample materials from Experiment 8 with adjective and adverb targets.

Example 14.
NA
++The crowd near the church indicates that an important funeral HAPPY and somber occasion
NR
The crowd near the church indicates that an important funeral HAPPILY and truly draws people
VA
The towers on the horizon indicate that the barriers isolate HAPPY and somber prisoners indeed
VR
The towers on the horizon indicate that the barriers isolate HAPPILY and truly unwilling people

As with the verb- and noun-target materials, the sents and targets in these materials are identical to those of Experiment 2, without exception.

The first two words of the ences for all four item versions are designed to be as grammatically uninformative as possible. This serves the general purpose of minimizing information from the right context as to the

intended structural role of the target adjective or adverb. A more particular purpose concerns version VA. If the ence for this version began in the most natural way, with a noun, this noun would be the head of the predictable object noun phrase. It is possible that this would be sufficient to yield a lowered RT based on HOPH, even though the target itself is not the predicted head of a phrase. Regardless of how likely this possibility is, it is safest to avoid the issue by withholding the phrase head until the third word of the ence.

The ences begin with a coordinating conjunction (and, or, but, or yet) followed by an adjective after adjective targets and an adverb after adverb targets. From the third word on, the ences continue differently for each version. In each version except NA, the ence completes the item as a sentence, though usually a rather awkward one.

Both the adjective-adverb section of Experiment 8 just described and the verb-noun section described earlier require balancing nonword items. The nonword items of Experiment 2 were provided with ences in just the same way as the word items, and included in Experiment 8. They will provide another opportunity to see whether sentence context can affect the rejection of nonwords marked by morphology with a grammatical category. It is possible that the modification of the paradigm will allow such an effect to emerge.

METHOD

For the Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 7.

Materials

The Experiment 8 materials with word targets are shown in Appendix H. Sample quadruples appear in Examples 13 and 14. The sents and targets of these materials are identical to the Experiment 2 materials. The ences after verb targets start with a noun phrase introduced by an adjective (or derived form serving as an adjective). The ences after noun targets start with a prepositional phrase. The ences after adjective targets start with a conjunction (and, or, but, or yet) and another adjective. The ences after adverb targets start with a conjunction and another adverb. Beyond the second word the structure of the ences was not carefully controlled, based on the assumption that these words could not consistently affect the lexical decision task. Often the end of the ence served to complete a sentence.

Ences for nonword items have the same properties as the ences for the corresponding word items.

Since the Experiment 2 items were often long, the target word in Experiment 8 is often the last word on its line, with the entire ence appearing on the next line. In general, as much of the ence was shown on the first line as was possible. The one constraint observed was that if any item in a quadruple had its entire ence on the next line, all other items in the quadruple did as well.

RESULTS AND DISCUSSION

Word Items With Verb and Noun Targets

Table 11 shows the results for the materials exemplified in Example 13.

Table 11. Experiment 8, Word Items, Verb/Noun Targets. Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word. Error percentages are in parentheses.

		Target: Verb	Target: Noun	
Sent-Final Word	Noun	752 (7)	767 (10)	760
	Verb	754 (13)	736 (2)	745
		753	751	

The expected effect was an interaction of the target type and sent type variables, with the NN and VV versions yielding higher RTs than the other two. The interaction effect of 33 msec was in the right direction but far from significance. Both main effects in the RT analysis were likewise far from significance.

The interaction in the error analysis showed that the NN and VV cases produced more errors, if not longer RTs. This error interaction was near significance for both subjects ($F(1,19) = 3.1$, $p = .095$) and items ($F(1,11) = 4.6$, $p = .056$). The main effect of sent type for errors was far from significance, but the main effect of target type was near significance (Note 10) for items ($F(1,11) = 4.7$, $p = .054$) but not for subjects ($F(1,19) = 1.5$, $p = .23$).

The failure to replicate the Experiment 2 interaction effect for RTs (Note 11) indicates that the modified paradigm is not behaving like the original one. The near significant error interaction suggests that syntactically unfavored cases are indicated in the modified paradigm by increased error rates rather than increased RTs. Recall that the main effect of sent type in Experiment 7 -- replicating the Experiment 1 result -- was observed for both RTs and errors.

Both RT increases and error increases are common indicators of

processing difficulty, and are frequently observed to cooccur. If the modification of the paradigm shifts the manifestations of expected effects from RTs to errors, this is a superficial change. It does not indicate that the involvement of linguistic processes is different in the two paradigms. In this light, the above results constitute both a replication of the Experiment 2 effect and a calibration indicating that the modified paradigm is likely to show effects in errors rather than RTs.

Word Items With Adjective and Adverb Targets

Table 12 shows the results for the materials exemplified in Example 14.

Table 12. Experiment 8, Word Items, Adj/Adv Targets. Reaction Time (Milliseconds) as a Function of Syntactic Category of Target and Sent-Final Word. Error percentages are in parentheses.

		Target: Adj	Target: Adverb	
Sent-Final Word	Noun	740 (17)	766 (13)	753
	Verb	723 (3)	753 (3)	738
		731	759	

For RTs, all differences are small and far from significance. For errors, the main effect of target type and the interaction are also far from significance, but the main effect of sent type is significant for both subjects ($F(1,19) = 9.2$, $p = .0068$) and items ($F(1,11) = 6.1$, $p = .032$) but not for min F' (min $F'(1,25) = 3.7$, $p < .10$).

In Experiment 8 (with verb and noun targets) and in Experiment 7, results from the original paradigm have been replicated in the error data. For this reason the present error effect is especially notable. It is unlikely that it results from a difference in the sents themselves, since 1)

the very same sents with verb and noun targets (above) showed no hint of such a main effect, and 2) similar sents with nonword targets (below) show no such main effect either.

If one considers only the results with the adjective target, the effect takes the form of a higher error rate for the highly unlikely continuation (NA) than the syntactically common continuation (VA). These two versions did not differ in Experiment 2. This could be taken as evidence that the modified paradigm is more sensitive than the original one, and that HOPH and CSH do not provide a complete account of effects with the modified paradigm. However, one cannot simply ignore the parallel error pattern of the adverb targets. There is no reason to expect the NR version to produce more errors than the VR version. A comparison with the error patterns of the parent materials in Experiment 2 is suggestive, though. Comparing Tables 3 and 12, one sees that the error percentages in the adverb column are similar, while those in the adjective column are dramatically reversed. If one took the Experiment 2 error results as a baseline, the Experiment 8 effect would be very dramatic. However, the justification for such baselining is highly questionable, since the modified paradigm is yielding different error patterns across the board.

All considered, this error effect is tantalizing but difficult to interpret.

Items With Nonword Targets

The left side of Table 13 shows the results for the items with "verb" and "noun" nonword targets; the right side shows the results for "adjectives" and "adverbs."

Table 13. Experiment 8, Nonword Items.
Reaction Time (Milliseconds) as a Function of
Syntactic Category of Target Nonword and Sent-Final Word.
Error percentages are in parentheses.

	Target "Verb"		Target "Noun"			Target "Adj"		Target "Adverb"		
Sent: Noun	797	(7)	804	(5)	800	778	(5)	834	(7)	806
Sent: Verb	776	(5)	799	(5)	788	800	(3)	870	(12)	835
	786		802			789		852		

There is a main effect of target type for the "adjectives" and "adverbs", with the "adverbs" taking longer to reject and producing more errors. The RT effect is significant for both subjects ($F(1,19) = 4.6$, $p = .046$) and items ($F(1,11) = 5.6$, $p = .037$) but not for min F' (min $F'(1,30) = 2.5$, $p > .05$). The error effect is significant for items ($F(1,11) = 11.0$, $p = .0069$) but not for subjects ($F(1,19) = 2.8$, $p = .11$). The same (RT) effect was observed for these nonwords in Experiment 2. A similar Experiment 2 main effect for "verbs" and "nouns" was not found here; in fact, all other effects (for both RTs and errors) with these nonwords are small and far from significance.

As in Experiment 2, the nonword results do not show any enlightening relationship to the word results. This strengthens the conclusion of Experiment 2 that nonwords cannot be endowed with grammatical category in a way that interacts with context.

EXPERIMENT 9

Experiments 6a and 6b were designed to distinguish the Heads-of-Phrases Hypothesis and the Complete Sentence Hypothesis. Neither produced the positive result which would have decided the issue in favor of HOPH. Though CSH was slightly favored by those results, there were reasons to doubt that the materials provided a fair test. While attention in this series of experiments has shifted to eliminating the task-specific incentives for a concern with complete sentences, it is still important to know whether HOPH or CSH is responsible for the observed effects.

To this end, Experiment 9 was constructed. As in Experiments 6a and 6b, a noun target appears either where it is a bad continuation or where it is a predictable head of a phrase, but not the completion of a sentence. Example 15 shows a sample quadruple.

Example 15.
PN It is believed that the fruit of ENTRIES in cities tastes awful
P- It is believed that the fruit of ENDOTES in cities tastes awful
MN *Fred thinks that the red plums could ENTRIES if they are not kept cold
M- Fred thinks that the red plums could ENDOTES if they are not kept cold

The head-of-phrase environment is following a preposition which is part of a subject noun phrase. The bad continuation environment is following a modal verb. The versions with nonword targets provide a baseline measure of the processing difficulty of the sent. The ences of Experiment 9 vary with the sent. They are designed to complete a sentence, provided an appropriate word is substituted for the target.

HOPH predicts a lower RT (or fewer errors) for version PN than MN, relative to the nonword baselines. CSH predicts no differences. Unlike Experiments 6a and 6b, the targets appear neither within adverbial clauses nor within incomplete WH-phrases. This is to reduce the possible processing backlog which might have arisen in 6a and 6b and masked the effect predicted

by HOPH.

METHOD

For Subjects, Filler Items and Randomization, Procedure, and Overall Results sections, see Experiment 7.

Materials

The complete set of Experiment 9 materials appears in Appendix I. A sample quadruple is shown in Example 15. The first sent of each pair ends after a preposition within a subject noun phrase. The second sent of each pair ends after a modal verb. The targets are plural nouns and superfically similar nonwords, which are the targets of Experiment 6 transformed into their plural forms. The ences in Experiment 9 vary with the sent. The ences after preposition sents begin with a prepositional phrase. The ences after modal sents are variable, but very often they are a conjoined clause. Both types are designed to complete a sentence, provided that the target is replaced with a suitable word. As in Experiments 7 and 8, part of the ence frequently appears on the following line.

Twenty such quadruples were constructed. Each subject could be presented with two items from each quadruple without showing him or her the same sent or target twice. This allowed 10 subjects to see each item, which was not possible in Experiments 7 and 8.

RESULTS AND DISCUSSION

The results of Experiment 9 are shown in Table 14.

Table 14. Experiment 9.
Mean Reaction Time (Milliseconds) as a Function of Target Type (Word or Nonword) and Sent-Final Word. Error percentages are in parentheses.

		Target: Noun		Target: Nonword		
Sent	Preposition	687	(3)	717	(1)	702
	Modal	710	(10)	722	(3)	716
		698		720		

For RTs, the effect score is 18 msec in the direction predicted by HOPH, but far from significance. Both main effects are far from significance. The failure to find RT effects is not surprising; Experiment 8 has primed us to expect effects in errors rather than RTs.

For errors, the main effect of target type is significant (min $F'(1,38) = 4.5$, $p < .05$). The main effect of sent type is significant for both subjects ($F(1,19) = 7.3$, $p = .014$) and items ($F(1,19) = 5.6$, $p = .028$) but not for min F' (min $F'(1,38) = 3.2$, $p > .05$). The interaction effect is near significance for items ($F(1,19) = 3.1$, $p = .094$), and not all that far from significance for subjects ($F(1,19) = 2.3$, $p = .15$).

The main effects assume no particular importance. The interaction, on the other hand, is exactly the effect (for errors) predicted by HOPH. The MN version (the bad continuation) is yielding more errors than the PN version (the predicted head-of-phrase), even relative to their baseline controls. The statistical significance of this effect is admittedly a bit weak. And in contrast to Experiments 7 and 8, Experiment 9 is based on 10 subjects per cell, so no apologies about small numbers of subjects are in order.

In sum, the results of this experiment provide some weak evidence that CSH is false and HOPH is true (Note 12).

* * * A MISCELLANEOUS EXPERIMENT: CLOSED CLASS TARGETS * * *

Experiment 10 was executed in the same experimental session as Experiment 2. Since it is somewhat tangential to the thesis, it appears here out of chronological order.

EXPERIMENT 10

The vocabulary of languages, and English in particular, can be divided roughly into function words and content words. Content words are the nouns, verbs, adjectives, and adverbs. They are the long words, they convey meaning, and they constitute the vast majority of the vocabulary. Words that are invented and rare words that one does not know are invariably content words. Function words are the short, frequent words which express grammatical relations. The function words are a fixed, closed class, while the content word categories are open. For this reason they are often called the closed class and the open class vocabularies.

All the word targets used in Experiments 1 and 2 were members of the open class. The behavior of closed class words as targets in this experiment is of interest for two rather disparate reasons.

Bradley (1978) has suggested that closed class words are accessed by a different system than open class words. She has proposed that the key role of closed class items in parsing is the purpose of the separate recognition system. If they are islands of reliability in a sea of open class syntactic ambiguity, it would make sense for their recognition to be impervious to syntactic influence. If the present paradigm shows this, it would constitute more evidence for a separate recognition system.

On the other hand, if the recognition for closed class items can be

influenced by sentence context, it will provide a larger vocabulary with which to ask questions of the human parser. For instance, if one wanted to determine the readiness of the parser to accept a prepositional phrase, one would like to present a preposition as the lexical decision target.

The four categories of closed class items chosen were objective case personal pronouns, prepositions, wh-words, and quantifiers. These were combined with the exact sents that were used in Experiment 1, i.e. those ending in modals and prepositions. These sents have a demonstrated ability to affect lexical decision for noun and verb targets. Example 16 shows samples of the resulting items, with markings as before to indicate rough intuitions of the expected processing difficulty. ("O" stands for "pronoun".)

Example 16.

MO *The man spoke but could IT
PO Just at the time of IT
MP +The man spoke but could TOWARD
PP ++Just at the time of TOWARD
MW *The man spoke but could WHO
PW +Just at the time of WHO
MQ ++The man spoke but could SOME
PQ Just at the time of SOME

METHOD

Experiment 10 materials were included with the materials of Experiment 2 in the execution of that experiment. Thus the Subjects and Procedure sections are identical to those of Experiment 2.

Materials

The materials used in Experiment 10 are shown in Appendix J; Example 16 shows a prototypical set. Twelve such sets were constructed. The 12 sent pairs were those taking word targets in Experiment 1a. Recall that the first member of each pair ends in a modal verb, the second in a preposition.

The target words were six of each of four types of closed class words. The objective case personal pronouns were IT, HIM, US, THEM, ME, and YOU. The prepositions were TOWARD, AMONG, WITH, FROM, UPON, and AGAINST, taken from the subset of prepositions that do not have uses as particles or adverbs. The wh-words were WHO, WHICH, WHERE, HOW, WHEN, and WHY. The quantifiers were SOME, ALL, EITHER, BOTH, EVERY, and ANY. These categories were selected because one can predict some difference in their acceptability after modals and prepositions, and because there are six exemplars of each. As it is, the wh-words and quantifiers are disturbingly heterogeneous. However, using categories with fewer members would invite the possibility that any differences obtained would be due to the particular items rather than the syntactic category.

The word items were formed by assigning one exemplar of each of the four closed class categories to a sent pair, producing a set of items such as that in Example 16. Each target word was assigned to two different sent pairs, compensating somewhat for the small target word sets. The items were counterbalanced among subject groups so that they saw each sent and each target exactly once. Nonword items were formed by assigning 24 short nonwords to the other 12 sent pairs of Experiment 1.

RESULTS AND DISCUSSION

Table 15 shows the reaction time results for Experiment 10. Sample items appear in Figure 16.

Table 15. Experiment 10.
Reaction Time (Milliseconds) as a Function of Grammatical Category of Target and Sent-Final Word. Error percentages are in parentheses.

		Sent-Final Word Category: Modal	Preposition	
Target:	Pronoun	596 (3)	567 (3)	582
	Preposition	629 (5)	621 (7)	625
	Wh-word	594 (3)	598 (2)	596
	Quantifier	586 (1)	576 (2)	581
		601	591	

For each of the four types of closed class target words, pairwise comparisons were made between the RT after modal sents and the RT after preposition sents. For prepositions, wh-words, and quantifiers, there is no hint of an effect of preceding sent. For pronouns, where the difference in syntactic acceptability is clearest, there is a 29 msec difference in the predicted direction. This is near significance for both subjects ($F(1,39) = 3.9$, $p = .054$) and items ($F(1,11) = 3.3$, $p = .098$).

The predicted contrast in processing difficulty for the preposition targets and for the WH-word targets is not great, so the finding of no difference in these two cases is not surprising. It certainly does not argue that decision times for closed class words cannot be influenced by sentence context.

For the pronoun and quantifier targets, though, the predicted differences in processing difficulty were substantial. We find no evidence of an RT difference for quantifiers, but good evidence for such a difference for pronouns. These results are completely in accord with HOPH. The pronoun in the PO version is a predictable phrase head, while it is a bad

continuation in the MO version. In the MQ version the quantifier is a highly unlikely continuation, while in PQ it is a likely continuation, but not the head of the predictable noun phrase.

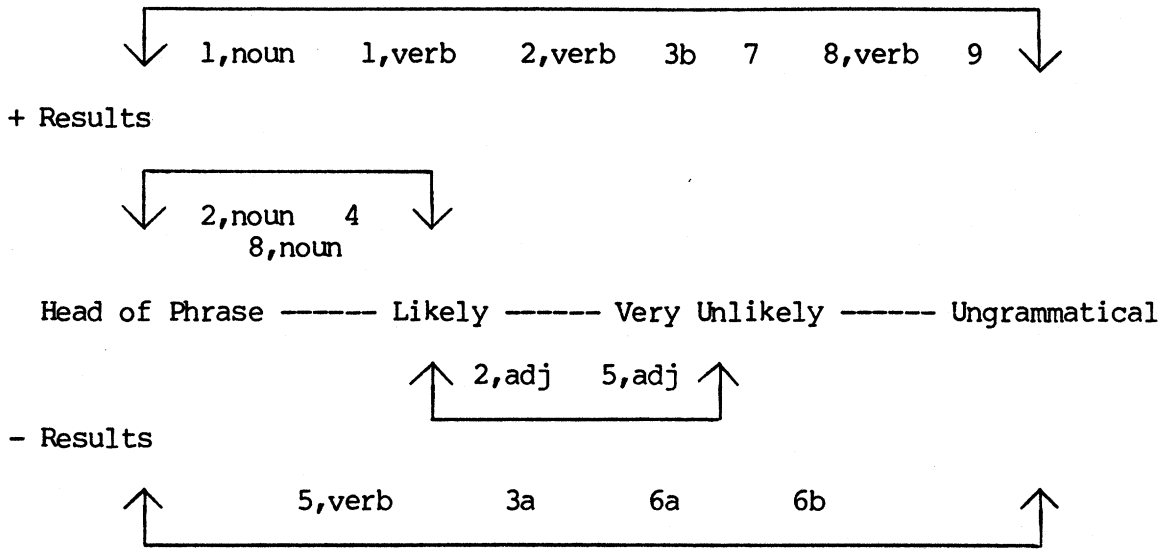
It appears that sentence context can affect the recognition of at least some kinds of closed class words (the pronouns). Because the closed class is so heterogeneous, this conclusion cannot be extended to the rest of the closed class vocabulary.

* * * THE HEADS OF PHRASES HYPOTHESIS: EVIDENCE AND IMPLICATIONS * * *

One useful way of categorizing the item versions of the different experiments is by the syntactic acceptability of the target word given the sent. I have distinguished four degrees of acceptability in this thesis: 1) Ungrammatical. The target word cannot grammatically be assimilated into the sent; the item has become ungrammatical. 2) Very Unlikely Continuation. The item can be grammatically completed, but the target word introduces a syntactically very unlikely construction. 3) Likely Continuation. The sent can be continued using the target word in a common syntactic construction. 4) Head of Phrase. The target can function (syntactically) as the head of a phrase which can be predicted, given a knowledge of the grammar, as of the end of the sent. Each experiment contrasts item versions which differ in their acceptability on this scale.

This scale is shown horizontally in Table 16. Target word types from experiments are grouped with horizontal lines, and the arrows on the ends of the line point to the levels of acceptability contrasted for that target word type. Contrasts that yielded differences (positive results) are shown above the scale, while findings of no difference (negative results) are shown below the scale. (Experiment 8 with adjective targets and Experiment 10 pose problems of classification and are omitted).

Table 16. Negative and Positive Experimental Results by Contrast in Acceptability of Continuation. Generally, there is one entry for each kind of target word.



Every positive result contrasts the Head-of-Phrase level with a lower level of acceptability. In particular, the 3 cases contrasting Head-of-Phrase with Likely all yielded positive results. This, then, is the evidence that Head-of-Phrase is the crucial relationship between sent and target from which RT differences arise.

Partial explanations have been offered for the negative results, with the exception of Experiment 2, Adjective. Experiments 5 and 6 have determiners in sent-final position, and Experiment 3a requires too complex a computation for the expected effect to emerge. In any event, all of these contrasts yielding negative results span the entire scale if they involve the Head-of-Phrase level at all, so they do not provide evidence that any other level on the scale is crucial.

It is possible that the paradigm could distinguish other levels. Three attempts have been made to contrast the Likely and Very Unlikely levels.

Two (2,adj and 5,adj) have yielded negative results and one (8,adj) has yielded uninterpretable results. Results to date provide only an absence of positive direct evidence on this question. There is, however, some indirect evidence that the contrast of Likely and Ungrammatical levels would yield a very small effect, if it were to yield one at all. The evidence is that the magnitude of the RT differences found when Head-of-Phrase is contrasted with Likely and with Ungrammatical are comparable. Because this line of reasoning is indirect, we cannot actually subtract the effects and conclude that Likely and Ungrammatical levels would not differ. It does seem safe to infer that the Head-of-Phrase level has a status apart from the other levels in the magnitude of its influence on RT.

Though (as Table 16 has shown) every positive result involves a Head-of-Phrase version, most often this version is also a Complete Sentence. The Complete Sentence Hypothesis maintains that this is the crucial property of versions yielding low RTs. But in Experiment 9 the Head-of-Phrase version is <u>not</u> a complete sentence, yet the experiment yields a (weak) effect. Of greater importance in ruling out CSH is the overall pattern of results from Experiments 7, 8, and 9. In these experiments there is little reason to think that the complete sentence is perceived by the subjects at all, since the target is embedded within a continuous stream of words. Yet we observe the same pattern of effects as in Experiments 1-6, arguing that CSH is not responsible for effects in either case.

A different (and more difficult) question is whether the lower RTs and reduced error rates arising when the target is a Head-of-Phrase are due to the normal operation of the human parser or emerge only because of the requirements of this experimental task. We have provided solid evidence (particularly in Experiments 2 and 8) that the effects are syntactic in the sense that syntactic category is the crucial difference between target word

conditions. Since subjects in this task are reading sentences and we are finding effects of syntactic category, we have a good prima facie case that it is the human parser which lies behind the effects. It is the human parser, after all, which is primarily concerned with syntactic category during language comprehension.

One argument against task-specific explanations is that the same pattern of results has been obtained with two different paradigms, though admittedly the paradigm differences are not great. In the original paradigm there is an alerting stimulus 300 msec before the target, while in the modified paradigm there is none. One cannot, for the modified paradigm, claim that parsing stops and task-specific predictions begin 300 msec before the lexical decision.

In arguing against CSH above, we noted that in the modified paradigm the target is embedded in a continuous stream of words. This argues not only against a task-specific concern with complete sentences, but complete phrases as well. Since the head of a phrase is the minimal requirement for a complete phrase, such a concern with complete phrases could lead to a prediction of phrase heads.

One possibility for an alternative explanation is that grammatical information is being harnessed for the (task-specific) purpose of improving lexical decision performance. This has some plausibility because word targets are quite often predictable Heads-of-Phrases. By predicting the syntactic category of the target, and being right much of the time, lexical decision performance is improved. One precondition of this explanation is that knowledge of syntactic category can actually improve lexical decision performance. Fay (1975) tested this hypothesis and found that prior knowledge of grammatical class decreased RTs to verbs, but not to nouns. We have in the present research found many RT contrasts for noun targets. If

an unreliable prediction of syntactic category in a sea of other variables can produce the effects we observe, a reliable prediction of category in isolation should produce very clear effects, but it does not. The reason that we, but not Fay, observe effects of syntactic category is that the effects are mediated by the human parser, a device which certainly is sensitive to syntactic category.

The potential supply of alternative explanations is inexhaustible, but I have considered all those which seem plausible to me. I leave the formulation of the correct alternative explanation, if there is one, to those with more fertile imaginations. In the next section I consider the implications of the Heads-of-Phrases Hypothesis for the human parser.

Implications

The Heads-of-Phrases Hypothesis has an obvious and direct translation in terms of the human parser: When a knowledge of the grammar allows the reliable prediction of a syntactic phrase, the parser predicts the phrase, and also predicts the head of the phrase. This will be known as "Parser Predicts Heads of Phrases" (PPHOP).

First some remarks on what PPHOP does not mean. "Prediction" of the phrase head does not mean that the parser predicts that the next word will be the phrase head, only that the phrase head will appear eventually. And "prediction" does not mean that the parser is priming the lexical access apparatus to receive the predicted part of speech. This is not impossible, but in my view it is unlikely. I suggest that the effects I have obtained in the experimental paradigm result from an interference of the parsing analysis with the lexical decision task. That is, the target is recognized and the parser incorporates it into its analysis of the sentence. If the target fulfills the Head-of-Phrase prediction, post-access processes in the

lexical decision task are facilitated. Restated, I propose that the prediction by the parser of the phrase head serves no function such as facilitating recognition of a word, but that the prediction is made by the parser for its own purposes. The effect of the prediction on the lexical decision task is only a fortunate epiphenomenon.

HOPH states that lowered RT results when the target can function as the head of a predictable phrase. Though intuitive enough, the inference of PPHOP from HOPH deserves some defense. The alternative view is that the parser operates bottom up, and somehow produces lowered RT when the target is a predictable head. In concrete terms, the challenge to the bottom-up parsing theorist is to explain why a noun target is facilitated but an adjective target is not, when each occurs after an object-taking verb (the result of Experiment 2). With sufficient assumptions, one might be able to concoct a bottom up system which would produce this result. But observe that the natural predictions of a bottom up parser are in line with the discredited Bad Continuation Hypothesis. One would expect to observe disruption if the target cannot be integrated into the sentence, and much smaller effects (if any) distinguishing cases where the target is integrated successfully in various grammatical roles.

Concluding that the parser actually does predict phrases (top down) which are predictable does not mean that the entire parsing process is top down; there may well be bottom-up components of the process. We are only maintaining that there are some (important and general) top down components to the parser.

PPHOP bears on the evaluation of existing parsing proposals to varying degrees.

Most recently-proposed parsers assume that a single analysis of the sentence is pursued at one time. However, some assume that many different

analyses are carried along in parallel. One can either hold that the parallel analyses are fundamentally unordered and of equal "strength", or that some are preferred to others. PPHOP militates against the former variety of parallel models. The parser of Thorne, Bratley, and Dewar (1968) would most naturally be construed as this kind of parallel parser, though it could easily be expanded to allow an ordering of parallel analyses. Essentially, if all possible analyses are given equal weight, then there is no reason to expect the parser to prefer any legal continuation of a sentence over any other. The differential prediction of phrase heads which is implicit in PPHOP is inconsistent with this view.

The special status of heads of phrases in PPHOP similarly argues against the Augmented Transition Network (ATN) parser, at least as formulated by Kaplan (1972). Kaplan proposes that parsing difficulty corresponds to the number of arcs traversed or attempted. Given that the ATN moves back and forth over the input string, this metric is inapplicable to any result showing differential processing difficulty which is internal to a sentence. One could propose a modified metric applicable to my experimental task: Let the ATN run until it calls for a word beyond the end of the sent. RT in the task then varies with the number of arcs traversed or attempted beyond this point for the ATN to incorporate the target into its parse. If in the networks the arc for the phrase head is tried before other arcs, the ATN-with-modified-metric will explain faster RTs for phrase heads than likely continuations. However, these cases would be distinguished by a small difference in arcs attempted, whereas an ungrammatical continuation would result in a very large number of arcs attempted. One would expect to find the primary RT differences differentiating the ungrammatical case from the others.

The fundamental problem of the ATN in accounting for the results is

that it has no way of distinguishing the prediction of the phrase head from any other prediction. One can venture that any parsing system which relies on a very simple and very general method of parsing sentences will have difficulty accounting for the present data. To account for a special processing consequence of finding predicted phrase heads, one needs a parser of sufficient complexity that a separable stage or process can be held accountable for the results.

The parsers outlined by Frazier and Fodor (1978), Marcus (1980), and Ford, Bresnan and Kaplan (in press) are all of sufficient complexity to accommodate the present experimental results. All are at least partially top-down, and all predict predictable phrases. None explicitly predicts the head of a predictable phrase, but all could be modified to accommodate this kind of prediction. By maintaining that information which can aid the parsing process is utilized optimally, Frazier and Fodor imply that the heads of phrases are predicted. However, it is one thing to assert that one's parser predicts heads of phrases, and another to construct an explicit model in which the interactions of different parsing principles can be tested. In the absence of such an explicit model, the Frazier and Fodor parser cannot claim the present PPHOP result as strong support.

* * * CONCLUSION * * *

The preceding discussion of Parser Predicts Heads of Phrase (PPHOP) could give the impression that I consider PPHOP to be the primary accomplishment of this thesis. This is not true. The major purpose and the major accomplishment has been to establish a new experimental paradigm that can reflect the operation of the human parser.

While PPHOP is the most general conclusion about the parser to emerge from this thesis, there is at least one other tentative conclusion. The contrast in results from Experiments 3a and 3b suggested that the parser initially assumes that a noun phrase following the verb is a direct object. If the NP must actually be the indirect object, reanalysis is required.

There are two phenomena which serve to raise interesting possibilities for future work. In Experiment 5, the syntactic environments created by articles (a and the) appeared to differ from those created by possessive pronouns (my, your, their, and our). This counterintuitive finding deserves further investigation. In Experiment 7, there was some evidence suggesting that context to the right of the lexical decision target could influence the decision. This in turn suggests a method for studying the parser's lookahead capabilities.

One final result, less directly relevant to parsing, is that derivational morphology on nonwords does not endow them with a grammatical category which is used by the parser. This was observed in Experiment 2 and again in Experiment 8.

The experimental paradigm described in this thesis appears to reflect the operation of the human parser. As such, it is a step in remedying the shortage of experimental tasks which can provide evidence about the parser. The variety of findings just described suggests that the paradigm could profitably be used to address a variety of questions about the human parser.

* * * REFERENCES * * *

Bradley, D.C. Computational distinctions of vocabulary type. Unpublished Ph.D. dissertation, MIT, 1978.

Bever, T.G. The cognitive basis for linguistic structures. In J.R. Hayes (Ed.), Cognition and the Development of Language. New York: John Wiley and Sons, 1970.

Clark, H.H. and Clark, E.V. Psychology and Language. New York: Harcourt, Brace, and Jovanovitch, 1977.

Cutler, A. and Norris, D. Monitoring sentence comprehension. in W.E. Cooper and E.C.T. Walker (Eds.), Sentence Processing. Hillsdale, N.J.: Erlbaum, 1979.

Fay, D.A. Some investigations of grammatical category in performance devices. Unpublished Ph.D. dissertation, U. Texas, Austin, 1975.

Fischler, I. and Bloom, P.A. Automatic and attentional processes in the effects of sentence contexts on word recognition. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 1-20.

Fischler, I. and Bloom, P.A. Rapid processing of the meaning of sentences. Memory & Cognition, 1980, 8, 216-225.

Fodor, J.A., Bever, T.G. and Garrett, M.F. The Psychology of Language. New York: McGraw-Hill, 1974.

Ford, M., Bresnan, J.W. and Kaplan, R.M. A competence-based theory of syntactic closure. In J. Bresnan (ed.), The Mental Representation of Grammatical Relations. Cambridge, Mass.: MIT Press, in press.

Frazier, L. and Fodor, J.D. The sausage machine: A new two-stage parsing model. Cognition, 1978, 6, 291-325.

Gleidman, J. Some experiments on the role of surface structures in memory. Unpublished Ph.D. dissertation, MIT, 1970.

Jackendoff, R. X-bar syntax: A study of phrase structure. Linguistic Inquiry Monograph Two, 1977.

Kaplan, R. Augmented transition networks as psychological models of sentence comprehension. Artificial Intelligence, 1972, 3, 77-100.

Kimball, J. Seven principles of surface structure parsing in natural language. Cognition, 1973, 2, 15-47.

Kleiman, G.M. Sentence frame contexts and lexical decision: Sentence-acceptability and word-relatedness effects. Memory & Cognition, 1980, 8, 336-344.

Kucera, H. and Francis, W.N. Computational Analysis of Present-Day American English. Providence: Brown University Press, 1967.

Limber, J. Syntax and sentence interpretation. In R.J. Wales and E. Walker, New Approaches to Language Mechanisms. New York: North Holland, 1976.

Marcus, M. A Theory of Syntactic Recognition for Natural Language. Cambridge, Mass.: MIT Press, 1980.

Posner, M.I. Cognition: An Introduction. Glenview, Illinois: Scott, Foresman, 1973.

Schuberth, R.E. and Eimas, P.D. Effects of context on the classification of words and nonwords. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 27-36.

Schuberth, R.E., Spoehr, K.T. and Lane, D.M. Effects of stimulus and contextual information on the lexical decision process. Memory & Cognition, 1981, 9, 68-77.

Simpson, G.B. Meaning dominance and semantic context in the processing of lexical ambiguity. Journal of Verbal Learning and Verbal Behavior, 1981, 20, 120-136.

Swinney, D.A. Lexical access during sentence comprehension: (Re)Consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 645-659.

Taft, M. and Forster, K.I. Lexical storage and retrieval of polymorphemic and polysyllabic words. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 607-620.

Thorne, J., Bratley, P., and Dewar, H. The syntactic analysis of English by machine. In D. Michie (ed.), Machine Intelligence, American Elsevier, New York, 1968.

Wanner, E., Kaplan, R., and Shiner, S. Garden paths in relative clauses. Unpublished paper, Harvard University, 1975.

Notes

Note 1. For instance, COMPETE and ENTRIES are a matched verb and noun. ENTRIES has at least two meanings (entries into a building, entries into a chart), but it is acceptable because in all its meanings it is used as a noun.

Note 2. In addition, each subject saw 12 word and 12 nonword items from a pilot experiment not discussed here.

Note 3. For all ANOVAs in this paper, subjects (or items, for the item analysis) are crossed with both the sent and target variables. For RT analyses, each subject's data are mean corrected for his or her mean on all nonpractice items (including fillers). This has no effect on the subject analysis, but reduces variability due to subjects within the item analysis. For the subject analysis, errors are replaced with the subject's mean for that item version, before collapsing over items. For the item analysis, errors are replaced with that item's mean over subjects, before collapsing over subjects.

The error data are also analyzed with ANOVAs. Scattered throughout Experiments 1 to 6 are error effects which approach significance, but none which actually reach significance by the min F' test. The error tendencies are not reported in the text, because they do not make theoretical sense. One generalization can be made. In no case is there evidence of a speed-accuracy tradeoff. To the extent that errors correlate with RTs, slow RTs are associated with high error rates. In Experiments 7 to 9, error effects do make theoretical sense and are reported in the text.

Note 4. Being related by derivational morphology does not guarantee that two words have the same meanings. For example, OCCUPIES is a verb meaning roughly "move into" or "take up", and the noun OCCUPATION has corresponding meanings as in "Germany's occupation of Poland was

brutal." But OCCUPATION has another meaning, "a job", for which the verb OCCUPIES has no corresponding meaning. For this experiment, pairs were chosen which do share all their meanings.

Note 5. The nonfiller items included materials from other experiments not discussed here, including Experiment 10.

Note 6. A direct comparison of the NA cell with the other three cells does not seem appropriate given the requirement that main effects of sent type and target type be factored out. In any case, the direct comparison is also far from significance.

Note 7. Though in this section I uniformly speak of the parser as predicting or not predicting a second NP, the same argument applies if the parser does not predict, and RT effects instead arise from the differential ease of integrating the target word into the syntactic structure determined by the sent.

Note 8. A subject analysis is not feasible for these post hoc comparisons, since the number of items of each type seen by subjects is small. The item analysis is based on an unequal n analysis of variance.

Note 9. The observant reader might notice that the post hoc variable of determiner type affects the pattern of results not only for the determiner sents, baut also for the noun phrase sents. Surely this is bizarre, since the determiner (regardless of its type) does not even appear in the noun phrase sents. The difficulty can be traced to two pairs of target words (CAPABLE/SATISFY and ORDINARY/IDENTIFY) in which the adjective is recognized much more slowly than the verb, regardless of the context and regardless of which of the two sent pairs it occurs with. In the 3 most extreme of the 4 item quadruples in which these target pairs occur, they occur with article determiners rather than possessive pronoun determiners. This is sufficient to produce an apparent effect of determiner type on noun

phrase sents.

Note 10. These matched verb/noun pairs have by now occurred in five experiments: 1) Experiment 2, 2) the baseline condition of Experiment 2 where they were presented without sent contexts, 3) Experiment 4a, 4) Experiment 4b, and 5) Experiment 8. In every case the error rate for the verb targets (TRANSLATES) has been higher than for the noun targets (TRANSLATION). Though never significant in any one experiment, it is virtually certain that this error effect would be significant if the data from all five experiments were combined. The effect could be taken as evidence that inflections (such as -s) must be stripped from a word during lexical access, while words with derivational morphemes (such as -ion) are accessed without the extra step of suffix-stripping.

Note 11. There is a single item with an interaction score 2.2 standard deviations from the mean, in the opposite direction from the predicted effect. If this item is eliminated, the interaction effect is significant for items ($F(1,10) = 6.8$, $p = .026$), but little closer to significance for subjects ($F(1,19) = 1.4$, $p = .26$). It seems overly liberal to eliminate an item that is barely 2 SDs from the mean, based on such a small set of data. Nonetheless, this stray item suggests that with more subjects an RT effect might emerge.

Note 12. As the final copy of this thesis was nearing completion, results from another execution of Experiments 7 through 9 became available. Generally, the theoretically interesting error effects described in Experiments 7 through 9 were not found, but the effects emerged in the corresponding RT comparisons. In Experiment 9, the 54 msec interaction effect was significant for both subjects ($F(1,19) = 8.2$, $p = .0099$) and items ($F(1,19) = 6.3$, $p = .021$) though not for min F'. In Experiment 8 with verb and noun word targets, the interaction of 86 msec was near significance

for subjects ($F(1,19) = 3.8$, $p = .066$) and items ($F(1,11) = 2.2$, $p = .16$). For Experiment 7, the 35 msec main effect of sent type was near significance for both subjects ($F(1,19) = 2.6$, $p = .12$) and items ($F(1,11) = 4.9$, $p = .049$). However, the suggestive effects of right context on RT disappeared entirely in the replication of Experiment 7.

When the paradigm was modified following Experiment 6, effects were observed for errors rather than RTs. In this execution of the same (modified) paradigm, the effects shift back from errors to RTs. This is further evidence that the shift in the informative response measure does not indicate any difference in linguistic processes. The shift is rather due to a superficial variation in task-specific response processes.

* * * APPENDICES * * *

A typical entry in these appendices consists of a pair of sents, followed by the target words and associated pairs of RT means (in milliseconds). The first RT in a pair is for the target after the top sent; the second is for the target after the bottom sent. In some Appendices the number of errors appears, separated from the RT by a " - ". This is not an error percentage; percentages could be derived by dividing by the number of presentations of that item as given in the Appendix caption. For Experiments 7 to 9, the plus sign " + " indicates where an item was continued on the next line. Throughout these appendices, RTs are from the item analysis. RT means in text are an average of item and subject analysis means. Due to differences in error replacement, these values differ slightly.

Appendix A, page 1 of 2
Experiment 1a.

The man spoke but could
Just at the time of
703/691 COMPETE 772/647 ENTRIES

Enforcement of the law could
Nothing can be said against
593/671 LOCATE 554/661 ERRORS

Life in dorms at Harvard would
Bill told me to go without
746/720 REPRESENT 821/627 QUALITIES

Everyone interested in this course can
The cold morning wind blew through
698/731 GOVERN 624/588 HABITS

If your bicycle is stolen, you must
For now, the happy family lives with
724/692 FORMULATE 732/683 BATTERIES

The seller agrees that the buyer shall
Suppose that the criminal escaped from behind
591/626 REALIZE 647/681 DEGREES

In the writings of Karl Marx, we can
There is only one person in charge of
592/667 BRING 536/537 FACTS

Within the framework of this theory, it should
There is an old Norwegian folk tale about
694/675 ORGANIZE 857/764 PASSAGES

When you buy a car, the owner's manual will
In the morning, the dog awakened and stared across
569/589 BELIEVE 644/642 ACTIONS

The essential beauty and order of the universe might
The nervous student closed his eyes and whispered into
676/738 ENTER 763/678 UNITS

If the warhead is detected during its flight, it may
Israel's national airline will open a commercial air route between
608/773 ADMIRE 833/780 OCEANS

In modern Japan elegance is sought in simple things, and this should
A man under the influence of this strange new drug passes into
616/801 CONTINUE 723/767 PURPOSES

Appendix A, page 2 of 2
Experiment 1b.

All people who agree should
The book was lying on
554/911 COMPETE 662/723 ENTRIES

The American political system can
The former general concluded with
628/692 LOCATE 709/573 ERRORS

All the players agreed they would
Nobody knew that the office of
717/810 REPRESENT 703/686 QUALITIES

A law requiring mandatory sentences could
There is an interesting similarity between
692/797 GOVERN 778/683 HABITS

Today, the concern of all people must
New York needs to deal vigorously with
743/775 FORMULATE 713/686 BATTERIES

Soldiers who refused to obey orders could
The typical alcoholic is apparently sincere in
703/710 REALIZE 825/689 DEGREES

The simple satisfaction of physical labor will
The company's decision affects all trains from
713/689 ORGANIZE 944/625 PASSAGES

Dinner at a fine restaurant in Boston would
This incident brings me to the method by
636/739 BRING 603/612 FACTS

The governor also believes that the state should
It is reported that a tornado swept through
669/674 BELIEVE 573/641 ACTIONS

After you have added the three eggs, you must
My grandmother looked at the floor, her hands beneath
636/559 ENTER 766/561 UNITS

There are ways in which both men and women might
The universe that we have learned about spreads out from
663/734 CONTINUE 644/826 PURPOSES

If a ball is set on an inclined plane, it will
Before the child got into bed, he put his shoes near
669/850 ADMIRE 794/762 OCEANS

Appendix B, page 1 of 4
Experiment 2. Word targets, Verb and Noun.

A very large pine forest
A few strange men devote
725/791 EXPELS 693/640 EXPULSION

When an airplane reports to the control tower that its right engine
If the situation at home becomes so bad that the boys accuse
800/757 BETRAYS 834/704 BETRAYAL

The angry woman stormed out, and now her exhausted husband
The great ship broke apart, and now the survivors prevent
674/719 ROTATES 798/731 ROTATION

The professor speaks gently, but a terrified student
The shots have helped, but my allergies include
664/782 DISCOVERS 617/676 DISCOVERY

The tallest building near the hotel
Some remote villages now can enjoy
624/639 ASSURES 709/612 ASSURANCE

I doubt that my small contribution
I doubt that his books characterize
748/702 PERSISTS 678/731 PERSISTENCE

In a move of great significance, Britain's new parliament
To the distress of the opposition, Germany's laws discourage
631/664 ACCELERATES 761/655 ACCELERATION

In this election, the leading candidate
In this serial, the women introduce
513/569 EXISTS 708/661 EXISTENCE

Some fanatics argue that each case of pollution
Most citizens feel that hard work can replenish
578/757 SUPPRESSES 698/758 SUPPRESSION

During inventory, the shopkeeper will find that an expensive camera
When proofreading, even famous authors find that the editors modify
702/780 GROWS 709/649 GROWTH

The crowd near the church indicates that an important funeral
The towers on the horizon indicate that the barriers isolate
688/723 TRANSLATES 710/672 TRANSLATION

Accompanied by some elves, a graceful fairy
Unaware of the risks, some people evade
587/604 REFLECTS 648/634 REFLECTION

Appendix B, page 2 of 4
Experiment 2. Word targets, Adjective and Adverb.

A very large pine forest
A few strange men devote
675/707 INFREQUENT 738/703 INFREQUENTLY

When an airplane reports to the control tower that its right engine
If the situation at home becomes so bad that the boys accuse
710/703 FOOLISH 700/687 FOOLISHLY

The angry woman stormed out, and now her exhausted husband
The great ship broke apart, and now the survivors prevent
722/714 SYSTEMATIC 765/779 SYSTEMATICALLY

The professor speaks gently, but a terrified student
The shots have helped, but my allergies include
710/837 VAST 765/766 VASTLY

The tallest building near the hotel
Some remote villages now can enjoy
650/536 CONTINUOUS 621/610 CONTINUOUSLY

I doubt that my small contribution
I doubt that his books characterize
748/730 RIGID 750/709 RIGIDLY

In a move of great significance, Britain's new parliament
To the distress of the opposition, Germany's laws discourage
723/759 FAVORABLE 813/706 FAVORABLY

In this election, the leading candidate
In this serial, the women introduce
622/547 SIMILAR 601/582 SIMILARLY

Some fanatics argue that each case of pollution
Most citizens feel that hard work can replenish
791/787 PERPETUAL 844/835 PERPETUALLY

During inventory, the shopkeeper will find that an expensive camera
When proofreading, even famous authors find that the editors modify
562/528 EASY 686/672 EASILY

The crowd near the church indicates that an important funeral
The towers on the horizon indicate that the barriers isolate
542/486 HAPPY 580/558 HAPPILY

Accompanied by some elves, a graceful fairy
Unaware of the risks, some people evade
651/589 SERIOUS 595/532 SERIOUSLY

Appendix B, page 3 of 4
Experiment 2. Nonword targets, "Verb" and "Noun".

A geologist believes that the presence of a valuable mineral
A physicist observed that new studies on the nucleus utilize
785/727 IRROLINATES 709/756 IRROLINATION

An estimate of the storage capacity
Her writings on the problem identify
705/705 FABLORATES 638/730 FABLORATION

A single piece of fabric
A strong person will attain
845/787 PROSURIZES 799/744 PROSURIZATION

Either someone is being funny, or the library's rarest volume
Either stolen goods are appearing, or the city's merchants obtain
780/758 DISCULATES 768/743 DISCULATION

The evidence suggests that the younger defendant
Farm statistics suggest that small farmers cultivate
691/658 ISOPIFIES 735/646 ISOPIFICATION

Despite the advances of the enemy, the small nation's defense
Despite a sharp rise in pressure, the vessel's walls contain
809/683 DISTOCIZES 801/651 DISTOCIZATION

History has shown that each legislative session
The custom is that all Republicans replace
719/683 CONJORIFIES 850/834 CONJORIFICATION

While conservatives favor a tax cut, at least one liberal
While adults argue about child abuse, more toys could satisfy
744/759 TERMORATES 799/860 TERMORATION

Experts say that during an earthquake, every single resident
Henry says that in his opinion, these changes simplify
686/662 INTRANIZES 819/715 INTRANIZATION

The press deadline is approaching, but the harried reporter
The promised elections are nearing, but the protests resemble
581/687 LUDIFIES 617/783 LUDIFICATION

A mother can sing a lullaby, but her small infant
The energy shortage can be filled, but we must invent
647/639 BECEPTIFIES 786/702 BECEPTIFICATION

The detective believes that the desk calendar
The engineer believes that the fans generate
690/676 CALSUMIZES 734/722 CALSUMIZATION

Appendix B, page 4 of 4
Experiment 2. Nonword targets, "Adjective" and "Adverb".

A geologist believes that the presence of a valuable mineral
A physicist observed that new studies on the nucleus utilize
796/798 MOMENABLE 793/828 MOMENABLY

An estimate of the storage capacity
Her writings on the problem identify
837/801 MAGNIFUL 888/826 MAGNIFULLY

A single piece of fabric
A strong person will attain
804/716 REPAROUS 786/749 REPAROUSLY

Either someone is being funny, or the library's rarest volume
Either stolen goods are appearing, or the city's merchants obtain
713/668 FOLLAINABLE 655/646 FOLLAINABLY

The evidence suggests that the younger defendant
Farm statistics suggest that small farmers cultivate
674/625 BLURKFUL 580/613 BLURKFULLY

Despite the advances of the enemy, the small nation's defense
Despite a sharp rise in pressure, the vessel's walls contain
854/674 SERICAL 736/787 SERICALLY

History has shown that each legislative session
The custom is that all Republicans replace
698/725 AMPORABLE 756/758 AMPORABLY

While conservatives favor a tax cut, at least one liberal
While adults argue about child abuse, more toys could satisfy
707/767 PRETERNFUL 779/825 PRETERNFULLY

Experts say that during an earthquake, every single resident
Henry says that in his opinion, these changes simplify
682/649 NERCUDICAL 692/690 NERCUDICALLY

The press deadline is approaching, but the harried reporter
The promised elections are nearing, but the protests resemble
720/693 EXARDABLE 804/774 EXARDABLY

A mother can sing a lullaby, but her small infant
The energy shortage can be filled, but we must invent
739/722 INTERREDIOUS 822/803 INTERREDIOUSLY

The detective believes that the desk calendar
The engineer believes that the fans generate
688/703 RELETUAL 788/832 RELETUALLY

Appendix C, page 1 of 1
Experiments 3a and 3b. Top sent of each pair is from 3a, bottom from 3b.
The slash separates results for the two target words.
Errors are out of 10.

Expt 3a	Expt 3b	
Passing the gift shop reminded the man to send his invalid mother		
The man will mail to his invalid mother		
809-4/861-3	724-0/763-0	EXPELS/EXPULSION
Since it was her birthday, the boss said he would get his secretary		
The boss will obtain for his secretary		
775-2/852-3	735-2/741-0	BETRAYS/BETRAYAL
The students decided that they would give their teacher		
The students gave to their teacher		
699-0/787-1	753-1/624-0	ROTATES/ROTATION
When the new dictator granted his old rival		
The dictator should grant to his old rival		
697-2/686-0	740-0/648-0	DISCOVERS/DISCOVERY
After winning the lottery, I decided to buy my daughter		
I purchased for my daughter		
696-1/821-0	799-0/787-1	PERSISTS/PERSISTENCE
The little girl at the zoo tossed the elephant		
The child at the zoo threw to the elephant		
659-4/771-1	724-2/725-0	ASSURES/ASSURANCE
With great arrogance the doctor handed the nurse		
The doctor is handing to the nurse		
740-1/641-0	772-0/702-1	ACCELERATES/ACCELERATION
The druggist asked the clerk to sell the young man		
The company has released to the press		
620-0/659-0	620-0/584-0	EXISTS/EXISTENCE
The exhausted father agreed to read his young son		
The minister was reciting to the dying woman		
781-2/665-0	789-1/626-1	SUPPRESSES/SUPPRESSION
The girl reached across the table to pass her father		
The girl is getting for her father		
688-5/759-4	791-2/618-1	GROWS/GROWTH
In his speech, the manager wished the retiring staff		
The manager has bought for the retiring clerk		
669-0/654-1	660-0/607-0	TRANSLATES/TRANSLATION
The bartender asked the musicians to play the assembled crowd		
The musician gave to the assembled crowd		
671-0/683-1	662-0/568-0	REFLECTS/REFLECTION

Appendix D, page 1 of 2
Experiment 4.

Hanging from the ceiling, the two strips of flypaper seemed to be very
Swarming out of the cavern, the colony of bats led us to appreciate
599/806 FOREIGN 662/714 DOLEIGN

The insides of an automatic banking machine can appear very
The weathered exterior of an old farmhouse can easily hide
576/647 RECENT 694/739 WACENT

A couple argued, and the man appeared barely
The sun shone, and the small puppy chased
531/606 FULL 702/622 FUTE

The chandelier hanging over the table appears to be highly
Water dripping from the ceiling convinced the man to fix
614/667 MEANINGFUL 605/612 NEANARDESH

The character trudging past seemed to be extremely
People passing by persuaded the workmen to modify
568/614 WIDE 664/615 NIRE

The pamphlets on the display rack seemed utterly
The new magazines in the library have caused
582/630 RELIABLE 647/610 REMIADIT

The interesting clock seems very
Your visiting friend should enjoy
641/749 TOLERABLE 612/643 RALORALET

The passengers in the airline terminal seemed to be absolutely
The program's history of large cost overruns serves to complicate
618/620 SUCCESSFUL 823/824 INAPESSIRE

Those two people were completely
Some recent authors have described
647/609 LAZY 646/618 AMER

Young codfish swim so that they remain absolutely
Migrating geese fly so that they can identify
788/686 IRRATIONAL 711/804 INDATACTER [item rejected]

The mechanism of airplane flight remains very
A knowledge of classical music must include
581/600 REALISTIC 663/661 PROMASTIN

Appendix D, page 2 of 2
Experiment 4 (continued).

Physicists say the first few seconds of the universe were very
Journals now say the ozone layer of the atmosphere rarely receives
650/722 DELICATE 658/658 TERICORE

The croutons in the salad appeared to be nearly
The people in that restaurant must try to consume
607/582 STUPID 629/675 SPONAD

Scraping off the old paint turned out to be highly
Turning over the rocky soil should be done to remove
653/807 OBVIOUS 660/701 RAVIRGE

The use of salt in prepared foods has become highly
The makers of baby food in glass jars also package
612/637 MODERN 733/668 LADERN

The old green shirt had become barely
The new plastic tape will actually fasten
648/624 AVAILABLE 621/644 NORSHELON

Appendix E, page 1 of 2
Experiment 5. Word-target data is from first execution, nonword-target data is from second execution.

Adjective	Verb	"Adj" Nonword	"Verb" Nonword
Heart attack victims almost always use a			
Ambulance drivers make sure that the wounds			
786/736 CAPABLE	618/598 SATISFY	627/697 MARDFUL	655/625 CREPENT
Modern food processors should chop a			
The repairman knows that many cracks			
559/664 MAGNETIC	693/643 RESEMBLE	627/730 BERNOUS	691/725 REFRALK
Pumpkins lay all over the ground and some horses ate the			
Eight small goats are in the field and two large cows			
670/682 HARMLESS	748/696 DISPLACE	729/802 DROOMIFUL	715/669 INDRAFUGE
Edward ran quickly when he wanted to reach our			
The little boy tells me that his awful parents			
640/646 NERVOUS	693/746 INSTALL	626/670 ACLIRE	662/698 OBDITE
New fertilizers allow farmers to cultivate their			
Farm experts believe that the hungry beetles			
611/682 CRUCIAL	673/688 EXCLUDE	548/691 REGNAPIC	639/693 INSERBLE
The spy found it necessary to deceive the			
The agent suggested that the most useful men			
714/706 ORDINARY	653/552 IDENTIFY	732/736 UNDORARY	749/713 COMBRIFY
In solving a difficult problem, you must utilize your			
In searching for a paper topic, some careless students			
720/626 TYPICAL	626/656 ISOLATE	770/692 SELFATURE	711/705 EXPRODISH
By saving their pennies, the elderly couple was able to acquire a			
By investing in the money market, most middle or upper income people			
704/757 PREMATURE	675/776 REPLENISH	712/687 TRUMIAL	824/798 REPRUTE
Some people have the tendency to adopt every			
Some religions propose that all of the animals			
546/592 CAREFUL	585/624 PREVENT	688/695 HORKLESS	649/727 DISGLANE
The innovation allows engineers to simplify their			
Progress implies that the old familiar methods			
565/537 USEFUL	609/606 EXCITE	625/717 ORMDISH	650/684 DEVORGE
The curious cat inspected my			
All equations in two variables			
654/617 SELFISH	586/630 DECEIVE	773/678 GRADOM	737/738 RESNIT
The sound of the construction excited the			
One engineer feels that some bad defects			
626/599 BEAUTIFUL	734/672 INTRODUCE	709/723 CACICLE	678/696 ROTRIFY

Appendix E, page 2 of 2
Experiment 5. Word-target data is from second execution, nonword-target data is from first execution.

Adjective	Verb	"Adj" Nonword	"Verb" Nonword
People usually fall asleep when they attempt to memorize the			
Silent movies are often boring when their most exciting scenes			
660/668 NERVOUS	674/687 INSTALL	659/686 ACLIRE	684/771 OBDITE
Careful planning can help you complete your			
Kids think that their brothers and sisters			
589/589 HARMLESS	572/660 DISPLACE	776/789 DROOMIFUL	671/687 INDRAFUGE
Though inexpensive, scotch tape accomplishes the			
Though rarely needed, several pencil sharpeners			
613/600 BEAUTIFUL	735/628 INTRODUCE	699/623 CACICLE	678/675 ROTRIFY
The treatment of diabetes requires that you inject a			
Any careful doctor will find that your sore muscles			
673/740 CAPABLE	632/640 SATISFY	766/692 MARDFUL	739/647 CREPENT
If forced to, I would characterize my			
If they are alarmed, large soldier ants			
624/564 CAREFUL	568/605 PREVENT	734/712 HORKLESS	719/803 DISGLANE
A good botanist can easily identify every			
The newspaper said that the unique fossils			
618/660 CRUCIAL	766/625 EXCLUDE	597/651 REGNAPIC	784/792 INSERBLE
The mechanic told us that we needed to replace our			
The car salesman told me that the four new tires			
687/685 ORDINARY	618/623 IDENTIFY	753/767 UNDORARY	673/663 COMBRIFY
The biology students found they had to dissect their			
The scientists believe that their long and hard efforts			
648/598 TYPICAL	679/668 ISOLATE	735/734 SELFATURE	726/741 EXPRODISH
By skillfully moving his pieces, the chess player created a			
By attending to the cards, most top notch bridge players			
794/788 PREMATURE	808/695 REPLENISH	717/666 TRUMIAL	850/819 REPRUTE
The confident man bought a			
The piles of empty boxes			
590/627 MAGNETIC	754/681 RESEMBLE	697/603 BERNOUS	593/696 REFRALK
An ocean of cars surrounded the			
All wooden chairs with four legs			
579/615 USEFUL	617/642 EXCITE	587/660 ORMDISH	810/763 DEVORGE
The pinball machine's silver ball hit a			
The merchant claims that all his profits			
604/661 SELFISH	814/707 DECEIVE	638/717 GRADOM	684/661 RESNIT

Appendix F, page 1 of 2
Experiments 6a and 6b. The sents shown are from Experiment 6a. To obtain 6b sents, remove first word of each sent; also substitute any bracketed word for preceding word.

6a Noun	6b Noun		6a Nonword	6b Nonword	
Among the many reasons[duties] which the Among things[Crimes] which the law might					
633/592	651/656	UNIT	632/619	582/572	UNAL
Because the large mob of angry students that the Because the number of commuters that the train must					
590/715	589/692	SITUATION	760/801	851/702	SITUAMERN
Because large sets of numbers which every Because simple tasks which every child can					
671/688	656/681	BASEMENT	720/651	623/652	TASEBORM
Since most of the animals that my Since many people that my dog can					
561/573	605/651	WINDOW	633/591	613/640	WENGOW
At the street corner which a At stores which[that] a person would					
584/694	556/603	QUALITY	622/595	656/638	RUALORM
Near the giant palace that our Near streets that our kids will					
550/562	650/562	ERROR	665/666	627/586	ELTOR
After seeing[The] several[interesting] monuments which our After investigations which our study should					
608/561	540/575	ACTION	628/653	660/612	ARDILT
Following an interview with the spectator that the Following questioning of people that the judge may					
664/710	645/725	PASSAGE	590/638	640/624	PASLANE
During the recent military exercises which every During the speech which every congressman must					
788/709	803/781	OCEAN	597/642	692/637	OCARL
Against the wall of the building which a Against stacks of boxes which a man could					
576/556	526/586	FACT	665/662	613/613	TALT
Besides the pieces of fruit which the Besides the straw which the horse could					
622/593	629/549	ENTRY	576/623	595/601	ENDOT

Appendix F, page 2 of 2
Experiments 6a and 6b (continued).

6a Noun	6b Noun		6a Nonword	6b Nonword	
Despite some[Those] advances in the field which every					
Despite lots of facts which every person would					
594/607	668/586	HABIT	652/615	587/699	RADIN
Under the cold and windy skies that the					
Under stacks of papers that the men will					
567/668	633/612	BATTERY	668/722	647/706	TAMMERY
Before any new class that a					
Before oil that a tanker can					
657/673	663/676	DEGREE	668/708	657/667	ROGREE
In a little cafe which my					
In houses which my wife can					
580/572	649/618	PURPOSE	716/811	706/733	PURPANG
Though the morning paper which the					
Though things which the girl should					
623/582	650/631	RIBBON	597/627	594/647	BIMBON
Since the piece of cake that every					
Since the dish that every guest will					
604/556	647/655	EVENT	634/637	624/606	EDORN
While many of the large insects that a					
While places to visit that a tourist can					
724/744	716/680	LEDGE	707/680	748/656	DELGE
Behind the old oaken door which the					
Behind the window which the bed will					
822/735	727/782	COWARD	660/677	666/658	COLURD
Of the great composers that our					
Of movies that our town might					
543/563	562/634	KITCHEN	641/622	644/614	KARTHEN

Appendix G, page 1 of 2
Experiment 7. Word Targets. Errors are out of 5.

Ence Type
Det /Verb

658-0/690-0 The man spoke but could FORGET
761-0/739-0 Just at the time of FORGET
a/be sentence only with difficulty

545-0/538-0 In the writings of Karl Marx, we can BRING
629-1/505-0 There is only one person in charge of BRING
a/be brilliant mind

650-0/661-0 Elegance is found in simple things, and this should CONTINUE
580-0/568-0 A man under the influence of alcohol passes into CONTINUE
the/have view + of all of us

611-0/755-0 Enforcement of the law could LOCATE
782-1/756-0 Nothing can be said against LOCATE
the/have problem effectively

677-0/622-0 Within the framework of this theory, it should ORGANIZE
630-0/732-1 There is an old Norwegian folk tale about ORGANIZE
the/be resemblance + to earlier results

831-1/714-0 If the warhead is detected during its flight, it may ADMIRE
898-2/758-1 The national airline will open a new air route between ADMIRE
a/be radio + signal

711-0/741-0 Life in dorms at Harvard would REPRESENT
831-0/899-0 Bill told me to go without REPRESENT
a/be difficult dilemma for me

596-0/787-0 The seller agrees that the buyer shall REALIZE
642-0/757-0 Suppose that the criminal escaped from behind REALIZE
the/have mortgage

555-0/582-0 When you buy a car, the owner's manual will BELIEVE
559-0/704-0 In the morning, the dog awakened and stared across BELIEVE
a/have section on + tires

668-0/881-0 Everyone interested in this course can GOVERN
722-0/807-0 The cold morning wind blew through GOVERN
the/be reading if they try

651-0/692-1 If your bicycle is stolen, you must FORMULATE
795-1/785-0 For now, the happy family lives with FORMULATE
a/have complaint with the + police

601-0/622-0 The essential beauty and order of the universe might ENTER
645-0/603-0 The nervous student closed his eyes and whispered into ENTER
the/have world's + problems someday

Appendix G, page 2 of 2
Experiment 7. Nonword Targets. Bracketed word is part of verb ence, but not determiner ence.

Ence Type
Det /Verb

688-0/738-0 A law requiring mandatory sentences could LESAN
697-1/700-1 There is an interesting similarity between LESAN
[be] the crime problem + substantially

741-0/802-0 Today, the concern of all people must NEBORTATE
698-1/705-0 New York needs to deal vigorously with NEBORTATE
[have] the people suffering + in Brazil

739-0/774-0 The governor also believes that the state should RELOSTEND
762-0/728-0 It is reported that a tornado swept through RELOSTEND
[have] a ban on + fishing

776-0/724-0 All the players agreed they would SMEATE
716-0/722-1 Nobody knew that the office of SMEATE
[be] a heroic effort

993-0/815-2 Soldiers who refused to obey orders could CONSURE
895-1/819-0 The typical alcoholic is apparently sincere in CONSURE
[have] a speedy + execution

717-0/749-1 After you have added the three eggs, you must GONITATE
714-1/664-0 My grandmother looked at the floor, her hands beneath GONITATE
[have] the + mixture for ten minutes

612-0/769-0 The American political system can FOSTOW
773-0/736-0 The former general concluded with FOSTOW
[have] the problems only at + great cost

790-1/770-0 The simple satisfaction of physical labor will SHROATEN
662-1/809-0 The company's decision affects all trains from SHROATEN
[be] a person's + health

793-0/812-0 If a ball is set on an inclined plane, it will PARDORM
682-0/699-0 Before the child got into bed, he put his shoes near PARDORM
[be] the bottom + eventually

820-0/776-1 All people who agree should PRELOCT
723-0/842-0 The book was lying on PRELOCT
[have] a friend to join us

760-0/655-0 Dinner at a fine restaurant in Boston would RELIMP
663-0/652-1 This incident brings me to the method by RELIMP
[be] the effort required + to get there

727-0/766-0 There are ways in which both men and women might ESTIP
557-0/661-0 The universe that we have learned about spreads out from ESTIP
[be] a better + attitude

Appendix H, page 1 of 8
Experiment 2. Word targets, Verb and Noun. Errors are out of 5.

A very large pine forest
775-0 EXPELS lumbering operations
778-1 EXPULSION to their trees
A few strange men devote
739-1 EXPELS lumbering operations
832-0 EXPULSION to their social lives

When an airplane reports to the control tower that its right engine
726-0 BETRAYS + odd noises, help is required
839-0 BETRAYAL + of smoking, help is required
If the situation at home becomes so bad that the boys accuse
808-1 BETRAYS + odd noises, help is required
781-0 BETRAYAL + of smoking, help is required

The angry woman stormed out, and now her exhausted husband
923-2 ROTATES his + weary bones
782-3 ROTATION of the + fire
The great ship broke apart, and now the survivors prevent
768-1 ROTATES his + weary bones
851-0 ROTATION of the + fire

The professor speaks gently, but a terrified student
806-0 DISCOVERS scribbled + notes quietly
726-0 DISCOVERY of old + fears
The shots have helped, but my allergies include
824-1 DISCOVERS scribbled + notes quietly
674-0 DISCOVERY of old + houses

The tallest building near the hotel
687-1 ASSURES major repairs
684-0 ASSURANCE of a new kind
Some remote villages now can enjoy
756-0 ASSURES major repairs
647-0 ASSURANCE of a new kind

I doubt that my small contribution
764-0 PERSISTS serious problems at all
748-1 PERSISTENCE in a novel way
I doubt that his books characterize
747-1 PERSISTS serious problems at all
830-0 PERSISTENCE in a novel way

Appendix H, page 2 of 8
Experiment 2. Word targets, Verb and Noun (continued).

In a move of great significance, Britain's new parliament
848-0 ACCELERATES + rising inflation in a new light
761-0 ACCELERATION + near military bases
To the distress of the opposition, Germany's laws discourage
782-0 ACCELERATES + rising inflation in a new light
706-1 ACCELERATION + near military bases

In this election, the leading candidate
634-0 EXISTS his positions carefully
700-0 EXISTENCE for later development
In this serial, the women introduce
612-0 EXISTS his positions carefully
620-0 EXISTENCE for later development

Some fanatics argue that each case of pollution
715-0 SUPPRESSES needed + economic growth
776-0 SUPPRESSION of a clean + environment
Most citizens feel that hard work can replenish
840-1 SUPPRESSES needed + economic growth
803-0 SUPPRESSION of a happier + era

During inventory, the shopkeeper will find that an expensive camera
823-0 GROWS + significant shoplifting
789-0 GROWTH + in unusual ways
When proofreading, even famous authors find that the editors modify
888-1 GROWS + significant shoplifting
697-0 GROWTH + in unusual ways

The crowd near the church indicates that an important funeral
709-0 TRANSLATES + traffic patterns seriously
820-1 TRANSLATION + from something they want
The towers on the horizon indicate that the barriers isolate
636-1 TRANSLATES + traffic patterns seriously
658-0 TRANSLATION + from something they want

Accompanied by some elves, a graceful fairy
580-1 REFLECTS tiny vegetables
799-0 REFLECTION of their tiny bodies
Unaware of the risks, some people evade
711-0 REFLECTS tiny vegetables
722-0 REFLECTION of their unearned income

Appendix H, page 3 of 8
Experiment 2. Word targets, Adjective and Adverb. Errors are out of 5.

A very large pine forest
746-0 INFREQUENT and large trees are unusual
860-1 INFREQUENTLY and gently comforts the mind
A few strange men devote
849-0 INFREQUENT and large sums to their yachts
768-0 INFREQUENTLY and gently their efforts to odd + causes

When an airplane reports to the control tower that its right engine
863-1 FOOLISH + and unfamiliar noises, something is wrong
889-2 FOOLISHLY + and ominously fails, something is wrong
If the situation at home becomes so bad that the boys accuse
680-0 FOOLISH + and unfamiliar people, something is wrong
761-0 FOOLISHLY + and ominously, something is wrong

The angry woman stormed out, and now her exhausted husband
880-0 SYSTEMATIC + or careful thought
742-0 SYSTEMATICALLY + or unthinkingly is calling the police
The great ship broke apart, and now the survivors prevent
758-2 SYSTEMATIC + or careful thought
731-0 SYSTEMATICALLY + or unthinkingly the best course

The professor speaks gently, but a terrified student
786-3 VAST or unknown is + leaving
912-1 VASTLY or truly is + leaving
The shots have helped, but my allergies include
822-0 VAST or unknown things
743-1 VASTLY or truly unknown + things

The tallest building near the hotel
698-0 CONTINUOUS but plentiful servants
654-0 CONTINUOUSLY but rarely blocks the sun
Some remote villages now can enjoy
693-0 CONTINUOUS but plentiful food
673-0 CONTINUOUSLY but rarely supplies + delivered by truck

I doubt that my small contribution
895-1 RIGID yet interesting help to anyone
750-2 RIGIDLY yet truly helps anyone
I doubt that his books characterize
799-0 RIGID yet interesting subjects
749-0 RIGIDLY yet truly differing subjects

Appendix H, page 4 of 8
Experiment 2. Word targets, Adjective and Adverb (continued).

In a move of great significance, Britain's new parliament
549-3 FAVORABLE and + harsh legislation
860-1 FAVORABLY and + harshly has cracked down
To the distress of the opposition, Germany's laws discourage
854-0 FAVORABLE and + harsh punishment
932-1 FAVORABLY and + harshly worded statements

In this election, the leading candidate
690-1 SIMILAR or meaningless statements
701-0 SIMILARLY or freely avoids + questions
In this serial, the women introduce
630-0 SIMILAR or meaningless topics only
782-0 SIMILARLY or freely constructed + themes

Some fanatics argue that each case of pollution
795-1 PERPETUAL but + lasting consequences
745-1 PERPETUALLY but + unnecessarily hurts us
Most citizens feel that hard work can replenish
811-0 PERPETUAL but + lasting consequences
779-0 PERPETUALLY but + unnecessarily destroyed values

During inventory, the shopkeeper will find that an expensive camera
687-0 EASY + yet correct item to stock
632-0 EASILY + yet truly sells well
When proofreading, even famous authors find that the editors modify
688-0 EASY + yet correct passages
791-0 EASILY + yet truly correct passages

The crowd near the church indicates that an important funeral
539-0 HAPPY and + somber occasion
625-0 HAPPILY and truly draws people
The towers on the horizon indicate that the barriers isolate
521-0 HAPPY and + somber prisoners indeed
693-0 HAPPILY and + truly unwilling people

Accompanied by some elves, a graceful fairy
673-0 SERIOUS or necessary work
830-0 SERIOUSLY or quickly dashed + away
Unaware of the risks, some people evade
578-0 SERIOUS or necessary work
673-0 SERIOUSLY or quickly changing + regulations

Appendix H, page 5 of 8
Experiment 2. Nonword targets, "Verb" and "Noun". Errors are out of 5.

A geologist believes that the presence of a valuable mineral
752-0 IRROLINATES + some cases of mining
781-0 IRROLINATION + of a serious danger
A physicist observed that new studies on the nucleus utilize
824-0 IRROLINATES + some cases of mining
810-0 IRROLINATION + of a hydrogen atom

An estimate of the storage capacity
696-0 FABLORATES this report completely
779-0 FABLORATION by greedy investors
Her writings on the problem identify
806-0 FABLORATES this report completely
857-0 FABLORATION by greedy investors

A single piece of fabric
967-0 PROSURIZES interesting possibilities
809-0 PROSURIZATION about the frame
A strong person will attain
789-0 PROSURIZES interesting possibilities
816-0 PROSURIZATION about the age of twenty

Either someone is being funny, or the library's rarest volume
946-2 DISCULATES + a clear forgery
692-0 DISCULATION + from generous donors
Either stolen goods are appearing, or the city's merchants obtain
784-0 DISCULATES + a clear forgery
715-0 DISCULATION + from generous donors

The evidence suggests that the younger defendant
907-0 ISOPIFIES innocent + people
704-1 ISOPIFICATION with + great efficiency
Farm statistics suggest that small farmers cultivate
678-0 ISOPIFIES innocent + people
709-0 ISOPIFICATION with + great efficiency

Despite the advances of the enemy, the small nation's defense
694-1 DISTOCIZES + strong forces
867-0 DISTOCIZATION + of an army
Despite a sharp rise in pressure, the vessel's walls contain
701-0 DISTOCIZES + strong forces
687-0 DISTOCIZATION + of an inert gas

Appendix H, page 6 of 8
Experiment 2. Nonword targets, "Verb" and "Noun" (continued).

History has shown that each legislative session
782-0 CONJORIFIES more people + than one would think
770-1 CONJORIFICATION without + problems
The custom is that all Republicans replace
798-1 CONJORIFIES more people + than one would think
840-0 CONJORIFICATION without + problems

While conservatives favor a tax cut, at least one liberal
853-0 TERMORATES + very different ideas
744-0 TERMORATION + in the schools
While adults argue about child abuse, more toys could satisfy
989-1 TERMORATES + very different ideas
968-2 TERMORATION + in the schools

Experts say that during an earthquake, every single resident
771-0 INTRANIZES + an enormous risk
786-1 INTRANIZATION + at every level
Henry says that in his opinion, these changes simplify
730-1 INTRANIZES + an enormous risk
858-0 INTRANIZATION + at every level

The press deadline is approaching, but the harried reporter
668-0 LUDIFIES + good stories in any case
963-0 LUDIFICATION + behind the scenes
The promised elections are nearing, but the protests resemble
745-0 LUDIFIES + good stories in any case
818-0 LUDIFICATION + behind the scenes

A mother can sing a lullaby, but her small infant
728-1 BECEPTIFIES little + benefit from it
878-0 BECEPTIFICATION on + every possible occasion
The energy shortage can be filled, but we must invent
678-0 BECEPTIFIES little + benefit from it
806-0 BECEPTIFICATION on + every possible occasion

The detective believes that the desk calendar
753-0 CALSUMIZES some important + and subtle things
871-0 CALSUMIZATION with traces + of a poison
The engineer believes that the fans generate
765-0 CALSUMIZES some important + and subtle things
730-1 CALSUMIZATION with traces + of a poison

Appendix H, page 7 of 8
Experiment 2. Nonword targets, "Adjective" and "Adverb".
Errors are out of 5.

A geologist believes that the presence of a valuable mineral
907-1 MOMENABLE + and immediate mining
847-1 MOMENABLY + and completely overrides other factors
A physicist observed that new studies on the nucleus utilize
1022-0 MOMENABLE + and immediate freezing
889-0 MOMENABLY + and completely cooled samples

An estimate of the storage capacity
1017-2 MAGNIFUL but false calculations
982-1 MAGNIFULLY but unscrupulously + allays fears which are justified
Her writings on the problem identify
890-0 MAGNIFUL but false difficulties
1048-1 MAGNIFULLY but unscrupulously + collected funds

A single piece of fabric
828-0 REPAROUS or perfect solution
835-0 REPAROUSLY or separately does the trick
A strong person will attain
822-0 REPAROUS or perfect abilities by age twenty
773-1 REPAROUSLY or separately needed accounts by + middle age

Either someone is being funny, or the library's rarest volume
803-0 FOLLAINABLE + yet valuable photos
701-0 FOLLAINABLY + yet mysteriously has reappeared
Either stolen goods are appearing, or the city's merchants obtain
701-0 FOLLAINABLE + yet valuable items very cheaply
714-0 FOLLAINABLY + yet mysteriously transported items

The evidence suggests that the younger defendant
652-0 BLURKFUL and costly + vacations abroad
643-1 BLURKFULLY and + deliberately withheld information
Farm statistics suggest that small farmers cultivate
766-0 BLURKFUL and costly + varieties that don't sell
783-0 BLURKFULLY and + deliberately ill-chosen crops

Despite the advances of the enemy, the small nation's defense
689-0 SERICAL but + heavy casualties
745-1 SERICALLY but + dangerously is holding out
Despite a sharp rise in pressure, the vessel's walls contain
866-0 SERICAL but + heavy metals
852-0 SERICALLY but + dangerously hot materials

Appendix H, page 8 of 8
Experiment 2. Nonword targets, "Adjective" and "Adverb" (continued).

History has shown that each legislative session
798-0 AMPORABLE or difficult + matters for later
814-0 AMPORABLY or slowly + addresses only a few problems
The custom is that all Republicans replace
741-0 AMPORABLE or difficult + opponents
840-0 AMPORABLY or slowly + their opponents

While conservatives favor a tax cut, at least one liberal
781-0 PRETERNFUL + and morose in his outlook
927-0 PRETERNFULLY + and seriously proposed an increase
While adults argue about child abuse, more toys could satisfy
693-2 PRETERNFUL + and morose children
1134-3 PRETERNFULLY + and seriously disturbed children

Experts say that during an earthquake, every single resident
627-0 NERCUDICAL + but troublesome symptoms
846-0 NERCUDICALLY + but quickly will adjust
Henry says that in his opinion, these changes simplify
638-0 NERCUDICAL + but troublesome problems
665-0 NERCUDICALLY + but quickly arising difficulties

The press deadline is approaching, but the harried reporter
736-0 EXARDABLE or + confused sources
857-0 EXARDABLY or + carelessly is writing
The promised elections are nearing, but the protests resemble
823-0 EXARDABLE or + confused debates
821-1 EXARDABLY or + carelessly organized parties

A mother can sing a lullaby, but her small infant
807-0 INTERREDIOUS and + clever will not sleep
919-0 INTERREDIOUSLY and + effectively rebels
The energy shortage can be filled, but we must invent
871-0 INTERREDIOUS and + clever solutions over time
927-0 INTERREDIOUSLY and + effectively managed technologies

The detective believes that the desk calendar
738-0 RELETUAL or unhealthy + chemical residues
896-0 RELETUALLY or reliably + indicates the criminal's plans
The engineer believes that the fans generate
780-0 RELETUAL or unhealthy + chemical residues
1041-1 RELETUALLY or reliably + chemical residues

Appendix I, page 1 of 3
Experiment 9. Target word or nonword appeared in the blank.
Errors are out of 10.

Noun		Nonword	
It is widely known that the people on _____ of this drug will get sick John fears that the pilot and crew might _____ and all will be lost			
754-0/645-1	UNITS	720-0/783-0	UNALS
At this time of year, lucky people with _____ in their yards are + in good shape On Labor Day, local shops and vendors must _____ or face financial + difficulty			
682-0/665-0	SITUATIONS	776-0/740-0	SITUAMERNS
In actuality, the number of _____ at hand is quite small If encouraged, the child can _____ and help himself immensely			
810-1/721-2	BASEMENTS	719-0/741-0	TASEBORMS
If you listen to them carefully, the new songs on _____ of known artists + are meaningless If a person trains them, small birds in cages can _____ very well			
691-0/684-2	WINDOWS	647-0/638-0	WENGOWS
Near here, the street corner in _____ by the park is a shambles Near the statue, a person would _____ or risk being mugged			
602-1/674-0	QUALITIES	702-0/696-1	RUALTORMS
KI believe that the giant palace near _____ in the shopping district + is now closed The woman feels that the children will _____ and then be very cross			
793-0/650-0	ERRORS	657-1/659-0	ELTORS
In this city, interesting monuments of _____ on horseback are curiously + abundant In this document, our assumptions should _____ or it is not worth writing			
622-0/633-1	ACTIONS	670-0/712-1	ARDILTS

Appendix I, page 2 of 3
Experiment 9 (continued).

Noun		Nonword	
The interviewer said that the spectator in _____ by the fence was hurt			
The judge decreed that the tenants may _____ or withhold rent, if they + prefer			
784-0/815-2	PASSAGES	717-0/781-0	PASLANES
During the recent military exercises, the difficulty with _____ on tanks has + become apparent			
During this speech on militarization, the hecklers must _____ or we will + look foolish			
767-2/795-3	OCEANS	694-0/698-1	OCARLS
On the wall of the ancient building, a person in _____ of crayon has + inscribed his opinions			
On the pieces of paper, someone in a hurry could _____ without needing + any more information			
575-0/593-0	FACTS	704-0/855-0	TALTS
It is believed that the fruit of _____ in cities tastes awful			
Fred thinks that the red plums could _____ if they are not kept cold			
711-1/842-0	ENTRIES	718-0/708-0	ENDOTES
Taking everything into account, the area of inquiry in _____ with humans + does not look promising			
In summary, the most important result of waiting would _____ more time + to think			
679-0/630-0	HABITS	700-0/712-1	RADINS
In autumn, the cold and windy skies of _____ without high humidity are + the rule			
In this case, the piles of papers will _____ but do no real harm			
756-0/829-0	BATTERIES	796-0/731-0	TAMMERIES
At first, the new class of _____ for data storage are incomprehensible			
Right here, the crude oil can _____ and await its eventual end use			
587-0/797-1	DEGREES	816-1/765-1	ROGREES

Appendix I, page 3 of 3
Experiment 9 (continued).

Noun	Nonword
The man finds that the little cafe beside _____ near the town line + has the best deal on breakfast	
I fear that these new cheap houses can _____ and then will be vulnerable + to termite damage	
603-0/680-0 PURPOSES	689-0/733-0 PURPANGS
As Sally walked in, the morning paper under _____ in the hall fell	
Seeing as Friday is nearing, the girl should _____ or risk embarrassing + herself	
709-0/708-1 RIBBONS	761-0/683-1 BIMBONS
Because of the coffee, the piece of cake with _____ of lemon sauce tastes + awful	
Because of the guests, the dish of candies will _____ on the living room + table	
594-0/645-0 EVENTS	670-0/709-0 EDORNS
There are still books claiming that the insects in _____ with wings + cannot bite	
There are some people saying that cancer victims can _____ when they are in + pain	
639-1/844-4 LEDGES	686-0/628-0 DELGES
If refinished, the old oaken door in _____ by the stairway will be lovely	
If it is repaired, the window will _____ and open properly	
807-0/810-4 COWARDS	771-0/829-0 COLURDS
The great composer said that the piece of _____ in our attic is of + some interest	
The town manager said that the movies might _____ but won't show any + longer	
570-0/629-0 KITCHENS	744-0/627-0 KARTHENS

Appendix J, page 1 of 1
Experiment 10.

Pronoun		Preposition		WH-Word		Quantifier	
Enforcement of the law could							
Nothing can be said against							
593/606	IT	664/678	TOWARD	632/570	WHO	570/609	SOME
The man spoke but could							
Just at the time of							
631/582	IT	715/726	TOWARD	726/650	WHO	654/603	SOME
The seller agrees that the buyer shall							
Suppose that the criminal escaped from behind							
587/560	HIM	729/656	AMONG	566/610	WHICH	509/567	ALL
If your bicycle is stolen, you must							
For now, the happy family lives with							
680/617	HIM	714/676	AMONG	627/609	WHICH	558/531	ALL
The essential beauty and order of the universe might							
The nervous student closed his eyes and whispered into							
593/592	US	566/597	WITH	521/568	WHERE	574/603	EITHER
When you buy a car, the owner's manual will							
In the morning, the dog awakened and stared across							
597/478	US	518/514	WITH	547/528	WHERE	581/591	EITHER
Life in dorms at Harvard would							
Bill told me to go without							
630/559	THEM	689/648	FROM	608/617	HOW	627/566	BOTH
Everyone interested in this course can							
The cold morning wind blew through							
567/546	THEM	585/595	FROM	662/690	HOW	645/677	BOTH
In the writings of Karl Marx, we can							
There is only one person in charge of							
645/563	ME	624/566	UPON	528/509	WHEN	548/508	EVERY
Within the framework of this theory, it should							
TKhere is an old Norwegian folk tale about							
562/519	ME	608/575	UPON	533/597	WHEN	573/525	EVERY
If the warhead is detected during its flight, it may							
Israel's national airline will open a commercial air route between							
565/671	YOU	629/687	AGAINST	536/597	WHY	624/575	ANY
In modern Japan elegance is sought in simple things, and this should							
A man under the influence of this strange new drug passes into							
511/525	YOU	560/536	AGAINST	624/607	WHY	564/570	ANY
A man under the iNfluence of this strange new drug passes into							
333/333	YOU	333/333	AGAINST	333/333	WHY	333/333	ANY