

**Experimental Measures of Syntactic and Semantic
Effects in Real-time Sentence Processes**

by
Janell Marie Schweickert

A.B., Swarthmore College, 1978

**Submitted to the Department of Psychology
in partial fulfillment of the requirements
for the degree of**

Doctor of Philosophy in Psychology

**at the
Massachusetts Institute of Technology
May 1985**

© Massachusetts Institute of Technology

Signature of Author *Janell Marie Schweickert*

**Department of Psychology
May 15, 1985**

Certified by *Merrill F. Garrett*

**Merrill F. Garrett
Thesis Supervisor**

Accepted by *Richard Held*

**Richard Held
Chairman, Department of Psychology**



Table of Contents

Chapter 1: Introduction	7
Chapter 2: Literature Review	15
Chapter 3: Development of Experimental Materials	38
3.1 General Assumptions and Overview of Experiments	38
3.2 Development of Materials	46
3.3 Experiment 1: Timed Cloze	54
Chapter 4: Lexical Decision Experiments	61
4.1 Experiment 2: Lexical Decision with a Slow Presentation Rate	61
4.2 Experiment 3: Fast Lexical Decision	75
4.3 Comparison of Experiments 2 & 3	83
Chapter 5: Naming Experiments	90
5.1 Experiment 4: Slow Name	90
5.2 Experiment 5: Fast Name	96
5.3 Both Name Experiments	102
Chapter 6: General Discussion and Summary	106
Appendix A: Stimulus Materials	121
A.1 Experimental Trials	121
A.2 Filler Trials	129
A.3 Practice and Lead-in Trials	131
A.4 Nonword Trials	135
Appendix B: Miscellaneous Analyses	142
B.1 Analyses over all the Experimental Trials with the Factors of Context, Plausibility, and Grammaticality	142
B.2 Analyses over Implausible Targets	149
B.3 Analyses with Version and Kind as Factors	153
B.4 Analyses of the Nineteen Discarde Subjects from Experiment 3: Fast Lexical Decision	159
Appendix C: Ratings as a Function of Kind of Ungrammaticality	163
C.1 Experiment 2: Slow Lexical Decision	164
C.2 Experiment 3: Fast Lexical Decision	165
C.3 Experiment 4: Slow Name	166
C.4 Experiment 5: Fast Name	167
References	168

List of Figures

Figure 1:	One version of an autonomy model of language processing	10
Figure 2:	Alternative models of language comprehension that preserve independence of semantic and syntactic processing	42
Figure 3:	Percentage of subjects' ratings for each sentence type. Experiment 2: Slow Lexical Decision	68
Figure 4:	Percentage of subjects' ratings for each sentence type. Experiment 3: Fast Lexical Decision	80
Figure 5:	Percentage of subjects' ratings for each sentence type. Experiment 4: Slow Name	93
Figure 6:	Percentage of subjects' ratings for each sentence type. Experiment 5: Fast Name	99

EXPERIMENTAL MEASURES OF SYNTACTIC AND SEMANTIC
EFFECTS IN REAL-TIME SENTENCE PROCESSES

by

Janell Marie Schweickert

Submitted to the Department of Psychology in May 1985
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Psychology

This thesis explored the relationship between semantic and syntactic processing during sentence comprehension. The experiments tested two opposing theories of linguistic processing, theories which differ in their conceptualization of processing stages and information exchange during language comprehension. One class of models, interaction models, postulates unrestricted information flow throughout the processing system. The other class of models, autonomy models, compartmentalizes processing. Independent subprocessors are dedicated to the formation of intermediate linguistic representations, and the flow of information among them is highly constrained.

The presence of processing stages during language comprehension was tested using Sternberg's (1969) additive factors logic. Sternberg's logic provides a framework for the interpretation of experimental results when multiple contextual variables are present in the stimulus materials. If the variables do affect independent processing stages, then when they are multiply present, their effects will be statistically independent. But if the experimental variables affect the processing of a common stage, then their effects will be empirically evident as a statistical interaction. In five experiments, discourse and syntactic context were independently varied. The results of the experiments refuted the claim for processing interaction, and provided support for autonomy models of sentence comprehension; when the variables of discourse and syntactic context yielded strong main effects, those effects were always additive. There was never any indication of an interaction.

The experiments relied on a set of stimulus materials that allowed for the independent variation of three variables: the strength of the discourse context, the grammatical character of a critical portion of the discourse and the plausibility of a critical word in the discourse. Extensive discussion of the development of the materials is presented in Chapter 3. For the experiments, the stimuli were presented visually, using rapid serial visual presentation (RSVP). Experiment 1 used a "Timed Cloze" paradigm to validate assumptions about the availability of contextual information at a critical point in the discourse. In this paradigm, subjects read a short discourse in which one of the nouns had been replaced with a string of Xs. Subjects were instructed to say outloud the first noun that came to mind when they encountered the string of Xs. Robust effects of the contextual variables were found, setting the stage for the use of those variables in subsequent experiments.

Experiment 2 used a lexical decision task. The critical nouns that had been Xed out of the stimulus sets for the Timed Cloze task, were used as lexical decision targets. The experiment tested the combined effects of the contextual variables on the lexical decisions. The variables of discourse and syntactic context showed strong main effects on lexical decision times, with no statistical interaction. This finding supports autonomy models.

Experiment 3 used a lexical decision task, but varied the rate of presentation of the stimulus materials in an attempt to dissociate semantic and syntactic processes in time. There was no strong evidence for such a dissociation. An increase of the presentation rates from 200 msec/word to 117 msec/word did not affect the pattern of results, replicating the additivity of discourse and syntactic factors found in Experiment 2.

Experiment 4 used a naming task in another attempt to dissociate semantic and syntactic processes and thus provide converging evidence that these processes correspond to distinct levels in the language processor. Previous studies had found that lexical decision and naming are differentially sensitive to contextual variables. Evidence for dissociability was found; the effect of discourse context on naming time was absent while the effect of grammatical context remained strong.

Experiment 5 used a naming task along with an increase in presentation rate corresponding to that used in Experiment 3. Surprisingly, the effect of discourse context reappeared, demonstrating that naming tasks can be sensitive to discourse variables. As in the lexical decision experiments, the combined effects of discourse and syntactic context were additive, providing further support for autonomy models of language processing.

ACKNOWLEDGMENTS

I am very grateful to Merrill Garrett, my advisor and mentor. His intellect and optimism were guiding lights throughout my graduate career. The creative and elegant aspects of the research are signs of his handiwork. Lorraine Tyler deserves much credit for her role in the conception of the experimental series, and the ingenious design of the stimulus format. She offered valuable advice during the piloting of Experiments 2 and 3, and helped to shape the mixed use of on- and off-line measures. I was very fortunate in having both Merrill and Lolly collaborating on this project.

I would also like to thank the members of my thesis committee; Molly Potter, Steve Pinker and Martin Chodorow, for their help and advice during the last year of the project. Molly Potter, who was my secondary advisor, contributed a great deal to my education by providing many stimulating discussions over the years. She is a whiz at playing devil's advocate and by doing so she helped me to formulate and articulate many ideas.

My training at MIT encompassed a number of research projects. I would like to acknowledge the excellent training I received while working with Ed Walker, Joan Ryder, David Swinney, Celia Jakubowitz and Janet Sherman. I thank Judy Kroll, my undergraduate thesis advisor, for the wonderful time I had working with her. It was she who instilled in me the desire to pursue scientific questions, as well as enthusiasm about the actual process of experimentation.

Most of the students and staff deserve thanks for their support. Carmen Egido, Mara Blofelds, Tom Zeffiro, Carla Kirmani, Peter Gordon, John Rubin, Kathy Murphy, Jim Hodgson, Allan Doyle, Sean True, Bart Wright, Kevin Ackley and Janet Sherman are only a few of the many whose interactions provided me with academic advice and off-hours diversions.

My non-MIT friends contributed in important ways, too, from being subjects, to cooking dinners. I shall always be grateful for all that Miriam Kadansky contributed during the early years of my graduate work; she shouldered many of the household responsibilities in order that I might attend to my research. My job at Marcella's, and the friends I made there, provided me with an important outlet from the academic world.

I am grateful for financial support through a National Institute of General Medical Sciences Traineeship awarded to the Department of Psychology, and to the Department for Teaching Assistantships.

Finally, I wish to thank my family for their enthusiasm and support. Special thanks go to my husband, Eric Wassermann, who helped out in every way that anyone could.

Chapter 1

Introduction

One of the most fascinating aspects of language behavior is how effortlessly language is produced and, especially, understood. The more that is learned about the complexity of language, the more amazing our facility with it appears. The usual perception of reading text or listening to someone talk is that understanding is instantaneous. During comprehension, the listeners' conscious thoughts are directed toward what is being communicated and the relationship between the message content and established beliefs. Rarely is a listener's attention focussed on the form of what is said, the structure in which the content is couched. And if the form is noticed, it is generally focussed on pragmatically, in ways that help to enrich the message content (e.g. She's speaking in very simple sentences; she's being patronizing.).

Perhaps it is this invisibility of form that has motivated theories regarding sentence structure as an epiphenomenon of semantic processes (e.g. Riesbeck and Schank, 1978). Yet the majority of psycholinguists who work at the levels of sentence or discourse processing agree that structural analyses are not derived from semantic analyses. Indeed, it is acknowledged that structure must sometimes (if not most times) determine the course of semantic analysis. This is easily shown by

contrasting sentences with the same lexical content yet different structural relations (Scholes, 1978):

He fed her dog the biscuits.
He fed her the dog biscuits.

As well, it is this dependency on structure that gives language its true creativity. That is, we are able to produce, and more importantly, to understand sentences that express relations contrary to those observed in the world. If it ever became necessary to express the thought "The peanut butter ate the man" it could be done, with the assurance that people would understand what was being said. (They may not believe it, but that is an entirely different matter.)

Although there is general agreement that the two domains or knowledge sources for structural and interpretive processing are distinct, there is disagreement about the architecture of the language processing mechanism. In particular, the disagreement centers on whether or not there are computational levels and/or mechanisms specific to each distinct knowledge source.

The autonomy model, argued for by Forster (1979), among others, claims independence of informational types in processing. The theory states that the different types of information are handled by distinct processing mechanisms. For example, a semantic processing component and a syntactic processing component are both postulated to exist in the language processor, each with characteristic input and internal structure. The output of one component serves as the input for another. This informational flow is unidirectional. A schematization of one, very

simple autonomy model is given in Figure 1. By this formulation of the autonomy model, lexical recognition is not affected by syntactic or integrative processing; the determination of the syntactic structure of the input is not affected by integrative processing. Each component is encapsulated in such a way that only its output can affect another component, and only by becoming its input. The internal workings of any component are insulated from those of any other component. In its strongest form the model states that there are absolutely no "top-down" effects in processing. Information that is derived from the separate processing mechanisms does not feed back (down) to affect the processing of incoming material in other components. Efficiency in this model is seen as a product of having initial special-purpose processors which, due to their limited scope, quickly churn out specific representations. The syntactic component builds upon the lexical output and then transfers its output to the integrative level which can assimilate the new information to all that the listener knows. Presumably this later stage is quite time-consuming and efficiency is achieved by having it work on already developed and specified representations.

In autonomy models, the distinction between inter- and intra-component action is important. Intralevel computations are allowed to interact with one another. Thus semantic priming¹ (Meyer and

1. In the semantic priming paradigm, the effects of semantically related single word contexts are studied. For example, subjects are faster to decide that "nurse" is a real English word (as opposed to a nonword such as "blart") when it follows a related word such as "doctor" than when it follows an unrelated word such as "tractor". The related "prime" word is said to facilitate lexical processing.

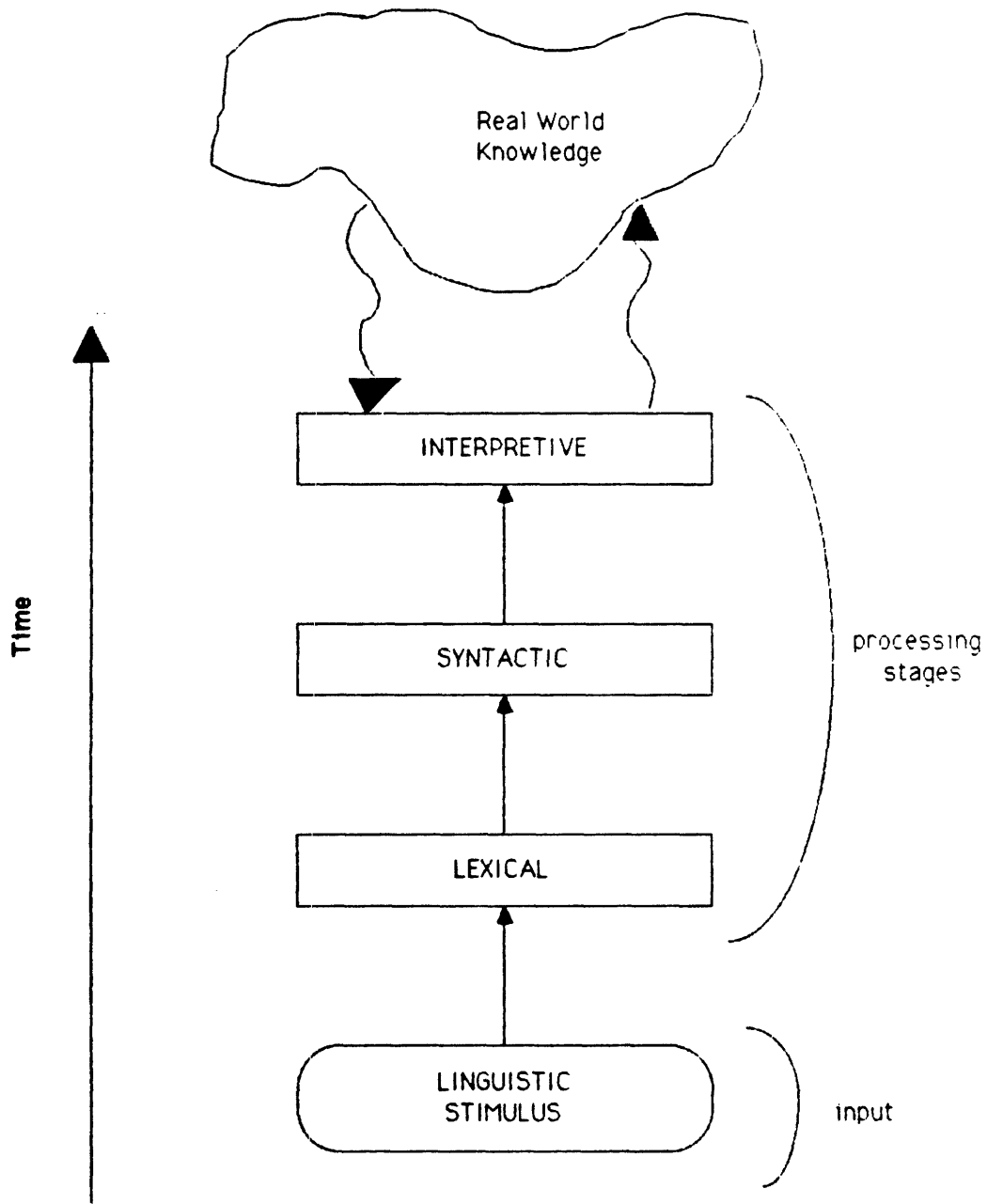


Figure 1: One version of an autonomy model of language processing.

Schwaneveldt, 1976) was at first taken to be a top-down effect; the observed facilitation was believed to be evidence that semantic processing can intervene in lexical processing. Most theories of word recognition postulate that whenever the lexical entry for a word is contacted (this process is referred to as "lexical access"), neighboring lexical entries are affected. These effects may appear in reaction time experiments as facilitation (or inhibition depending on the model) in ways that reflect the organization of the lexicon. Since all the effects are assumed to be a consequence of the actual structure of the mental lexicon, intralevel interaction is occurring. What is not allowed to happen within the framework of autonomy models is an alteration in the process of lexical recognition resulting from intervention of higher-order knowledge sources. That is, sentence context cannot influence lexical access above and beyond any lexically-based interaction. Of course, much of the computational work accomplished in the course of sentence processing is the integration of the information present in the stimulus with what the listener knows or believes. In the autonomy model this integrative process occurs high-up in the language mechanism, at a level where there are no constraints on information exchange. At this level, sentence context, and the way it relates to the listener's world-view, can very well affect the ease with which individual lexical items (or phrases) can be incorporated into the developing meaning representation. Notice that only those processes operating on meanings of lexical items or phrases are influenced by context; the processing steps that determine those meanings are not

influenced by the effects of conceptual processing.²

The contrasting view, developed most completely by Maralen-Wilson and Tyler (1980), is that lexical, structural and conceptual information are not processed independently. Processing is assumed to take place in an "interactive" manner whereby any process may affect any other process. In some formulations of the model there are no subprocessors, no "components" within the language processing mechanism. Different informational types are acknowledged, yet each is not accorded its own computationally unique processor and it is viewed as misleading to think in terms of levels of analysis. Processing levels are seen as artifacts borrowed from linguistic theory with no relation to psychological computations.

According to interaction models, at any point during comprehension, any step that is completed must influence the processing of newer, incoming material whether that new material is of the same informational type or not. The basis of this model is the belief that the language processor is capable of using various types of information to differing degrees, depending on what is available from the context. Thus, words presented in isolation would be recognized solely from bottom-up (stimulus-driven) information. Words presented within a context need not be recognized by purely bottom-up analyses, since the syntactic and

2. All of this assumes that a normal course of analysis is proceeding. All bets are off if extreme circumstances are forced upon the system. In this case, many types of strategies may be used in order to achieve comprehension, strategies which bear little resemblance to normal processing.

semantic information contained in the context would help to narrow the field of possible word candidates. This would reduce the system's reliance on bottom-up analysis and facilitate processing. The language processor is able to take advantage (indeed, must take advantage) of all available information at any level of processing and there are no limitations on when information can be used.

Studying the structure of the language processor has theoretical implications beyond the field of psycholinguistics; it may yield information on the structure of other cognitive systems and relate to a theory of mind in general. That is, a general question that may be addressed is: Are different mental processes (visual perception, memory and language, to name a few) subserved by computationally distinct mechanisms, or can all of mental (cognitive) activity be described by one set of processing algorithms? (See, e.g., Fodor, 1983.) Although this question of "mental modularity" is normally asked about the relationship between cognitive domains, knowledge of the architecture within a cognitive domain may constrain possible models of mental architecture. If it can be shown that there are distinct types of processing in language corresponding to linguistic informational types, then it appears more plausible that language processing differs from the processing in visual perception. In other words, if constraints on information exchange can be demonstrated in one cognitive domain, then the whole notion of constraints becomes more tenable. If subprocessors within a domain, dedicated toward achieving a common representation, are unable to interact with one another, then it seems unlikely that subprocessors within separate domains, working on different representations, would be

able to communicate freely with one another.

There is no necessary connection, however, between the question of mental modularity and that of autonomous levels of language processing. For example, it could very well be that within the language processing system, there are no constraints on information exchange, and yet there might still be a distinct computational boundary between language and other cognitive domains.

The next chapter is a review of recent experimental evidence addressing the question of how the language processor is structured.

Chapter 2

Literature Review

In general, the interaction/autonomy debate has centered on two issues: (1) What is the nature of the intermediate representations developed during the course of language analysis? and (2) Are language processing operations invariant across variations in contextual information? It is the second issue that is most often addressed empirically. Some of the language operations that researchers have explored are word identification, determination of syntactic relations, and semantic/conceptual interpretation. These operations correspond to basic stages in autonomy models. It is generally assumed that word identification must proceed prior to syntactic processing, in part because grammatical form-class information must be accessed from a lexical entry. In addition, autonomy models assume that the determination of structural relations proceeds without reference to the developing conceptual representation, and that in fluent language understanding the structural representation is a basic component of the integrative process. The autonomy models impose constraints on information flow during processing, thereby limiting contextual effects. In contrast, the interaction models do not (or at least not to the same extent). Experimental work has generally been based upon the autonomy

models' delineation of levels, motivated either by a desire to support an autonomy model or in order to refute one. The comments that follow are organized in terms of claims for or against constraints on lexical and syntactic processes.

Evidence for Constraints on Lexical Processing

Perhaps because word recognition has an extensive experimental history with accepted findings and widely used paradigms, it has been the focus of most of the experiments designed to contrast autonomy and interaction models. Most early experimental work seemed to support interactive models of language comprehension: Words were recognized with increasing facility as context became more supportive (e.g. Tulving and Gold, 1963; Morton, 1969). One of the first contrasting proposals (Rubenstein, Garfield, and Millikan, 1970) conceived word recognition as spanning several types or levels of processing. These proposals were elaborated and modified in a number of subsequent proposals by Forster (1976, 1979, 1981). These levels included the encoding of the stimulus, the contacting of the mental lexicon, the subsequent retrieval of lexical information, the integration of the retrieved information with prior context as well as with general conceptual structures, and the realization by the comprehender that a word has been perceived. Presumably context could be acting at some or all of these levels to facilitate word recognition. Whereas earlier work sought to determine whether context affected word recognition, current research is focussed

on determining the locus of contextual effects. The various models of language comprehension place the effects of context at different levels of processing.

Autonomy models predict that lexical access is not affected by syntactic or integrative processing. Note that although this formulation "encapsulates" the lexical component in that stimulus information is the only information allowed to be input to the lexical processor, it does not entail that lexical access be a purely data-driven process. That is, as discussed in the introduction, not only stimulus properties but also properties resulting from the organization of the mental lexicon, properties internal to the "capsule", can influence lexical access. Thus, when contextual information is said to influence the course of lexical access, it matters what the character of that information is, whether it is information internal or external to the lexical component. The most extreme interaction models hold that there are no constraints on lexical access; any form of information is allowed to alter the course of access (Norton, 1969; Grosjean, 1980; Marslen-Wilson, 1975). This lack of constraint has been variously modified in other interaction models, most notably the cohort models (eg. Marslen-Wilson and Welsh, 1978). In cohort versions of interaction models, a pool of candidates is initially specified by form-based analysis. However, in fluent language comprehension, those candidates which do not mesh with contextual information are dropped from further analysis. (Contextual information can range from semantic priming effects to discourse effects to the comprehender's set of beliefs about the world.) Stimulus information continues to be monitored and used, but after the initial stage it no

longer has priority in directing lexical access. Recognition occurs when only one lexical candidate remains in the cohort pool. The implication of this model is that reliance on stimulus information is modulated by variations in the contextual information available to the system. All interaction models predict that as syntactic and discourse-level context become stronger (more highly constraining), lexical access will rely less upon stimulus information. In contrast, autonomy models stress that the access process is invariant across all types of context (with the exception of lexically based contextual effects).

Interestingly, most of the research that supports an interactionist position has been conducted with auditory stimuli. The main tasks employed have been shadowing (Jakimik, 1979; Marslen-Wilson and Welsh, 1978), gating (Cotton and Grosjean, 1984; Grosjean, 1980; Salasoo and Pisoni, 1985; Tyler and Wessels, 1983), mispronunciation detection (Cole, 1973; Marslen-Wilson and Welsh, 1978), and various monitoring tasks: phoneme (Blank and Foss, 1978; Foss, Cirilo and Blank, 1979; Morton and Long, 1976; Swinney and Hakes, 1976), rhyme and word monitoring (Marslen-Wilson & Tyler, 1980). Exceptions are work done within visual priming/lexical decision paradigms (Goodman, McClelland and Gibbs, 1981; Lukatela, Kostic, Feldman and Turvey, 1983). In contrast, the research that supports an autonomy position has been conducted with purely visual (Forster, 1979, 1981) or cross-modal stimuli (Onifer and Swinney, 1981; Seidenberg, Tanenhaus, Leiman and Bienkowski, 1982; Swinney, 1979; Tanenhaus, Leiman and Seidenberg, 1979; Tanenhaus and Donnenwerth-Nolan, 1984). There has also been support from auditory tasks, (Cairns, Cowart and Jablon, 1981; Foss, 1982). It is possible, therefore, that some of

the controversy over constraints on early stages of lexical processing could be settled through a better understanding of modality differences and how modality interacts with psycholinguistic tasks. However, it is unlikely that modality affects lexical processing much beyond the stage of stimulus encoding, at levels where there are major differences between the autonomy and interaction models.

Some of the earliest research demonstrating contextual effects on lexical access was done with the semantic priming paradigm (e.g. Meyer and Schwanefeldt, 1971; Meyer, Schwanefeldt and Ruddy, 1975; Warren, 1972; 1977) as discussed in the Introduction. In this paradigm it was shown that naming, lexical decisions, and same-different judgments were facilitated when the target words followed related single word contexts. The reaction time effect was hypothesized to be a consequence of the structure of the mental lexicon and most of the subsequent work within the paradigm was carried out to determine the lexicon's storage and access characteristics. One important line of inquiry has concerned lexical ambiguity (see Simpson, 1984, for a review). The original focus of the research was on the storage of ambiguous words: Were multiple meanings associated with one lexical node, or did each meaning have a unique lexical entry? In time, however, the focus of the research turned to lexical access, the process of contacting semantic representations: Could context alter the number of meanings accessed when an ambiguous word is encountered? Or: Could context alter the order in which meanings are accessed? (e.g., Schwanefeldt, Meyer, and Becker, 1976; Simpson, 1981; See also: Conrad, 1974; Onifer and Swinney, 1981; Seidenberg, Tanenhaus, Leiman, and Bienkowski, 1982; Swinney, 1979; Tanenhaus,

Leiman, and Seidenberg, 1979). From debate on this question many distinctions in lexical processing have been made (see also Seidenberg, Waters, Sanders and Langer, 1984). Thus, the access of an entry may be distinguished from the retrieval of information from an entry and that process may in turn be distinguished from the incorporation of lexical information into a meaning representation. Another complicating distinction is that made between the computations that occur without conscious awareness and those that occur with it (e.g., Posner and Snyder, 1975; Neely, 1976; 1977). All questions about lexical processes must be distinguished from questions about knowing that lexical contact has been made. From the perspective of these distinctions, the issue of contextual influences on comprehension is significantly altered. Context might be affecting the course of lexical access, but it might also be operating at some time after access has occurred. Such a "postaccess" context effect would not be in violation of an autonomy model since context would have its effect at the level of lexical output. It is therefore imperative to determine at what level context is influencing lexical recognition.

The clearest empirical work to date supports a postaccess context effect in the processing of ambiguous words. For example, in what is called the "cross-modal priming paradigm", Swinney (1979) (see also Onifer and Swinney, 1981; Seidenberg, Tanenhaus, Leiman and Bienkowski, 1982; Tanenhaus, Leiman and Seidenberg, 1979)³ presented subjects with

3. See Glucksberg, Kreuz and Rho, 1985 for an alternative explanation of the cross modal priming results.

an auditory sentence context that strongly biased the interpretation of an ambiguous word contained in the sentence. As subjects heard the sentence contexts, they were asked to make lexical decisions to words that were simultaneously presented visually (one decision per sentence). From the subjects' point of view, the lexical decision targets bore no relationship to the words in the sentence. However, on a small proportion of trials, subjects made lexical decisions to words that were related to either the contextually biased or the contextually inappropriate meaning of the ambiguous word. Swinney found that if the lexical decision targets were presented at the offset of the ambiguous word, the responses to both of the related words were facilitated compared to unrelated control words. In other words, the contextual bias had no effect. However, if the lexical decision targets were presented a few syllables after the ambiguous word, then only the responses to the contextually biased meaning were facilitated. These studies suggest that, regardless of context, all meanings of an ambiguous word are initially accessed. Context then quickly acts upon the pool of retrieved meanings, causing only the relevant meaning of the word to be activated within 750 msec. Context does have an effect, but only after lexical access has occurred and therefore it appears as though it cannot determine the course of lexical access. This interpretation is clearly in accordance with the constraints on information flow imposed by the autonomy models.

Another theoretical distinction derived from work on lexical processing was discussed in Chapter 1: A taxonomy of contextual types has been developed. It is acknowledged that single word contexts might not

affect lexical processing in the same way that a discourse context might. It matters a great deal, theoretically, what type of context is capable of influencing lexical access (see Forster, 1981, for a discussion of this issue). Specifically, it is critical whether the context effect is internal to the lexical module. Schubert, Spoehr and Lane (1981) tested the effects of semantic congruity, stimulus quality and word frequency on lexical decision times to words following incomplete sentence contexts. They found that the pattern of results they obtained with sentence contexts differed from the results that other experimenters had obtained with single word contexts. Schubert, et al. (1981) proposed that priming with sentence contexts does not affect the lexical decision process in the same way that priming with single word contexts does. Foss (1982) demonstrated that the time course of semantic priming in word lists differed radically from the time course of the priming effects elicited by whole sentences. Sentence context had a much longer lasting facilitatory effect than did single word semantic priming. It appears therefore that the context effects demonstrated with isolated words provide an inadequate base for models of contextual effects during language processing.

The discussions initiated by the work on lexical processing have also created a climate for questioning experimental paradigms. Lexical tasks (such as naming, lexical decision, phoneme or other kinds of monitoring) are probably sensitive to extralexical influences. Patterns of response in such tasks may reflect, for example, the integration of a word's meaning into the developing interpretation of a sentence, rather than the process of lexical access (c.f. Cairns, 1983; Foss, 1982). Results might

also be task dependent. Forster (1981) examined contradictory studies of sentence context effects on word recognition (i.e., Fischler and Bloom, 1979; Schubert and Eimas, 1977; Stanovich and West, 1979, 1981; West and Stanovich, 1978). Taking the results from these studies together with his own, Forster argued for a principled distinction between lexical decision and naming tasks based on the presence of a postaccess decision or "checking" process in lexical decision. In lexical decision, one must execute arbitrary motor responses (i.e., keypresses) based upon metalinguistic judgments (i.e., Is this a word?). In naming tasks, subjects read words out loud. This type of response is naturally determined by information encoded directly into the lexical entry. Forster proposes that before the subjects signal their lexical decisions, they use all available information to assess their decisions. This assessment occurs more rapidly when the context is highly informative. There is no such decision process in naming, and therefore naming is not as affected by contextual strength as is lexical decision.

Making the same distinction but from a different perspective, Seidenberg, et al. (1984) reviewed a number of findings that support interaction models (i.e., Goodman, et al., 1981; Koriat, 1981; Lukatela, et al., 1983). All of these results had been obtained with a lexical decision task. Seidenberg, et al. (1984) postulated that the "signal detection" character of the lexical decision task (i.e. "Has a word occurred?") leads subjects to use contextual information in postaccess processing. Seidenberg, et al. (1984) presented the same contextual variables in a non-signal detection paradigm (i.e. a naming task), and found that most of the effects of context disappeared. They argued that

the contextual effects of semantic and associative priming, which remained stable throughout the paradigm shifts, are the result of intralexical processing. Their interpretation explains the pattern of results of a number of studies across the two paradigms and is consistent with the autonomy thesis. (See also: Balota and Chumbley, 1984; Seidenberg and Tanenhaus, in press; West and Stanovich, 1982).

The search for "pure" measures of lexical processing has led to the development of many interesting paradigms. Proponents of interaction models have devised tasks that are perceptual in nature and are therefore purported to be sensitive to subjects' reliance on form-based information. Most of these paradigms employ auditory stimuli. Yet ingenious as these tasks are, they are not immune to effects of subject strategies. For example, in monitoring tasks (Blank & Foss, 1978; Foss, Cirilo and Blank, 1979; Marslen-Wilson & Tyler, 1980; Morton and Long, 1976; Swinney and Hakes, 1976), subjects are given the target before they hear the stimulus. Thus they may be able to enlist predictive mechanisms, which on many accounts, are thought to have little or no place in normal language comprehension. In gating tasks (Cotton and Grosjean, 1984; Grosjean, 1980; Salasoo and Pisoni, 1985; Tyler and Wessels, 1983), subjects are asked to guess what the target is on the basis of a foreshortened auditory stimulus; but presumably we normally consciously guess at only a fraction of the words we hear. Similar considerations may be raised with regard to shadowing tasks (Jakimik, 1979; Marslen-Wilson and Welsh, 1978). In all of these tasks it is quite plausible that more effortful, perhaps conscious, inferential processing may contribute to the determination of subjects' responses. It is

possible that context is affecting these more central computations rather than the actual process of lexical access.

In another paradigm, Tyler and Wessels (1983) have concluded that, counter to the claims of a strong interaction model, syntactic context does not exert a powerful influence over the process of narrowing down a cohort field prior to lexical recognition. Instead, syntactic information in their amended interaction model is envisioned as having a unique role. Tyler and Wessels (1983) used the gating paradigm in which gated tokens were presented in contexts that varied in their semantic as well as syntactic constraints. They found that although both semantic and syntactic context reduced subjects' reliance on phonetic information in their determination of the target word, their effects did not interact. That is, semantic and syntactic constraints exerted independent effects on subjects' task performance. Tyler and Wessels concluded "Perhaps one moral of the present study is that an information source that is important in one aspect of the system does not need to be important in every aspect." (p. 418) This seems to be a move to differentiate the process of language analysis. In an unstructured interaction model, all information is potentially relevant to the language processor. In the Tyler and Wessels (1983) interaction model, distinctions are drawn between different types of information and their relevance to different "aspects" of language comprehension. In an autonomy model, "aspects" becomes "stages" and each information source is assumed to be relevant to only a particular stage; it is not allowed to influence to other stages in the analysis. The differentiated interaction model incorporates an autonomy feature, that of distinct

types of analyses within the language processor.

A coherent metric of the processing demands that different tasks impose has not yet emerged, even though there is experimental evidence supporting a decomposition of tasks into components that are variously affected by context. There is, moreover, no consensus on the range of strategies available to subjects as they participate in the experiments or the role of those strategies in everyday language use. Without a firm base from which to evaluate the various experimental paradigms, it is difficult to arrive at definitive conclusions about the presence and nature of constraints on lexical processing.

Evidence for Constraints on Syntactic Processing

A test of the autonomy and interaction models in the domain of syntactic processing centers on the issue of whether syntactic representations for linguistic input are computed directly from the meaning of the input. This research will be reviewed in more detail than the literature on lexical processing as it is directly related to the experiments reported in this thesis.

One well-known study providing support for a model in which semantic properties of the input affect the course of syntactic analysis is by Slobin (1966). That work concerns what has become known as "reversibility", namely, the extent to which the noun phrases in a sentence are semantically and pragmatically constrained to particular

syntactic roles. Slobin used a picture verification task. In this task subjects were read a sentence and then were shown a picture. They were asked to decide whether the sentence was a true or false description of the picture and their time to respond was recorded. The following stimulus examples represent a subset of the materials used in the experiment:

- Active Reversible: The dog is chasing the cat.
- Passive reversible: The cat is being chased by the dog.
- Active Nonreversible: The girl is watering the flowers.
- Passive Nonreversible: The flowers are being watered by the
 girl.

Nonreversible sentences were easier to process (i.e. responses were faster) than reversible sentences. Moreover, and most importantly, Slobin observed what was taken to be an interaction between syntactic and semantic variables. It was generally accepted that active sentences are easier to process than passive sentences; but Slobin found that this was only true of reversible sentences. When the roles of subject and object were not interchangeable on semantic grounds (as in the nonreversible sentences), the difference in active and passive sentences disappeared. The conclusion drawn was that part of the processing difficulty of passive sentences (making them harder than active sentences) was due to determining which noun was the subject and which was the object. When the sentence expressed nonreversible relations, the syntactic operations were simplified (or eliminated) because the semantic constraints could guide functional assignments. Herriot (1969) reached a similar conclusion using pragmatically nonreversible sentences (eg. The doctor

treated the patient.) in a task where subjects had to listen to a sentence and repeat the logical subject and the logical object. This conclusion is counter to the claim of autonomy models that syntactic analysis of sentences is computed without respect to the meaning of the sentence.

The work by Slobin (1966) and Herriot (1969) was taken to be definitive until Forster and Olbrei (1973) conducted a "meaning latency" task in which subjects were asked to decide whether a given sequence of words forms a meaningful sentence. Using Slobin-type stimuli with this paradigm and equating the stimuli for plausibility⁴ they found that active sentences were judged to be meaningful more quickly than were passive sentences for both reversible and nonreversible sentences, contrary to Slobin's (1966) result. They also employed another technique, sentence report, after Forster and Ryder (1971). Sentences were shown to subjects using Rapid Serial Visual Presentation (RSVP), a word-by-word presentation of a sentence in which subjects are given very little time to process each word. Forster and Olbrei (1973) found that subjects were better able to report active than passive sentences, and this was true for both reversible and nonreversible sentences. They concluded that syntactic analysis was accomplished without reference to the meaning of the sentence. (See Forster and Ryder (1971) for a similar argument.)

Forster (1979) reported further work with the RSVP report paradigm.

4. The stimuli were actually equated for predictability more than plausibility.

The stimuli in this experiment were sentences that varied in their semantic quality (that is, they were either plausible or anomalous). Minor changes in word order were introduced into some of the stimuli, yielding ungrammatical sentences (e.g. "The speech was..." would become "The was speech..."). These order changes had similar effects across the semantic conditions, suggesting that the same syntactic processes were applied to plausible and anomalous sentences.

An experiment by Tyler and Marslen-Wilson (1977) suggested a different conclusion. Subjects in this task heard a sentence fragment ending in a structurally ambiguous phrase and were then asked to name a visually presented probe word that was either an appropriate or an inappropriate continuation of the fragment. After the naming response subjects were asked to indicate whether or not the probe word was a good continuation of the sentence fragment. Tyler and Marslen-Wilson argued that such a task would enable them to tap the immediate perceptual processing of the sentence as opposed to previous tasks that typically probed at relatively late stages in the process, such as after the sentence had been completely understood. An example of their stimuli follows:

(a) If you know how to handle sudden gusts of wind,
flying kites...

(b) As they glide gracefully over the city, flying kites...

Verb probe: IS [appropriate for (a); inappropriate for (b)]

ARE [inappropriate for (a); appropriate for (b)]

If, in violation of autonomy models, previous semantic context is

able to dictate which syntactic description is given to a structurally ambiguous phrase, then subjects should have an easier time processing the appropriate verbs relative to the inappropriate verbs. If, however, semantic context does not affect the parsing of an ambiguous phrase, then subjects should not differentiate between appropriate and inappropriate completions. Results supported the view that syntactic and semantic processes interact during sentence comprehension; appropriate verbs were responded to more quickly than were inappropriate verbs.

Recently, however, Cowart and Cairns (1984) produced evidence that the Tyler and Marslen-Wilson (1977) result is not due to a semantic bias effect. Rather, they noted that in most (24 out of 28) of the semantic contexts used by Tyler and Marslen-Wilson (1977) to bias the structural interpretation in favor of a plural verb, the pronoun "they" occurred in the biasing clause. Cowart and Cairns argued that the need to find a referent for "they" affected the processing of the ambiguous phrase by guiding its interpretation as a plural noun phrase. They rewrote the Tyler and Marslen-Wilson materials to eliminate this "pronoun bias". For example:

- (c) As the birds soar gracefully above the field,
flying kites...

Putting these materials, along with stimuli containing the pronouns, in the Tyler and Marslen-Wilson (1977) paradigm, they replicated the Tyler and Marslen-Wilson results only in the sentences containing "they". In the sentences containing a lexical nounphrase as a replacement for "they", there was no difference in naming time to the two verb types.

This result alone does not rule out an interpretation based on interactive processes. Rather what it demonstrates is that reference relationships can influence syntactic processing. It is known that reference relations are themselves sensitive to nonsyntactic and nonlinguistic information as well as syntactic information. If it could be shown that any type of information affects the initial processing of the ambiguous phrase in order to arrive at the semantically or pragmatically coherent reference relation, then this would contradict models based on assumptions of the autonomy of processing. On the other hand, if it could be shown that only syntactic information can constrain the reference assignment, then this counters the assumption that all relevant information must immediately and interactively be employed in the syntactic analysis.

In order to test these two sets of predictions, Cowart and Cairns (1984) constructed three types of biasing context: structural, selectional and pragmatic. In each type of context, there was information which either permitted or disallowed anaphoric assignment between the ambiguous phrase and the contextual pronoun. In the examples below, the (a) cases correspond to meaningful interpretations based on the reference relation and the (b) cases correspond to anomalous interpretations if "they" is taken to be coreferential with the ambiguous phrase.

Structural (a) If they want to save money, visiting uncles ...
(b) If they want to believe that, visiting uncles ...

Selectional (a) Even though they use very little oil, frying eggs ...
(b) Even though they eat very little oil, frying eggs ...

- Pragmatic (a) Whenever they smile during the procedure, charming babies ...
(b) Whenever they lecture during the procedure, charming babies ...

Predictions were as follows: According to autonomy models, only processes operating within the syntactic component should be able to block the anomalous coreference relation, and only the Pronoun Bias Effect in the structural condition will be eliminated. According to interaction models, all contextual types should be able to block the anomalous anaphor assignment and thus eliminate the Pronoun Bias Effect in all three contextual conditions. The results supported the predictions of the autonomy model. A Pronoun Bias Effect was observed in both the Selectional condition (selectional constraints are treated as syntactic phenomena in some linguistic theories and semantic ones in others) and Pragmatic context conditions. There was no effect in the Structural condition.⁵

Townsend and Bever (1982) also provided an alternative explanation for the Tyler and Marslen-Wilson (1977) evidence. They proposed a different type of confound in the Tyler and Marslen-Wilson stimulus materials than the one explored by Cowart and Cairns (1984). Townsend and Bever (1982) noted that the presence of plural and singular morphological cues in the biasing context varied regularly with the appropriateness of plural and singular probe verbs. They also predicted that the size of the Tyler and Marslen-Wilson effect would vary with such

5. There was no effect perhaps due to an elevation of responses in a control condition. However, the important result is the failure of selectional and pragmatic context to override the Pronoun Bias Effect.

factors as the clause type of the biasing context (e.g. causal or adversative). In an experiment controlling for these and other factors they found patterns counter to the predictions of interaction models. The model they support is a modified autonomy model in which interactions occur at specified points ("natural units") at each level of analysis. That is, interactions are constrained; they are not continuous.

Using a similar paradigm but a completely different set of materials, Tyler and Marslen-Wilson (1982) continued to explore the relationship between discourse context and structural assignment; in this case the assignment of surface pronouns to logical functions. They predicted, on the basis of an interaction model, that discourse constraints could affect the process of assignment of anaphor. An example of the stimuli is:

Auditory stimulus: As Philip was walking back from the shop, he saw an old woman trip and fall flat on her face.

- a) Philip ran towards ...
- b) Running towards ...

Visual Probe Words: Appropriate target: HER
Inappropriate target: HIM

Subjects' task was to listen to the discourse; at the offset of the sentence fragment a visual probe word appeared that subjects were to name as quickly as possible. They found that for both fragment types a) and b) naming times were faster for the appropriate target (e.g. HER) than they were for the inappropriate target (e.g. HIM). In fact, naming times for the appropriate target were equally fast for both of the sentence fragment types even though fragment type a) contained an overt

reference to the agent in the surface structure whereas in fragment type b) the agent role must be inferred. Tyler and Marslen-Wilson (1982) suggested that the response times for the inappropriate target (HIM in the example above) were elevated both because the prior assignment of Philip to the agent role for "running" must be rejected as incoherent once the pronoun HIM is encountered, and because when Phillip is assigned to the target pronoun, the only candidate available for assignment to the implicit agent of "running" in sentence fragment b) is the old woman, and this interpretation is implausible. Such reasoning suggests that not only does context direct lexical role assignment, but it also supports early prediction of what the referent is for a missing argument.

However, Fodor and Garrett (personal communication) argued that the experimental design employed by Tyler and Marslen-Wilson (1982) did not permit a test of an alternative explanation of the robust HIM/HER effect. Fodor and Garrett suggested that subjects do not predict the referent for the missing argument, but rather make an assignment after overt lexical pronouns have been anaphorically linked. These two positions generate different predictions for the following stimulus continuations:

Auditory stimulus: As Philip was walking back from the shop, he saw an old woman trip and fall flat on her face.

- a) Philip ran towards ...
- b) Running towards ...

Visual Probe Words: HER / HIM / THEM

On the interactive predictive model, naming times to THEM should be

elevated compared to HER, because if Philip is assigned to be the referent of the missing argument then use of Philip (and the old woman) to fill the plural anaphoric needs of "them" will generate a contradiction, just as for the HIM target. By contrast, on the nonpredictive model, the missing argument is not determined until after a referent has been found for the overt pronoun. On this account, response times to HIM targets will still be elevated compared to HER targets because an implausible analysis (assignment of old woman to the agent role for "running") must be rejected, but the response times to THEM targets will not be elevated. Since referents can be found in the discourse for both HER and THEM without generating either a contradiction or an implausible role assignment, naming times should be equivalent for both of these pronouns. Experimental evaluation of this alternative is not presently available, but is in progress (Fodor and Garrett, personal communication).

Support for autonomous syntactic processing has been obtained in an eye-movement study by Rayner, Carlson and Frazier (1983). In these experiments subjects were asked to read visually presented sentences for comprehension and occasional paraphrasing. The sentence contained local structural ambiguities. For example:

The florist sent the flowers...

People generally would interpret this fragment as an active sentence depicting a florist sending flowers to someone else. However, a possible structural interpretation would be that of a reduced relative clause, corresponding to a pragmatically implausible reading of "The florist who

was sent the flowers...". By manipulating the meaning of the subject noun phrase and resolving the syntactic ambiguity later in the sentence, Rayner et al. (1983) were able to explore whether pragmatic plausibility would alter subjects' syntactic analysis, as inferred from eye movement measures such as reading time and regressions. Their results indicate that syntactic processing is insensitive to semantic contextual effects; subjects appeared to favor the same structural analysis (i.e. an active sentence rather than a reduced relative clause) even when that analysis led to an implausible interpretation. In their second experiment, Rayner et al. presented subjects with completely ambiguous sentences such as:

The spy saw the cop with binoculars but the cop didn't see him.
The spy saw the cop with a revolver but the cop didn't see him.

If a syntactic representation is computed without regard to plausibility, then subjects should experience difficulty with one of the sentences, since pragmatic considerations eventually force different syntactic descriptions on the sentences. If however, pragmatic and semantic context can help direct a parse, then subjects should be able to easily compute different structural representations for the same ambiguous sentences. That is, performance on the two sentences should not differ. In fact, subjects' performance on the two sentences of a pair was not equivalent. The results suggest that initial structural analyses are performed in the absence of semantic contextual influences.

As in the studies of contextual influences on lexical processing, a wide range of paradigms and methods has been applied to the study of contextual influences on syntactic processing. The existing evidence seems to favor an autonomous account of syntactic processing; however,

the data are by no means clear-cut. The next sections of this thesis describe five experiments that address the relationship between syntactic and contextual computations. Specifically, the experiments seek to explore whether syntactic processing varies as a function of the strength of semantic contextual support.

Chapter 3

Development of Experimental Materials

3.1 General assumptions and overview of experiments

Any test of the autonomy hypothesis must, of course, work within statistical conventions. Thus, even if it were possible to place a word in context and devise a method for examining whether the sentence context influences lexical access, autonomy models predict that context would have no effect (i.e. the null hypothesis). A finding of no effect is uninformative and might merely reflect the insensitivity of the dependent measure, not the structure of the language processor.

The experiments reported here were the result of a different approach. If a methodology could be found that was sensitive to a number of independent variables, each of which was postulated to have an effect on language processing, then when those variables were simultaneously present in stimulus materials one could observe whether or not their effects interacted. According to the additive factors logic described by Sternberg (1969), if demonstrated effects interact, then one may argue that the effects result from the same level of processing. If the effects are additive, that is, there is no statistical interaction, then

one may assume that they are consequences of different stages in the process. It is assumed that if the dependent measure is sensitive to the independent variables, then it will also be sensitive to any interaction of those variables. It is important to note that in this way one avoids the position of arguing for the null hypothesis. Sternberg also states that if experimental conditions produce artifactual results, those results will be most likely to produce spurious interactions, not accidental additivity. A finding of additivity is therefore quite persuasive.

The experimental series to be described tested the interaction of syntactic and semantic variables during sentence processing. Our aim was to demonstrate separately the consequences of manipulation of semantic and syntactic context and then to examine their combined effect. The predictions of the two models may be stated in the following way. In the autonomy model, information is processed in a very particular, constrained manner. If, for example, a language-user encounters a difficult-to-process syntactic structure or a (mildly) ill-formed input such as a typographical or speech error, it is the role of the syntactic processor to make sense of the input by assigning it the best available representation or representations.⁶ The semantic processor may reject a particular syntactic formulation in favor of a second analysis, but in this conception the semantic system would not control either the nature or the order of the analyses provided by the syntactic processor. This

6. By discussing only relatively mild difficulties, the cases in which the syntactic processor rejects the input as ill-formed are not considered.

means that for a given syntactic structure, no matter how strong the semantic environment, the routine followed by the syntactic processor would not vary.

Interaction models, on the other hand, demand that all sources of information have equal ability to exert their effects throughout the language process. When one form of information is degraded or missing or wrong, the information from the processing of other features in the input or context will support the assignment of the structure required to attain comprehension smoothly. This prediction then follows: If the language processor encounters a problematic syntactic configuration in its input, the stronger the information from other knowledge sources, the less consequential the problem will be for developing the final representation. This avoids a cost in extra processing time. Variation in semantic support will determine the degree to which syntactic problems are manifest. Interaction models predict that when semantic support is weak, the problem will take longer to resolve than when the semantic support is strong. Autonomy models predict that the time consequences of a syntactic difficulty will be constant across all degrees of semantic support.

It is important to note that the experiments reported here test the independence of syntactic and semantic processing and not their relationship to lexical processing. Our working hypothesis was that neither semantic nor syntactic context directly influences lexical recognition. However, it is logically possible (though a violation of the autonomy theory) that only one knowledge source and not the other

intervenes in lexical recognition. If no interaction were found, the logic of our experimental design would not help us differentiate between some alternative models. For example, if we find that the effects of semantic context and syntactic context are additive, it may be that the best model is a serial one such as that diagrammed in Figure 1 (on page 10). In this model, a data-driven process initiates an autonomous lexical processor, which delivers an output to an autonomous syntactic processor, which in turn delivers a structural representation to a semantic component (which happens in this formulation to communicate with other knowledge sources). But the additivity of our results would not support only that particular model; the results are equally compatible with the models shown on the next page in Figure 2. In each of these models, semantic and syntactic context are processed autonomously, but their relation to lexical processing differs. With the additive factors logic one may show that the two contexts are not processed at the same level, but one cannot identify the levels over which they operate or their relation to other processes.

The original investigation used a lexical decision task (Experiment 2). The reason for doing so was that lexical decision is quite sensitive to contextual manipulation; by using it one is most likely to find evidence in support of the interaction model. Following the initial test of the autonomy model, two further variations tested the independence of syntactic and semantic processing. These were a rate variation (Experiments 3 & 5) and a task variation (Experiments 4 & 5). If it is the case that the two different types of information are handled by separate and functionally distinct mechanisms, then they should be

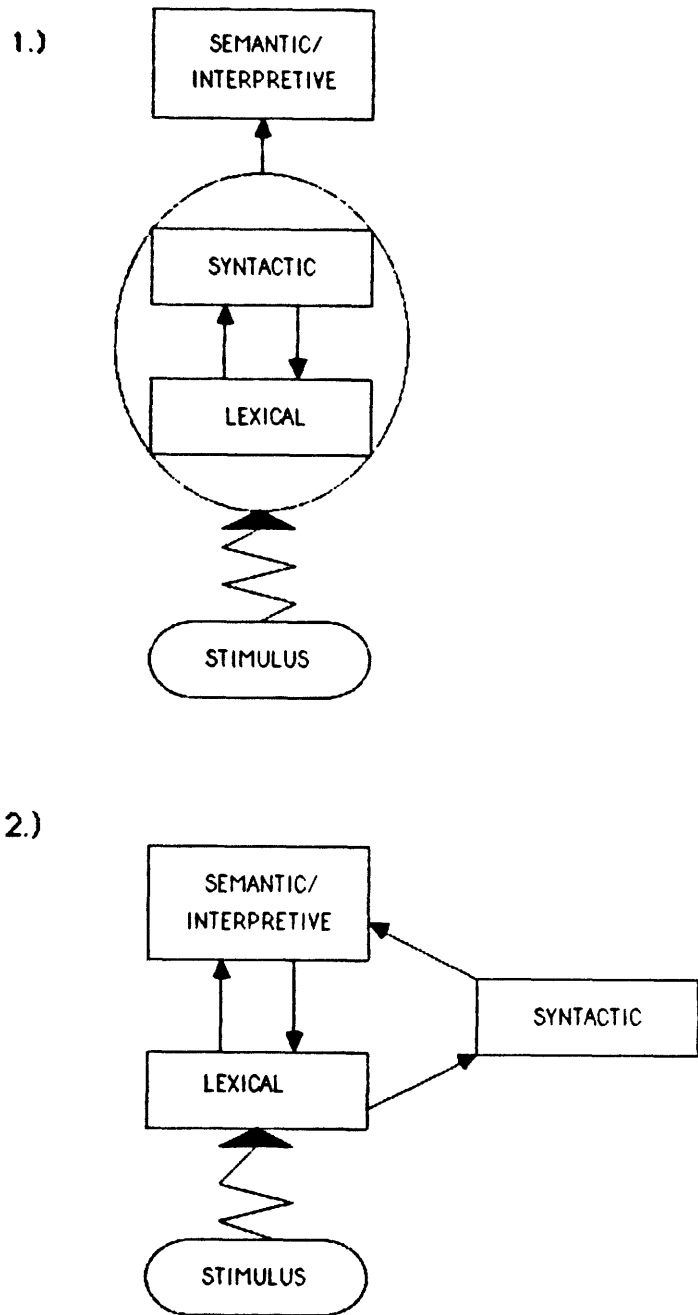


Figure 2: Alternative models of language comprehension that preserve independence of semantic and syntactic processing.

functionally dissociable. In particular, if a serial model of autonomy is invoked,⁷ the processes ought to be differentially affected by the time available for computation; with less time available the slower developing process should have a reduced impact on performance. On the opposite view that the two processes are interactive, they should show a common response to a rate (or any other) variation. We addressed this possibility in a rate variation of the stimulus presentation.

By the same logic, if semantic and syntactic representations arise from different levels of the linguistic computation, one might expect, given the appropriate task selection, that differences in task sensitivity would appear. The literature suggests the possibility of such an effect. Lexical decision seems to be sensitive to both sentential and lexical context, whereas naming seems primarily sensitive to lexically-based context (Seidenberg, et al., 1984; West and Stanovich, 1972). Accordingly, the last two experiments tested the same stimulus materials, with the same rate variation, but in a naming paradigm.

All of the experiments reported here were visual. This provides several methodological advantages (see below) but also raises a problem of the generality of the results. It is, of course, possible that the architecture of the language processor is modality-specific. Certainly the eye and ear differ. Stimulus processing and perhaps the form-based route into the lexicon may be unique to sensory modality. But after lexical contact is achieved, the character of linguistic interpretation

7. n.b. Strict seriality is not a necessary condition on autonomy.

is probably amodal. So, for example, the way in which anaphoric relations are computed or the subject of a relative clause is determined presumably has little to do with the modality of the input. On balance, for the variables at issue in this set of experiments, generalization across modalities seems plausible.

One other reason for using a visual task lies in the difficulty of controlling prosodic variables in manipulations of syntactic structure. The richness of auditory stimuli complicates the task of stimulus control. Coarticulation, intonation, volume, rate, and pausing can potentially provide cues ranging from the phonetic character of an upcoming word, to the eventual length of the sentence. Our knowledge of these processes is sketchy. This makes the independent manipulation of lexical variables somewhat problematic and raises uncertainty about what information subjects are relying on, and about what time in the course of the stimulus presentation they have access to relevant information. Many of these problems are avoided in a visual task. The information needed for understanding the sentence is contained solely in the words used and their sequential arrangement.

However, there are other problems in the use of the visual modality. The major problem in presenting subjects with normal text is that in the absence of monitoring eye movements, it is impossible to know where subjects are in the processing of the stimulus or whether they are following the given stimulus order (i.e. readers often backtrack). In addition, in a task using normal text the amount of time devoted to each word is completely under the subject's control, whereas in the auditory

case the presentation time is determined by speaking rate. Further, auditory stimuli are serially presented, with only one opportunity to hear the input. These differences between the visual and auditory domains are reconcilable. With the use of the RSVP (rapid serial visual presentation) paradigm (see Forster, 1970; Potter, 1984) the experimenter can control the presentation and duration of each stimulus word; in RSVP, subjects see only one word at a time, typically at presentation speeds that either approximate, or are faster, than the average reading rate. In this way, one can approximate the serial, time-constrained nature of speech using visual stimuli. With the RSVP methodology it is possible to test subjects at particular points during sentence comprehension with full knowledge of what information they've seen. This knowledge is important in investigating effects on real-time processing. Another advantage of the RSVP technique is that presentation rate can be easily manipulated. Thus it is easy to push subjects to the limits of their processing capabilities where subtle manipulations of linguistic input will have exaggerated (and hence measurable) consequences.

A final point on the differences between auditory and visual processing. One of the often-cited justifications for the interaction model is that speech is "degraded," i.e., the stimulus information in the speech signal is not complete enough to drive a form-based analysis. This point is primarily based on the difficulty encountered by listeners in analyzing isolated (excised) lexical samples from continuous speech. It is not clear that this observation is properly taken as proof of the degradation of speech, however. Everyday speech is normal speech--to think of the normal speech signal as being degraded seems odd.

Presumably our speech processing mechanisms have evolved to deal with the everyday case. There is, as noted above, much (prosodic) information contained in speech which is not contained in written text. That information is part of the form-driven base. Comparison of the informational quality in speech and in text is, therefore, less than straightforward.

In any event, if "perceptual adversity" is believed to be conducive to the appearance of interaction then the use of RSVP is an appropriate vehicle; it is a perceptually demanding task at moderate presentation rates. Thus, the use of the RSVP paradigm for visual presentation of a lexical decision task provides a favorable set of circumstances for a test of interaction models and a rigorous challenge to autonomy models.

All of the experiments in this thesis used the same stimulus materials. Therefore a discussion of the materials development precedes the presentation of the experiments.

3.2 Development of Materials

The manipulation of semantic constraint. The stimulus materials were developed with the lexical decision task in mind.⁸ The experiment was designed to test for the interaction of semantic and syntactic information during word recognition. Sentence level contextual

8. A version of a materials set developed by Lorraine K. Tyler while she was at the University of Chicago was adapted for this set of experiments.

influences rather than interlexical influences were chosen for focus. Moreover, since word identification was to be tested under different contextual conditions, it was important that the local environment of the target words be held constant, even when the strength of the constraining context varied. These aims were accomplished by having each stimulus consist of two sentences. The first sentence of the pair, the "Context" sentence, provided discourse-level context while the second, the "Target" sentence, contained the target word. For each target sentence there were alternate versions of the context sentence. One of the two context sentences was strongly predictive of a target word which was presented in the second (target) sentence. The predictive context was highly constraining, so this type of sentence is referred to as "High Context". The alternate context sentence ("Low Context") was not predictive of the target word. Note that although the Low Context sentence did not provide a lexically-specific constraint on the interpretation of the target sentence, it still provided a reasonable frame for interpretation. The following is an example of a stimulus set, with the target word shown in all capital letters:

High Context: My friends must call the exterminator again.

Low Context: My friends are thinking of moving again.

Target: The large number of BUGS in their apartment is driving them crazy.

The Low Context sentences were constructed to be nonpredictive of the target word as well as nonpredictive of any other specific word. This was because we wished to test the effects of a weak context versus a strong context, not to test the effects of two competing contexts, one of

which would be inappropriate to the target word. Similarly, we wished to ensure that the target sentences were compatible with the Low Context sentences. For a given stimulus set, the High and Low Context sentences had similar numbers of words and syllables and were matched in syntactic structure. Length and syntactic structure varied across stimulus sets. In most cases the material predictive of a particular target word appeared in the middle of the context sentence rather than at the end. Variations in presupposition or other factors that might trigger a change in the intonation of the target sentences were avoided. We also did not use homophones (e.g. road/rode) since we wished the materials to be appropriate for eventual testing in both auditory and visual tasks; homophones would be ambiguous for auditory tasks.

In order to determine if the intuitions guiding the construction of High and Low predictable contexts were valid, we ran a "Cloze" procedure on the materials (Taylor, 1953). In this task two groups of subjects were asked to complete unfinished sentences; groups were counterbalanced for High and Low Contexts. For example, Group One received:

Police statistics show that this is an unsafe neighborhood.
The amount of _____

and Group Two received:

Government statistics show that this is an unfortunate trend.
The amount of _____

The target sentences were all broken off just before the intended location of the target word. In all there were 46 High/Low Context pairs. The different versions were typed and the pages were randomized within each version. 15 to 19 written responses were collected per

context condition. Subjects were unpaid volunteers who were members of the MIT Psychology Department Staff, graduate students, or students in Psychology courses.

The responses were divided into those that were semantically related to the target word and those that were unrelated for each sentence pair. All but one of the tested sets showed a higher percentage of responses related to the target word under High Context as compared to Low Context sentences. The results support the contextual differentiation in the materials: The High Context led the subjects to respond with the target word, or one closely related to it semantically, in 75% of the cases. By contrast, the Low Context did not lead the subjects to the target word or a related response very often; such responses accounted for only 17% of the cases. Note that the occasional occurrence of the target word or a related word in the Low context condition shows that target sentences are indeed appropriate completions for that context. Some changes were made in the materials based on the results of the Cloze task. A few High Context sentences were changed to make the predictability stronger, and a few Low Context sentences were changed because they were too highly predictive of non-target words. Overall, there were few such changes in the materials.

The manipulation of syntactic structure. The second step in materials construction was to find a way of varying the syntactic processing load without interfering with semantic variation. This is not strictly possible since each variation of form will carry some interpretive significance. However, disrupting the syntactic processing

of the target sentence without increasing the number of alternative syntactic interpretations is a good approximation. This was done by changing closed class elements in the target sentences to produce ungrammatical forms. Target sentences were made ungrammatical by a variation just prior to the target word while still retaining the semantic force of the sentence. Three variants were used:

Note: G = Grammatical target sentence
U = Ungrammatical target sentence

(1) Deletion of a preposition

G: The large number of BUGS in their apartment is driving them crazy.

U: The large number \emptyset BUGS in their apartment is driving them crazy.

(2) Addition of an adverb⁹

G: Later, she found some TOOLS lying on the sidewalk.

U: Later she found some really TOOLS lying on the sidewalk.

(3) Change of an article or possessive pronoun to a nominative pronoun with gender kept constant.

G: Some of the PEARLS that she lost were very expensive.

U: Some of it PEARLS that she lost were very expensive.

It was intended that these variations should all render the string ungrammatical upon presentation of the target word, and most of the sentences are of this sort.¹⁰ In addition to the stimulus sets tested by

9. Although from the experimenter's point of view this is an addition of an adverb, notice that from the subject's point of view it could be the omission of an adjective.

10. A few, however, become patently illformed only a word or two subsequent to the target, e.g.:

the cloze procedure, two more sets were constructed, for a total of 48 item sets. This provided 16 sets for each of the three violation types.

The variation in plausibility. The manipulations just discussed permit us to examine the effect of sentence level contextual constraint upon syntactic processing. A corresponding disruption that suggests itself is a semantic one. If a target sentence does not follow semantically from a context sentence, both the autonomy and interaction models predict that processing time should be lengthened. Also, both models predict that processing time should be more affected under the High Context condition because having an odd word instead of a word that is strongly predicted should be more disruptive than having an odd word when no particular word is predicted. The disruption was intended to be at the level of individual word recognition and at the same serial locus as the contextual and syntactic manipulation, so an "odd" word was substituted for the target word. For example, in the exterminator/moving sentence where the target word was BUGS we now have:

Note: P = plausible target sentence
I = implausible target sentence

IG: The large number of HIPS in their apartment is
driving them crazy.

The odd word matched the original target word as much as possible in

The amount ~~Ø~~ CRIME has increased dramatically...

ungrammatical here

Since only a couple of sentences have this character, we elected to retain them.

frequency, length and in being a count or a mass noun. It did not start with the same sounds as other likely candidates for target words. For example, if a piece of furniture is predicted by the context, BEACH is not a good semantic substitution since it is similar to BED in its beginning sounds and therefore may delay the consideration of the semantically incompatible candidates. The implausible lexical target condition was fully crossed with the context and syntax conditions.

In summary; each experimental stimulus set was composed of six sentences: two kinds of context sentences and four kinds of target sentences (encompassing two types of lexical targets). An example of a full set of sentences is:

HIGH CONTEXT: Jack has finally decided to get a divorce.
LOW CONTEXT : Jack has finally decided to get
professional advice.

PG: He found that his WIFE had lost most of his money at
the racetrack.
PU: He found that he WIFE had lost most of his money at
the racetrack.
IG: He found that his TOWN had lost most of his money at
the racetrack.
IU: He found that he TOWN had lost most of his money at
the racetrack.

Each subject in an experiment saw only one of the eight possible pairings of context and target sentences. Therefore, the pairings were distributed among eight lists with equal representation of all combinations of all three variables: context, plausibility and grammaticality. Each list contained 48 experimental pairs, 12 of which were "normal" sentence pairs (PG) split between High and Low contextual types. Three-quarters of the experimental sentences in a list were "odd"

in some way, having a semantic and/or syntactic violation (PU, IG, IU). Therefore, in order to dilute the proportion of odd/normal sentence pairs, a set of 24 "filler" sentence pairs were constructed to complement the experimental pairs. The target sentences of the filler pairs were always plausible, grammatical continuations of the context sentence, which, in turn, provided either strong (High) or weak (Low) interpretive context. With the addition of this group of filler sentences to each experimental list there were 36 "normal" sentence pairs and 36 "odd" pairs. Because the same filler sentences appear in every experimental list, they can be used to evaluate performance variability across the lists. A set of practice sentence pairs¹¹ introduced the experimental series, and "lead-in" pairs were placed at the beginning of the experimental block and after a break halfway through the experiment to stabilize subjects' performance before they encountered the experimental trials.

Nonword foils and target location. For the lexical decision experiments, a set of 72 additional sentence pairs containing nonword targets was also constructed. They were constructed along the same lines as the experimental sentence pairs; that is, there were equal numbers of High and Low contextual types with equal numbers of "normal" and "odd" sentence types. A noun in each of the target sentences was replaced by a nonword starting with the same consonant cluster. (Since a third of the words replaced were implausible targets and half of the replaced

11. The practice sentences varied in their contextual strengths, grammaticality and plausibility in ways similar to the experimental sentences.

plausible targets follow a nonpredictive context, there was often no planned connection between the form of the nonword and the sentence context.)

Some changes were made in nonexperimental sentence pairs in order to make the location of the target word less predictable. If all of the targets in experimental, filler and nonword sentence pairs were to appear midway in the second sentence, subjects might ignore the material presented in the first sentence and focus their attention on the second sentence where their speeded response was sought. Placing lexical decision targets in the first (context) sentence of a small number of filler and nonword trials was intended to mitigate this possibility. The full list of stimulus materials is presented in Appendix A.

3.3 Experiment 1: Timed Cloze

Introduction

In spite of the demonstrated contextual constraint provided by the cloze procedure, we were concerned about the context predictability variation in our materials. In particular, we were concerned about the time at which the contextual constraint became effective. Thus, we decided to explore the on-line availability of potential target items experimentally. One possibility is that the cloze score obtained for our Low Context sentences reflected only the response variability between subjects, and not the relative ease of generating each individual

response. Our assumption had been that a low Cloze score represented not only a weak constraint on the range of possible completions, but also on the relative accessibility of each individual completion. We expected that when there were strong contextual constraints it would be easier for subjects to generate a response (i.e. they would be faster) , and that all subjects would generate the same response. That assumption was tested in a timed cloze task. If the subjects were placed under time pressure while they were performing a cloze procedure, it could be determined whether the time it took them to generate a response was dependant upon the strength of the contextual information.

Subjects. Thirty-seven volunteers from the MIT subject pool were paid for their participation. All were native English speakers, had normal (corrected) vision and were under 35 years of age. Five subjects were dropped from the analysis because of high error rates or exceedingly long responses.

Materials. The stimuli described in the Materials Development section were used. Only the experimental sets (n = 48) and the filler sets (n = 24) were used for this experiment. All sentence pairs had target words in the second sentence. They were adapted as follows. Each letter in every target word was replaced with an "x". An example of an experimental set is:

High: Susan was making coffee when she realized she had no milk.

Low : Susan had left everything to the last minute so she was very disorganized.

G: She had to rush to a xxxxx before her guests arrived.

U: She had to rush to a very xxxxx before her guests arrived.

Note that in this experiment there are only two target sentences per stimulus set since by removing the target word we also removed the plausibility variation. The four possible pairs of context and target sentences from each stimulus set were apportioned to four experimental lists so that context and target types were equally represented in each list. Filler sentences were the same in all four lists. Eight subjects saw each list.

Procedure. The stimuli were presented with a Terak microcomputer to subjects in a dimly-lit room. Subjects fixated on a plus sign and initiated each trial with a footpress. The sentences were presented in RSVP, retaining normal capitalization patterns and punctuation. The stimulus words appeared one right after another in the center of the CRT screen, centered over where the fixation point had been. In order to enhance comprehension of the context sentence, the average presentation rate for it was slower than that for the target sentence. The presentation rate of the context sentence accelerated, with the first word shown for 200 msec and the last three words for 83 msec each. (The rate of acceleration of the middle words depended on the length of the sentence; shorter sentences sped up faster than longer ones.) This encouraged full processing of the context sentence, while still providing a smooth transition to the target sentence where the intention was to push language processing to its limits. A blank interval of 200 msec intervened between the context and target sentences. The target sentence was presented at a constant rate of 83 msec per word. Subjects' instructions were to read the sentences silently as they appeared on the CRT screen, paying close attention to both sentences. They were told that somewhere in the second sentence of each pair there would be a string of x's. As soon as they saw the x's they were to say the first noun that came to mind. (They were assured this was not a personality test.) They were told that the x's would be embedded in the second sentence but that they were to base their response on what they saw before the x's and not after. Subjects were also warned that some of the sentences would be ungrammatical. In this as in all other cases they were instructed to fill-in the x's with a noun. Speed was stressed as well as the need to fit the noun in with the preceding context.

Latencies to begin a response were collected via a voice trigger connected with the Terak real-time clock. The experimental session was tape-recorded, enabling the experimenter to transcribe the cloze responses and check the response latency if needed. Subjects saw 15 practice trials, after which the experimenter reminded them of the important aspects of the task. The 72 trials were interrupted by a break halfway through. Subjects' comments were also recorded.

Results. The results are summarized in Table 1.

Table 1: Mean subject response latencies (RL) in msec and percentage of "good" responses (PGR), Timed Cloze.

Grammaticality	<u>Contextual Constraint</u>			
	HIGH		LOW	
	RL	PGR	RL	PGR
Grammatical	1340	80	1842	21
Ungrammatical	1585	72	2119	18

The Cloze scores reflected the pattern of the earlier paper-and-pencil Cloze task. Subjects generated the target word or one related to it more frequently under High contextual constraint than under Low contextual constraint (80% vs 21% good responses). There was as well a small effect of the grammaticality of the target sentence: Ungrammatical contexts lowered cloze scores by 6% on average. The Response Latencies (RL) were calculated as the time from the presentation of the x's to the onset of the subject's response, without taking the meaning of the response into account (i.e., RLs to begin uttering both "good" and "bad" responses were entered into the analysis). These data show that it took over half a second longer (517 msec) for subjects to generate any response after weak contexts than after strong ones. Moreover, subjects were slower (by 261 msec) to generate a response in an ungrammatical context than in a grammatical one.

Analyses of variance on response latencies were based on subject and item medians; the percentage-of-"good"-responses analysis was based on item values. The factors in the analyses were Context (high,low) and

Grammaticality (grammatical, ungrammatical). A group of analyses was performed on the subject data which included an additional factor, Versions, representing the different stimulus lists. However, since there was no effect of this factor nor any interaction involving it, it was dropped from the analyses. Similarly, stimulus items were nested under the factor Kind, the nature of the ungrammaticality in the target sentences. Again, there was no main effect of Kind and no interactions involving Kind, so it was dropped from the item analysis. All effects that are reported to be significant here, and in subsequent experiments, have a probability of chance occurrence of .05 or less.

In the latency analyses there were main effects of Context [min $F'(1,75) = 35.1715$] and Grammaticality [min $F'(1,70) = 17.6155$] with no interaction between the two [both subject and item $F < 1$]. In the percentage-of-good responses analysis there were main effects of Context [$F(1,47) = 216.046$] and Grammaticality [$F(1,47) = 7.116$], again with no interaction [$F(1,47) = 1.579$]. Thus cloze responses were more easily generated and were closer in meaning to the intended target word under strong contextual constraints. The grammaticality of the surrounding context also affected both the choice of and the ease of making the cloze response.

Discussion. The results from the Timed Cloze task justify the assumption that context affects not only which items are generated but also the ease with which they are generated. We found substantial differences in response latency as a function of semantic and syntactic context although the effects of the two contexts did not interact. Here,

then, is a demonstration that words are "less available" in weak contexts than in strong ones. This is also true of syntactic context. Syntactic context, like semantic context, affects subjects' ability to generate a response, but unlike semantic context, it has little effect on the meaning of the particular response given. Most particularly these results give us confidence that the contextual influences we are investigating are strong and immediate in their availability. They do not require a reflective judgment on the part of subjects.

When subjects were asked for comments at the end of the experiment, many said that although the instructions had prepared them for seeing ungrammatical sentences, they hadn't noticed any. Other subjects stated they noticed a few ungrammatical sentences, or sentences for which it was inappropriate to fill in a noun. When they were asked to estimate what proportion of trials contained ungrammatical sentences most subjects replied "Oh, just a few" or "Maybe one out of ten". (In reality one out of three trials was ungrammatical.) Only one subject thought that as many as 10 or 20 percent of the trials were ungrammatical. Given such low awareness of the ungrammatical character of many of the sentences, it becomes even more interesting that the time to generate a response was reliably longer in ungrammatical contexts.¹² Subjects do not appear to be aware of syntactic processing in this presentation condition.

12. In related pilot experiments using a repetition paradigm with the same materials, we found a similar mismatch of conscious report and processing outcomes. Subjects were not able to differentiate grammatical and ungrammatical sentences with a rating scale. In addition, their post-experiment comments reflected a lack of awareness of ungrammatical trials. Nonetheless, subjects were worse at repeating the ungrammatical sentences.

The additivity of the semantic and syntactic contextual effects is intriguing as it suggests that the two contexts act independently in determining an appropriate response. If syntactic and semantic constraints apply at the same processing stage, one might have expected to see very strong facilitation effects of the High context condition compensate for the syntactic processing problem. However, since this task involves production of a lexical item as well as comprehension, we cannot take the results to be a strong test of our language processing models. Additionally, the statistical analysis of the response latencies was performed over medians; Sternberg (1969) states that only arithmetic means may be used to determine whether factors interact or are additive.

Having demonstrated the effects of contextual differences in our materials in the presentation mode to be used, we now turn to a direct test of the autonomy and interaction models.

Chapter 4

Lexical Decision Experiments

4.1 Experiment 2: Lexical Decision with a Slow Presentation Rate

The aim of these studies was to determine whether syntactic and semantic effects interact during sentence comprehension. Therefore, it was very important to ensure that subjects attended fully to the semantic content of the sentences. As noted in Chapter 3, in the Materials Development section, target items were placed in the first sentence of some filler and nonword distractor sentence pairs in order to make the position of the target word less predictable and thereby prevent subjects from "relaxing" during the presentation of the context sentence. Subjects' attention to the context was further encouraged by two additional measures: First, subjects were periodically given comprehension probes during the experiment and, following each stimulus, they were required to rate the coherence of the two-sentence passage. Instructions for the rating task and descriptions of it and the comprehension task are included in the procedures section. Second, to make processing easier, a 200 msec presentation rate was used. This was a slower rate of presentation than that used in Experiment 1. This is certainly slow enough to allow the meaning representation to proliferate,

yet fast enough to allow the measurement of real-time sentence processes. At 5 words per second it is just faster than the normal speech rate, and within the normal reading rate range for text. Much of the visual work on context effects during sentence comprehension has been done at rates of 300-500 msec per word; those are slow rates given the known capabilities of human language processing. We were impressed in Experiment 1 and related pilot experiments by how much subjects were able to retain at speeds of over 10 words per second. (See Potter, 1984 for a relevant review.)

Methods

Subjects. Forty three volunteers from the MIT subject pool were paid for their participation. They were all native English speakers, had normal (corrected) vision and were under 35 years of age. One subject's data file was incompletely written to the disk and was therefore lost; two other subjects had median reaction times to plausible targets in grammatical sentences well beyond the preestablished cutoff point of 1 sec and were dropped from the analysis.

Materials. The stimulus materials described in the Materials Development section were used. There were 20 practice pairs and no lead-in pairs. Subjects were assigned to one of 8 materials lists. Therefore, 5 subjects saw each list.

Procedure. The stimuli were presented with a Terak microcomputer in a dimly lit room. Subjects fixated on a plus sign and initiated each trial with a footpress. The sentences were presented in RSVP, retaining normal capitalization patterns and punctuation. The stimulus words appeared sequentially in the center of the CRT screen, centered over the spot where the fixation point had been. The presentation rate was 200 msec per word, with a 400 msec interval between the two sentences of each pair. Subjects were instructed to read silently and to understand the sentences as they appeared. Lexical decision targets were cued by a case shift: Whenever a subject saw a string of letters written entirely in capital letters, s/he was to make a lexical decision for that item as quickly as possible. Responses were manual; subjects pressed one of two microswitches indicating the lexical status of the target (right switch = word; left switch = nonword) while continuing to read the words appearing on the screen after the target. When each stimulus pair had finished, subjects were required to rate how closely the two sentences were related. In addition, on occasional trials (20 out of 144) subjects had to answer comprehension questions. The following is an excerpt from the

subject instructions covering the rating scale and comprehension questions:

It's very important that you read and understand the sentences that you see. In order for you to do a good job, you must be reading the sentences carefully. Some of the sentences you'll be seeing are a little peculiar and, in fact, some of them aren't real sentences. Some of the sentences have strange words inserted in them and others have a word or two deleted. This means that some of the sentences you'll be seeing are ungrammatical.

After you've made the word/nonword decision and both sentences have finished, I want you to give me a simple rating of how well the two sentences go together. The scale is from 1 to 3. You should rate only those sentences which contain real English words. If the second sentence followed very nicely from the first, if it is strongly related to the first sentence, give it a rating of "one". To do that, just say out loud "one". If the second sentence follows from the first somewhat, but not very strongly, give it a "two". If the second sentence doesn't seem to follow at all, or if either sentence seems to be peculiar or ungrammatical, give it a "three". In other words, if you strongly expected the topic of the second sentence from the first, say "one", if you didn't have any strong expectations of the topic but the second sentence made perfect sense following from the first, then say "two", and if the two sentences don't go together at all or if there is anything a bit bizarre or ungrammatical about the trial, then say "three". ...

The last thing I'm asking you to do is to answer some comprehension questions. At random points throughout the experiment I'll ask you a question about the pair of sentences you've just seen. I'll ask this question after you've given your rating. So please wait for me to say "okay" after you've given your rating (or said "nonword") so you don't start the next trial before I have a chance to ask you a question.

It was anticipated that the rating data produced by the subjects would mirror the stimulus variations even if their latency results did not. Thus, it was expected that plausible, grammatical sentences following a High Context would be rated "1", those following a Low Context would be rated "2", and all implausible or ungrammatical sentence pairs would be rated "3".

Subjects were given 20 practice trials. Between these trials the subjects were encouraged to attend to the meanings of the sentences and to make their word/nonword decisions quickly while they continued to read the sentences. There were more comprehension questions during the practice trials than during the experiment (8 out of 20 trials). Subjects were given short breaks after the practice trials and halfway

through the experiment. Their comments were collected at the end of the experiment.

Results.

General performance levels were good. The mean reaction time for "normal" sentences was 784 msec and the mean percentage of errors was 3.5% . Subjects typically enjoyed this experiment and found it easy to perform.

Statistical analyses for this and subsequent experiments were performed with the Perlman and McClelland UNIX-based ANOVA package developed at the University of California at San Diego. The analyses of variance were based on a method described in Keppel (1973). Individual subject or item scores are the input for the analysis and errors need not be replaced as long as there are no empty cells. Subjects' errors were omitted from the response latency analysis and latencies longer than two standard deviations were replaced with the cutoff values from each condition. Analyses were not conducted on the nonword foils; these items were of no interest in relation to the hypotheses tested in the experiments.

Both subject and item analyses were run and $\min F'$ values were calculated for response latency and percent error. All reported values were significant at the 0.05 level unless otherwise noted. When subjects were treated as the random factor, there were fixed factors of Version (subjects were nested in 8 materials lists), Context (high and low), Grammaticality (grammatical and ungrammatical) and Plausibility (plausible or implausible target word). When items were treated as the

random factor, there were fixed factors of Kind (items were nested within three types of ungrammaticality), Context, Grammaticality and Plausibility. Separate analyses were performed without the nesting factors of Version or Kind to provide F values for the minF' calculation. There was never any main effect of Version or any interpretable interactions involving Version. The F values for Version interactions are therefore not presented in the body of the paper, but rather are relegated to Appendix B. Whenever there were main effects or interpretable interactions involving the factor of Kind, they are reported after the results of the minF' calculation, otherwise the values are reported in Appendix B. The investigation of Plausibility was seen as being independent of the investigation of the combined effects of discourse context and syntactic environment. Therefore, separate analyses were conducted for plausible targets only and implausible targets only (each including the factors of Context and Grammaticality), as well as grammatical targets only (including the factors of Context and Plausibility). Predictions had not been generated on the combined effects of Plausibility and Grammaticality. However, an inclusive analysis was performed with the factors of Context, Grammaticality and Plausibility. The results of this analysis are reported in Appendix B. Chi-square statistics were computed for the rating data. Except where noted, the data for the chi-square analyses were the frequencies with which each of the three rating scale points was used.

The data on the effects of semantic and syntactic context on the processing of plausible targets bear most directly on the evaluation of the autonomy and interaction models. Those results are presented first.

Response latencies and percent errors, plausible targets. Mean subject response latencies and percentage of errors for the different experimental conditions are summarized in Table 2. When sentences contained plausible targets, both the Context and Grammaticality manipulations dramatically influenced response latencies. Their effects are additive. The same items were responded to faster (by 88 msec) after a strong context than after a weak one [$\text{minF}'(1,85) = 12.346$]. The processing of items was disrupted in ungrammatical contexts; subjects' responses were slowed, on average, by 111 msec [$\text{minF}'(1,81) = 16.996$].

Table 2.

Mean subject response latencies in msec (and percent errors) to make lexical decisions to plausible targets under different contextual conditions. Slow Lexical Decision.

Grammaticality	<u>Contextual Constraint</u>	
	High	Low
Grammatical	745 (2.9)	823 (4.2)
Ungrammatical	846 (3.8)	944 (6.7)

There was no statistical interaction between Context and Grammaticality (both F_1 & $F_2 < 1$), and there is, therefore, no evidence for a common processing stage for semantic and syntactic context. Subjects were hampered just as much by a syntactic disruption under a strongly supportive context (101 msec) as under a weak one (122 msec), the 21 msec difference not being reliable. Error rates did not differ significantly across conditions.

In the item analysis including Kind of ungrammaticality as a factor, there was no main effect of Kind or interactions involving Kind in the

response latency or error data. This indicates that the processing consequences of the three different types of ungrammaticality were equivalent.

Ratings, plausible trials. Although the primary objective of including the rating scale was to force subjects to attend to the meaning of the sentences, the rating data themselves provided interesting results. A set of histograms representing the relative frequency with which each rating point was used under the different contextual conditions is presented in Figure 3. On the basis of the outcome of the cloze procedures used in the development of these materials, one might expect ratings of "1" for plausible, grammatical sentences after High context, and ratings of "2" for the same target sentences following Low context. Although there was a strong effect of Context [$\chi^2 = 171.15$; $df = 2$], subjects showed some tendency to rate any grammatical sentence "1" rather than "2". This was reflected in a "1" being assigned after Low context in 36% of the trials, trials on which subjects ideally would have assigned a rating of "2". There was also a strong effect of grammaticality [$\chi^2 = 121.00$; $df = 2$] reflecting a higher proportion of ratings of "3" when the target sentence was ungrammatical than when it was grammatical. Again, subjects judged the ungrammatical sentences closer to normal than expected. Inspection of the histogram shows that fewer than a third of the ungrammatical trials were actually rated as ungrammatical. These findings, taken together with subject interviews conducted after the experiment, suggest that the subjects were frequently unaware of the ungrammatical nature of the stimulus materials.

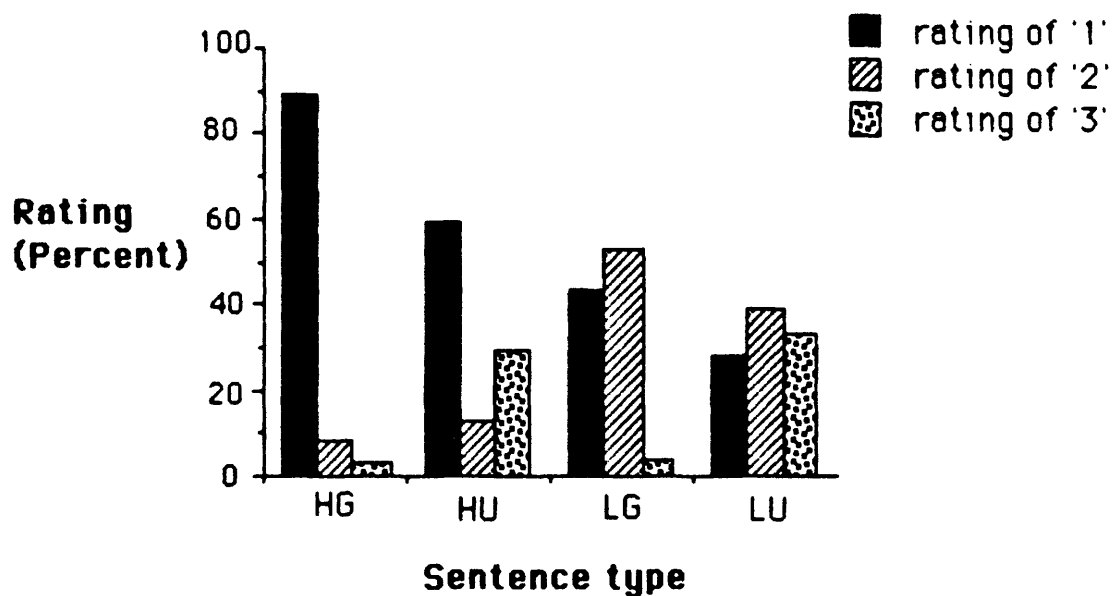


Figure 3: Percentage of subjects' discourse ratings for each sentence type. Experiment 2: Slow Lexical Decision. (H: High Context sentence; L: Low Context sentence; G: Grammatical Target sentence; U: Ungrammatical Target Sentence.)

A correlational analysis was performed to assess the relation between awareness of an ungrammaticality and its effect on lexical decision time. For both the subject and item data, differences in response latencies under grammatical and ungrammatical contexts were paired with the differences in the rating scores under the two contexts, the correlation coefficient was determined and then checked for significance using a calculated t-value. The results from the two sets of means are different. In the analysis over subject means there seems to be no statistical relation between how perceptive subjects were of the grammaticality of the target sentence and the magnitude of the processing effect which results from that ungrammaticality: subject $r = .206$, $t = 1.30$, $p > 0.05$, one-tailed. In the analysis over item means, there was a significant relation: item $r = .4022$, $t = 2.98$, one-tailed.

The ratings were also tabulated as a function of Kind of ungrammaticality. The grammatical sentences of each item type were rated similarly [$\chi^2 = 5.241$, $p < 0.10$; $df = 4$]. However, the ungrammatical sentences were not [$\chi^2 = 36.990$; $df = 4$]. There were big differences in how apparent the ungrammaticalities were: Violations of the "addition of adverb" type were rated as a "3" 47% of the time; when an article or pronoun was altered, that was rated a "3" 30% of the time; and the least detectible error was the omission of a preposition, which was rated a "3" only 16% of the time. These differences are especially interesting when they are taken in conjunction with the absence of any latency or error differences as a function of Kind of ungrammaticality. That is, ungrammaticalities which varied widely in their noticeability had

equivalent processing consequences. Tables of the rating percentages as a function of Kind for this and the other experiments are presented in Appendix C.

Response latencies, percent error, and ratings; implausible trials

The same analyses reported for plausible targets can be run for implausible targets. However, the results so generated are only of interest if the main effects are evident. The actual results were not very intriguing, but are presented in Appendix B for this and other experiments.

Response latencies and percent error, grammatical trials. In the next set of analyses our aim was to examine the effect of plausibility. Therefore we analyzed only grammatical trials. MinF's were calculated based on subject and item analyses over the factors of Context and Plausibility. Subject means for response latencies and percent error are displayed in Table 3.

Table 3.
Subject means for targets appearing in grammatical sentences under different degrees of contextual support. Response latencies in msec (percent error). Slow Lexical Decision.

Plausibility	<u>Contextual Constraint</u>	
	High	Low
Plausible	745 (2.9)	823 (4.2)
Implausible	953 (17.1)	977 (8.3)

By minF' the 51 msec effect of context just missed significance

[$\text{minF}'(1,86) = 3.455$] even though both subject and item analyses revealed a main effect of Context [$F_1(1,39) = 7.282$; $F_2(1,47) = 6.57$]. There was a powerful main effect of Plausibility, with lexical decisions on implausible items being completed 181 msec more slowly than on plausible items [$\text{minF}'(1,86) = 26.016$]. There was no reliable interaction between the two factors although there was a trend toward significance in the item analysis [Context X Plausibility: $F_2(1,47) = 3.248$, $p < 0.10$].

The Context effect on error rates was apparent in both subject and item analyses [$F_1(1,39) = 4.794$; $F_2(1,47) = 5.742$] although the 4.7% difference was not significant by the minF' calculation ($p > 0.10$). There was a substantially greater number of errors made on implausible than on plausible targets (12.7% vs 3.5%), [$\text{minF}'(1,69) = 14.178$]. There was also a Context X Plausibility interaction for percent error in both subject and item analyses, though again the interaction was not strong enough to reach significance by minF' [$F_1(1,39) = 6.411$; $F_2(1,47) = 9.724$; $\text{minF}'(1,79) = 3.8637$, $p < 0.10$]. This interaction is due to subjects having made more errors to implausible targets when those targets followed a High Context than when they followed a Low one. [Newman-Keuls on subject means: $q(2,39) = 4.47$, $p < 0.05$]. In contrast, subjects' errors to plausible targets did not significantly vary with contextual constraint [$q(2,39) = 0.66$, $p > 0.05$].

Ratings, grammatical trials. Sentence pairs containing plausible targets were rated lower (i.e. more related) than pairs containing implausible targets [$\chi^2 = 351.52$]. Subjects were asked to differentiate context on plausible trials yet rate all implausible trials uniformly.

Thus, effects of sentential context on the rating of implausible targets were not expected to be evident. Almost all (86 %) implausible targets were rated "3" irrespective of the preceding semantic context. However, there was more of a tendency to rate implausible targets a "2" after Low context (20 %) than after High context (11%). [$\chi^2 = 7.32$]. This probably reflects the same bias subjects had to rate normal sentences a "1". Perhaps in order to use all points on the rating scale subjects rated sentences which were a little odd a "2", and really weird sentences a "3". Together with the error data, these data suggest that implausible targets appearing after High context were perceived as more extreme than when the same targets followed Low context.

Discussion.

The results clearly support the autonomy theory. When plausible sentences are being processed, the strength of the preceding context as well as the sentence's grammatical integrity has substantial effects on lexical processing as measured in a lexical decision task. When these two factors are both in play, their effects are additive; they do not interact. In terms of Sternberg's additive factors logic, semantic/inferential processing and structural processing are operating at different levels or stages during sentence comprehension.

Taken with the response latency results, the rating data reflect the effects of the independent variables. However the rating data suggest a dissociation between the conscious awareness of a context and its effect

on processing. Subjects did not seem fully aware of the ungrammatical nature of many of the stimulus pairs, yet those ungrammaticalities significantly slowed response latencies. In the post-experiment interviews, subjects estimated the percentage of ungrammatical trials to be approximately 16%, on average. The percentage was actually 33%; if subjects' postexperiment estimates are taken at face value, they noticed only about half of the ungrammaticalities contained in the experiment. The discourse ratings themselves indicated that subjects were aware of only a third of the ungrammatical sentences. From the results of the correlational analysis, it appeared that subjects who differed on how frequently they noticed ungrammaticalities did not differ in how they responded to those ungrammaticalities in the lexical decision task. Type of syntactic violation had a substantial impact on subjects' ratings of grammaticality. This occurred even though the processing consequences of the violations, as measured by the latency data, were equivalent across all three types. However, to a very limited extent, the more noticeable an ungrammatical item was, the more slowly it was responded to, relative to its grammatical counterpart.

Though the ratings were, on the whole, sensitive to the semantic contextual manipulation, subjects tended to rate any plausible sentence pair "1". It was decided therefore, to amplify the subject instructions for the rating task in subsequent experiments in order to mitigate the effects of that bias.

Effects of sentential context were not apparent in the analysis comparing plausible and implausible targets in grammatical sentences.

The plausibility difference on the other hand was very secure. This difference was so large (181 msec) that we feel confident it was not due to the fact that there were two different sets of target items contributing to the means, but was rather a product of the sentence comprehension process. Contrary to expectations, there was no evidence of a Context X Plausibility interaction in the latency data.

This experiment was successful in demonstrating two classes of contextual effects on sentence processing and enabled us to test for an interaction of those effects. However, having established that semantic and syntactic context operate at different levels of analysis, the additive factors logic does not permit the conclusion that those levels are independent of one another; it cannot determine whether the levels correspond to separate processes. In order to establish the postulated independence of semantic and syntactic processing, convergent evidence is needed. Sternberg (1969), in fact, discusses an exception to the additive factors logic; two factors might additively influence the same stage in processing. Although he states his belief that this would be a rare occurrence, it can be checked with the addition of a third factor to the experiment. If this third factor interacts with both of the original factors or with neither, then it seems likely that the original two factors are not independent, and are probably operating at the same stage in the analysis.¹³ In the same way, if a third factor can be found that

13. However this is not proof of interdependence, as it is possible that the third factor may be involved independently with both of the two separate stages or acting at a completely different level in the analysis. Presumably, thinking about the whole process under study would allow one to avoid this pitfall of choosing factors which are either too

interacts with only one of the two original factors and not the other, this strengthens the argument that the factors correspond to independent levels. This, then, was the motivation for the next three experiments. Experiment 3 employed a rate variation; Experiments 4 and 5 were experiments in a different paradigm, the Naming paradigm, with the same rate variation. The dissociation of syntactic and semantic processing was the goal in each of these experiments.

4.2 Experiment 3: Fast Lexical Decision

If semantic and syntactic processing are subserved by independent mechanisms, it is likely that they have different time-courses. On a model such as the one diagrammed in Figure 1 (on page 10), for example, this would be especially likely since semantic processing occurs at a later stage in the linguistic analysis of the input than does syntactic processing. It was conjectured that if the subjects were put under more time pressure, semantic/inferential analyses would necessarily lag behind and thus show reduced influences at the time of our probe. If syntactic processes are faster/earlier, the rate change should have less impact. If on the other hand, syntactic and semantic factors operate at a common level they ought to show similar consequences of the rate change. Specifically, the prediction was that at faster speeds of presentation semantic context would no longer have an effect on the lexical decision

general or too far removed from the postulated stages to be useful in teasing apart those stages.

task, whereas the effects of syntactic context would be preserved. This prediction would only hold if subjects have access to the output of the syntactic processor for use in the lexical decision task. If that output is not available for such a nonlinguistic application, but instead is immediately used as a semantic, interpretive base (cf. Cairns, in press); then effects of the discourse context would still be evident. To the extent that main effects of discourse and syntactic manipulations remained, no interaction was expected.

Methods

Subjects. Fifty-nine volunteers from the MIT subject pool were paid for their participation. They were all native English speakers, had normal (corrected) vision and were under 35 years of age. Seventeen subjects had mean reaction times for "Yes" trials longer than 1300 msec, error rates for "Yes" trials greater than 15 % and/or error rates for "No" trials greater than 25 %. These subjects were excluded from the analysis. In addition, two subjects were dropped to even out the design. Both of these subjects had mean reaction times for "Yes" trials of over 1200 msec.

Materials. The stimulus materials from Experiment 2 were used. Ten additional practice items and four sets of lead-in sentences were constructed. Subjects were assigned to one of eight materials lists, therefore five subjects saw each list.

Procedure. The procedure was identical to that of Experiment 2 with the following exceptions : 1) The rate of presentation was changed from 200 msec per word to 117 msec per word with a 234 msec blank interval between the two sentences of each pair. 2) The number of practice items was increased from 20 to 30 items. This was done because subject performance was much more variable under the faster speed and subjects needed more time to become proficient at the task. In addition, two sets of two lead-in sentence pairs were constructed. These sets were inserted at the beginning of the experiment and after the half-way break to allow subjects' performance to stabilize before they encountered the experimental trials.¹⁴ 3) The instructions for the rating scale were

14. These changes were actually made after the first nineteen subjects were run under the fast presentation rate. Seven of these subjects' data had to be dropped from the analysis, therefore twelve of the forty subjects contributing to the means had less practice and less experience

expanded. More emphasis was placed on the distinction between ratings of "1" and "2", in an effort to eliminate subjects' bias to rate any normal sentence "1". Immediately after the instructions for the ratings, subjects were given six sentence pairs on index cards and asked to rate each one. If the subject's ratings differed from those of the experimenter, the disagreement was discussed until one or the other was convinced (this was usually, but not always, the subject). 4) During and after practice, some subjects were questioned for the reasoning behind certain of their ratings in order to emphasize the difference between the ratings of "1" and "2".

Results.

The same analysis approach was used as for Experiment 2. In addition to the analyses run on the forty "experimental" subjects a separate set of analyses was run on the nineteen "discarded" subjects. These subjects had a mean reaction time approximately 400 msec longer than the experimental subjects and a mean error rate of 12.4 %. They showed strong effects of discourse Context and Plausibility, but no effect of Grammaticality. We interpret the performance of the discarded subjects to indicate an inability to linguistically process the input at this rate of presentation. The results of the analyses are presented in Appendix B.

Response latencies and percent error for plausible trials. When plausible target items were analyzed it was immediately apparent that, contrary to the predictions, both context effects were still in force. The data are summarized in Table 4.

with the rating scale before starting the experiment than the other twenty-eight subjects.

Table 4.
 Mean subject response latencies in msec (and percent errors)
 to make lexical decisions to plausible targets in different contexts.
 Fast Lexical Decision.

Grammaticality	<u>Contextual Constraint</u>	
	High	Low
Grammatical	761 (0.8)	819 (4.3)
Ungrammatical	806 (2.9)	880 (8.3)

Strong contexts produced faster lexical decision times than did weak contexts; the 72 msec difference is reliable [$\min F'(1,79) = 9.9911$]. The Grammaticality effect of 50 msec was also significant [$\min F'(1,86) = 6.192$]. The interaction was not significant in either subject or item analysis ($p > 0.10$). In the item analysis over response latencies with Kind of ungrammaticality as a factor there was no main effect of Kind or any interaction involving it.

With the faster presentation, error rates were more sensitive to the experimental variations. There was a main effect of Context (High = 1.9%, Low = 6.3%) supporting the response latency difference [$\min F'(1,85) = 4.239$]. The main effect of Grammaticality was significant in both F1 and F2, although $\min F'$ was only marginal [$\min F'(1,85) = 54.91, p < 0.10$; $F_1(1,39) = 7.205$; $F_2(1,47) = 4.784$]. In the separate item analysis including Kind as a nesting factor, there was no main effect of Kind. The Kind X Grammaticality interaction was marginal [$F(2,45) = 2.803, p < 0.10$]; subjects made more errors on targets following Low contexts than High contexts on only one type of syntactic violation, that of addition of an adverb. The relevant Newman-Keuls values are: Omission of

preposition, $q(2,45) = 0.209$, $p > 0.05$; Addition of adverb, $q(3,45) = 4.388$; Article or pronoun change, $q(4,45) = 2.507$, $p > 0.05$. There was also a significant Kind X Context interaction; these data are presented in Appendix B.

Ratings, plausible trials. Histograms of the rating point frequencies are shown in Figure 4. There was a strong effect of Context [$\chi^2 = 186.96$; $df = 2$]. In contrast to Experiment 2, the histograms for the grammatical sentences for this experiment show that there was less of a subject bias toward assigning any "normal" sentence a rating of "1". Inspection of Figure 4 also reveals that awareness of the ungrammatical sentences diminished sharply with the fast presentation rate, although they were still rated differently than the grammatical sentences [$\chi^2 = 11.10$; $df = 2$]. The histogram shows that subjects only rated 11 % of the ungrammatical sentences as such (i.e. assigned them a "3"). As in Experiment 2, a correlation coefficient was determined for ratings and response latencies. The low correlation coefficients (subject $r = .13$, $t = 1.00$, $p > 0.05$, one-tailed; item $r = .29$, $t = 2.06$, $p > 0.05$, one-tailed) provided additional evidence that processing difficulties did not depend upon the awareness of grammatical deviance.

The ratings were examined with regard to Kind of syntactic violation. There were no differences between the item sets for the ratings of grammatical sentences [$\chi^2 = 4.100$, $p > 0.10$; $df = 4$]. For ungrammatical sentences, however, the types of ungrammaticality were rated differently [$\chi^2 = 34.125$, $p < 0.001$; $df = 4$]. "Addition of an adverb" was the most noticeable violation (20 % of the violations were

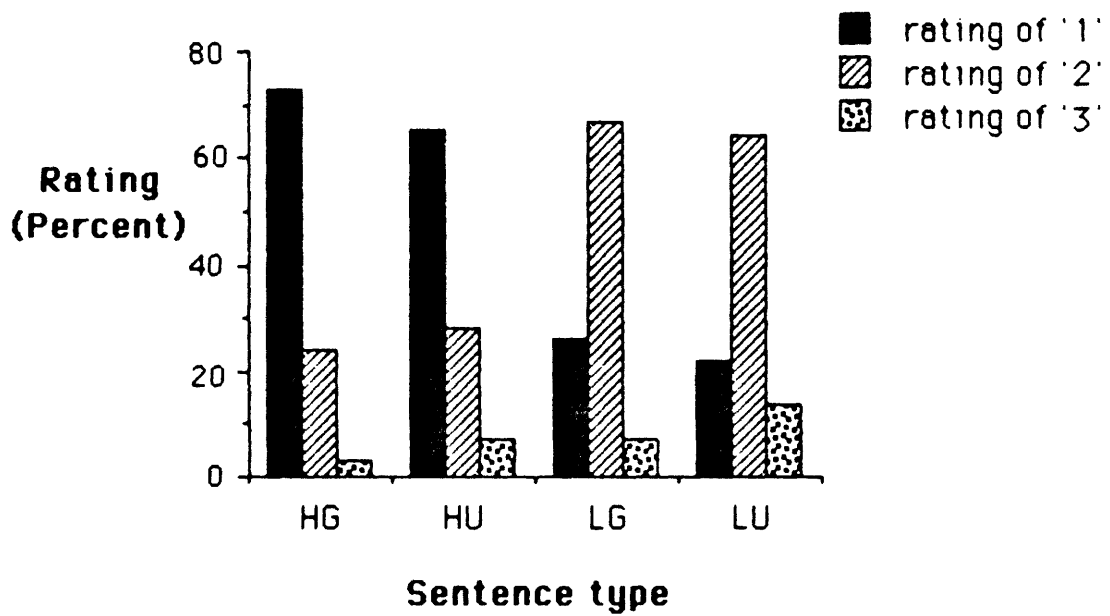


Figure 4: Percentage of subjects' discourse ratings for each sentence type. Experiment 3: Fast Lexical Decision. (H: High Context sentence; L: Low Context sentence; G: Grammatical Target sentence; U: Ungrammatical Target sentence.)

rated a "3") followed by, in turn, "alteration of an article or pronoun" at 10 %, and "omission of preposition" at 2 %. This pattern is identical to that found in Experiment 2; and as in that experiment, there were no latency differences between the levels of Kind to correspond to the differences in the ratings.

Response latencies and percent error, grammatical trials.

Grammatical trials analyzed with factors of Context and Plausibility showed no main effect of Context in either response latency [F_1 and $F_2 < 1$] or percent error ($p > 0.10$). The effect of Plausibility was strong in both analyses: latency [$\text{min}F'(1,86) = 41.912$], and percent error [$\text{min}F'(1,85) = 22.981$]. The Context X Plausibility interaction was not reliable in the combined statistic for latency data [$\text{min}F'(1,85) = 3.086$, $p < 0.10$], although the 239 msec difference between plausible and implausible targets under High context was statistically larger than the 141 msec difference under Low context in both subject [$F(1,39) = 7.835$] and item [$F(1,47) = 5.091$] analyses. Newman-Keuls performed on the subject means revealed that although subjects were faster to respond to plausible targets after High contexts than Low contexts [$q(2,39) = 3.46$, $p < 0.05$], contextual constraint had no reliable effect on subjects' responses to implausible targets [$q(2,39) = 2.19$, $p > 0.05$].

The Context X Plausibility interaction in the error data was also significant for subject-based [$F(1,39) = 7.000$] and item-based [$F(1,47) = 8.586$] analyses, but did not hold up in the $\text{min}F'$ calculation [$\text{min}F'(1,83) = 3.873$, $p < 0.10$]. Newman-Keuls on the subject means suggests that subjects made approximately the same number of errors on the

plausible targets across the variation in contextual support [$q(2,39) = 1.80, p > 0.05$]. However, when implausible targets appeared following High contexts, subjects committed more errors than when they followed Low contexts [$q(2,39) = 3.50, p < 0.05$]. These data are summarized in Table 5.

Table 5.
Subject means for plausible and implausible targets in grammatical sentences following various strengths of context. Response latencies in msec (percent error). Fast Lexical Decision.

Plausibility	<u>Contextual constraint</u>	
	High	Low
Plausible	760 (0.8)	820 (4.2)
Implausible	999 (18.3)	961 (11.7)

Thus, although there were no main effects of Context in these latency and error data, there are indications that the two forms of Context are not equivalent.

Ratings, grammatical trials. Sentence pairs containing plausible targets were rated lower than pairs containing implausible targets [$\chi^2 = 662.45; df = 2$]. Almost all (92%) implausible targets were rated a "3" irrespective of the preceding semantic context. With the more extensive instructions on the use of the rating scale, subjects rated implausible targets uniformly across contexts: under High context the percentages of "1" and "2" ratings were 2% and 6%, respectively; under Low context they were 3% and 6%, respectively.

4.3 Comparison of Experiments 2 & 3

No difference in the pattern of results was observed for Experiment 3 as compared to Experiment 2, except that with the faster rate of presentation the ungrammatical trials seemed less apparent to the subjects. In order to examine more carefully the differences in the two sets of results, an analysis was performed over both sets of data. The fixed factors were Rate, Context, and either Grammaticality or Plausibility.

Results of combined analyses.

Response latencies and percent errors, plausible trials. In the subject analysis of plausible targets there were no main effects of Rate in latency or percent error [both $F_s < 1$]. There was an effect of Rate in the item latency data [$F(1,47) = 4.161$] reflecting a trend for item latencies to be smaller with faster rates of presentation. There was no main effect of rate in the item percent error data [$F < 1$]. When both experiments are considered together, there were strong main effects of Context and Grammaticality in the latency data [Context: $\text{MinF}'(1,87) = 18.743$; Grammaticality: $\text{minF}'(1,121) = 22.426$] with no interaction of the two factors [$\text{minF}' < 1$]. For the error analysis there was only a main effect of Context as measured by the minF' statistic [$\text{minF}'(1,100) = 4.008$], although the effect of Grammaticality was significant in both subject [$F(1,78) = 8.796$] and item analyses [$F(1,47) = 6.377$]. These data

are presented in Table 6.

Table 6: Mean subject response latencies in msec (and percent error) to plausible targets in sentences varying in their grammatical status and with different degrees of contextual support. Both Lexical Decision Experiments.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	753 (1.9)	822 (4.2)
Ungrammatical	823 (3.4)	913 (7.5)

In none of the types of data did Context interact with Rate. That is, there was no change in the size of the Context effect as a function of the speed of presentation [all p 's > 0.10]. There was a Rate X Grammaticality interaction in the latency data; the magnitude of the Grammaticality effect diminished with the increase in the presentation rate. These data are shown in Table 7.

Table 7: Response latencies in msec to plausible targets in grammatical and ungrammatical sentences presented at slow (200 msec) or fast (117 msec) rates. Both Lexical Decision Experiments.

Grammaticality	<u>Speed of Presentation</u>	
	SLOW	FAST
Grammatical	785	790
Ungrammatical	896	840

This interaction missed significance with $\min F'$ [$\min F'(1,120) = 3.3508$, $p < 0.05$] but attained it in the subject and item analyses [Rate X

Grammaticality: $F_1(1,78) = 6.409$; $F_2(1,47) = 7.022$].

Ratings, plausible trials. In order to determine whether or not subjects were more aware of ungrammatical sentences with slow presentation than fast presentation, chi-square was computed over a 2 X 2 table. In this table, the ratings for the ungrammatical sentences in each of the experiments (Slow and Fast) were classified as either a "grammatical" rating ("1" or "2") or an "ungrammatical" rating ("3"). These frequencies are shown as percentages in Table 8.

Table 8: Percentage of grammatical and ungrammatical ratings for ungrammatical sentences presented at either slow (200 msec) or fast (117 msec) rates of presentation. Both Lexical Decision Experiments.

Rating	<u>Presentation Rate</u>	
	Slow	Fast
Grammatical	69 %	89 %
Ungrammatical	31 %	11 %

Ungrammatical sentences were more perceptible when they were presented at the slower speed [$\chi^2 = 56.83$; $df = 1$].

Response latencies and percent error, grammatical trials. When grammatical sentences only were considered in the analyses of latency and error data there were no main effects of Rate [F_1 and $F_2 < 1$]. Effects of Context were apparent only in a subject analysis [latency: $F_1(1,78) = 6.072$; $F_2(1,47) = 3.561$, $p > 0.05$; error: $F_1(1,78) = 5.766$; $F_2(1,47) = 3.673$, $p > 0.05$]. There were strong main effects of Plausibility [latency: $\min F'(1,14) = 64.213$; error: $\min F'(1,108) = 26.255$]. The

Context X Plausibility interaction reached only trend status in the latency data by minF' [$\text{minF}'(1,101) = 3.355, p < 0.10$], although it was significant by both the subject analysis [$F(1,78) = 9.113$] and the item analysis [$F(1,47) = 5.830$]. The Context X Plausibility interaction was quite reliable in the error data [$\text{minF}'(1,123) = 7.464$]. Subject means for these conditions are displayed in Table 9.

Table 9: Mean subject response latencies in msec (and percent error) for plausible and implausible targets in grammatical sentences, under differing contextual conditions. Both Lexical Decision Experiments.

Plausibility	<u>Contextual Constraint</u>	
	HIGH	LOW
Plausible	753 (1.9)	822 (4.2)
Implausible	976 (17.7)	969 (10.0)

Newman-Keuls tests on the latency data revealed that subjects' responses to plausible as well as implausible targets did not reflect the differences in contextual constraint [plausible trials: $q(2,78) = 1.75, p > 0.05$; implausible trials: $q(2,78) = 0.18, p > 0.05$]. Newman-Keuls tests on the error data revealed a positive pattern of contextual effects. Subjects made errors equally often on plausible targets under the different contextual conditions [$q(2,78) = 1.69, p > 0.05$]. However, they made more errors on implausible targets following High contexts than Low contexts [$q(2,78) = 5.66$]. Perhaps there was no difference in the latency data for implausible targets as a function of context because subjects had reached a time-out threshold, that is, they were unwilling to consider the items any longer; certainly the error data signify that the

implausible targets were harder to classify as real words following a highly constraining context than following a minimally constraining context.

Ratings, grammatical trials. No analysis was conducted to compare the ratings of plausible and implausible trials for the two experiments. Differences in the ratings instructions between the two experiments probably changed subjects' criteria for rating implausible sentences; subjects rated implausible sentences a "2" more often in Experiment 2 (16% versus 6% for Experiment 3). However, this is not a theoretically interesting result given the change in instructions emphasizing that a "2" rating should be given to plausible sentences only.

Discussion.

Experiment 3. In Experiment 3 the presence of strong effects of integrative context did not support the prediction that contextual processing would be highly vulnerable to an experimental change in the speed of linguistic input. Instead, the results of the Slow Lexical Decision experiment (Experiment 2) were replicated. Main effects of contextual strength as well as grammaticality of the target environment were observed, with no statistical interaction between the two. The main difference between the two experiments resided in the change in subjects' perception of the grammatical status of the target sentences; with the faster rate, subjects were less aware of ungrammaticalities.

Experiments 2 & 3. In a statistical comparison of the two experiments there was no evidence that the effects of sentence-level (inferential) context were diminished with an increase in presentation speed. There was evidence for just such a diminution in the effect of the grammatical environment of a target word as well as strong evidence that under faster speeds of presentation subjects were less aware of ungrammaticalities in the stimulus materials. The pattern of the latency results was contrary to that predicted. The attenuation of the processing effects and the perceptibility of the ungrammaticalities might have been due to our insistence on comprehension during the task; in tasks where subjects are required to attend to the grammatical properties of the materials, awareness of ungrammaticalities is quite high (Sherman, Schweickert and Garrett, work in progress). The hypothesis that subjects' conscious strategies might have been influencing some aspects of task performance is supported by the differences in the correlational analyses performed in the two experiments. With slow speeds of presentation, the processing consequences of a grammatical violation depended, in some small part, on how aware the subjects were of it. With fast speeds of presentation, subjects were generally unaware of the violations; the size of the effect of an ungrammaticality bore no relation to how noticeable it was. It was also the case that although not all of the types of syntactic violation were perceived as ungrammatical equally often, there were no differences in the processing consequences of the types apparent in the latency data.

In neither of the two experiments nor in the combined analysis did we find evidence for the interaction of syntactic and semantic contexts

during sentence processing. This result provides strong support for the autonomy model.

Given that the semantic and syntactic effects survived the rate change, there is little to be said about their dissociability. An explanation of the failure of the time variable to parcel out the two effects could take one of two forms: Either semantic and syntactic processing effects are not dissociable in real time or the speed of presentation used in Experiment 3 is still too slow to effect that dissociation. Almost a third of the subjects who participated in the Experiment 3 had to be excluded from the data analysis because their results did not meet the criteria for acceptable data. If speed of presentation is increased yet again, the task would undoubtedly become too difficult. The limits of our subjects' performance were approached with a presentation speed of 117 msec per word in this methodology. Therefore, we decided to try another approach, that of changing the task to one that might selectively tap earlier processing stages.

Chapter 5

Naming Experiments

5.1 Experiment 4: Slow Name

There is a rapidly growing body of evidence on the different sensitivities of lexical decision and naming experiments to sentential context effects. Essentially, the results from a number of experiments indicate that whereas lexical decision tasks are sensitive to sentential contexts, naming tasks are insensitive to those contextual effects (Forster, 1981; Seidenberg et al., 1984). If syntactic processing is independent of semantic processing, then a change in task might functionally dissociate the two stages. The next two experiments were designed to be parallel to the two lexical decision experiments. The same experimental materials were used, and a rate variation was employed.

Methods

Subjects. Forty-seven volunteers from the MIT subject pool were paid for their participation. One subject's data file was lost; six subjects had error rates beyond a preestablished cutoff value of 10% and were removed from the analysis.

Materials. The stimulus materials described in Chapter 3 were used. Only real-word stimulus sets were used for a total of 72 trials in the experiment. There were 15 practice pairs and two lead-in sentence pairs

which were added at the beginning of the experimental session after the practice break and after the halfway break. The forty subjects were each assigned to one of eight materials lists, therefore, each list was seen by five subjects.

Procedure. The procedure was identical to that of Experiment 3 with the following exceptions: The rate of presentation of the sentences was 200 msec per word with an interval of 400 msec between the two sentences of a pair. The instructions were changed to be appropriate for the naming task. Subjects were told to read the target word (signalled by being written entirely in upper case letters) aloud as soon as they saw it. They were forewarned that at times the word would not "fit in" with the sentence context but they were to read it without thinking about how or if it fit in. A real-time clock was started with the presentation of the target item and was stopped with a voice trigger when the subject began to articulate the target word.

Results.

The results comported well with our predictions about the naming task.

Response latencies and percent error, plausible targets. There was no main effect of Context in either the latency or percent error data (all F values < 1). In contrast to the disappearance of the on-line effects of Context with the change in tasks, there continued to be a healthy (41 msec) main effect of Grammaticality in the latency data [$\ln F'(1,86) = 15.496$]. There was no effect of Grammaticality in the error data [$p > 0.10$]. Moreover, in neither the latency nor the percent error data was there a Context X Grammaticality interaction [F_1 & $F_2 < 1$]. These data are presented in Table 10.

Table 10: Mean subject onset-to-naming latencies in msec (and percent errors) for plausible target appearing in a variety of contextual conditions. Slow Name.

Grammaticality	<u>Contextual Constraint</u>	
	High	Low
Grammatical	595 (1.3)	589 (1.4)
Ungrammatical	630 (2.1)	636 (2.5)

In the item analysis with Kind of ungrammaticality as a nesting factor there were no main effects of Kind nor any interactions involving it in either latency or error data [all F values < 1]. A correlation coefficient was calculated similar to the ones in the lexical decision experiments. The very low coefficients suggest there is no relation in the naming task between awareness of an ungrammaticality and that ungrammaticality's processing effect (subject means: $r = -.07$; item means: $r = 0.01$).

Ratings, plausible trials. Histograms of the rating data are shown in Figure 5. There was a stable effect of Context in the rating data [$\chi^2 = 157.37$; $df = 2$]. In this experiment, although subjects reflected the contextual manipulation in their sentence ratings, their on-line naming performance was unaffected by the degree to which the identity of the naming target was constrained by the context. There was also an effect of Grammaticality [$\chi^2 = 297.82$; $df = 2$]. Comparison of the histograms from Experiments 3 & 4 (Figures 4 & 5; Figure 4 may be found on page 80) show marked similarity in the rating patterns of plausible, grammatical trials [$\chi^2 = 3.47$, $p > 0.10$; $df = 2$]. It seems that the training

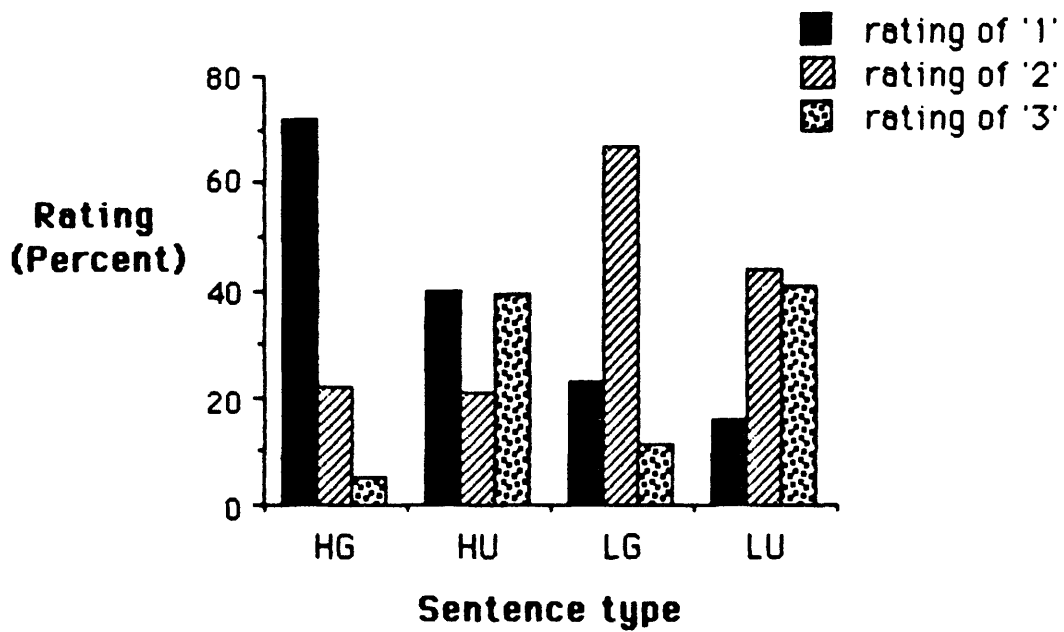


Figure 5: Percentage of subjects' discourse ratings for each sentence type. Experiment 4: Slow Name. (H: High Context sentence; L: Low Context sentence; G: Grammatical Target sentence; U: Ungrammatical Target sentence.)

procedure on the use of the rating scale, instituted in Experiment 3, decreased subjects' reliance on the rating of "1" for any normal sentence. Subjects differentiated the contextual types with their ratings. Thus, even in the absence of differences in the latency or error data there is still an indication that subjects were attending to and fully processing the sentence context.

An analysis of ratings by type of syntactic violation revealed that the grammatical sentences from the three item sets were not rated differently [$\chi^2 = 3.372$, $p > 0.10$, $df = 4$]. The ratings for the ungrammatical sentences did vary, however [$\chi^2 = 49.104$; $df = 4$]. The addition of adverb type was often correctly rated as ungrammatical (60% of ungrammatical trials were rated "3") followed by alteration of an article or pronoun (38%) and omission of preposition (23%). In the naming paradigm as well, it is apparent that some types of ungrammaticality were more noticeable than others, yet the effects on naming latency did not differ as a function of Kind.

Response latencies and percent error, grammatical sentences. In the group of analyses over grammatical trials testing for effects of Context and Plausibility there was no main effect of Context in either the latency data [F_1 & $F_2 < 1$] or the error data [F_1 & $F_2 < 1$]. Variations in the Plausibility of the target item significantly affected latency patterns [$\min F'(1,86) = 5.945$]. Subjects were slow to name implausible targets relative to plausible targets, although the 25 msec difference is much smaller than that seen in the lexical decision experiments. Naming errors tended to be made more often on implausible trials than plausible

ones, although this trend did not reach reliable levels of significance in either subject or item analyses [$F(1,39) = 3.928$; $F(1,45) = 3.092$, $p < 0.10$]. The latency and error rate data are listed in Table 11.

Table 11: Subject means of response latency in msec (and percent error) to plausible and implausible targets in grammatical sentences following different degrees of contextual support. Slow Name.

Plausibility	<u>Contextual Constraint</u>	
	High	Low
Plausible	595 (1.3)	589 (1.4)
Implausible	620 (2.4)	615 (3.3)

Ratings, grammatical trials. The effect of plausibility on subjects' ratings was significant. Subjects differentially rated contextual types containing plausible targets as "1" or "2", but rated most (91%) implausible trials a "3" [$\chi^2 = 627.23$; $df = 2$].

Discussion. The slow rate of presentation in the naming task produced results which fit in well with the literature on differential sensitivities of tasks to contextual types. There was no effect of the strength of the preceding context on the onset-to-naming latencies or error rates to plausible targets even though subjects' post-sentence ratings proved that they were aware of contextual differences. In contrast, the effect of grammatical context on the naming of plausible targets was very secure in both the latency and rating data. When the ratings for the ungrammatical trials were classified as either "grammatical" ("1" or "2") or "ungrammatical" ("3") and a χ^2 calculated

comparing the Slow Lexical Decision experiment with the Slow Name experiment, there was a significant difference. Subjects seemed a bit more aware of the ungrammatical sentences in this experiment than in the Slow Lexical Decision experiment, rating 40% of them as a "3" compared to 31% . However, all the evidence suggests that in the naming task there was no conscious contribution to the magnitude of the processing effect of the ungrammaticalities. Although some types of ungrammaticality were more noticeable than others (as measured by the rating scale), there were no corresponding latency differences among the kinds of ungrammaticality. In the naming task as well as the lexical decision task there was a discrepancy between linguistic processing and conscious awareness.

When only grammatical trials were considered, the naming latencies to implausible targets were slow relative to plausible ones. The magnitude of the effect seemed much smaller than those found in the other experiments. The results from this experiment show that even when the naming task is not sensitive to some semantic effects, it can be sensitive to others. Discussion of this point is delayed until the last experiment has been presented.

5.2 Experiment 5: Fast Name

In this experiment the same rate variable that was applied to the lexical decision task was applied to the naming task. It was expected that the rate change would not alter the basic pattern of results found

in Experiment 4; it was predicted that, if anything, awareness of the ungrammatical sentences would drop, with perhaps an attenuation of the grammaticality effect on response latency.

Methods

Subjects. Forty-nine volunteers from the MIT subject pool were paid for their participation. They were all native English speakers, had normal (corrected) vision and were under 35 years of age. Nine subjects' data were excluded from the analysis: One forgot his glasses, four subjects' data were incompletely written to disk after the experiment, and four subjects had over 10% errors, which was beyond the pre-established cutoff value.

Materials. The materials used in Experiment 4 were used here, with no changes. The forty subjects were each assigned to one of eight materials lists, therefore five subjects saw each list.

Procedure. The procedure was identical to that of Experiment 4 with the exception of the presentation speed. The sentence pairs were shown at a rate of 117 msec per word, with a 234 msec interval between the two sentences of a pair.

Results.

The predictions were not borne out completely.

Response latencies and percent error, plausible trials. The data for Plausible trials are presented in Table 12.

Table 12.
Mean subject onset-to-naming latencies in msec (percent error) for plausible targets appearing in different contextual conditions.
Fast Name.

Grammaticality	<u>Contextual Constraint</u>	
	High	Low
Grammatical	583 (0.4)	594 (4.2)
Ungrammatical	605 (1.7)	627 (2.5)

Subjects were an average of 17 msec faster to name a target following a strong context than following a weak one. This small difference was significant on both subject and item analyses, but did not attain significance in the minF' computation [F1(1,39) = 8.472; F2(1,47) = 4.446]. Supporting the latency difference, there was an effect of Context in the error data revealing that fewer errors were made on targets following High than Low contexts (1.0% versus 3.3%) [F1(1,39) = 4.921; F2(1,47) = 4.160]. The 27 msec main effect of Grammaticality was reliable in the latency data [minF'(1,85) = 5.881]. Effects of Grammaticality were not apparent in an analysis of error rate [F1 & F2 < 1]. There was no indication of a statistical interaction of Context X Grammaticality in the latency data [F1 & F2 < 1]. There were no main effects of Kind in the item analyses that included it as a nesting factor, nor did Kind interact with any of the other factors. This was true for both the latency (all F values < 1) and error analyses (all p values > 0.10).

Ratings, plausible trials. A sizeable main effect of Context was found in the rating data [$\chi^2 = 207.71$; df = 2]. The effect of Grammaticality was still significant [$\chi^2 = 11.80$; df = 2] although as Figure 6, the rating histogram, shows, subjects were rating only a small (12%) proportion of trials as ungrammatical. The types of syntactic violation had a familiar effect on the rating patterns. The grammatical sentences for the different types were rated equivalently [$\chi^2 = 5.283$, p > 0.05; df = 4]. The rating scores for the ungrammatical sentences varied according to the type of ungrammaticality, [$\chi^2 = 17.875$, p < 0.01; df = 4], with the same ordering of noticeability that was found in the other three experiments. The percentages of "3" ratings for each of the item

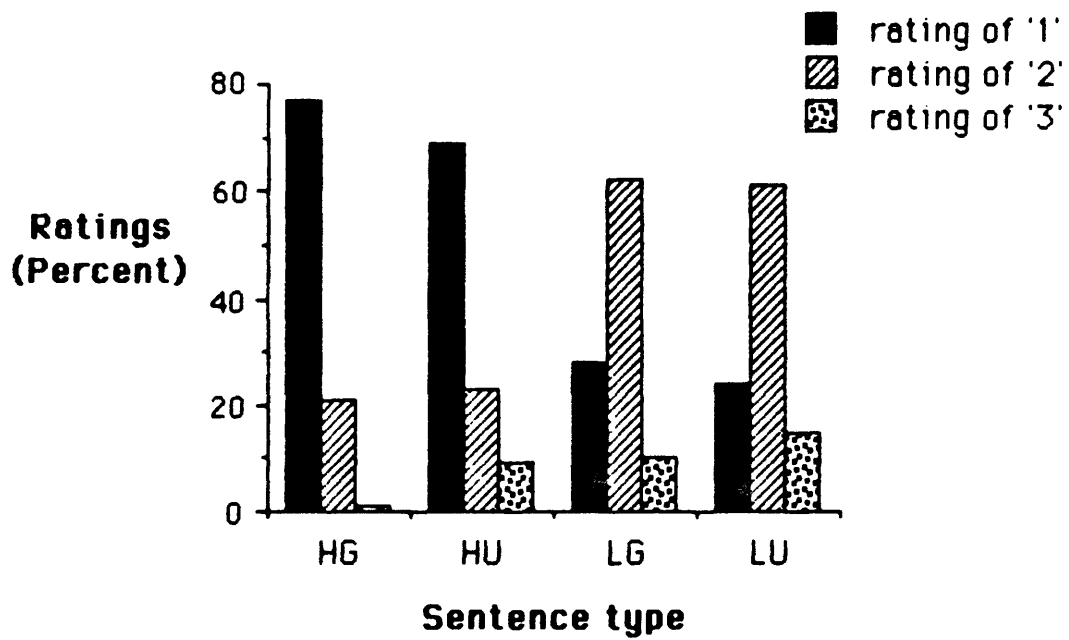


Figure 6: Percentage of subjects' discourse ratings for each sentence type. Experiment 5: Fast Name. (H: High Context sentence; L: Low Context sentence; G: Grammatical Target sentence; U: Ungrammatical Target sentence.)

types were: Addition of adverb, 20%; alteration of an article or pronoun, 11%; and omission of preposition, 5%. A correlation coefficient was computed for subject and item means to examine the relation between the noticeability of an ungrammaticality and its effect on response latency. As in the previous naming experiment, the correlation coefficient was extremely low (subject: $r = -.12$; item: $r = -.05$). Here again is the pattern of a strong effect of grammaticality, without a correspondingly strong awareness by the subjects.

Response latencies and error rates, grammatical trials. When the analyses were performed on only the grammatical trials to examine the effects of the Context and Plausibility manipulations, there were no main effects of Context in the latency data ($p > 0.10$). There was a small effect in the error data, but this was only statistically apparent in the item analysis [$F_1(1,39) = 1.545, p > 0.10$; $F_2(1,47) = 5.062, p < 0.05$]. Subjects were slower (by 62 msec) to name implausible targets than plausible targets [$\text{min}F'(1,83) = 14.549$] and although the error pattern seemed to mirror the latency data (5.6% error on implausible targets versus 2.3% on plausible targets) this was significant only in an analysis of subjects [$F_1(1,39) = 9.043$] and not items [$F_2(1,47) = 3.450, p < 0.10$]. The data for the grammatical trials are presented in Table 13.

Table 13: Mean subject response latencies in msec (and percent errors) for plausible or implausible targets in grammatical sentences under different contextual conditions. Fast Name.

Plausibility	<u>Contextual Constraint</u>	
	High	Low
Plausible	583 (0.4)	594 (4.2)
Implausible	665 (4.6)	635 (6.7)

The Context X Plausibility interaction in the latency data was significant in a subject-based latency analysis [$F(1,39) = 5.733$] but just missed significance in an item-based analysis [$F(1,47) = 3.934$, $p < 0.10$]. Newman-Keuls on the subject means revealed that although plausible targets were named equally fast under the two strengths of context [$q(2,39) = 1.307$, $p > 0.05$], implausible targets following High context were named more slowly than the same targets following Low context [$q(2,39) = 3.208$].

Ratings, grammatical trials. As in the other experiments, there was a strong Plausibility effect in the rating data [$\chi^2 = 691.54$; $df = 2$].

Discussion.

It was expected that contextual manipulations would not influence onset-to-naming latencies. Contrary to expectations, there were undercurrents of contextual effects in these data; on subject and item analyses context did reach significance in both latency and percent error data, with subjects having relatively short latencies and making relatively few errors on plausible targets following very supportive

contexts. The convergence of these two pieces of information suggests that the context effect observed is probably reliable. It suggests that lexical decision tasks might not be different in principle from naming tasks. That is, when there is a heavy emphasis on comprehension, as in this experiment, integrative processing might influence the naming task as well as the lexical decision task. And relevant to the question of the additivity of sentence-level semantic and syntactic contexts, there were no traces of an interaction between these factors. Thus, given that effects of context may have been present in naming, they were operating independently of syntactic effects.

In the analysis of grammatical trials there was no effect of context, although there was a strong (62 msec) plausibility effect.

With the results from Experiment 4 in hand, the results from the Experiment 5 are even more puzzling. A methodological assumption was that increasing the speed of stimulus input decreases the proliferation of higher-order representations. Therefore, effects dependent upon those representations should suffer. The context effect did not suffer, rather the opposite occurred. To examine more closely the two naming experiments a joint analysis was performed which included Rate of presentation as a factor.

5.3 Both Name Experiments

Response latencies and percent error, plausible trials. In analyses

over plausible trials there were no significant effects in any of the error data. In the latency data there was only a significant effect of Grammaticality [$\text{min}F'(1,105) = 18.538$], not effects of Rate or of Context or of a Rate X Context interaction (all p values > 0.10) although these latter effects involving Context were significant on a subject analysis [$F(1,78) = 4.393$ and $F(1,78) = 4.071$ respectively]. That is, subjects were 9 msec faster to respond to targets after High contexts than after Low contexts. However, when the rate of presentation is taken into account, context affects response latencies only under the faster rates of presentation. There was no Context X Grammaticality interaction effect in either subject or item analyses (both p values > 0.10).

Ratings, plausible trials. The ratings for the grammatical trials from the two experiments were compared to check for differences in subjects' use of the rating scale under different speeds of stimulus presentation. The results of the chi-square analysis [$\chi^2 = 7.47$; $df = 2$] indicated that subjects did rate grammatical sentences in the naming experiments a bit differently. There seemed to be a greater tendency for grammatical sentences to be rated a "3" in the Slow Name experiment (10%) than in the Fast Name (6%). There was also some indication that subjects in the Fast Name experiment used a "1" rating more often than subjects in the Slow Name experiment (53% versus 47%).

The ratings for the ungrammatical trials were classified as either grammatical ("1" or "2") or ungrammatical ("3") and a chi-square statistic calculated comparing the two naming experiments. There was a strong effect of Rate [$\chi^2 = 87.01$; $df = 2$]. A higher proportion of

ungrammatical trials were rated as such when the sentences were presented more slowly (40% for the Slow Name experiment; 13% for the Fast Name experiment).

Response latencies and percent error, grammatical trials. In a set of analyses over grammatical trials the combined statistic revealed no main effects of Rate for any of the dependent measures. However, there were a couple of Rate effects in analyses based on item means that were significant: i.e., latency [$F(1,47) = 5.338, p < 0.05$] and error rate [$F(1,47) = 4.380, p < 0.05$]. Items were named more quickly and accurately at slower rates of presentation. Context was not significant in the latency data, but was significant in the error data by both subject and item analyses [$F_1(1,78) = 5.827; F_2(1,47) = 4.083, p < 0.05$]. There were main effects of Plausibility evident in both analyses: latency [$\text{min}F'(1,91) = 17.787$]; and error [$\text{min}F'(1,88) = 4.243$]. The only interaction to reach significance in the latency data was that of Rate X Plausibility [$\text{min}F'(1,114) = 4.566$]; these data are shown in Table 14.

Table 14: Mean subject onset-to-naming latencies to plausible and implausible targets in grammatical sentences under different rates of sentence presentation.

Plausibility	<u>Rate of Sentence Presentation</u>	
	Slow (200 msec)	Fast (117 msec)
Plausible	592 (2.3)	589 (1.3)
Implausible	617 (5.6)	650 (3.2)

There was a much larger effect of Plausibility when Rate of presentation

is faster (61 msec versus 25 msec). The similarity of the means of the plausible targets is striking; by a Newman-Keuls test the means of the plausible targets were equivalent [$q(2,78) = 0.51, p > 0.10$]. In responding to implausible targets, subjects were significantly slower under the fast speed of presentation than under the slow speed [$q(2,78) = 5.60, p < 0.01$]. In other words, subjects in the Fast Name experiment exhibited more interference in their responses to implausible targets than did subjects in the Slow Name experiment.

Ratings, grammatical trials. The rating data also revealed a difference in the processing of implausible trials as a function of rate. When the data from the implausible trials were divided into plausible ("1" or "2") or implausible ("3") ratings, it was apparent that grammatical, implausible trials were rated as such more often under the faster speed of presentation than under the slow speed (96% versus 88%, respectively) [$\chi^2 = 8.30; df = 1$].

Discussion.

The weak main effect of Context was not enhanced by analyzing the two naming experiments together. We conclude that sentence-level semantic context has no reliable influence on naming latencies, contrary to the strong influence of syntactic context. We believe that the effects of these two contexts would not be differentially affected by a change in task if they were operating at the same level in the language process. The pattern of results in the lexical decision and naming experiments shows that sentential-level semantic and syntactic contexts are subserved by autonomous mechanisms.

Chapter 6

General Discussion and Summary

Discourse constraints and syntactic processes.

The preceding experimental series tested two opposing theories of linguistic processing. The theories differ in their conceptualization of processing stages and information flow during language comprehension. One class of models, interaction models, assumes that there are no processing levels beyond the initial form-driven determination of a pool of candidates for word recognition and a final level of discourse representation that is not linguistically specific (i.e. it is a general cognitive representation). Interaction models postulate unrestricted information flow throughout the processing system and, in fact, demand that any relevant information (potentially any and everything a person knows) be brought to bear upon the comprehension process as early as possible. The other class of models, autonomy models, compartmentalizes processing. This class assumes that although many different domains of knowledge (including extralinguistic knowledge) contribute to language processing, the language and general cognitive systems are distinct. That is, independent subprocessors are dedicated to the formation of linguistic representations. Several possible levels of analysis are contemplated and the flow of information among them is highly

constrained.

The presence of processing stages during language comprehension was tested using Sternberg's (1969) additive factors logic. The strategy was to demonstrate separate effects of different types of contextual manipulation during sentence processing. Sternberg's logic provides a framework for the interpretation of experimental results when more than a single type of contextual variable is present in the stimulus materials. To the extent that these experimental manipulations correspond to distinct informational types, they should, on the autonomy models' assumption, tap separate processing stages. If the variables do affect independent processing stages, then when they are multiply present, their effects will be independent, i.e., their combined effect will equal the sum of the individual effects. But if the experimental variables affect the processing of one common stage, as they must on the interactive model's assumption that there are no distinct levels in the language processing mechanism, then their effects will not be additive. This result should be empirically evident as a statistical interaction.

The results of five experiments refute the claim for processing interaction. The materials for the experiments contained three types of contextual manipulation: The first type of manipulation was in discourse context. This variation used interpretive frames that affected the predictivity of critical "target" words.

In Experiment 1, a Timed Cloze experiment, the discourse variation affected subjects' generation of suitable target words. When the context was highly predictive, it facilitated performance;

subjects were faster to generate a response and their responses matched the intended target word more often after the stronger context than after the weaker context.

In Experiments 2 and 3, the lexical decision experiments, plausible targets were responded to faster following strongly predictive as opposed to weak contexts. An off-line measure, a rating scale, was used to assess subjects' conscious ability to discriminate the contextual types. Subjects were quite successful in differentiating strong context from weak context.

Experiments 4 and 5 were naming variants of the lexical decision tasks. Here the on-line dependent measures of response latency and percentage error were much less sensitive to the discourse variation. This was true even though the off-line measure did not vary when the experimental paradigm was changed from lexical decision to naming. That is, subjects' ratings of the discourse variation were unchanged.

The second type of materials manipulation varied syntactic context. Changes in closed-class words caused ungrammaticalities immediately prior to the presentation of the target words.

In Experiment 1, the Timed Cloze experiment, the presence of an ungrammatical frame resulted in longer generation times.

In Experiments 2 and 3, the lexical decision experiments, the result of the ungrammatical context was to slow subjects' responses relative to the same targets in grammatical contexts. The off-line

ratings measure showed the syntactic variation to be detectible, but less so than the discourse variation.

In Experiments 4 and 5, the naming experiments, the syntactic variables again showed significant effects. Ratings were equivalent to those observed in the lexical decision task.

The third type of materials manipulation was one of plausibility. An alternate set of target words was constructed to match the original set as much as possible in frequency, length and phonotactic structure. This set of implausible targets appeared in the same discourse and syntactic environments as the original set of plausible targets.

In all experiments (except Experiment 1 which did not include a plausibility variation), implausible targets were harder for subjects to respond to; response latencies and the percentage of errors increased for the implausible targets. Implausible targets were also highly noticeable; subjects consistently rated them as implausible.

The manipulation of discourse context and syntactic context was the crux of the test of the processing models. The major assumption underlying the use of the additive factors logic was that a dependent measure sensitive to changes in the independent variables would also be sensitive their interaction.

In Experiments 2 and 3, robust effects of (discourse) Context and Grammaticality were observed without any concomitant interaction. This is certainly a result which is counter to the

assumptions of interactive models of language processing.

The naming paradigm was used to provide convergent evidence on the question of independent processing stages. Naming was chosen because it is purported to be immune to the influences of discourse-level effects. (Whether this is true or not is immaterial to the logic of the next step, but see the discussion on task differences below.) If a change in the experimental design differentially affects discourse and syntactic processes, then it may be assumed that those processes are independent, that they correspond to different levels in the language analysis.

The naming paradigm was not sensitive to the discourse manipulation for the latency data in Experiment 4 (Slow Name). The rating data, however, demonstrated that subjects were just as aware of the discourse manipulations as the subjects were in the lexical decision experiments. There were indications of a sensitivity to the discourse variable in Experiment 5 (Fast Name). Both experiments 4 and 5 showed strong effects of the grammatical variable. But again, there was no hint of an interaction between Context and Grammaticality.

Over all, the conditions of this experimental series should have permitted the detection of effects of interaction: lexical decision is a task presumed to be open to all kinds of contextual influences; the design of the task allows real-time processing to be investigated; the stimulus construction successfully embodied different types of contextual variation. The uniform absence of an interaction effect strongly favors a conception of semantic and syntactic processing as being subserved by

separate and independent stages in the language process; this is the main conclusion to be drawn from these experiments. There are other issues to be discussed, however.

Effects of plausibility.

The inclusion of the plausibility manipulation was motivated more by sheer curiosity than by detailed theory; it was intended to provide another perspective on the semantic contextual manipulation and a fail-safe demonstration of semantic processing effects. At first blush, the additive factors logic is applicable to an investigation of the locus of the (in)plausibility effect: Do the processing consequences of discourse manipulations and plausibility manipulations stem from the same stage in the analysis?

The interaction position is, naturally, that there is a common stage for these two types of semantic processes. The autonomy position on this issue is not as straightforward as it was on the examination of discourse and syntactic context. If semantic processing is seen as being solely interpretive, then all semantic effects should arise from a common stage. However, if there is a "local" semantic representation, e.g. a level of analysis of the meaning of the sentence itself without reference to the discourse or conceptual relations, then it is possible that one could find evidence of two types of semantic effects.

The results from these experiments do not provide a clear answer to an investigation of plausibility. The working predictions were that a) implausible targets should be harder to reject than plausible targets and

b) if there is any interaction, it should be that implausible targets are harder to process under the highly constraining context than under the more general context. Notice that this latter prediction removes the plausibility manipulation from the domain of additive factors logic. That is, the normal effect of strong context is to facilitate processing. But when there is an implausibility, the strong context works against the interpretation; the direction and perhaps magnitude of the context effect is now changed. If the effect of context is no longer equivalent, then it cannot be entered into the additive factors logic. Nonetheless, if prediction b) is supported, then the hypothesis that there is a level of semantic analysis separate from the conceptual level must be discarded.

In the experiments, prediction a) was always borne out. The support for prediction b) was not nearly as neat.

In Experiment 2 (Slow Lexical Decision) there was no interaction of the Context and Plausibility factors in the latency data, although there was in the error data. It is of note that errors followed the predicted interaction pattern even when the latency data did not.

In Experiment 3 (Fast Lexical Decision) there was an interaction, although in the latency data the predicted pattern of greater inhibition following High context compared to Low context did not occur. As in Experiment 2, the error data did follow the predicted pattern.

In any lexical decision task it is always possible that when response latencies are long, subjects may use a "time-out" strategy. That is, they may wait only so long before they execute a response, even if they are unsure of their response. Under this scenario the predicted pattern of results could match those obtained: Subjects' decision times to implausible targets could very well be equivalently long in the two discourse conditions, although error rates should not be. The alternative explanation is that local plausibility and discourse are processed independently and so the decision times to an implausible target would be constant under different contextual conditions.

A "time-out" strategy effect should be less apparent in a naming task; a response can only be executed after contact with a lexical representation.¹⁵

In Experiment 4 (Slow Name), there was no evidence of any Context X Plausibility interaction in either latency or error data.

However, in Experiment 5 (Fast Name) there was an interaction in an analysis of the latency data over subject means, but not over item means. This interaction followed the predicted pattern: onset-to-naming latencies were longer for implausible targets following High Context than for the same targets following Low context. The error data from the naming experiments were, in general, uninformative.

In summary, the mixed patterns in the data which resulted from the plausibility manipulation are not easily explained. The only clear result was

15. This is begging the question of rule-generated as opposed to lexically-based responses.

a simple one: implausible targets consistently caused processing difficulties.

Task differences and subject strategies

Making a lexical decision, deciding whether a letter string is a real word or not, would seem to be a fairly uncomplicated task. All that is required is that the stimulus match an entry in lexical memory. For some time in psycholinguistics, variations in lexical decision times were seen as being a function of the ease of contacting a lexical entry (contributions of stimulus encoding and response execution to the overall latency were seen as being invariant within a particular experiment). In recent years, however, it has become apparent that subjects do not rely on the minimum amount of processing to accomplish the lexical decision response. The prevailing opinion is that most variations in lexical decision times are a consequence of "post-access" processing; after the lexical entry has been contacted ("accessed") it is believed to be immediately integrated into whatever developing linguistic representation(s) are present and/or is entered into a "decision" stage which is separate from the language mechanism. What is relevant to this discussion is not the changing characterization of the lexical decision task per se, but rather its relationship to another paradigm, the naming task. (For discussions of both the lexical decision and naming tasks as well as their relationship to one another, see Forster, 1981; Seidenberg and Tanenhaus, in press). Among the furor over the lexical decision task, the naming task has been viewed as a "pure" measure of lexical access, uncontaminated by "decision" level processing because there is no decision to be made. In addition, naming is thought to be based on the level of lexical output; the

naming response is determined by the lexical entry. Naming is typically faster than lexical decision. In the lexical decision task, responses are not determined by the content of the lexical entry; rather the fact that contact has been made is used to drive the response. Presumably, this awareness, which is not a usual part of the language process, takes time to develop. During the delay, the lexical entry that has been contacted has ample time to enter the developing meaning representation and this process may actually be able to direct a response sooner than the awareness of a successful access. Because naming taps a natural language process, there is no intrinsic delay in its execution. Therefore, the articulatory information in the lexical entry is used before the entry becomes involved in integrative processes. The implication is that naming is, in principle, different from lexical decision. The naming response is based purely upon language, and specifically lexical, processes while lexical decision is influenced by both linguistic and extralinguistic processes. Moreover, even within the influence of linguistic processes, the naming response is seen as a result of access processes whereas lexical decision is a product of mostly post-access processes.

The results from the naming experiments in this thesis do not support such a strong and principled distinction between lexical decision and naming tasks (see also Schweickert and Kroll, in preparation; Hodgson, 1985). First of all, there were strong effects of two contextual manipulations which most likely influence processes beyond the lexical level: those were variations in syntactic environment and plausibility. Moreover, with the fast speed of presentation (117 msec), naming times did vary as a function of the discourse manipulation, a result clearly in conflict with earlier studies (e.g., Forster, 1981). It is interesting to note that the onset-to-naming latencies

for the fastest condition (plausible, grammatical trials) were virtually identical across the two naming experiments (592 msec for Slow Name; 589 msec for Fast Name). Therefore an explanation of the presence of discourse-level influences on performance in the Fast Name experiment cannot be readily based on a slowness in the development of a response.

An alternative explanation is based on subject strategies. There was a heavy emphasis on comprehension in all of these experiments. Perhaps the integrative processes were relied on more heavily in these experiments than in other researchers' experiments. If this is the correct explanation it is especially interesting that integrative processing does not interact with syntactic processing in these experiments. But whatever the explanation, since discourse and syntactic effects emanate from different stages in the comprehension process at least one of them must be post-lexical; naming therefore is sensitive to some post-lexical processes.

Dissociation of ratings and latency effects

On-line dependent measures such as response latencies yield clues to the operation of real-time processes. In these experiments, an off-line measure of a rating scale was also used to explore the conscious products of such processes. The results from the two types of measures did not mirror one another.

The discourse variation was consistently reflected in the rating scores for plausible targets. This was true both when that variation was apparent in the latency data (Experiments 2, 3 and 5) and when it was absent (Experiment 4). That is, subjects' perception

of the discourse variation was constant across conditions in which processing consequences of the variation changed.

In contrast, the grammaticality variation was not always manifest in the rating scores, even though all of the experiments were sensitive to the syntactic manipulation. Subjects differentiated ungrammatical from grammatical sentences only when the presentation rate was relatively slow (200 msec; Experiments 2 and 4). Moreover, even though the different types of ungrammaticality were rated differently in these experiments, their processing consequences were equivalent. Under a faster rate of presentation (117 msec) the awareness of the ungrammaticalities was greatly diminished although the processing effects were still robust (Experiments 3 and 5).

There was very good correspondence between the ratings of implausible trials and the processing consequences of implausible targets. Implausibilities were extremely noticeable and had dramatic effects on processing in all experiments.

The dissociation of processing and the conscious awareness of that processing is especially interesting to find in the case of syntactic variation. It also makes sense in terms of the goals of the language processing system. If comprehension is the goal, then processing hurdles are immaterial to the final analysis, as long as they are overcome. Conscious awareness of local difficulties could only detract from the achievement of comprehension.

Rate variation

The manipulation of the speed of presentation never yielded predicted results. It was expected that the effects of semantic context variations would disappear under fast speeds of presentation. This did not happen. Perhaps this expectation was unrealistic, given the emphasis on comprehension in the experiments. Strategic or conscious influences on processing do appear to have slightly affected the magnitudes of the effects of the experimental manipulations. There were statistical differences in magnitude as a function of speed of presentation for the Grammaticality effect in the lexical decision experiments and the Plausibility effect in the naming experiments. Recall that as the speed increased, conscious awareness of the ungrammaticalities decreased, as measured by the rating scale. However, the processing effect of the ungrammaticalities was strong at all rates of presentation. In the lexical decision experiments, but not the naming experiments, the magnitude of the Grammaticality effect diminished with the faster rate of presentation. Conversely, in the naming experiments, but not the lexical decision experiments, with the faster presentation rate the magnitude of the Plausibility effect increased. The effect of discourse (Context) was apparent at both rates in the lexical decision experiments, but was only apparent in naming when the stimuli were presented at the fast rate.

Summary Statement

In this series of experiments, effects of plausibility, discourse context and syntactic integrity have been shown to influence linguistic processing. Using experimental tasks stressing comprehension, variations

in the speed of stimulus presentation had only a limited impact upon the pattern of results. Changing the task from lexical decision to naming did affect the experimental variables in different ways, although there was evidence that naming, as well as lexical decision, can be sensitive to post-access processing. In each experiment there was clear evidence in support of the view that syntactic processing is distinct from discourse-level processing. These results are contrary to claims for interactive processes during language comprehension; such a pattern of results better comports with a model of language comprehension which embodies autonomous, special-purpose subprocessors within the language processing mechanism. The results of the experiments may be succinctly put as follows:

1. Discourse level constraints produced robust effects on lexical decision performance.
2. The same discourse constraints did not show strong or uniform effects for a naming task. There is, however, evidence to suggest a claim for the sensitivity of naming measures to supralexic variables.
3. Grammatical violations produced strong effects in both lexical decision and naming tasks.
4. A proportional stimulus presentation rate change of .6 did not have a marked impact on patterns of performance for the experimental variables.
5. Awareness of the discourse levels of constraint and the syntactic violations was not closely linked to effects of those variables on

response latencies.

6. Whenever the variables of discourse context and syntactic violation yielded strong main effects there was never any indication of an interaction; the effects were additive.

Appendix A

Stimulus Materials

A.1 Experimental Trials

Each sentence group is arranged as follows:

- n. High context sentence.
- Low context sentence.
- Plausible, grammatical target sentence.
- Plausible, ungrammatical target sentence.
- Implausible, grammatical target sentence.
- Implausible, ungrammatical target sentence.

The target word is entirely written in uppercase letters.

SENTENCE SETS ARE ARRANGED BY TYPE OF SYNTACTIC MANIPULATION.

-----omission of preposition-----

1. Police statistics show that this is an unsafe neighborhood.
Government statistics show that this is an unfortunate trend.
The amount of CRIME has increased dramatically in the last few years.
The amount CRIME has increased dramatically in the last few years.
The amount of GOLF has increased dramatically in the last few years.
The amount GOLF has increased dramatically in the last few years.
2. Marcy tries not to give in to her alcoholic cravings.
Marcy tries not to give in to her compulsive desires.
But she's so fond of WINE that she can't stop drinking it.
But she's so fond WINE that she can't stop drinking it.
But she's so fond of DUST that she can't stop drinking it.
But she's so fond DUST that she can't stop drinking it.
3. Astronauts are pessimistic about the future of their program.
Federal officials are pessimistic about the future of many programs.
The government has cut back the funding for SPACE research in the new budget.

The government has cut back the funding SPACE research in the new budget.
The government has cut back the funding for VOICE research in the new budget.
The government has cut back the funding VOICE research in the new budget.

4. We're hoping to spend some time skiing in Vermont this winter.
We're hoping to spend some time vacationing with the family this year.
But unless there's a lot of SNOW it won't be much fun.
But unless there's a lot SNOW it won't be much fun.
But unless there's a lot of GOLD it won't be much fun.
But unless there's a lot GOLD it won't be much fun.

5. There is some talk that fuel shortages may have wide effects on the economy.
There is some talk that factories may be forced into a four day work week.
If the supplies of OIL keep decreasing, the problems will only get worse.
If the supplies OIL keep decreasing, the problems will only get worse.
If the supplies of FILM keep decreasing, the problems will only get worse.
If the supplies FILM keep decreasing, the problems will only get worse.

6. My friends must call the exterminator again.
My friends are thinking of moving again.
The large number of BUGS in their apartment is driving them crazy.
The large number BUGS in their apartment is driving them crazy.
The large number of HIPS in their apartment is driving them crazy.
The large number HIPS in their apartment is driving them crazy.

7. Great Britain commanded one of the largest naval fleets in the world before World War II.
South America attempted to export more coffee to the western countries after World War II.
But recently the number of SHIPS used in that capacity has decreased.
But recently the number SHIPS used in that capacity has decreased.
But recently the number of PAINS used in that capacity has decreased.
But recently the number PAINS used in that capacity has decreased.

8. Seafood can be an excellent source of protein.
Food can be prepared in interesting ways.
But some people think that any kind of FISH is disgusting
But some people think that any kind FISH is disgusting.
But some people think that any kind of BOMB is disgusting.
But some people think that any kind BOMB is disgusting.

9. Their contractor said he'd have the roof completed two days ago.
Their handyman said he'd have the job completed two days ago.
But, it seems that he can't find the right kind of SHINGLES for the job.
But, it seems that he can't find the right kind SHINGLES for the job.
But, it seems that he can't find the right kind of CRATER for the job.
But, it seems that he can't find the right kind CRATER for the job.

10. Sidney really enjoyed working in the greenhouse last summer.
Sidney really enjoyed working for the town last summer.
He found that taking care of PLANTS was very relaxing.
He found that taking care PLANTS was very relaxing.

He found that taking care of TESTS was very relaxing.
He found that taking care TESTS was very relaxing.

11. A growing number of people are moving from urban areas to the country.

A growing number of people are searching for better places to live.
But some people refuse to move out of CITIES despite their drawbacks.
But some people refuse to move out CITIES despite their drawbacks.
But some people refuse to move out of WATERS despite their drawbacks.
But some people refuse to move out WATERS despite their drawbacks.

12. Pharmaceutical companies are required by law to run each of their new products through extensive tests.

Manufacturers are required by law to run each of their new products through extensive tests.
Even so, a number of DRUGS are not entirely safe.
Even so, a number DRUGS are not entirely safe.
Even so, a number of WINGS are not entirely safe.
Even so, a number WINGS are not entirely safe.

13. Nate bought a very expensive briefcase while he was in the city.

Nate bought a very expensive outfit while he was in the city.
It was made out of LEATHER that had been imported from Italy.
It was made out LEATHER that had been imported from Italy.
It was made out of HONEY that had been imported from Italy.
It was made out HONEY that had been imported from Italy.

14. Freshly squeezed juice is a real treat.

Seasonal foods are a real treat.
But if the cost of ORANGES increases, it/they will soon be too expensive.
But if the cost ORANGES increases, it/they will soon be too expensive.
But if the cost of BISHOPS increases, it/they will soon be too expensive.
But if the cost BISHOPS increases, it/they will soon be too expensive.

15. John left the confessional feeling much better.

John left the meeting feeling much better.
Whenever he spoke to a PRIEST his problems seemed to disappear.
Whenever he spoke PRIEST his problems seemed to disappear.
Whenever he spoke to a CHORUS his problems seemed to disappear.
Whenever he spoke CHORUS his problems seemed to disappear.

16. Ross was mailing a heavy sweater to his sister in Canada.

Ross was carrying a heavy package to his sister in Boston.
It had about ten pounds of YARN in it and was very bulky.
It had about ten pounds YARN in it and was very bulky.
It had about ten pounds of HAWK in it and was very bulky.
It had about ten pounds HAWK in it and was very bulky.

-----addition of adverb-----

1. Carol saw the mechanic drop his equipment as he ran from the garage.
Carol saw some people working in the public park as she drove by the entrance.

Later, she found some TOOLS lying on the sidewalk.
Later, she found some really TOOLS lying on the sidewalk.
Later, she found some MILLS lying on the sidewalk.
Later, she found some really MILLS lying on the sidewalk.

2. They think that the sale of firearms ought to be restricted in this country.

They think that the sale of certain things ought to be restricted in this country.

If anyone can buy a GUN the number of accidental deaths will increase.
If anyone can buy a very GUN the number of accidental deaths will increase.
If anyone can buy a POOL the number of accidental deaths will increase.
If anyone can buy a very POOL the number of accidental deaths will increase.

3. Jane has had to call in a plumber twice because of the flooding.

Jane has had to complain to her landlord because of his negligence.
She's had incredible amounts of trouble with the PIPES in the last few weeks.
She's had incredible amounts of trouble with the newly PIPES in the last few weeks.
She's had incredible amounts of trouble with the FORKS in the last few weeks.
She's had incredible amounts of trouble with the newly FORKS in the last few weeks.

4. When Delila moved into her new apartment there wasn't even anything to sit on.

When Delila moved out on her husband she couldn't take anything with her.
So, one of the first things she bought was a CHAIR which had been painted blue.
So, one of the first things she bought was a really CHAIR which had been painted blue.
So, one of the first things she bought was a GOAL which had been painted blue.
So, one of the first things she bought was a really GOAL which had been painted blue.

5. There have been many changes in the world's monarchies during the last century.

There have been many changes in the way people live during the last century.
In the few countries which continue to have a KING the distribution of power is unequal.
In the few countries which continue to have a really KING the distribution of power is unequal.
In the few countries which continue to have a SONG the distribution of power is unequal.
In the few countries which continue to have a really SONG the distribution of power is unequal.

6. Hilary decided that her only jacket is too light to keep her warm during the winter.

Hilary decided that her personal comfort is too important to sacrifice for fashion.
This weekend she's going to find a COAT suitable for sub-zero temperatures.

This weekend she's going to find a more COAT suitable for sub-zero temperatures.

This weekend she's going to find a SNAKE suitable for sub-zero temperatures.

This weekend she's going to find a more SNAKE suitable for sub-zero temperatures.

7. Bob hates the long hauls he has to make for the new company.

Bob hates the headaches he gets while working on his new job.

It is so noisy in his TRUCK that it's hard for him to relax.

It is so noisy in his badly TRUCK that it's hard for him to relax.

It is so noisy in his DREAM that it's hard for him to relax.

It is so noisy in his badly DREAM that it's hard for him to relax.

8. Simon couldn't get into his office until he found the janitor.

Simon didn't realize his mistake until he saw his mother.

He'd left his KEYS in the jacket which he'd lent to his brother.

He'd left his rather KEYS in his jacket which he'd lent to his brother.

He'd left his BARS in the jacket which he'd lent to his brother.

He'd left his rather BARS in the jacket which he'd lent to his brother.

9. The hospital employees are discussing whether to go on strike.

The union members are discussing whether to go on strike.

Most of the NURSES feel that they work too many hours per shift.

Most of the badly NURSES feel that they work too many hours per shift.

Most of the SUSPECTS feel that they work too many hours per shift.

Most of the badly SUSPECTS feel that they work too many hours per shift.

10. We have to get to the railroad station early this morning.

We have to cancel our dinner meeting for this evening.

We're taking a TRAIN which leaves at six o'clock.

We're taking a newly TRAIN which leaves at six o'clock.

We're taking a STOP which leaves at six o'clock.

We're taking a newly STOP which leaves at six o'clock.

11. Susan was making coffee when she realized that she had no milk.

Susan had left everything to the last minute so she was very disorganized.

She had to rush to a STORE before her guests arrived.

She had to rush to a very STORE before her guests arrived.

She had to rush to a JUDGE before her guests arrived.

She had to rush to a very JUDGE before her guests arrived.

12. Julie has decided to change the color of the walls in her apartment.

Julie has decided to expand the range of products in her store.

She's going out today to order some PAINTS that sparkle in the dark.

She's going out today to order some very PAINTS that sparkle in the dark.

She's going out today to order some BLOCKS that sparkle in the dark.

She's going out today to order some very BLOCKS that sparkle in the dark.

13. No one could see the bank robber's face.

No one could describe the old man exactly.

He was wearing a MASK which covered everything except his eyes.

He was wearing a very MASK which covered everything except his eyes.

He was wearing a CORK which covered everything except his eyes.
He was wearing a very CORK which covered everything except his eyes.

14. Michael has always been an insatiable reader.

Michael has always had one big weakness.

He spends so much money on BOOKS that it's a wonder he's not bankrupt.
He spends so much money on really BOOKS that it's a wonder he's not bankrupt.
He spends so much money on RATES that it's a wonder that he's not bankrupt.
He spends so much money on really RATES that it's a wonder he's not bankrupt.

15. The government is providing free vaccine to counteract the flu epidemic.
The government is providing free medical care for children in rural areas.

But, unless everyone has their SHOTS the prevention program won't work.
But, unless everyone has their very SHOTS the prevention program won't work.
But, unless everyone has their FARMS the prevention program won't work.
But, unless everyone has their very FARMS the prevention program won't work.

16. Stereotypically, elephants are supposed to be frightened of small animals.

Stereotypically, children are supposed to be frightened of many things.

When they see a MOUSE they should get very upset.
When they see a rather MOUSE they should get very upset.
When they see a COIN they should get very upset.
When they see a rather COIN they should get very upset.

-----article or possessive changed to pronoun-----

1. My aunt said that the string on her favorite necklace broke at the party last night.

My aunt said that the zipper on her favorite purse broke at the party last night.

Some of the PEARLS that she lost were very expensive.
Some of it PEARLS that she lost were very expensive.
Some of the CONES that she lost were very expensive.
Some of it CONES that she lost were very expensive.

2. Jack has finally decided to get a divorce.

Jack has finally decided to get professional advice.

He found that his WIFE had lost most of his money at the racetrack.
He found that he WIFE had lost most of his money at the racetrack.
He found that his TOWN had lost most of his money at the racetrack.
He found that he TOWN had lost most of his money at the racetrack.

3. Bart has been able to walk more easily since his cast was removed.

Bart has been in a much better mood since I last saw him.

But he still finds that his LEG hurts him occasionally.
But he still finds that he LEG hurts him occasionally.
But he still finds that his PAGE hurts him occasionally.
But he still finds that he PAGE hurts him occasionally.

4. A bad storm hit the boy scouts' camp last night.

A bad storm hit the shore last night.
Most of the TENTS were blown down by the wind.
Most of it TENTS were blown down by the wind.
Most of the MONKS were blown down by the wind.
Most of it MONKS were blown down by the wind.

5. The conservationists are doing what they can to save the most famous American bird.

Some concerned citizens are doing what they can to stop the horrible destruction.

But unless we make a big effort, the EAGLE will soon become extinct.
But unless we make a big effort, it EAGLE will soon become extinct.
But unless we make a big effort, the WITCH will soon become extinct.
But unless we make a big effort, it WITCH will soon become extinct.

6. Max can play most wind instruments very well.

Max can learn most new skills very easily.

But he has always found that the FLUTE is difficult to control.
But he has always found that he FLUTE is difficult to control.
But he has always found that the SNAIL is difficult to control.
But he has always found that he SNAIL is difficult to control.

7. Chicago is famous for its special kind of music.

Chicago is worth visiting for a number of reasons.

It has many places where good JAZZ can be found.
It has many places where they JAZZ can be found.
It has many places where good CLAY can be found.
It has many places where they CLAY can be found.

8. Harry hasn't visited his expensive tailor for quite a while.

Harry hasn't budgeted his income very well recently.

His SUITS are looking shabby.
He SUITS are looking shabby.
His WOODS are looking shabby.
He WOODS are looking shabby.

9. William had been planning to go cycling this weekend.

William had been planning to have a relaxing weekend.

Unfortunately, his BIKE was stolen last night.
Unfortunately, he BIKE was stolen last night.
Unfortunately his PIE was stolen last night.
Unfortunately, he PIE was stolen last night.

10. The doctor told Sam that his blood pressure is way too high.

The manager told Sam that his workload averages fifty hours per week.

If he's not careful, his HEART might collapse.
If he's not careful, he HEART might collapse.
If he's not careful, his RANGE might collapse.
If he's not careful, he RANGE might collapse.

11. The American electorate doesn't seem to be interested in the election.

The American people don't seem to be enthusiastic about many causes.

Most of them feel that their VOTE isn't worth anything in the long run.

Most of them feel that it VOTE isn't worth anything in the long run.
Most of them feel that their SIGN isn't worth anything in the long run.
Most of them feel that it SIGN isn't worth anything in the long run.

12. It's becoming increasingly difficult for the independent researcher to get financial backing.

It's becoming increasingly difficult for the average student to do original work.

These days, if you want to get a GRANT you have to be affiliated with a big university.

These days, if you want to get it GRANT you have to be affiliated with a big university.

These days, if you want to get a COOK you have to be affiliated with a big university.

These days, if you want to get it COOK you have to be affiliated with a big university.

13. Karen really hurt herself when she tried to lift the heavy chest.

Karen really hated herself after she tried to go cross-country skiing.

She thinks that her BACK might have been strained.

She thinks that she BACK might have been strained.

She thinks that her WORK might have been strained.

She thinks that she WORK might have been strained.

14. The firemen couldn't get the raging fire under control for several hours.

The press couldn't get past the temporary blockade for several hours. They had a difficult time approaching the building because the FLAMES were so hot.

They had a difficult time approaching the building because it FLAMES were so hot.

They had a difficult time approaching the building because the PLATES were so hot.

They had a difficult time approaching the building because it PLATES were so hot.

15. Although Marvin says he likes to be punctual, he rarely arrives on time.

Although Marvin says he likes his new schedule, he rarely attends his classes.

Because he doesn't have a WATCH he misses a lot of his meetings.

Because he doesn't have it WATCH he misses a lot of his meetings.

Because he doesn't have a PRICE he misses a lot of his meetings.

Because he doesn't have it PRICE he misses a lot of his meetings.

16. Sally decided that she doesn't like contact lenses at all.

Sally decided that she doesn't like long movies at all.

They make her EYES itch and water.

They make she EYES itch and water.

They make her HANDS itch and water.

They make she HANDS itch and water.

A.2 Filler trials

---with High context sentences

1. Our friends have decided to go CAMPING this summer.
They've heard good reports on a beautiful in Michigan. (sic)
2. Ed has been very short of CASH lately.
He spent a lot of money on his girlfriend and now he can barely afford to eat.
3. Jess was nearly in an accident at a bad intersection last night.
A brown Pontiac went right through the LIGHT and he stopped just in time.
4. Louise told David to lie down and relax.
She convinced him that if he took a NAP his headache would go away.
5. Sarah went to a single's bar last night after all.
She was looking for a really nice MAN to dance with.
6. Everyone thinks that Wade should shave.
He looks like a goat with that sparse BEARD he's grown.
7. Rather than being sold during international disputes, surplus farm PRODUCTS often go to waste.
Sometimes even perfectly good grain is allowed to rot.
8. My friend came to stay with me last night after her husband had beaten her.
She had a swollen LIP and was feeling very angry.
9. We had to take Sophie to the emergency room after she was knocked unconscious last night.
A brass sculpture fell on her HEAD while she was dusting the shelves.
10. Bob had to walk home last night after he got a flat TIRE in the suburbs.
He didn't have a spare so he had to abandon the car.
11. If THIEVES ever break into Hank's house, they'll have to come in through the window.
He has so many safety gadgets on the door that even he has trouble opening it.
12. The children were very excited about the robin's egg they found.
A robin must have a NEST in our front yard.

---with Low context sentences

1. The BUILDING inspectors are coming today around three in the afternoon. We'll have to get the garbage cleared out before they arrive.
2. Ben had to return home earlier than his friends. He told us that he had lost a SANDAL in a swamp and couldn't go on.
3. Larry was worried about his oldest DAUGHTER when I spoke to him. Yesterday some tea spilled over her and she was badly burned.
4. Violet wants to borrow the car. She wants to pick up a BED she bought last night.
5. Cynthia already said that she doesn't want ANY, so don't ask her again. She hates any kind of liquor, no matter how it's mixed.
6. They were talking to Ricky last night over a beer at the corner bar. He said that since he's been playing BALL, his wrist has been killing him.
7. My friends always seem to have TROUBLE with everything they own. Even their new car has a broken light.
8. Lisa is becoming quite concerned with the wide use of pesticides. Today she was complaining that most of the APPLES in the market are unfit to eat.
9. Tracy will be back at her DESK in a minute. She's gone to wipe off the coffee she spilled on her dress.
10. Though Lance is delightful to be with, he is rather eccentric. He always raises his dirty little HAT whenever he meets a woman.
11. Pete will have to start all over again. There was too much RED in the paint he was mixing.
12. Mitchell went around all day biting his nails. He's sure he'll have to cancel the PICNIC if there's any rain.

A.3 Practice and Lead-in trials

Real word trials

---with High context sentences

1. With so many brands of stereo equipment on the market it's hard to choose the right components.
But for good party music it's important to have large SPEAKERS which have a crisp sound.
2. Carl's sister has a birthday soon and he can't afford to buy her a present.
He'll probably just mail her a CARD with a nice message.
3. The hardware STORE down the street is selling lighting fixtures at fantastic savings.
Joan just bought a new lamp to go beside her sofa.
4. James told Sandra that she couldn't expect to see her parakeet again.
Once she'd let it out of its CAGE it was bound to fly away.
5. If an electrical appliance isn't working, there are a few things to check before it is repaired.
First, one should see whether the PLUG is in the outlet.
6. Sally suprised us with the NEWS of her engagement.
We hadn't even known that she was seeing anyone.
7. Molly was looking very elegant last NIGHT in a black mohair pullover.
She told me she bought we sweater on sale last week.
8. Stewart is thinking about buying a computer to write his papers with.
He's going to need one with a really good very EDITOR for that reason.
9. My grandmother still doesn't use postal zip-codes.
She doesn't understand why MOMENTS are returned to her.
10. Crossword puzzles often use antonyms as clues.
So, if hot is the clue, then HAIR is probably the answer.

---with Low context sentences

1. It's going to be difficult to do the experiment now.
The sudden theft of our CLOCK is going to delay us.
2. Betay went on a big shopping spree yesterday.
She came home with twenty new PANS for the kitchen.
3. Millie has been receiving twenty phone CALLS a day.
Ever since she put a want-ad in the paper she hasn't had any peace.
4. The CHURCH was broken into by amateurs around midnight.
They stole a lot copper off the roof.
5. It is funny to see DOGS trotting along the street alone.
They seem to have a real sense of very purpose as if they were running an important errand.
6. Doug becomes furious over the smallest things.
Kevin forgot to return they CAMEL and Doug nearly exploded.
7. The local newspaper has been reporting some questionable stories.
This morning they ran a headline about the drop in CARROT consumption.
8. Donna had trouble getting out of the house this morning.
First she got a very CAR in her stocking and then a button came off her skirt.
9. Carl has been unhappy since the party he gave last weekend.
It turns out that his favorite ANSWER was stepped on.

Nonword trials

---with High context sentences

1. A blackboard is not necessarily colored black.
Some of them are really GRALT (green) while others are gray.
2. The Susan B. Anthony coin wasn't very popular because it was heavy and bulky.
People would rather carry very SOLPS (bills) which are lighter and more compact.
3. Charles hit a pothole on his bicycle this morning.

When he got to work he noticed that a WHOIN (wheel) was bent out of shape.

4. The TOINDS (sounds) that insects make in the country can be almost deafening.

If one isn't used to the noise it can be very hard to sleep when the crickets are singing.

5. Driving over holiday weekends is no FOL (fun) at all. The amount traffic is intolerable.

6. We can never decide what to watch on television at night. We always seem to switch from one really CHIRRUM (channel) to another.

7. Wanda wasn't waering PRAMES (gloves) when she pruned her rose bushes so she hurt her hands.

The thorns on the flowers made it job very unpleasant.

8. In winter, people fall down a lot.

They slip on the rather BAW (ice) is probably the answer.

---with Low context sentences

1. When people live in the city they rarely see animals. But they do seem some SQUAMMENS (squirrels) running about the parks.

2. Asking little kids what they want to be when they grow up is fun. Most want a professional job but a few say they want to be a rather LUAN (poet) or an artist.

3. Fairy TONKS (tales) can be very fanciful. Dwarfs and giants are able to escape with smiles that they've stolen.

4. Neal was late coming YOLE (home) from school. He got involved in games at a friend's house and lost track of time.

5. Bruce hasn't seen anyone for days. He recently got a case of the MENKS (mumps) and has been quarantined.

6. Roger has every imaginable kind of office doo-dad in his desk. Not only that, but he also keeps enough TEEMIES (cookies) to feed a kindergarten class.

7. One of my high school teachers encouraged everyone to go to college. He felt that one should go to experience the FROIDAL (freedom) of college life even more than the education.

8. We almost went to the movies last night. If the hardly LEPES (lines) hadn't been so long we would have gone.

9. One can study prejudice by looking at JURPS (jokes) and cartoons. Political views are often expressed in that way also.

A.4 Nonword Trials

The following sentence pairs are arranged by type of syntactic manipulation.

---with High context sentences

1. OMISSION OF PREPOSITION

---grammatical sentences

1. Shower curtains now come in designer patterns.
The most dramatic ones have bright graphic designs stamped on PLINKET (plastic).
2. When going to a popular tourist spot it is important to make hotel reservations early.
Some people end up without RUEPS (rooms) because they waited too long.
3. More and more people are seeking out the counselors in career planning offices for advice.
Everyone wants to find the sort of JAD (job) that has a high salary.
4. The SAWN (soil) in parts of New Mexico makes great pottery, but it's not so good for farming.
It is hard to till because there is a high proportion of clay in it.
5. Wherever Clarissa walks she leaves behind a sweet smell that lasts for minutes.
She puts on PELDUME (perfume) so lavishly that she could almost be used as a walking sachet.
6. Grocery and liquor stores are overrun by people who are moving in the spring and fall.
Everyone is looking for BICES (boxes) to pack their things in.
7. Toni is a wonderful seamstress and she cannot go past a fabric store without going in.
Then she'll always buy yards of FINELIAR (material) that she'll never have time to sew.
8. Claire's mother has been very depressed by the deaths of many of her relatives.
She feels that she's gone to DEPINERS (funerals) more times than she's gone to the movies lately.

---ungrammatical sentences

9. Christina has bought some hens.
She hopes they will keep her family supplied ECTS (eggs) for the next few years.
10. The plants in the Southwestern deserts are remarkable.
The thorns CARPIN (cactus) really contrast with its delicate flowers.
11. Denise has to go to the drugstore before she can wash her hair.
It seems that she's all out PHACTIE (shampoo) and doesn't want to borrow any.
12. The upper class is quite pleased with Reagan's budget plan.
They're looking forward to a cut in the amount MISES (taxes) they must pay.

11. ADDITION OF ADVERB

---grammatical sentences

1. The army is looking for volunteers but they aren't having much success recruiting young people.
Most people don't consider being a SOWNIOT (soldier) when they plan their careers.
2. Mike's sister broke a down pillow over his head last month.
They've been finding FEAMPERS (feathers) in his room ever since.
3. Stacy is going to start medical BLOIRD (school) next fall.
Already everyone is telling her what a great doctor she's going to be.
4. Dentists claim that sweets are very bad for the teeth.
They advise that people should not eat COLKY (candy) or chew bubblegum.
5. Visitors to Yellowstone Park are asked not to feed the animals.
Yet it is hard to refuse a BOIN (bear) when it stands over your car.
6. Carmen hasn't had to buy any vegetables all FENMOR (summer) because she's grown her own.
She's had such a productive garden that she's even canned some food.
7. Fran wants to take a photography CAINSE (course) next semester, but the cost is high.
She would have to rent a camera as well as buy a lot of film.
8. Crash diets may make people lose weight but they can also cause nutritional problems.
To prevent this, many doctors recommend that people take BEROTENS (vitamins) to supplement their diets.

---ungrammatical sentences

9. Pam was trying to find a way to warn the birds in her yard when her cat is around.

She finally decided to attach a rather BIST (bell) to his collar.

10. Honey KOOS (bees) are not very aggressive.

They usually don't sting unless their likely hive is being threatened.

11. Jody thinks MEPHS (moths) got into her closet over the summer.

A lot of her wool skirts have very holes in them that weren't there last year.

12. Last Halloween Keith needed some make-up to put red splotches on himself. After he used his girlfriend's very PILSPOCK (lipstick) he looked very scary indeed.

iii. ARTICLE OR POSSESSIVE PRONOUN CHANGED TO NOM. PRONOUN

---grammatical sentences

1. Felt-tip pens last a long time unless the point dries out. So it's important to put the CUG (cap) back on tightly.

2. Joseph always makes sure he takes six pencils with him to drawing CLORD (class). Whenever his teacher watches him he gets so nervous that he breaks the lead in his pencil.

3. It is generally agreed that people with long hair shouldn't use a bristle brush on it. The hair will break less if a COLP (comb) with wide teeth is used instead.

4. Wilma has been learning how to weave and she loves it. Her mom thinks she'll buy her a LOUK (loom) for her birthday.

5. When people are superstitious they often believe that certain actions will cause them to have bad luck for years. They believe they'll have seven years of bad luck if they break a MALLYR, (mirror) for instance.

6. Western Europe is unhappy with the American decision to place nuclear warheads in their countries. They believe that the presence of the MUTTLES (missiles) makes nuclear war more likely to happen.

7. As soon as Gary gets home he always asks if any GADE (mail) came for him. He's expecting a letter from his girlfriend.

8. Sophia has loved the ballet ever since she took her first ballet lesson. She was thinking about becoming a TOPHER (dancer) but some injuries prevented that.

---ungrammatical sentences

9. It's good that the WOILS (heels) on women's shoes are becoming lower. Maybe women won't sprain they ankles so much anymore.
10. It's always good to have a few candle stubs around in case the power goes off unexpectedly. Then, if the batteries are dead in they FLOMPGICKT (flashlight), there will still be something to produce light.
11. Frankie's dog is constantly scratching. He's worried that she GLOYS (fless) might get into the carpet.
12. Patty couldn't eat anything after she rode the ferris wheel at the county fair. The movement of the ride made she WRIDARD (stomach) feel upset for hours afterward.

---with Low context sentences

1. OMISSION OF PREPOSITION

---grammatical sentences

13. There are often unique CRAFTS (sic) displayed at county fairs. There can be sculpture made from fungus and jewelry made from nails or cork.
14. Mr. Evans isn't the usual tourist. He wants to see all of the FOWMPOITS (fountains) in a city, instead of museums.
15. Statistics are often used to support wild claims. Someone recently tried to relate the increase of MOLLIATE (marriage) to a rise in stock values.
16. Carla's department has been brainstorming for the last week. They're trying to come up with a TIBLY (title) for their annual play.
17. Americans are becoming fonder of outdoor sports. Stores report that the number of CABAYS (canoes) sold increases every year.
18. Steve's aunt and INKLE (uncle) are looking for a spot to build a house. They'd like to be near a beach but they aren't insiating on it.
19. Very different THAMPS (things) sell in department stores, depending on the season. In September, furniture departments sell a lot of trunks for students to take to school.
20. Jan had a silly excuse for not getting her homework done.

She ran out of LOCER (paper) and couldn't think of where to find more.

---ungrammatical sentences

21. Woody is very particular when it comes to caring for his own car. He even has a certain type SPUDGE (sponge) he uses to wash it.

22. The Westons have a small house. But since the house came PORST (porch) they have enough living space.

23. It is sometimes hard to find ingredients for gourmet FIESES (sauces). Often they are thickened blood and seasoned with exotic mushrooms.

24. A local church had the oddest raffle in June. The first prize was a piece TIKE (cake) big enough for four people.

11. ADDITION OF ADVERB

---grammatical sentences

13. When people provided all their own food, nothing went to waste. They always used the BYDES (bones) of an animal to make a healthy stock.

14. Rose hates most party SEEBS (foods). In fact, she'll only go to a party if there is a nice cheese to eat.

15. The conference room is always a mess. There are always NOGRIMS (napkins) and ashtrays lying around.

16. Many children's playgrounds get vandalized. Sometimes only a SHOBE (slide) remains standing.

17. The managers of a local restaurant decided to change its image. They started using a HIRCE (horse) as its trademark instead of a bee.

18. With building costs soaring, people are looking for ways to save money. They are using more CAMSROTE (concrete) than they used to, for example.

19. There are many ways to tell if a new neighborhood might be sprouting up. If a new SWIGHT (church) has been built then people must be interested in the area.

20. Colleges depend on contributions to continue improving their facilities. Money for a new MIGRALY (library) is certain to come quickly.

---ungrammatical sentences

21. Surgeons must be careful not to catch diseases from their patients.

Especially when handling very BREEMS (brains) they must be cautious of certain viruses.

22. Lois' dog is very jumpy.
It is even scared of its own barely SHATEY (shadow) when it goes outside.

23. Public high school programs are not what they used to be.
In many schools the students can't learn a rather SOCHIAGE (language) or play football anymore.

24. Ellen has a lot of strange things displayed in her apartment.
On one shelf all she has are ugly very DIRNS (rocks) which are supposed to contain gold.

iii. ARTICLE OR POSSESSIVE PRONOUN CHANGED TO NOM. PRONOUN

---grammatical sentences

13. Kurt is very clumsy around mechanical things.
Only yesterday he banged his THORF (thumb) in the wind tunnel.

14. Cindy wants to make her office look more homey and warm.
A friend suggested that she place a BLUCKEN (blanket) over her file cabinets.

15. Japanese poetry is often not very easy to understand.
The MEEG (moon) can be a symbol for many different things.

16. Lou was excited by his VONT (find) at the rummage sale.
He bought a bottle that was over seventy years old.

17. People from different cultures organize their time differently.
Europeans usually have their DILLET (dinner) later than Americans do.

18. Amy hates to run errands.
The BIRT (bank) is her least favorite spot to go to.

19. There are certain questions that are impolite to ask.
One should never inquire the ONPYLE (income) of another person.

20. Nat is a talented bricklayer.
He built the TROCKEY (chimney) in our house.

---ungrammatical sentences

21. People are becoming more adventurous when it comes to planning vacations.
Even she JINKLES (jungles) are drawing lots of tourists.

22. Stuart had his (OLN) arm in a sling today.
He tripped on he rug and twisted it.

23. Don has been impossible to be around since he quit smoking cigarettes. He's always got they HUSP (mint) in his mouth, and he's grumpy besides.

24. Ramona came back from her YILK (walk) all shaken up. She had fallen asleep in the sun and then found we skunk next to her when she woke up.

Appendix B

Miscellaneous Analyses

In all of the following analyses, reported effects reached significance at the $p < 0.05$ level at least.

B.1 Analyses over all the experimental trials with the factors of Context, Plausibility and Grammaticality

Experiment 2: Slow Lexical Decision

In the analysis over the latency data there were main effects of all three factors: Context [$\text{minF}'(1,85) = 5.153$], Plausibility [$\text{minF}'(1,78) = 20.953$], and Grammaticality [$\text{minF}'(1,86) = 10.864$]. The Context X Plausibility interaction was significant in the separate analyses over subjects and items, although it was only marginally so in the minF' calculation [$F_1(1,39) = 11.498$; $F_2(1,47) = 5.468$; $\text{minF}'(1,81) = 3.706$, $p < 0.10$]. In the analysis over the error data there was a main effect only of Plausibility [$\text{minF}'(1,86) = 16.775$]. Context was marginally significant on both subject and item analyses [$F_1(1,39) = 3.003$, $p < 0.10$; $F_2(1,47) = 3.548$, $p < 0.10$]. Grammaticality missed significance in all analyses [all p 's > 0.10]. There was also an interaction of Context X

Plausibility in the error data [$\min F'(1,84) = 6.516$]. The latency and percent error interaction data are presented in Table 15.

Table 15: Mean subject response latencies in msec (and percent error) for plausible and implausible targets following different strengths of contextual constraint. Note: These means are collapsed over the factor of Grammaticality.
Experiment 2: Slow Lexical Decision.

Plausibility of Target	<u>Contextual Constraint</u>	
	HIGH	LOW
Plausible	796 (3.3)	884 (5.4)
Implausible	970 (17.3)	986 (10.4)

Newman-Keuls performed on the latency data revealed that although plausible targets were responded to more quickly after High context than after Low context [$q(2,39) = 8.27$], the strength of the discourse context had no effect on the processing of implausible targets [$q(2,39) = 1.50, p > 0.05$]. The post hoc tests on the error data showed a complementary pattern; that is, errors were made equivalently often on plausible targets no matter what the preceding context was [$q(2,39) = 1.64, p > 0.05$]. However, more errors were made to implausible targets when the context was highly predictive (of another target word) than when the context was less predictive [$q(2,39) = 5.39$].

There was a significant interaction of Plausibility X Grammaticality in the latency data [$\min F'(1,86) = 4.047$] although this interaction was not significant in the error data [both F_1 and $F_2 < 1$]. The latency data are

shown in Table 16.

Table 16: Mean subject response latencies in msec for plausible and implausible targets in different grammatical environments. Note: These means are collapsed over the factor of Context. Experiment 2: Slow Lexical Decision.

Grammaticality	<u>Plausibility of Target</u>	
	Plausible	Implausible
Grammatical	784	965
Ungrammatical	895	991

According to a Newman-Keuls analysis, the effect of Grammaticality is only apparent with plausible targets [$q(2,39) = 8.02$], it is not apparent with implausible targets [$q(2,39) = 1.88, p > 0.05$].

Experiment 3: Fast Lexical Decision

In the latency data there were main effects of Context [$\min F'(1,79) = 4.205$] and Plausibility [$\min F'(1,86) = 48.952$]. Implausible targets were responded to 160 msec more slowly than were plausible targets; the respective means were 975 and 815 msec. At the faster rate of presentation, there was a marginal main effect of Grammaticality in the subject analysis [$F(1,39) = 3.680, p < 0.10$] although it reached reliable levels of significance in the analysis over item means [$F(1,47) = 4.105$]. In the subject analysis there were three statistically significant interactions: Context X Plausibility [$F(1,39) = 6.702$]; Plausibility X Grammaticality [$F(1,39) = 5.426$]; Context X Grammaticality [$F(1,39) =$

6.413]. There was only one interaction significant with an item-based analysis, that of Context X Grammaticality [$F(1,47) = 4.122$]. This interaction did not reach significance in a minF' calculation [$\text{minF}'(1,85) = 2.510, p > 0.10$]. These data are shown in Table 17.

Table 17: Mean subject response latencies in msec for targets appearing in grammatical and ungrammatical environments, under different degrees of contextual constraint. Note: Both plausible and implausible targets contributed to these means. Experiment 3: Fast Lexical Decision.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	880	891
Ungrammatical	867	942

Newman-Keuls suggests that the interaction is due to the selective lengthening of responses to targets appearing in the Low Context, Ungrammatical condition; this mean differs from all the others (e.g. the difference between 942 and 891 corresponded to a q value of: $q(2,39) = 4.02$). However, none of the other means differed significantly from one another; for example, in the remaining means the largest range was from 867 msec to 891 msec which corresponded to a q value of: $q(3,39) = 1.89, p > 0.05$. These data are collapsed over the factor of Plausibility, and as such, are not within the domain of the additive factors logic as it had been defined for the experimental series. That is, responses to implausible targets followed a different pattern than did the responses to plausible targets. The additive factors design was used to explore

discourse context variables and syntactic variables when other variables (e.g. plausibility) were held constant.

In the error data there was only one main effect apparent, that of Plausibility [$\text{minF}'(1,80) = 25.105$]. (Subjects made, on average, 4.1 % errors on plausible trials compared to an error rate of 15.6 % on implausible trials.) In neither subject nor item analyses were there effects of Context [both F_1 and $F_2 < 1$] or Grammaticality [both p 's > 0.10]. Although the interaction of Context X Plausibility reached significance in both subject and item analyses, it had only the force of a trend in a minF' calculation [$F_1(1,39) = 7.023$; $F_2(1,47) = 8.529$; $\text{minF}'(1,83) = 3.852$, $p < 0.10$]. These error data are displayed in Table 18.

Table 18: Mean percentage of errors made by subjects to plausible and implausible targets following different degrees of contextual constraint. Note: These means are collapsed over the factor of Grammaticality.
Experiment 3: Fast Lexical Decision.

Plausibility of Target	<u>Contextual Constraint</u>	
	HIGH	LOW
Plausible	1.9	6.3
Implausible	17.7	13.5

According to a Newman-Keuls post-hoc test, context did not significantly affect the number of errors subjects made to targets. The comparison of the error rates for the plausible targets as well as the comparison of

the rates for implausible targets revealed no significant differences as a function of context [plausible targets: $q(2,39) = 2.72, p > 0.05$; implausible targets: $q(2,39) = 2.60, p > 0.05$]. The interaction appears to be based on the shift in the error pattern as a function of plausibility. That is, subjects made both the least number and the greatest number of errors under high contextual constraint, depending on whether the target was plausible or implausible.

Experiment 4: Slow Name

In the latency data there was no main effect of Context in either the subject or item analyses [F_1 & $F_2 < 1$]. There were main effects of Grammaticality [$\min F'(1,85) = 16.612$] and Plausibility [$\min F'(1,85) = 6.210$]. Plausible targets were named more quickly than were implausible targets, 612 versus 631 msec. There were no interactions present in the item data. There was, however, an interaction of Context X Grammaticality in the subject analysis [$F(1,39) = 5.026$]. These data are shown in Table 19.

Table 19: Mean subject latencies in msec for targets appearing in a variety of contextual conditions. Note: These data are collapsed over the factor of Plausibility.
Experiment 4: Slow Name.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	607	602
Ungrammatical	632	645

A Newman-Keuls test performed on the means revealed that although Context did not affect the responses to targets in grammatical sentences [$q(2,39) = 1.190, p > 0.05$], Low context slowed subjects' naming responses to targets appearing in ungrammatical sentences [$q(2,39) = 3.094$].

The error data were uninformative [all p 's > 0.10].

Experiment 5: Fast Name

In the latency data there was no main effect of Context in either subject or item analyses (both p 's > 0.10). There were main effects of Grammaticality [$\text{minF}'(1,86) = 7.968$] and Plausibility [$\text{minF}'(1,71) = 22.253$]. Plausible targets were named, on average, 63 msec faster than implausible targets (602 and 665 msec were the respective means). An interaction of Context X Grammaticality was significant in both subject and item analyses, but only marginally so by minF' [$F(1,39) = 12.566$; $F(1,47) = 4.699$; $\text{minF}'(1,76) = 3.420, p < 0.10$]. The interaction data are shown in Table 20.

Table 20: Mean subject response latencies in msec for targets appearing in a variety of contextual conditions. Note: These data are collapsed over the factor of Plausibility. Experiment 5: Fast Name.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	624	615
Ungrammatical	635	662

The result of a Newman-Keuls test suggested that when both plausible and implausible targets are considered together, only the combination of an ungrammatical environment and a weakly constraining context caused a significant lengthening of naming times [615 vs 662 : $q(4,39) = 9.08$; 635 vs 662 : $q(2,39) = 5.22$].

In the error data separate subject and item analyses showed main effects of Context [$F(1,39) = 5.442$; $F(1,47) = 4.225$] and Plausibility [$F(1,39) = 14.793$; $F(1,47) = 4.796$] with only the latter reaching a marginal level of significance with a minF' calculation [$\text{minF}'(1,74) = 3.622$, $p < 0.10$].

B.2 Analyses over Implausible targets

In this set of analyses over implausible targets there were factors of Context and Plausibility. In the subject analyses there was also the nesting factor of Version while in the item analyses there was the nesting factor of Kind of Ungrammaticality. Interactions involving these nesting factors are reported in the next section of this Appendix.

Experiment 2: Slow Lexical Decision

There was quite a discrepancy between subject and item analyses over implausible targets. No main effects or interactions reached levels of significance in the item analysis for either latency or error data [all p 's > 0.10]. However in the subject analysis over the latency data there

were main effects of both Context [$F(1,39) = 18.233$] and Grammaticality [$F(1,39) = 6.089$]. In addition, the Context X Grammaticality interaction was significant [$F(1,39) = 6.089$]. In the error data there was no Context effect [$F < 1$] and trends only of Grammaticality [$F(1,39) = 4.042$, $p < 0.10$] and the Context X Grammaticality interaction [$F(1,39) = 3.186$, $p < 0.10$]. These data are displayed in Table 21:

Table 21: Subject latency means in msec (and percent error) for implausible targets appearing in a variety of contextual conditions. Experiment 2: Slow Name.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	866 (9.6)	886 (7.9)
Ungrammatical	861 (10.0)	940 (11.8)

Results from a Newman-Keuls analysis suggest that when the targets were implausible, only the combination of a weak contextual constraint and an ungrammatical sentence frame caused significant slowing of the lexical decision response; e.g. the q-value for a comparison of 940 msec with 886 msec is $q(2,39) = 4.330$ whereas the q-value for a comparison of 886 msec with 866 msec is $q(2,39) = 1.604$, $p > 0.05$.

Experiment 3: Fast Lexical Decision

For the latency data, in neither the subject nor the item analysis were there main effects of Context or of Grammaticality [all p 's > 0.10].

There was, however, a Context X Grammaticality interaction in the subject data [$F(1,39) = 5.636$] with a corresponding trend in the item data [$F(1,47) = 3.893, p < 0.10$]. These data are shown in Table 22.

Table 22: Subject latency means in msec for implausible targets in a variety of contextual conditions.
Experiment 3: Fast Lexical Decision.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	999	961
Ungrammatical	937	1004

Even with the significant interaction, a post-hoc Newman-Keuls test did not differentiate any pair of means; e.g. the q-value for the comparison of 1004 and 937 was: $q(4,39) = 3.042, p > 0.05$.

In the error data there were marginal effects of Context in both analyses [$F_1(1,39) = 3.058, p < 0.10$; $F_2(1,47) = 2.888, p < 0.10$] reflecting a trend for more errors to be made after High context than after Low context. There was no effect of Grammaticality [both F 's < 1] or of an interaction [both p 's > 0.10].

Experiment 4: Slow Name

In the latency data there was no main effect of Context [both subject and item $p > 0.10$] although there was a main effect of Grammaticality [$\min F'(1,86) = 5.567$]. Implausible targets in ungrammatical sentences

were responded to more slowly than when they appeared in grammatical sentences (644 versus 617 msec). There was no interaction of Context X Grammaticality in the item analysis [$F(1,47) = 1.583, p > 0.10$], although there was a trend towards such an interaction in the subject analysis [$F(1,39) = 2.977, p < 0.10$]. There was nothing of interest in the error data.

Experiment 5: Fast Name

In the latency data there was no main effect of Context [both F_1 and $F_2 < 1$]. There was a main effect of Grammaticality in the separate subject and item analyses [$F_1(1,39) = 7.766; F_2(1,47) = 5.812$] although the effect was only marginal in a minF' calculation [$\text{minF}'(1,86) = 3.324, p < 0.10$]. The same pattern was true of the Context X Grammaticality interaction [$F_1(1,39) = 11.038; F_2(1,47) = 4.217; \text{minF}'(1,77) = 3.051, p < 0.10$]. The interaction data are displayed in Table 23.

Table 23: Subject latency means in msec for implausible targets appearing in a variety of contextual conditions.
Experiment 5: Fast Name.

Grammaticality	<u>Contextual Constraint</u>	
	HIGH	LOW
Grammatical	665	635
Ungrammatical	665	696

A Newman-Keuls test suggested that subjects responded more slowly to implausible targets following High context than they did following Low

context [$q(2,39) = 3.272$]; but only when the targets were presented in grammatical sentences. When the target sentences were ungrammatical, the pattern was reversed [$q(2,39) = 3.381$]. Another way of describing these results would be that the variable of Grammaticality only had an effect on implausible targets when the context did not strongly predict another (plausible) target.

There were no effects apparent in the error data [all F 's < 1].

B.3 Analyses with Version and Kind as factors

The results reported here are from subject analyses run with Version of experimental materials as a nesting factor and item analyses with Kind of ungrammaticality as a nesting factor. Only those effects or interactions that reached at least marginal levels of significance are reported. The results are arranged by experiment, and within experiment by the subset of data used as input to the analyses. Aside from the nesting factors, the common factors for the different subsets of data analyses are:

- All experimental trials: Context, Plausibility, Grammaticality
- Plausible trials: Context, Grammaticality
- Grammatical trials: Context, Plausibility
- Implausible trials: Context, Grammaticality

Experiment 2: Slow Lexical Decision

All experimental trials.

Subject analysis. In the latency data there was no main effect of Version. There were two interactions with Version, however. There was a three-way interaction of Version X Context X Plausibility [$F(7,32) = 3.025$]. There was also a four-way interaction of Version X Context X Plausibility X Grammaticality [$F(7,32) = 2.361$]. In the error data there was no main effect of Version nor any interactions involving it.

Item analysis. There was no main effect of Kind nor any interactions involving Kind in either the latency or error data.

Plausible trials.

Subject analysis. There was no main effect of Version on response latency or error rate (although this factor did approach significance, $F(7,32) = 2.250$, $p < 0.10$). For response latencies there was a three way interaction between Version, Context and Grammaticality [$F(7,32) = 5.892$]. In the error data there was the same three way interaction between Version, Context and Grammaticality [$F(7,32) = 2.769$, $p < 0.05$].

Item analysis. There was no main effect of Kind nor any interaction involving it in either latency or error data.

Grammatical trials.

There were no main effects of either Version or Kind in the analysis over grammatical trials. In addition, the nesting factors were not part

of any interactions.

Implausible trials.

There were no main effects or interactions of the nesting factors in subject and item analyses over implausible targets.

Experiment 3: Fast Lexical Decision.

All experimental trials.

Subject analysis. There were no main effects of Version or any interactions involving it in either the latency or error data.

Item analysis. There was no main effect of Kind, although there was a three-way interaction of Kind X Context X Plausibility in both the latency and error data [latency: $F(2,45) = 3.577$; error: $F(2,45) = 4.525$].

Plausible trials.

Subject analysis: The only effect of Version was an interaction appearing in the error analysis, Version X Grammaticality [$F(7,32) = 3.355$].

Item analysis. There were no effects of Kind.

Grammatical trials.

Subject analysis. There was no main effect of Version, nor any interactions involving Version.

Item analysis. Kind interacted with Plausibility in the latency analysis [$F(2,45)=3.328$].

Implausible trials.

Subject analysis. There was no main effect of and no interactions with Version.

Item analysis. In the error analysis there was a Kind X Context interaction [$F(2,45) = 3.773$].

Experiment 4: Slow Name

All Experimental Trials.

Subject analysis. There was only an interaction of Version X Plausibility in the error data [$F(7,32) = 2.980$].

Item analysis. There was a marginal four-way interaction in the latency data: Kind X Context X Plausibility X Grammaticality [$F(2,45) = 2.642, p < 0.10$].

Plausible trials.

Subject analysis. There were no effects of Version in the latency data. However, in the error data there was a main effect of Version [$F(7,32) = 2.539$] as well as two interactions involving Version [with Context: $F(7,32) = 2.614$; and with Grammaticality: $F(7,32) = 2.914$].

Item analysis. There were no effects of Kind.

Grammatical trials.

There were no effects of or interactions involving either of the nesting factors when the database for the analysis consisted of the responses to grammatical trials.

Implausible trials.

Subject analysis. In the error data there was a marginal interaction of Version X Context [$F(7,32) = 2.156, p < 0.10$] and a more stable interaction of Version X Context X Grammaticality [$F(7,32) = 2.584$].

Item analysis. There was a marginal effect of Kind interacting with Grammaticality [$F(2,45) = 2.535, p < 0.10$].

Experiment 5: Fast Name.

All Experimental trials.

Subject analysis. In the latency data there was no main effect of Version, although it was involved in two interactions: A two-way interaction between Version and Plausibility [$F(7,32) = 5.114$] and a three-way interaction of Version X Context X Plausibility [$F(7,32) = 3.680$]. In the error data there were marginal interactions involving Version X Plausibility [$F(7,32) = 1.991, p < 0.10$] and Version X Grammaticality [$F(7,32) = 2.007, p < 0.10$].

Item analysis. There were no main effects of Kind, nor any interactions with it in the latency data. However, when the analysis was conducted on the error rates, there was a trend of a main effect of Kind [$F(2,45) = 2.655, p < 0.10$].

Plausible trials.

Subject analysis. In the latency data there was no main effect of Version, although there were interactions. Version interacted with Grammaticality [$F(7,32) = 2.433$] and then both of those factors interacted with Context [$F(7,32) = 2.361$]. In the error data there was a Version X Grammaticality interaction [$F(7,32) = 3.136$].

Item analysis. There were no effects of Kind.

Grammatical trials.

Subject analysis. In the latency data Version interacted with Context [$F(7,32) = 1.212$]. Version also interacted with Context and Plausibility [$F(7,32) = 3.132$]. In the error data there were three marginal interactions involving Version: Version X Context [$F(7,32) = 1.950, p < 0.10$]; Version X Plausibility [$F(7,32) = 2.140, p < 0.10$]; and Version X Context X Plausibility [$F(7,32) = 1.941, p < 0.10$].

Item analysis. There were no effects of Kind.

Implausible trials.

Subject analysis. In the latency data there was a significant three-way interaction of Version X Context X Grammaticality [$F(7,32) = 2.490$] and a trend for the same interaction in the error data [$F(7,32) = 2.210, p < 0.10$].

Item analysis. There was a marginal main effect of Kind in the error data [$F(7,32) = 2.609, p < 0.10$].

B.4 Analyses of the nineteen discarded subjects from
Experiment 3: Fast Lexical Decision

In conjunction with the analyses run over the subjects for the Fast Lexical Decision Experiment, the same set of analyses was performed on the data from the nineteen subjects who were excluded from the analyses in Experiment 3: Fast Lexical Decision. These analyses are reported here according to the data subset that provided the input for the analysis, i.e. all experimental trials, plausible trials, grammatical trials, implausible trials.

All Experimental trials

The factors of the analyses run on the latency and error data for all the experimental trials were Subject(Version), Context, Plausibility and Grammaticality for the subject analysis and Item(Kind), Context, Plausibility and Grammaticality for the item analysis. In addition, another set of analyses was run on both subject and item data excluding the respective nesting factors of Version and Kind. MinF' was calculated based on these latter analyses; the results from these analyses are reported first, then main effects of the nesting factors, or interactions involving them are subsequently reported.

In the latency data there was no main effect of Context [all p's > 0.10] or Grammaticality [F1 and F2 < 1]. There was a strong main effect of Plausibility [minF'(1,42) = 24.925]. The Context and Plausibility

interaction was significant in both subject and item analyses, but was only marginally so in the minF' calculation [$F_1(1,18) = 11.815$; $F_2(1,47) = 5.728$; $\text{minF}'(1,65) = 3.858$, $p < 0.10$]. These data are presented in Table 24.

Table 24: Subject latency means in msec for plausible and implausible targets appearing after contexts differing in their predictive strength. Note: These data are collapsed over the factor of Grammaticality.
Experiment 3: Fast Lexical Decision; Discarded Subjects.

Plausibility of target	<u>Contextual Constraint</u>	
	HIGH	LOW
Plausible	1076	1219
Implausible	1495	1429

A Newman-Keuls test showed that although the degree of contextual constraint had a significant impact on reaction times to plausible targets [$q(2,18) = 4.706$], it was ineffectual on reaction times to implausible targets [$q(1,18) = 2.172$, $p > 0.05$].

In the error data there were also no main effects of either Context [F_1 and $F_2 < 1$] or Grammaticality [all p 's < 1]. There was a strong effect of Plausibility [$\text{minF}'(1,54) = 10.585$]; the error rate for implausible targets was approximately 19 %, substantially higher than the 6 % error rate observed for plausible trials. There were no significant interactions in the error data.

In the analyses including Version as a factor, there was no main

effect of Version in either the latency or error data. In the latency data, Version interacted with Plausibility [$F(7,11) = 4.578$]. In the error data there were interaction trends involving Version [Version X Context: $F(7,11) = 2.391$, $p < 0.10$; Version X Context X Plausibility: $F(7,11) = 2.373$, $p < 0.10$]. In the analyses including Kind of ungrammaticality as a factor, there were no main effects of Kind, nor any interactions involving it.

Plausible trials

In this set of analyses, besides the nesting factors of Version and Kind, there were factors of Context and Grammaticality. In the latency data there was a main effect of Context [$\min F'(1,61) = 5.161$]; subjects were 143 msec faster to respond to plausible targets following High contexts than Low ones (1076 vs 1219 msec). There was no main effect of Grammaticality [both F_1 and $F_2 < 1$] and no Context X Grammaticality interaction [all p 's > 0.10].

The error data were uninformative; all F 's were less than one.

There was no apparent influence of the nesting factors of Version or Kind in the data from plausible trials.

Grammatical trials

The factors for the analyses over grammatical trials were the relevant nesting factors of either Version or Kind, Context, and Plausibility. In the latency data there was no main effect of Context [both F_1 and $F_2 < 1$]. There was a robust Plausibility effect [$\min F'(1,58)$]

= 15.578]; subjects responded to plausible targets sooner than they did to implausible targets (1142 msec vs 1447 msec). There was no Context X Plausibility interaction [all p's > 0.10].

The pattern in the error data was similar to that in the latency data. There was no effect of Context [all p's > 0.10]; a healthy effect of Plausibility [$\text{min}F'(1,58) = 5.905$] reflecting that not only were subjects slower to respond to implausible targets, they also made more errors on them (there were approximately 17 % errors made on implausible targets compared to 5 % on plausible targets); and there was no interaction between the two factors [all p's > 0.10].

In this set of analyses there were no effects of either Version or Kind.

Implausible trials

For implausible trials, the factors of Context and Grammaticality were analyzed along with the nesting factors of Version and Kind. There were no significant effects or interactions in either the latency or error data. However, there was an interaction that reached trend status in the subject error data: Version X Context [$F(7,11) = 2.742, p < 0.10$].

Appendix C

Ratings as a Function of Kind of Ungrammaticality

This appendix contains the percentage of time that each rating point was used for grammatical and ungrammatical sentences as a function of type of ungrammaticality. These data are arranged by experiment and are from trials containing plausible targets only.

In each of the tables, the three types of ungrammaticality are abbreviated as follows:

1. Omission of a Preposition = PRP
2. Insertion of an Adverb = ADV
3. Alteration of an Article or a Pronoun = PRO

C.1 Experiment 2: Slow Lexical Decision

Table 25: Rating percentages of each of the three kinds of ungrammaticality. Grammatical sentences.
Experiment 2: Slow Lexical Decision.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	72	67	60
2	25	29	37
3	3	4	3

Table 26: Rating percentages of each of the three kinds of ungrammaticality. Ungrammatical sentences.
Experiment 2: Slow Lexical Decision.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	56	29	44
2	28	24	26
3	16	47	30

C.2 Experiment 3: Fast Lexical Decision

Table 27: Rating percentages of each of the three kinds of ungrammaticality. Grammatical sentences.
Experiment 3: Fast Lexical Decision.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	49	51	49
2	49	42	45
3	2	7	5

Table 28: Rating percentages of each of the three kinds of ungrammaticality. Ungrammatical sentences.
Experiment 3: Fast Lexical Decision.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	48	41	42
2	50	39	47
3	2	20	10

C.3 Experiment 4: Slow Name

Table 29: Rating percentages of each of the three types of ungrammaticality. Grammatical sentences.
Experiment 4: Slow Name.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	48	50	43
2	45	41	50
3	6	10	7

Table 30: Rating percentages of each of the three kinds of ungrammaticality. Ungrammatical sentences.
Experiment 4: Slow Name.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	41	17	25
2	36	23	37
3	23	60	38

C.4 Experiment 5: Fast Name

Table 31: Rating percentages of each of the three kinds of ungrammaticality. Grammatical sentences.
Experiment 5: Fast Name.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	48	56	51
2	48	38	41
3	4	5	8

Table 32: Rating percentages of each of the three kinds of ungrammaticality. Ungrammatical sentences.
Experiment 5: Fast Name.

Rating Scale Point	<u>Kind of Ungrammaticality</u>		
	PRP	ADV	PRO
1	51	45	42
2	44	35	47
3	5	20	11

References

- Balota, D.A. and Chumbley, J.I. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. Journal of Experimental Psychology: Human Perception and Performance, 1984, 10, 340-357.
- Blank, M. and Foss, D.J. Semantic facilitation and lexical access during sentence processing. Memory and Cognition, 1978, 6, 644-652.
- Cairns, H.S. Current issues in research in language comprehension. In R. Naremore (Ed.), Recent Advances in Language Sciences, San Diego: College Hill Press, 1983.
- Cairns, H.S., Cowart, W., and Jablon, A.D. Effects of prior context upon the integration of lexical information during sentence processing. Journal of Verbal Learning and Verbal Behavior, 1981, 15, 681-689.
- Cole, R.A. Listening for mispronunciations: A measure of what we hear during speech. Perception and Psychophysics, 1973, 11, 153-156.
- Conrad, C. Context effects in sentence comprehension: A study of the subjective lexicon. Memory and Cognition, 1974, 2, 130-138.
- Cotton, S. and Grosjean, F. The gating paradigm: A comparison of successive and individual presentation formats. Perception and Psychophysics, 1984, 35, 41-48.
- Cowart, W. and Cairns, H. The influence of reference relations on syntactic processing. Manuscript, 1984.
- 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.
- Fodor, J.A. The Modularity of Mind, Cambridge, Massachusetts: Bradford Books at the MIT Press, 1983.
- Forster, K.I. Visual perception of rapidly presented word sequences of varying complexity. Perception and Psychophysics, 1970, 8, 215-221.
- Forster, K.I. Accessing the mental lexicon. In R.J. Wales and E.C.T. Walker, (Eds.), New Approaches to Language Mechanisms, Amsterdam: North-Holland, 1976.

- Forster, K.I. Levels of processing and the structure of the language processor. In W.E. Cooper and E.C.T. Walker (Eds.), Sentence Processing, New York: Wiley, 1979.
- Forster, K.I. Priming and the effects of sentence and lexical contexts on naming times: Evidence for autonomous lexical processing. Quarterly Journal of Experimental Psychology, 1981, 33A, 465-495.
- Forster, K.I. and Olbrei, I. Semantic heuristics and syntactic analysis. Cognition, 1973, 2, 319-347.
- Forster, K.I. and Ryder, L.A. Perceiving the structure and meaning of sentences. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 285-296.
- Foss, D.J. A discourse on semantic priming. Cognitive Psychology, 1982, 14, 590-607.
- Foss, D.J., Cirilo, R.K. and Blank, M.A. Semantic facilitation and lexical access during sentence processing: An investigation of individual differences. Memory and Cognition, 1979, 7, 346-353.
- Glucksberg, S., Kreuz, R.J. and Rho, S. Context can constrain lexical access: Implications for models of language comprehension. Manuscript, 1985.
- Goodman, G.O., McClelland, J.L., and Gibbs, R.W. The role of syntactic context in word recognition. Memory and Cognition, 1981, 9, 580-586.
- Grosjean, F. Spoken word recognition processes and the gating paradigm. Perception and Psychophysics, 1980, 28, 267-283.
- Herriot, P. The comprehension of active and passive sentences as a function of pragmatic expectations. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 166-169.
- Jakimik, J. Word recognition and the lexicon. In J.Wolf and D.Klatt (Eds.), Speech communication papers presented at the 97th meeting of the Acoustical Society of America, New York: Acoustical Society of America, 1979.
- Keppel, G. Design and Analysis: A Researcher's Handbook, Englewood Cliffs: Prentice-Hall, 1973.
- Koriat, A. Semantic facilitation in lexical decisions as a function of prime-target association. Memory and Cognition, 1981, 9, 587-598.
- Lukatela, G., Kostic, A., Feldman, L.B., and Turvey, M.T. Grammatical priming of inflected nouns. Memory and Cognition, 1983, 11, 59-63.
- Marslen-Wilson, W.D. Sentence perception as an interactive parallel process. Science, 1975, 189, 226-228.

- Marslen-Wilson, W.D. and Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.
- Marslen-Wilson, W.D. and Tyler, L.K. The temporal structure of spoken language understanding. Cognition, 1980, 8, 1-71.
- Meyer, D.E. and Schvaneveldt, R.W. Meaning, memory structure, and mental processes. Science, 1976, 192, 27-33.
- Meyer, D.E., Schvaneveldt, R.W., and Ruddy, M.G. Loci of contextual effects on visual word recognition. In P.M.A. Rabbit and S. Dornic (Eds.), Attention and Performance V, London: Academic Press, 1975, pp. 98-118.
- Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.
- Morton, J. and Long, J. Effect of word transitional probability on phoneme identification. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 43-51.
- Neely, J.H. Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. Memory and Cognition, 1976, 4, 648-654.
- Neeley, J.H. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. Journal of Experimental Psychology: General, 1977, 106, 226-254.
- Onifer, W. and Swinney, D.A. Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. Memory and Cognition, 1981, 9, 225-236.
- Posner, M.I. and Snyder, C.R.R. Attention and cognitive control. In R.L. Solso (Ed.), Information processing and cognition: The Loyola Symposium, Hillsdale, NJ: Erlbaum, 1975, pp. 55-85.
- Potter, M.C. Rapid Serial Visual Presentation (RSVP): A Method for Studying Language Processing. In D. Kieras and M. Just (Eds.), New Methods in the Study of Immediate Processes in Reading Comprehension, Hillsdale: Erlbaum, 1984, pp. 91-118.
- Rayner, K., Carlson, M. and Frazier, L. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. Journal of Verbal Learning and Verbal Behavior, 1983.
- Riesbeck, C.K. and Schank, R.C. Comprehension by computer: Expectation based analysis of sentences in context. In W.J.M. Levelt and G.B. Flores d'Arcais (Eds.), Studies in the Perception of Language, New York: Wiley, 1978.

- Rubenstein, H., Garfield, L., and Millikan, J.A. Homographic entries in the internal lexicon. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 487-494.
- Rumelhart, D.E. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and Performance VI, Hillsdale, NJ: Erlbaum, 1977.
- Salasoo, A. and Pisoni, D.B. Interaction of knowledge sources in spoken word identification. Journal of Memory and Language, 1985, 24, 210-231.
- Scholes, R.J. Syntactic and lexical components of sentence comprehension. In A. Caramazza and E. Zurif (Eds.), The Acquisition and Breakdown of Language: Parallels and Divergences, Baltimore, MD: The Johns Hopkins University Press, 1978.
- Schuberth, R.F. and Eimas, P. Effects of context on the classification of words and nonwords. Journal of Experimental Psychology: Human Perception and Performance, 1977, 33, 27-36.
- Schuberth, R.E., Spoehr, K.T. and Lane, D.M. Effects of stimulus and contextual information on the lexical decision process. Memory and Cognition, 1981, 9, 68-77.
- Schvaneveldt, R.W., Meyer, D.E., and Becker, C.A. Lexical ambiguity, semantic context, and visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 243-256.
- Seidenberg, M.S. and Tanenhaus, M.K. Modularity and lexical access. In I. Gopnik (Ed.), McGill Studies in the Cognitive Sciences, Norwood, N.J.: Ablex, in press.
- Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., and Bienkowiaki, M. Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 1982, 14, 489-537.
- Seidenberg, M.S., Waters, G.S., Sanders, M. and Langer, P. Pre- and postlexical loci of contextual effects on word recognition. Memory and Cognition, 1984, 12, 315-328.
- 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.
- Simpson, G.B. Lexical ambiguity and its role in models of word recognition. Psychological Bulletin, 1984, 96, 316-340.
- Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adult. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.

- Stanovich, K.E. and West, R.F. Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. Memory and Cognition, 1979, 7, 77-85.
- Stanovich, K.E. and West, R.F. The effect of sentence context on ongoing word recognition: Tests of a two-process theory. Journal of Experimental Psychology: Human Perception and Performance, 1981, 7, 658-672.
- Sternberg, S. The discovery of processing stages: Extensions of Donners' Method. In W.G. Koster (Ed.), Attention and Performance II, Amsterdam: North Holland, 1969, pp. 276-315.
- Swinney, D.A. Lexical access during sentence comprehension: (Re)Consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 645-660.
- Swinney, D.A. and Hakes, D.T. Effects of prior context upon lexical access during sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 681-689.
- Tanenhaus, M.K. and Donnenwerth-Nolan, S. Syntactic context and lexical access. Quarterly Journal of Experimental Psychology, 1984, 36A, 649-661.
- Tanenhaus, M.K., Leiman, J.M. and Seidenberg, M.S. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 427-440.
- Taylor, W.L. "Cloze" procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.
- Townsend, D.J. and Bever, T.G. Natural units of representation interact during sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 1982, 21, 688-703.
- Tulving, E. and Gold, C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. Journal of Experimental Psychology, 1963, 66, 319-327.
- Tyler, L.K. and Marslen-Wilson, W.D. The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 683-692.
- Tyler, L.K. and Marslen-Wilson, W.D. Speech comprehension processes. In J. Mehler, E.C.T. Walker, and M. Garrett (Eds.) Perspectives on mental representation: Experimental and theoretical studies of cognitive processes and capacities, Hillsdale, NJ: Erlbaum, 1982, pp.169-185.
- Tyler, L.K. and Wessels, J. Quantifying contextual contributions to word-recognition processes. Perception and Psychophysics, 1983, 34, 409-420.

- Warren, R.E. Stimulus encoding and memory. Journal of Experimental Psychology, 1972, 94, 90-100.
- Warren, R.E. Time and the spread of activation in memory. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 450-466.
- West, R.F. and Stanovich, K.E. Automatic contextual facilitation in readers of three ages. Child Development, 1978, 49, 717-727.
- West, R.F. and Stanovich, K.E. Source of inhibition in experiments on the effect of sentence context on word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1982, 8, 385-399.