

INSTITUTE OF COGNITIVE SCIENCE



*Technical Report*

University of Colorado, Boulder

**The Role of Lexical Information and  
Discourse Context in Syntactic Processing:  
A Review of Psycholinguistic Studies**

Evelyn C. Ferstl

Institute of Cognitive Science  
University of Colorado  
Boulder, Colorado 80309-0344

Technical Report #93-03

---

The Role of Lexical Information and  
Discourse Context in Syntactic Processing:  
A Review of Psycholinguistic Studies

by

Evelyn C. Ferstl  
University of Colorado at Boulder

Technical Report 93-03

---

\_\_\_\_\_

\_\_\_\_\_

---

**The Role of Lexical Information and Discourse Context in Syntactic Processing:  
A Review of Psycholinguistic Studies**

**Evelyn C. Ferstl  
University of Colorado at Boulder**

**Abstract**

Syntactic processes are an important component of the language processing system. While semantic and pragmatic information has been shown to influence the eventual interpretation of an utterance, psycholinguistic theories have not come to an agreement on how this information is combined with syntactic knowledge during the initial parsing process. This paper reviews recent psycholinguistic research on the effects of lexical information and discourse context on syntactic processing. The results of these studies are often contradictory, and do not allow to draw firm conclusions on how syntactic and non-syntactic processes interact. Nevertheless, it is suggested that the psychological evidence points towards a weakly interactive, parallel model of parsing. Methodological implications and further research directions are discussed, and an evaluation of text comprehension theories with respect to syntactic processing is attempted.

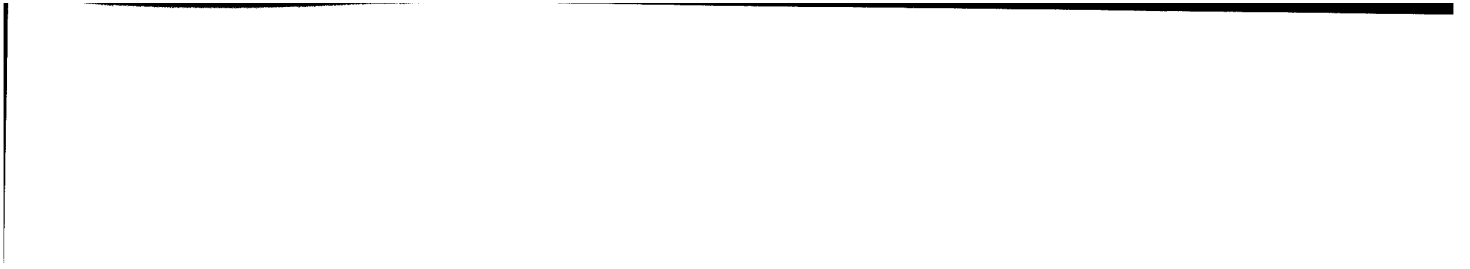
\_\_\_\_\_

\_\_\_\_\_

L

## Table of Contents

1. INTRODUCTION .....	3
2. ISSUES AND THEORIES IN SYNTACTIC PROCESSING RESEARCH .....	6
1. Issues in psycholinguistic research on parsing .....	6
1. Rule-based versus lexical representation of syntactic knowledge .....	6
2. Parallel versus serial processing .....	8
3. Informationally encapsulated versus interactive processing .....	9
2. Theories of syntactic processing .....	10
1. Syntax-first models .....	10
2. Interactive models .....	11
3. Summary .....	12
3. METHODS IN SYNTACTIC PROCESSING RESEARCH .....	13
1. Materials .....	13
1. Unambiguous materials .....	13
2. Ambiguous materials .....	14
2. Paradigms .....	15
1. Eye movement monitoring .....	15
2. Self-paced reading paradigms .....	16
3. Dual-task paradigms .....	18
4. Priming paradigms .....	19
4. LEXICAL EFFECTS ON SYNTACTIC PROCESSING .....	21
1. Verb complexity .....	21
2. Subcategorization preference .....	23
3. Argument preference .....	27
4. Selection restrictions .....	29
5. Summary .....	31
5. DISCOURSE EFFECTS ON SYNTACTIC PROCESSING .....	32
1. Pragmatic knowledge and sentence context .....	32
2. Discourse context .....	37
3. Summary .....	48
6. CONCLUSIONS .....	50
1. Methodological implications .....	50
1. Comparison of experimental paradigms .....	50
2. Selection of materials .....	51
3. Further research directions .....	52
2. Implications for parsing theories .....	52
1. Conclusions on the architecture of the syntactic processor .....	53
2. Alternative research directions .....	53
3. Implications for text comprehension theories .....	54
1. NL-SOAR .....	54
2. The construction-integration model of text comprehension .....	56
4. Summary .....	58
References .....	59



## CHAPTER 1 INTRODUCTION

The interdisciplinary study of language has undergone considerable change. Early psycholinguistic research was mostly concerned with identifying syntactic mechanisms in sentence processing, in particular with respect to linguistic grammar theories (for reviews, see Fodor, Bever & Garrett, 1974; Clark & Clark, 1977). After the attempt to provide evidence for the psychological reality of transformational grammar had failed, however, psycholinguistics became "a little depressed" (Ford, 1987). At the same time, experimental studies addressing the use of semantic and pragmatic knowledge in language comprehension provided strong evidence against the representation of sentences on a linguistic or structural level. For instance, Bransford and Franks (1971) found that after hearing the sentence

- 1) Three turtles rested on a log and the fish swam beneath them.

subjects could not distinguish which of the following sentences they had heard:

- 2) The fish swam beneath the log.
- 3) The fish swam beneath the turtles.

The linguistic properties of the sentence 1) cannot account for this distortion which, instead, is due to the subjects' forming a situational interpretation of the sentence. Similarly, Bever (1970) pointed out that a garden-path as occurring in sentence 1) is avoided in sentence 2), because appropriate world knowledge can be used to facilitate syntactic processing.

- 1) The authors read in the garden stank.
- 2) The articles read in the garden stank.

Results like this have led to the appreciation of the use of non-syntactic knowledge in language understanding. Text comprehension theories have been developed which stress the importance of discourse influences and assume representations of a text on different levels (e.g., Kintsch & vanDijk, 1978). While some representation of the surface structure is still postulated, the importance of more global representations of the gist of a message is also recognized. The pragmatic knowledge of the reader or listener is assumed to be integrated with the current input to form what is called a mental model (Johnson-Laird, 1983), discourse model (Crain & Steedman, 1985), or situation model (vanDijk & Kintsch, 1983) of the text. This representation contains general information on "what the text is about." Consequently, issues with which earlier psycholinguistic research was concerned, such as the psychological validity of grammatical principles, or the resolution of syntactic ambiguities, did not seem as important as a more ecologically valid description of how language serves its communicative purpose. Indeed, as Kintsch (1974) pointed out, the interpretation of many ambiguities studied in the sentence processing literature becomes an almost trivial task if they are embedded in a context which strongly suggests one particular reading over its alternatives. In spoken language, in addition, prosodic cues, such as intonation or pauses, can often be used for the correct interpretation of structurally ambiguous utterances. The rather strong conclusion that sentence processing research does not provide much insight in language comprehension processes occurring in natural contexts was formulated by Kintsch (1984) in this way (p. 128):

"The problems posed by the analysis of sentences out of context are exceedingly complex, as a glance into the current or past philosophic and linguistic literature shows. Much of that complexity is simply irrelevant for text studies. This does not mean that these ancient problems are solved, but that we need not solve them for our purposes."



A similar shift in focus could be observed in the artificial intelligence community, albeit for different reasons. Linguistic theories, in particular theories of syntax, did not enable researchers to successfully develop natural language processing systems. The correct parsing of syntactic structures could not be accomplished without a mechanism to incorporate knowledge from non-syntactic sources. As a result, more effort was spent on developing representational systems for general world knowledge, and alternative approaches to the parsing of natural language. Riesbeck and Schank (1978), for example, proposed that a grammatical analysis is necessary only after other comprehension heuristics failed. In their framework, a conceptual analysis of the relationships among the entities of an utterance is the goal of the comprehension process. This goal is best accomplished using mainly semantic and pragmatic information. Schank and colleagues (Schank & Birnbaum, 1980; cited in Marcus, 1984, p.254) put forward the thesis that "no independent level of syntactic representation is being constructed, operated upon, or output by the language analysis process."

As Marcus (1984) shows convincingly, however, linguistic cues are necessary even for the comprehension of simple sentences. In the absence of context, word order, for instance, is crucial for establishing the conceptual relationships in an utterance (e.g., "Bill insulted John"). A purely semantic language processing mechanism, which is assumed to operate using mainly relationships between content words, cannot account for such phenomena.

Similarly, given psychological evidence on language processing, it cannot be denied that comprehenders do take advantage of linguistic cues. Gernsbacher and Shroyer (1989) has shown that the rather subtle difference of using the function word "this" instead of a definite or indefinite article has a strong effect on which concepts are foregrounded in discourse. Similarly, Givón (1992) provided evidence for the effectiveness of particle movement as a cue to anticipating the focus of the subsequent discourse (also, Clifford, 1990).

Thus, discourse context and general world knowledge facilitate syntactic processing, and syntactic cues are useful in establishing the meaning of discourse. A complete, psychological theory of language comprehension cannot neglect the syntactic properties of the language. The crucial question, therefore, is not whether a syntactic processing mechanism exists, or whether a mechanism using pragmatic knowledge exists, but rather how these components of the language processing system interact.

Psycholinguistic research has only recently begun to address this question. Three influences can be identified which revived the study of sentence processing. The first influence comes from the development of two competing linguistic theories of syntax. The debate about the validity of Government Binding Theory (Chomsky, 1981) and Lexical Functional Grammar (Bresnan & Kaplan, 1982) has been accompanied by experimental work attempting to find psychological evidence for the constructs of these theories (see Sells, 1985, for an introduction). The second influence comes from Fodor's (1983) modularity hypothesis. According to this thesis, language has a special status within the cognitive system, and general world knowledge cannot directly affect syntactic mechanisms. Finally, the attempt to formulate comprehensive psychological theories of language understanding led to the need to investigate the role of syntactic mechanisms in building representations of discourse (Just & Carpenter, 1992; Lehman, Lewis, & Newell, 1991a).

In this paper, I review experimental studies which investigate how lexical information and discourse context affect syntactic processing, (in contrast to how linguistic information is used to form representations of the meaning of discourse). Although the lexical information is considered part of the syntactic processor, the distinction between semantic and pragmatic knowledge on one hand, and lexical knowledge on the other is highly dependent on the underlying theory of grammar. Because the purpose of this review is not to find evidence for a particular theory of grammar, but rather to be able to outline the general architecture of the cognitive language processing system, lexical effects as well as context effects are considered. The goal is to collect psychological phenomena observed during parsing in order to evaluate comprehensive theories of text understanding.

Since most of the experimental studies attempt to distinguish between psycholinguistic theories of parsing, these theories, as well as the most important issues addressed, are summarized in Chapter 2. The on-line methodologies used to investigate syntactic processes are the topic of Chapter 3. Chapter 4 reviews the experimental results on lexical effects in parsing, and Chapter 5 reviews the results on pragmatic knowledge and discourse context effects. Finally, in Chapter 6, these results are evaluated in three ways. First, methodological conclusions are drawn. Second, the results are summarized with respect to which effects seem to hold across paradigms and across studies and further research directions are outlined. Third, text comprehension theories which include a parsing step are evaluated with respect to their psychological validity in light of the results from the psycholinguistic literature.

## CHAPTER 2 ISSUES AND THEORIES IN SYNTACTIC PROCESSING RESEARCH

Although parsing can be studied on a variety of levels, three issues have evolved as most influential in guiding experimental psycholinguistic research. The focus does not as much lie on testing hypotheses about how a particular grammatical construction is understood, as on establishing facts about the general architecture of the human language comprehension system. Assuming one part of this system is a syntactic processor, three questions arise. The first question which has to be addressed is what kind of grammatical knowledge the syntactic processor uses in comprehension and how this grammatical knowledge is represented. The second question concerns the processes operating on this syntactic knowledge, and the third involves the interaction between the syntactic processor and other cognitive mechanisms. This last issue, in particular the interaction between discourse context and syntactic processing, is in the focus of this paper; however, the interpretation of experimental results pertaining to it depend crucially on assumptions about the syntactic mechanisms. To give an example, consider the verb "intend" which requires a human subject. If this fact is assumed to be stored in the lexicon, and thus, part of linguistic knowledge, processing difficulties for sentences in which this constraint is violated, can be interpreted as reflecting purely syntactic processes. In contrast, if the proposed lexicon does not contain such selection restrictions, the same processing difficulties are evidence for semantic effects in parsing.

Therefore, the interaction of syntactic and non-syntactic mechanisms cannot be studied without reference to the general architecture of the syntactic processor. In the first section of this chapter, I will summarize the proposals which have been brought forward in an attempt to answer the aforementioned questions. Specifically, I will discuss the difference between a rule-based and a lexically based parsing mechanism, the issue of parallelism in syntactic processing, and hypotheses about the interactions between syntax and other knowledge sources. The second section of the chapter contains an overview of the most important psycholinguistic theories of parsing. While a wide variety of parsing models have been proposed (in particular for machine language processing), only the few which have had an impact on experimental research are considered here.

### 1. Issues in psycholinguistic research on parsing

Linguistic grammar theories are useful descriptions of the properties of the language. Many processing models are based on syntactic theories, and possible relationships between the grammar and the processing system provide an interesting field of study (Frazier, 1985; Crain & Fodor, 1985). The intuitive difficulties of processing certain sentences can be used to distinguish between syntactic theories analytically (e.g., Pritchett, 1988). For a psychological theory of parsing, it is necessary to provide additional, experimental evidence for the psychological validity of the theoretical constructs. The actual experimental work, however, does not match the complexity of the grammar theories. With respect to knowledge representation, for instance, testing specific proposals about the contents of the lexicon and on the rules which combine lexical items into constituent structures have not been in the focus of research. Considerable attention has been paid, on the other hand, to the more fundamental question if initial parsing decisions are primarily guided by grammatical rules which are independent of individual lexical entries, or if they are guided by the grammatical knowledge stored in the form of specific features of lexical items.

#### 1. Rule-based versus lexical representation of syntactic knowledge

Rule-based or principled parsing models (Berwick, Abney, & Tenny, 1991) propose that syntactic processing is accomplished through the application of a set of rules. These rules make use only of the lexical class of the incoming words and phrases (e.g., noun, verb, noun phrase) and of

the properties of the constituent structure resulting from combining these entities into a parse tree. The rules do not necessarily correspond to the actual phrase structure rules of the underlying grammar; instead they are often formulated to represent strategies (Kimball, 1973; Clark & Clark, 1977; Frazier & Fodor, 1978; Frazier & Rayner, 1982).

For the current psycholinguistic research most influential are the Principle of Minimal Attachment and the Principle of Late Closure, suggested by Frazier (Frazier, 1978; Frazier & Rayner, 1982). The Principle of Minimal Attachment states, in simple terms, that a new item is to be attached to the partial parse tree in a way which results in a minimal number of nodes. For instance, in the syntactically ambiguous sentence

- 1) The spy shot the cop with the revolver.

the parse tree for the interpretation of the prepositional noun phrase "with the revolver" as an instrument of the action "shot" contains fewer nodes than the parse tree for the interpretation of the prepositional phrase as an adjunct of the noun phrase "the cop." The Principle of Minimal Attachment has been shown to account for many preferences, such as the tendency to interpret a noun phrase after the verb as a direct object rather than the beginning of a complement clause (e.g., in the sentence fragment "We knew the girl..."), or the tendency to analyze a clause as a main clause rather than a subordinate clause (Frazier & Fodor, 1978).

The Principle of Late Closure states that new items are to be attached to the phrase or clause postulated most recently. This principle accounts for the preference of interpreting the relative pronoun as referring to the teacher instead of to the husband:

- 1) This is the husband of the teacher who is a communist.

It is important to note, however, that the application of these strategies is dependent on the rules of the grammar. For instance, after the application of one of these principles, the processor has to be able to evaluate the grammaticality of the resulting attachment. Similarly, the construction of parse trees, whose parsimony is used to apply the Principle of Minimal Attachment, cannot be accomplished without the use of more basic syntactic knowledge, such as knowledge about what constitutes a phrase. Konieczny, Hemforth, and Strube (1991), for instance, point out that the preferences for prepositional phrase attachments, as defined by the Principle of Minimal Attachment, are qualitatively different depending on the specific rules for combining phrases. Using the recursive rules

- 1a) NP --> det N
- 1b) NP --> NP PP

the principle predicts preference for verb phrase attachment. If these rules are substituted with the simple "flat" rule

- 2) NP --> det N PP

then the two attachment alternatives result in parse trees containing the same number of nodes, and the Principle of Minimal Attachment does not make predictions for a preferred interpretation. Thus, the postulation of parsing strategies of the kind described in this section is not sufficient as a description of the syntactic knowledge used in comprehension.

The most important feature of these general strategies is their context independence. The alternative approach assumes that grammatical knowledge is not represented in general rules, but that each lexical entry (including words of different lexical classes, but also lexical items such as affixes; Speas, 1990) contains information about the syntactic structures in which it can appear (Ford, Bresnan & Kaplan, 1982; Bresnan & Kaplan, 1982; Ford, 1989; Small & Rieger, 1982). The specific form of the lexical information is dependent on the underlying grammar theory, but there is

considerable agreement as to which kind of knowledge is represented. In particular, verbs are assumed to carry information about the arguments and complements with which they are associated. For example, the verb "know" allows for either a direct object (as in "Brian knows Boris") or a complement sentence (as in "Brian knows Boris is blond."), while the intransitive verb "sleep" does not take a direct object. This structural information is contained in so-called subcategorization frames of the verb (Chomsky, 1981) or in thematic grids (also called predicate-argument structures) (Williams, 1984). The subcategorization frames contain information on the syntactic structures the verb allows. For instance, the verb "donate" is associated with the two subcategorization frames (Shapiro, Zurif, & Grimshaw, 1989)

- 1) [ \_ NP ]
- 2) [ \_ PP ]

reflecting the grammaticality of the sentences

- 1) Renoir donated the painting.
- 2) Renoir donated the painting to the museum.

The thematic grid of this verb, on the other hand, contains the information that a grammatical sentence consists either of two arguments (agent, theme), or three arguments (agent, theme, recipient) which are from the specified thematic role (Fillmore, 1968; Jackendoff, 1972). Thus, in this particular case, and in many others, the subcategorization and the predicate-argument structures are highly redundant. As in the rule-based models, however, additional information, most importantly knowledge about the simple phrase structure rules, is needed to complete the parsing process.

In lexically based models, the parsing decision is assumed to be guided by specific lexical information of this kind. In contrast to rule-based models, therefore, the particular verb used in a sentence is predicted to influence the parsing process immediately. For example, for the syntactically identical sentences, which the absence of the complementizer "that" renders locally ambiguous,

- 1) Joe agreed leaves change in the fall.
- 2) Joe taught leaves change in the fall.

a rule-based parsing theory does not predict processing differences. A lexically based parser, in contrast, uses the fact that "agree" does not take a direct object to postulate a sentential complement upon encountering "leaves." Thus, processing difficulties are predicted for sentence 2), but not for sentence 1).

## **2. Parallel versus serial processing**

While the distinction between rule-based and lexically guided parsing models is a rather recent consequence of the development of lexical-functional grammar, parallelism within the syntactic processor has been a topic in the psycholinguistic literature for some time (see Fodor, Bever, & Garrett, 1974, for a review of early studies). Nevertheless the issue has not been resolved.

Three proposals for the processing of syntactic ambiguities have been brought forward. Serial models (e.g., Frazier & Rayner, 1982; Ford, Bresnan & Kaplan, 1982) suggest that only one interpretation is built. The syntactic processor is not affected by the ambiguity unless this first interpretation turns out to be incorrect in the given context. If inconsistent disambiguating information is encountered a reanalysis is initiated. Thus, processing differences are expected for sentences which require different final resolutions. If reanalysis is necessary, processing difficulties should be observed in the disambiguating region of a sentence. Moreover, serial models predict that the processor is not aware of the ambiguity, unless the initially chosen interpretation turns out to be incorrect later in the sentence.

Parallel models, or multiple representation models, on the other hand, postulate that all alternative interpretations of an ambiguous syntactic construction are activated upon encountering the ambiguous information (Altmann & Steedman, 1988; Gorrell, 1989; Just & Carpenter, 1992; Shapiro, Zurif & Grimshaw, 1987, 1989). Compared to unambiguous sentences, in this account, processing difficulties should arise during the ambiguous region due to demands of keeping several structural possibilities in working memory; and processing difficulties should arise during the disambiguating region due to the demands of the selection process. However, the parallel view does not predict that reanalysis is needed, because the appropriate interpretation is still available at the time of disambiguation.

Delay models, finally, postulate that the parsing mechanism postpones a commitment to one structure until disambiguating information is encountered (e.g., Marcus, 1980; Perfetti, 1990). While this is also the case in parallel models, delay models do not assume that alternative structures are activated before this commitment is made. The parser "notices" the ambiguity and thus delays the attachment decision, but it does not build hypotheses about the possible structures immediately. In this approach, no reanalysis processes are predicted because postulating a delayed decision ensures that the correct structural assignment is always computed (therefore, delay models have also been called deterministic parsers).

Although it seems like these three approaches make differential predictions for the processing of syntactic ambiguities, it is rather difficult to distinguish between them experimentally. If processing difficulties arise during comprehension of disambiguating material, they could be due to reanalysis processes (as predicted in serial models), they could be due to selection processes (as predicted in parallel models), or they could be due to the need of a deterministic parser to make attachment decisions which have been delayed earlier in the sentence (Fodor, 1985, cited in Mitchell, 1987b). Backtracking, although seemingly supporting a serial model which makes a reanalysis prediction, can also be interpreted as the need to reinstate information which is not any more available in working memory. However, only parallel and delay models predict that processing load increases during the comprehension of locally ambiguous material.

### **3. Informationally encapsulated versus interactive processing**

This issue has received much attention since Fodor (1983; 1987) put forward the modularity thesis. Following a proposal by Forster (1979), Fodor elaborated the idea that the language processing system is an independent module and serves as an input system to the general cognitive system. It has therefore a special status within the cognitive system which is comparable to the perceptual system. In particular, he postulates that syntactic parsing is automatic and that the syntactic processor is, as he argues, informationally encapsulated from other knowledge sources. The syntactic processing is not, and cannot in principle, be influenced by pragmatic knowledge or contextual factors. An example for the impenetrability of the perceptual system which best illustrates this point is the perception of optical illusions. Even conscious knowledge about the illusion does not prevent the perceptual system from the wrong analysis. This argument is also supported in research on lexical access of homonyms. Initially, all meanings of an ambiguous word are activated, and the appropriate one is chosen in a later stage (Swinney, 1979). This effect is even found if the ambiguous word is embedded in a strongly biasing context (Seidenberg, Tanenhaus, Leiman & Bienkowsky, 1982; Till, Mross & Kintsch, 1988). For syntactic processing, the modularity hypothesis predicts analogously that even with strong contextual bias or under the strong influence of pragmatic knowledge, parsing processes initially rely on syntactic information only.

The initial syntactic analysis is thus seen as independent of other processes involved in language comprehension. The syntactic processor is considered a module in the sense that it does not have access to external information, and that there is no direct interaction with other parts of the cognitive system. The output of the syntactic processor, consisting of a structural analysis of the utterance, however, is in turn used as the input to higher level processes. This view predicts, consequently, that the initial processing of syntactic ambiguities which are disambiguated through pragmatic

knowledge or discourse context proceeds in the same way as for isolated sentences. The effects of the pragmatic information can only be observed in a later stage of comprehension.

One argument in favor of this assumption is that the syntactic processor can operate more efficiently if the knowledge available to it is limited. Interestingly, efficiency considerations also led to the alternative approach. Theorists who propose interactions between different kinds of knowledge sources (e.g., Crain & Steedman, 1985; Altmann & Steedman, 1988; Steedman, 1989; Taraban & McClelland, 1990; Perfetti, 1990; McClelland, 1987; St. John & McClelland, 1990; Just & Carpenter, 1992; MacDonald, Just & Carpenter, 1992) argue that immediately incorporating all available information prevents the language processing system from misanalyses. In this view, syntactic knowledge does not have a special status within the cognitive system. The language processing system attempts to extract meaning from an utterance, and it does so by using all types of available knowledge in parallel. The importance of one kind of information depends on its relative usefulness in a given context. Thus, in sentence processing experiments which present sentences in isolation, syntactic information might be used to a larger degree than in natural language understanding in which contextual knowledge or communicative goals play a role.

## 2. Theories of syntactic processing

These three issues have been combined in different ways in models of the parsing process. Although they are in principle independent of each other, not all combinations of answers have been brought forward. Moreover, some parsing models contain components which are not directly related to these three issues. In this section, I will summarize the most influential theories of human parsing in terms of what they propose about the grammatical knowledge representation, about the parallelism of the processor, and about its interactions with non-syntactic processes. (for more extensive reviews of psycholinguistic parsing models see Altmann, 1988; Altmann, 1990; Foss, 1988; Frazier, 1987a; Frazier, 1987b; Frazier, 1990a; King, 1983; Konieczny, Hemforth, & Strube, 1991; Mitchell, 1987b; Norris, 1987).

### 1. Syntax-first models

Syntax-first models are based on the assumption that the syntactic processor is informationally encapsulated, and that the processing proceeds serially. The main distinction between these models concerns the knowledge representation issue. The so-called garden-path model suggested by Lyn Frazier (Frazier & Fodor, 1978; Frazier & Rayner, 1982) holds that syntactic principles guide the parsing process. In the case of syntactic ambiguities, only the preferred analysis, suggested by the application of syntactic principles, is computed. In a second stage, a so-called thematic processor uses conceptual knowledge to check the consistency of this analysis with the discourse context and semantic constraints. In particular, the thematic processor compares the constraints specified in the predicate-argument frame of the verb with the constituent structure output by the syntactic processor. The mechanisms for evaluating the plausibility in the discourse context are not explicitly spelled out, but they are assumed to involve the use of a mental model of the discourse and sentence topics (Ferreira & Clifton, 1986).

Lexical models in contrast, in particular the model based on lexical-functional grammar (Ford, Bresnan, & Kaplan, 1982; Bresnan & Kaplan, 1982; Ford, 1989), postulate that lexical information associated with the heads of phrases guides the parsing process. In this model, a preference rank ordering is assumed to be part of the lexical entry. If a verb can be used in several syntactic structures, the subcategorization frame with the highest preference ranking, which corresponds to a frequency index (Connine et al. 1984), is chosen for the initial parsing decision. If the resulting interpretation turns out to be incorrect, the sentence is reanalyzed using the subcategorization frame with the next highest preference rank.

In both the garden-path model and the lexical model, the informational encapsulation is a crucial feature, whereas serial processing is not. In both models a mechanism which considers multiple representations in parallel is conceivable. In the garden-path approach these interpretations would correspond to alternative syntactic structures, whereas in the lexical model the interpretations would correspond to the different subcategorization frames.

## 2. Interactive models

Interactive models propose the parallel use of several types of information, and thus do not assume that the syntactic processor is an informationally encapsulated module. Crain & Steedman (1985) and Altmann & Steedman (1988) distinguish weak interaction and strong interaction. In strongly interactive models, the language comprehension process is considered a constraint-satisfaction process in which multiple sources of information are used in parallel. No restrictions as to the flow of information are postulated (e.g., McClelland, 1987; St. John & McClelland, 1990; Marslen-Wilson, 1975; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Tyler, 1987; Waltz & Pollack, 1985). In this account, it is possible that pragmatic knowledge and contextual constraints propose an analysis to the syntactic processor. For instance, the order of application of grammatical rules could be influenced, the rules could contain conditions which are context-dependent, (and thus be applicable only in certain contexts but not in others), or the different knowledge sources could be used in a constraint-satisfaction network, in which the contextually inappropriate interpretation is quickly deactivated.

In weakly interactive models, on the other hand, syntax is autonomous and initially proposes structural interpretations in parallel, but independent of context. Thus, support for weakly interactive models does not disconfirm the modularity hypothesis. The pragmatic and semantic knowledge is then used to dispose of inappropriate analyses offered by the syntactic processor. The difference to informationally encapsulated models, such as the garden-path model discussed earlier, is *when* the pragmatic information is taken into account. In an interactive approach, the interplay between syntactic and non-syntactic processes is assumed to take place immediately after a constituent has been constructed. This assumption corresponds to the immediacy hypothesis (Just & Carpenter, 1980) which states that comprehenders interpret each word on several levels (e.g., lexical access, syntactic attachment) as they encounter it. In non-interactive models, on the other hand, the thematic and pragmatic plausibility check is postponed until a larger syntactic structure has been assembled. In this account, effects of pragmatic plausibility or discourse context are observable only at the point of disambiguation or syntactic closure, but not earlier. This approach is in-line with Fodor, Bever and Garrett's (1974) proposal that clause boundaries are the locations within a sentence at which syntactic integration takes place.

The most important interactive model in the psycholinguistic literature on parsing is the referential support model proposed by Steedman and colleagues (1989; Crain and Steedman, 1985; Altmann and Steedman, 1988). This model makes specific predictions on the circumstances in which discourse context influences the parsing of syntactic ambiguities; and all experimental studies which investigate context effects refer to this approach. After the proposal of multiple representations in parallel, the model applies two principles which are based on a mental model, the discourse model, which incorporates conceptual information from the prior context. The Principle of Parsimony states that "a reading which carries fewer unsupported presuppositions will be favoured over one that carries more." (Altmann & Steedman, 1988, p. 203). This principle accounts, for instance, for the garden-path in the sentence (Bever, 1970)

- 1) The horse raced past the barn fell.

The interpretation of "the horse raced" as a main clause involves establishing one referent for the concept "horse." The interpretation of "raced past the barn" as a reduced relative clause, in contrast, requires establishing several referents for the concept "horse," only one of which was raced past the barn. This principle is based on the assumption that language serves a functional purpose in communication, and only information is included in an utterance which allows the formation of a



coherent discourse model. A special case of the Principle of Parsimony is the Principle of Referential Support, which states that "a noun phrase which is referentially supported will be favoured over one that is not." (Altmann & Steedman, 1988, p. 201). In a locally ambiguous sentence fragment such as

- 1) The psychologist told the woman that ..... .

which can either be continued with a relative clause modifying the noun phrase "the woman" or with a complement clause, the former interpretation is more likely if the prior discourse contained several referents. If only one woman was mentioned before, the complement reading is more likely, because a reference specification is not needed.

Obviously, these principles cover only a subset of the possible configurations in discourse models or situation models. The pragmatic information contained in these representations does not only consist of the cardinality of the sets of possible referents of a word, but also of conceptual information about the relationships between the objects in the discourse. The parsimony principle, formulated in a rather general way, can be applied to other types of contextual influences. However, it is necessary to define more explicitly what a "parsimonious" discourse model entails.

Moreover, the theory does not contain an explicit description of the mechanisms which give rise to the application of the principles. The syntactic rules (as before, the parser must recognize noun phrases before being able to apply the principle of referential support) which the parser uses are not spelled out, and the model does not make assumptions about the representational properties of the discourse model. Nevertheless, this approach provides a concrete starting point for experimental tests of the interactivens of the syntactic processor.

### 3. Summary

Theories of parsing can be classified as syntax-first models and interactive models. Syntax-first models postulate an informationally encapsulated syntactic module within the cognitive system and predict that initial syntactic processing cannot be influenced by semantic, pragmatic, or contextual information. In this framework, a detailed characterization of the contents of the syntactic processor and the lexicon is necessary. Because only syntactic knowledge is assumed to be taken into account initially, the validity of these models can only be tested if an appropriate description of the syntactic knowledge is available.

Interactive models, in contrast, allow the immediate interaction of syntactic and non-syntactic knowledge in the comprehension process. A clear classification of information into lexical, semantic, and pragmatic components is thus not needed to be able to make predictions about its effects on parsing. Moreover, in contrast to syntax-first models, interactive models allow for parallel activation of several alternative structural analyses.

An experimental distinction between these types of models is only possible if the time course of processing can be observed accurately. Both approaches can account for syntactic preferences and global processing difficulties, and both approaches allow the eventual use of non-syntactic information. Their differential predictions can only be tested using an on-line analysis of the processing as it occurs during natural reading. Thus, the experimental paradigms used in the psycholinguistic study of parsing must employ tasks which resemble natural comprehension and they must allow for a fine-grained description of the comprehension process.

## CHAPTER 3 METHODS IN SYNTACTIC PROCESSING RESEARCH

To make the following sections more easily readable, it is helpful to give an overview over the methods and materials employed in the various experiments. In particular, it is important to be clear about the assumptions which are necessary for interpreting a certain pattern of results, and to point out the processing differences induced by using a particular paradigm. As will be elaborated later in this paper, the relatively small effects obtained in many sentence processing studies are often not invariant across tasks and often strongly dependent on the materials, so that an evaluation of experimental results must take into account the task specific processes as well as a careful analysis of the materials. In the first section of this chapter, I will give a brief overview over the types of materials found in sentence processing studies, and in the second section I will summarize the on-line paradigms which are used to measure the processes employed in comprehension.

### 1. Materials

Syntactic processing usually proceeds smoothly and automatically. The strategies involved in comprehension can therefore only be identified if processing difficulties are induced, or if the sentences contain syntactic ambiguities. The following sections provide an overview over the types of materials used in psycholinguistic studies on parsing.

#### 1. Unambiguous materials

Unambiguous materials are used in two ways. First, comparisons of two different, syntactically correct structures provide information about their relative complexity. Examples for this type of comparison include the finding that sentences containing two clauses are more difficult to process than sentences of similar length containing only one clause, or the finding that object-relative clauses (1) are generally more difficult to process than subject-relative clauses (2). (see Fodor, Bever, & Garrett, 1974).

- 1) object-relative:  
The girl whom John kissed blushed.
- 2) subject-relative:  
The girl who kissed John blushed.

Similarly, this type of materials is used to investigate how specific lexical or pragmatic information influences the syntactic processing. In this case, the syntactic structure of the experimental sentences is held constant while specific content words are replaced (e.g., Rayner & Duffy, 1986). Examples include the comparison of the processing of sentences such as

- 1) The student denied the allegations.
- 2) The student heard the allegations.

Second, the comprehension of syntactically or semantically anomalous, but unambiguous sentences provides information on the interaction between different knowledge sources. For instance, it has been shown that certain types of syntactic errors (e.g., the omission of a function word) do not substantially disrupt comprehension (Flores d'Arcais, 1987). On the other hand, syntactically coherent, but semantically uninterpretable sentences are still comprehensible to some extent, as the classical example "Colorless green ideas sleep furiously." (Chomsky, 1957) demonstrates. Identifying the class of syntactic or semantic distortions which still allow the subject

to form a representation of the sentence can provide evidence on the relative importance of various linguistic cues.

## **2. Ambiguous materials**

Ambiguous materials are used to study the processes involved in assigning one of several alternative interpretations. Globally ambiguous sentences do not contain disambiguating information, so that either of two structural assignments yields a syntactically and pragmatically acceptable reading. An example is the sentence

- 1) The parents discussed the problem with the math teacher.

in which it is not clear if the parents and the teacher discussed the problem, or if the parents discussed, amongst themselves, the problem they had with the teacher. Sentences like this are used to study the preference for one syntactic interpretation (e.g., Holmes, 1984), and thus to identify preferred syntactic strategies. However, this type of materials is not very useful for studying on-line comprehension processes because it is difficult to distinguish which of the two interpretations an individual computes. Even if subjects are asked to indicate which interpretation they computed after reading the sentence, the answer might not coincide with their initial syntactic analysis.

More informative are sentences which are only locally ambiguous, i.e., sentences in which subsequent information disambiguates the initially ambiguous structure. If disambiguation can be accomplished using purely syntactic considerations, the sentences are called garden-path sentences (see the garden-path theory discussed in Chapter 2). Classical examples are sentences containing embedded reduced relative clauses (Bever, 1970), such as

- 1) Locally ambiguous:  
The horse raced past the barn fell.
- 2) Unambiguous:  
The horse which was raced past the barn fell.

If initially the parser interprets "raced" as a past tense verb, then the verb "fell" at the end of the sentence cannot be successfully integrated into a grammatically correct constituent structure. Only the interpretation of "raced" as a past participle leads to the correct reading of the sentence. To study the effects of the ambiguity, the processing of locally ambiguous sentences, such as 1), is compared to their disambiguated forms, such as 2). The regions in which processing differences might be observed are the locally ambiguous region ("raced past the barn") and the disambiguating region ("fell").

Instead of the syntactic structure, semantic or pragmatic information can be used for disambiguating local ambiguities. In this case, the sentences whose processing is compared have identical grammatical structures. In terms of their syntax, these sentences are globally ambiguous because both alternative interpretations yield a grammatically correct constituent structure. However, pragmatic plausibility considerations strongly suggest one alternative. For instance, in the sentences (cf. Taraban & McClelland, 1988)

- 1) Verb phrase attachment:  
The janitor cleaned the room with the broom.
- 2) Noun phrase attachment:  
The janitor cleaned the room with the windows.

the last noun ("broom" or "windows") disambiguates the prepositional phrase in sentence 1) as the instrument of "cleaning," whereas it is interpreted as a modification of the "room" in sentence 2). The comparison of processing times for the disambiguating region can in this case provide evidence for eventual initial syntactic decisions.

## 2. Paradigms

Since the goal is to provide experimental evidence for a hypothesized time course of processing, only on-line measures are appropriate. Early studies on sentence processing and syntactic processing often used global comprehension measures, such as sentence reading times, grammaticality judgments, paraphrasing responses, or answers to forced choice decisions on the correct interpretation of an ambiguous sentence (see Fodor, Bever & Garrett, 1974). Results from tasks like these only show which sentence representation is built eventually, or which syntactic structures are considered more difficult, but they do not shed light on the question of how this is accomplished. The goal of more recent studies on syntactic processing is not only to identify processing difficulties for certain syntactic structures, and to assess their final interpretation, but to assess in detail the time course of processing. Thus, the perfect method should allow the researcher to observe the comprehension process in real time without superimposing unusual demands on the subjects. The reading (or listening) situation should be as similar as possible to natural reading conditions, in particular with respect to presentation mode and instructions. Moreover, the units of analysis should be small enough to yield information on the exact location at which processing difficulties might arise.

The on-line paradigms used in research on syntactic processing fall into four main classes (for comprehensive reviews see Just & Carpenter, 1980; Rayner & Pollatsek, 1987; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989; Kieras & Just, 1984, in particular the chapters by Aaronson & Ferres; Mitchell; McKoon & Ratcliff; Rayner & Carroll; Just & Carpenter): eye movement monitoring, self-paced reading time measures ("button-press tasks"), dual-task paradigms, and priming methods. These classes are distinct in terms of the way processing difficulties are operationalized and measured. However, there are many other factors which influence results from on-line studies. In particular, the presentation mode, the instructions, the way to test for comprehension, and the statistical analyses have a large impact on the outcome of the experiments. Since the many variations and combinations of these factors are often confusing, I will describe each of the main paradigms and their relation to these factors in some detail. However, the purpose of this section is to give a general overview. Studies which explicitly compare different paradigms, or studies which provide evidence for task specific effects are not considered here, but included in the review in Chapter 4 and Chapter 5.

### 1. Eye movement monitoring

Eye movement monitoring is often considered the preferred method for studying on-line comprehension processes. While the subject is reading a sentence which is presented on a computer screen, his or her eye movements are recorded. Thus, except for the bite bar which restrains subjects' head movements, the task is unintrusive; i.e., the reading task resembles natural reading in which the subject is highly practiced. Although the sentences are presented on a computer screen, they are typed in the usual fashion. The subjects can use highly practiced reading strategies, for instance, skip words, or reread unclear portions of the sentence. Since the reading situation is relatively normal, the instruction to "read for comprehension" is sufficient, while other instructions, such as memorization, can also be used (which does influence the pattern of eye movements; Carroll & Slowiaczek, 1987). That subjects in fact follow these instructions is verified using comprehension questions or paraphrasing tests after a subset of the materials.

Under the so-called eye-mind assumption (Just & Carpenter, 1980) eye movements directly reflect the location of attention, and are not highly susceptible to strategic processes. It has been shown that fixation times and number of regressions (i.e., looking back at a word that occurred earlier in the text) are indeed related to the difficulty of the presented material. As is the case for other measures, however, the reading times reflect not only syntactic processes, but also lexical access effects and semantic and pragmatic integration processes. In fact, word frequency and word length account for most of the variance in the pattern of eye movements, and there are a noticeable sentence wrap-up effects. The fixation times of a given word, in addition, are not perfectly coupled

with the processing time of the word. For instance, function words are often skipped, but it is not safe to conclude that they are not processed at all. In contrast, these function words seem to be processed within the fixation on the immediately preceding words. On the other hand, increases in fixation times on a given word sometimes reflect processing difficulties on the preceding words, which is called a spill-over effect (Rayner & Pollatsek, 1987). Thus, although eye fixations provide a rather detailed picture of the on-line reading processes, they do not perfectly mirror the processing times for each individual word.

An important advantage of eye movement monitoring is that the pattern of eye fixations can be used to separate initial processing effects from backtracking effects. For the analysis, the sentences are usually divided into experimentally relevant regions (for instance, a locally ambiguous phrase), and reading times are divided by the number of characters to normalize for length. Related to reading time measures in other tasks, the total time spent processing a given word (or region), is reflected in the so-called gaze duration, which is defined as the sum of all fixations on the word. The initial processing, which is particularly important in distinguishing between parsing theories, is reflected in the first pass reading times. This measure incorporates the fixation durations in the target region which were made in a left to right fashion, but does not include the fixations during regressive eye movements. These regressive fixations are captured in the second-pass reading times (with the gaze duration being the sum of the first- and second-pass reading times), and also in the probabilities of regressions out of the target region or into the target region. The first-pass reading times together with the pattern of regressive eye movements shed light on eventual reanalysis processes.

As Rayner et al. (1989) point out, however, the definition (and the size) of the target region is crucial for the results of eye movement experiments. If, for instance, regressions are likely from the first word in a region consisting of several words, then the first pass reading times are shorter than predicted by the length of the region. If the region is relatively long, then an eventual effect which manifests itself in a subregion might be hidden in noise.

Nevertheless, eye movement data yield the most complete description of the reading processes in time. It is important to note, however, that this completeness comes with high costs. The eye tracking equipment is expensive, and the experimental procedure requires initial tuning and constant monitoring of the calibration. Finally, the eye movement records contain a wealth of data which can only be analyzed efficiently if appropriate computer programs facilitate the task (Rayner et al., 1989). In fact, in most eye movement studies, only a small subset of the possible analyses is reported. If the experimental question is focussed, it might therefore be preferable to employ a slightly less sensitive paradigm which is not as costly.

## **2. Self-paced reading paradigms**

Self-paced reading paradigms, (also called subject-paced), (Mitchell, 1984) are technically easier to administer than eye movement monitoring. The materials are again presented on a computer screen, but the presentation rate is determined by the subject. After reading each display, the reader is asked to press a computer key to go on to the next display. The reading times for each display are recorded. Processing difficulties, in this paradigm, are indicated by increases in reading times. The size of the displays is predefined by the experimenter. If each display consists only of a single word, the paradigm is called a word-by-word self-paced reading task. If the displays consist of several words, it is called a phrasal reading task. The word-by-word presentation has the advantage that reading times can be measured for each word, which allows for a more fine-grained analysis than phrasal presentation. However, it also forces the subject to spend a minimum amount of time reading each word of the sentence. Generally, but in particular for function words, this leads to higher processing times than observed in eye movement experiments. One partial remedy for this effect is the so-called moving-window technique (Just & Carpenter, 1992). To provide at least suggestive information about upcoming function words, the initial display consists of dashes in the locations where the characters of the sentence will appear during the trial. Successively, the dashes are replaced by the current display, which in turn is replaced by dashes after it has been read. The

length of the words can in this way be used as a cue for retrieval of previously displayed information and as a cue for predicting the location of function words. In this way, the reading task resembles natural reading more than if the displays are presented individually in the center of the screen.

Phrasal presentation, on the other hand, allows the subject to skip words, but the chosen segmentation of the sentences has an impact on the reading processes. An example which illustrates this phenomenon (see Chapter 4) is a sentence similar to the ones used by Mitchell and Holmes (1985) (the phrasal segmentation is indicated by /):

- 1) After the manager had phoned his wife / started preparing dinner.

In this example, the subjects are led to an interpretation in which the clause boundary coincides with the display boundary (Fodor, 1988). After reading the second display, however, this misanalysis has to be corrected. The second issue which requires comment is the size of the phrases used. If a display contains too many words, the reading time measures are not truly on-line. (The extreme case would be to present the sentence in one display which yields only global sentence reading times). If the sizes of the displays vary widely within the experimental sentences, in particular if the target display is short, while the other displays are much longer, the subjects' might alter their comprehension strategies. An example of this type of segmentation comes from Mitchell, Corley & Garnham (1992), with the middle display being the one of interest:

- 1) The headmaster told the boy that / had been / standing there to go and wait outside his room.

Therefore, if using a phrasal presentation, the segmentation of the materials has to be carefully screened to detect eventual confounds.

Independent of the display size and the segmentation is the mode of presentation. The main distinction is between cumulative and non-cumulative presentation. In cumulative mode, the previously read words remain on the screen, while in non-cumulative mode each display is replaced by its successor. The cumulative presentation has the advantage that rereading of prior material is possible, as it is in normal reading. However, since the reading times are measured for the last display only, increases in reading times can either be caused by longer processing of the current display or by regressions. The method does not allow to distinguish these two factors. Moreover, in some studies it has been observed that subjects tend to adopt the strategy of requesting the next display indiscriminately until they can read the sentence in its entirety. Reading times collected under these circumstances cannot be considered to reflect local processing difficulties. The non-cumulative presentation has the opposite effect. As only one display is visible at a time, the subjects cannot use backtracking in the case of comprehension difficulties. Since subjects realize that they have to rely on their memory in this task, they tend to process each display more thoroughly. Thus, the reading times are longer overall, and comprehension strategies might be invoked which are not utilized under normal reading conditions.

Consequently, in all variations of the self-paced reading task, subjects adjust to task demands. The processing does not proceed as naturally as in eye movement monitoring experiments, and the requirement to request the next display interferes with comprehension processes. Nevertheless, many studies have used self-paced reading paradigms successfully to document processing differences caused by experimental manipulations.

Another important variation within the self-paced reading paradigm concerns the instructions. Besides the instructions to read for comprehension at the usual pace, instructions involving grammaticality judgments have been used. Instead of to indicate a request for the next display, the subject is asked to make a yes/no decision after which the next display is presented. Ford (1989) calls this paradigm the Continuous Syntactic Decision Task (also called "stop-making-sense task" if the presentation is aborted after the subject indicated ungrammaticality), and argues that the decision times reflect syntactic processes more purely than simple reading time measures. An additional

advantage of this paradigm is that the analysis of the grammaticality judgments can provide converging evidence. However, it is not likely that the grammaticality judgments can be made independently of lexical or integrative processes. An important drawback of this method is that the notion of grammaticality is not easily communicated to subjects so that the demands of the judgment task affect the primary task of comprehension.

The analysis of reading time data is more straightforward, and at the same time less informative, than the analysis of eye movement data, since only total reading times for each display are measured. However, the adjustment for the length of the display is trickier. While in eye movement studies a simple division of the reading time by the number of characters is assumed to be sufficient, this is not the case for self-paced reading time data. Since the processing time for the display is added to the reaction time for pressing the key, the appropriate adjustment is to regress each subject's data on the number of characters of the display and use the residuals in the statistical analysis. This procedure estimates the reaction time for requesting the next display as the intercept of the regression equation, and the increase in processing time for each additional character as the slope in the regression equation. In particular if the displays vary in size and the regions compared are not identical in content, a correct analysis must be based on the adjusted reading times.

Although self-paced reading tasks do not reflect all facets of the time course of comprehension, the paradigm emerged as the most important one next to eye movement monitoring. While it is undisputed that eye fixation patterns draw a more differentiated picture of reading comprehension patterns than reading times, there is a trade-off between technical considerations and the sensitivity of the task. Non-cumulative word-by-word presentation, in particular the moving-window technique, allow to analyze the processing times for each word, and the task is not considerably more demanding than reading a printed text. Results from experiments using phrasal presentation or cumulative presentation, on the other hand, are to be taken with caution. For the former, the particular segmentation of the sentences is likely to induce effects which are not present in normal reading, and for the latter, task specific strategies are likely to contaminate the reading time data.

### **3. Dual-task paradigms**

Dual-task paradigms are based on the assumption that the cognitive system has only limited resources available for which several concurrent processes have to compete (Baddeley, 1986). If comprehension proceeds smoothly, more capacity can be allocated for performing a secondary task. If, on the other hand, processing difficulties arise, less resources are available for performing the secondary task.

In dual-task paradigms, subjects are required to respond to a stimulus at a predefined location in the sentence, and the reaction time for these responses are measured. Processing difficulties are reflected in longer reaction times in the secondary task. Although not in principle tied to the oral modality, this paradigm is most often used in a cross-modal fashion. While subjects listen to a tape recorded sentence, they are asked to perform a secondary task, such as a monitoring task, a lexical decision task, or a naming task. In monitoring tasks, subjects have to respond as soon as they detect a target stimulus in the sentence. The target can be, for instance, a phoneme (e.g., "b"), a specific word (e.g., "room"), a word that rhymes with a given word (e.g., the target "head" to the cue "lead"), or a word from a given category (e.g., the target "robin" to the cue "birds"). In all these cases, the targets are part of the sentences, but not predictable from the context. This ensures that differential effects are not due to intra-sentential priming, but to the processing load of the comprehension task.

For lexical decision tasks or naming tasks, a word appears on a computer screen and the subject has to respond with a yes/no answer or name the word. Once more, the probes are not predictable from the context or semantically related to other words in the sentence, in order to avoid priming effects. Unlike in the monitoring tasks, however, the probes are not part of the experimental sentence whose presentation is continued concurrently.

Dual task paradigms have also been used for studying reading comprehension in uni-modal procedures (e.g., Healy, Oliver, & McNamara, 1987, for letter detection). In this case, the sentence is presented on the computer screen, and the lexical decision or naming probe also appears on the screen. The trials in which a response is required are marked in some way (e.g., by inserting asterisks before the probe, or by presenting the probe in a different location), so that they can easily be distinguished from the text presentation. In this case, however, the main advantage of the method, namely the naturalness of the comprehension process is lost. If the text is presented at a fixed rate, for instance in a rapid serial visual presentation paradigm (RSVP), the subjects cannot use their usual comprehension strategies, and the presentation rate cannot be adjusted to the individuals' reading skills. While data from this type of experiment might shed light on the relative difficulties of different materials, they cannot provide information on the syntactic processes used under normal reading conditions. If the presentation is subject-paced, on the other hand, keeping track of the appropriate responses to the two tasks is so difficult that the small effects of increased syntactic processing might be hidden. For both types of presentation, the sentence context has to be interrupted for the secondary task trials.

In contrast to reading time methods, the dual task paradigm does not provide information on the time course of processing throughout the sentence; rather, it is used to test specific predictions on the demands for the processing of different types of sentences at a predefined location. Thus, no information is available for other locations within the sentence (see Marslen-Wilson & Tyler, 1980, for an exception). It is possible, therefore, that the processing differences arise earlier, or continue after the point of test. In most applications of this method, however, the stimulus is presented at the first location at which differences in processing can appear (for instance, immediately after the first word at which two initially identical sentences diverge).

In summary, the dual-task paradigm fulfills the criteria for a useful method to study normal syntactic processing on-line - as long as a cross-modal procedure is used. Listening to speech is the most natural condition in which people comprehend language. The demands of the secondary tasks are negligible, and reaction times in target detection, lexical decision or naming tasks have been proven to be very sensitive. Unfortunately, results from studies on speech comprehension do not always generalize to reading comprehension (Flores d'Arcais, 1990; Shapiro, Zurif, & Grimshaw, 1991). For written language comprehension, in contrast, the method does not seem appropriate because the reading conditions are distorted.

#### **4. Priming paradigms**

Priming paradigms (McKoon & Ratcliff, 1984; Tanenhaus, Carlson, & Seidenberg, 1985) are closely related to the dual-task paradigm with respect to the tasks employed and the methodological issues which have to be addressed. In particular, the same differences hold between cross-modal priming paradigms, in which the sentences are presented orally, while the lexical decision or naming probes are presented visually, and uni-modal priming paradigms, in which reading of a sentence has to be interrupted by the priming trials. The assumptions underlying the interpretation of the results are very different, though. While using lexical decision tasks and naming tasks (also, but rarely, Stroop tasks) in similar experimental procedures, priming experiments employ probes which are carefully selected to be closely related to the sentence context. The questions addressed do not as much concern the processing difficulties of the materials, as the particular interpretation entertained at a given point during comprehension. Assumptions which allow the interpretation of priming data rest on two basic findings. In syntactic priming, (Wright & Garrett, 1984; West & Stanovich, 1986) it has been shown that the lexical class of a probe affects the latency in naming and lexical decision tasks. Targets which are syntactically impermissible in the given sentence context are inhibited compared to grammatical continuations of the sentence.

In associative priming studies, on the other hand, the facilitation of responses, which is indicated in lower reaction times, is considered evidence for the activation of the target (e.g. Swinney, 1979). The experimental questions which can be successfully answered in an associative priming study are thus highly dependent on the nature of the syntactic phenomenon. Only comparisons of sentences



whose different structural properties can be captured in assumptions about activation levels of specific target words are feasible. The applications of this method to parsing issues have therefore been mostly concerned with pronoun reference assignments, anaphoric processes, and filler-gap assignments. In these cases, the candidate referents are assumed to be differentially activated depending on the specific assignment made.

The on-line methods described in this section are all suitable for studying syntactic processing. Eye fixation paradigms and reading time paradigms provide data over the entire course of comprehension. While eye movement recording does not require the subjects to adjust to specific task demands, reading time data are easier to collect and analyze. Paradigms employing monitoring, lexical decision or naming have the advantage that reaction times in these tasks are extremely sensitive. However, experiments using these tasks can only be used to test very specific hypotheses, and the experimental design has to be carefully controlled in order to avoid confounding factors. Ideally, results from the different types of on-line methods should provide converging evidence. As will be seen in the following two chapters, this is not always the case. Thus, to be able to evaluate the empirical results, it is important to keep in mind the differences between the paradigms.

## CHAPTER 4 LEXICAL EFFECTS ON SYNTACTIC PROCESSING

The influence of lexical information, in particular, information about the lexical properties of the predicate in a sentence has received some attention in psycholinguistic research. Although other lexical items are also assumed to carry lexical information, the experimental research has focussed on verb effects. Lexical entries of a verb are considered to consist of its syntactic category, its subcategorization frame, possible argument structures, and thematic information. For example the lexical entry of the verb "give" might look like this (from Shapiro et al., 1991):

<i>give</i>	syntactic category:	verb
	strict subcategorization:	[NP PP] [NP NP]
	argument structure:	(x, y, z)
	thematic grid:	(agent, theme, goal)

The subcategorization frames are also assumed to contain information about their relative frequencies. In addition, selection restrictions which specify if a certain type of argument is required for a slot in the thematic grid can be part of the lexical entry. For instance, "intend" requires a human (or at least animate) subject, and "drink" only takes liquids as direct objects.

In a model of how lexical information is used in parsing, several questions have to be addressed. Thus, the following review of psycholinguistic studies concerned with how these lexical entries influence parsing processes is divided into four sections. The first section summarizes work on verb complexity as defined by the number of structures a verb allows. The second section includes experiments which address the question of whether subcategorization preferences are immediately taken into account to facilitate syntactic decisions. The third section tries to answer the question of whether interpreting a phrase as an argument in a verb frame has precedence over other syntactic interpretations. And finally, the importance of selection restrictions in guiding parsing decisions is discussed.

The goal of this chapter is to provide the necessary background for the review of studies concerned with discourse influences in parsing. Thus, the review of lexical effects is restricted mainly to studies which employ reduced relative clauses, reduced complement clauses, and prepositional phrase attachment ambiguities. While other syntactic phenomena have also been studied (e.g., gap-filling, Boland, Tanenhaus & Garnsey, 1990, Tanenhaus, Garnsey, & Boland, 1990; long-distance dependencies, Clifton & Frazier, 1989, Tanenhaus & Carlson, 1989; anaphoric reference, Clifton & Ferreira, 1987, Garnham, 1989; pronoun reference, Greene, McKoon, & Ratcliff, 1992), the experiments concerned with discourse influences make use only of these three types of ambiguities.

### 1. Verb complexity

A series of studies by Shapiro and colleagues (Shapiro, Zurif & Grimshaw, 1987, 1989; Shapiro, Brookins, Gordon & Nagel, 1991) is concerned with processing difficulties induced by verb complexity. In this work, syntactic verb complexity is defined in terms of either the number of subcategorization frames, or the number of predicate-argument structures. If it is assumed that processing of a verb includes the access of the lexical information associated with it, the verb complexity should be related to processing difficulties.

In their first study, Shapiro, Zurif and Grimshaw (1987) investigated whether subcategorization or predicate-argument structure affected verb processing. Using a secondary lexical decision task for an unrelated target word in the vicinity of the verb, they concluded that the number of different argument structures had a larger impact on the lexical decision times than the number of subcategorization frames.

In a follow-up study, Shapiro, Zurif and Grimshaw (1989) used sentence contexts which biased towards one of the possible argument structures. If only the appropriate argument structure became activated in these contexts, increased processing times should disappear. On the other hand, if all of the argument structures were activated initially, the sentence bias should not facilitate verb processing. In analogy to findings on the lexical access of polysemous nouns, Shapiro et al. (1989) expected that even in the biasing contexts the number of argument structures would predict processing complexity. Sentences were written using a syntactic structure which unambiguously provided the arguments before the verb was encountered. For instance, the sentence "It was the girl to whom the letter was sent." specifies the direct and indirect objects prior to the verb. The results supported the notion of exhaustive parallel activation of the argument-structure possibilities. Shapiro et al. interpret this exhaustive activation as evidence for the contextual impenetrability of the grammatical subsystem which processes verb information.

Schmauder (1991) pointed out several methodological problems with these studies. For instance, each of the verbs used in occurred in more than five sentences. This repetition might have confounded the results. The sentences in which the verbs were embedded were not constant, so that the context could have given rise to processing differences. Moreover, the frequency within the verb classes were not controlled, and only two sets of probes for the lexical decision task were used.

In four experiments which avoided these confounds, Schmauder (1991) failed to replicate Shapiro et al.'s (1987) results. As an additional independent variable, Schmauder used two types of sentence frames which differed in their syntactic complexity (the verb was preceded by a prepositional phrase or an embedded relative clause). Eye fixation times for the verbs could be predicted by frequency and length, but neither subcategorization complexity, nor argument-structure complexity yielded additional effects. In dual-task paradigms using lexical decision and naming tasks, the type of syntactic structure of the sentence frames influenced latencies, but verb class did not. Schmauder (1991) concluded that the Shapiro et al.'s findings were an artifact of their procedure.

The issue has not been settled, though. Shapiro, Brookins, Gordon, and Nagel (1991) attempted to replicate their results once more, again using a cross-modal lexical decision task. Specifically addressing Schmauder's criticisms, they found shorter latencies for purely transitive verbs (which appear in only one argument frame) than for dative verbs (which have two argument frames). Similarly, the reaction times were different for two-complement verbs (two subcategorization frames and two argument frames) as compared to four-complement verbs (two subcategorization frames and two argument frames). Shapiro et al. (1991) explain the differences of these two studies with the relative difficulties of the secondary task. While Schmauder used only short, frequent probes, Shapiro et al. used probes of various lengths and frequencies. Thus, in the former study, the secondary task was too easy to capture processing difficulties caused by verb complexity. This account was confirmed in further experiments.

In summary, this series of studies did not yield a conclusive answer to the question if access of a verb with a more complex lexical entry increases processing demands. While the cross-modal dual task paradigm seems to provide evidence for verb complexity effects, these effects could not be replicated using an eye-movement paradigm (for a similar failure to find verb complexity effects, using a semantic definition, see Rayner and Duffy, 1986). Thus, results from sentence processing studies using oral presentation might not be generalizable to written language comprehension.

## 2. Subcategorization preference

The hypothesis that the complexity of the lexical entry of a verb influences its processing rests on the assumption that access of the verb involves activation of all features of its lexical entry. However, lexical parsing models do not always postulate this parallelism. Instead, Bresnan and Kaplan (1982) argue that lexical information enters parsing processes through the activation of the most frequent, preferred subcategorization frame. In this account, processing difficulties should not depend on verb complexity, but on the need of the parser to postulate an unpreferred subcategorization frame.<sup>1</sup>

Clifton, Frazier, and Connine (1984) used a dual-task paradigm to study this hypothesis. Transitive or intransitive usage of verbs was crossed with normative frequency of either transitive or intransitive subcategorization. An example of the four resulting conditions is (p. 698):

- 1) transitive preference, transitive sentence:  
The babysitter read the story to the sick child.
- 2) transitive preference, intransitive sentence:  
The babysitter read to the sick child.
- 3) intransitive preference, transitive sentence:  
The babysitter sang the story to the sick child.
- 4) intransitive preference, intransitive sentence:  
The babysitter sang to the sick child.

The dependent variable was reaction time in a lexical decision task. The sentences were presented word-by-word, and the unrelated probe was presented after the first word following the verb. This word unambiguously indicates the structure of the sentence (a preposition follows the verb only in intransitive sentences). The results showed that in the two congruent conditions, in which verb preference and sentence frame matched, the secondary task was easier than in the incongruent conditions. Processing, therefore, was more difficult if the expectations induced by subcategorization information were violated.

However, as pointed out in Chapter 3, a uni-modal dual task paradigm cannot be considered reflecting normal reading processes. The words were presented for 350 msec each, and the words immediately preceding the probe were short function words, such as personal pronouns, articles, or prepositions. The relatively long processing of these words, induced by the experimental procedure, could have led to integrative processes which might not, or not at this location, be observed in natural comprehension.

Mitchell and Holmes (1985) tested the influence of the subcategorization preference using a phrasal self-paced reading task. In each of 24 locally ambiguous sentence frames, two verbs were inserted. The verb preference allowed to disambiguate the structure in one case but not in the other. For example, in the sentences

- 1) optionally transitive:  
As soon as he had *phoned* his wife started to prepare dinner.
- 2) intransitive:  
As soon as he had *arrived* his wife started to prepare dinner.

the correct localization of the clause boundary is facilitated in sentence 2), if the verb information is taken into account. The verb bias was pretested in a study in which appropriate paraphrases of the ambiguous sentence fragments had to be selected.

<sup>1</sup> In the remainder of this paper I make a distinction between subcategorization and predicate-argument frames only when necessary. For the purpose of most of the studies reviewed, both types of lexical information capture the experimental variables.

Reading times from a phrasal self-paced task clearly reflected the verb preference patterns in all four types of structural ambiguities used (relative vs. adverbial, early vs. late closure, relative vs. main adverb, complement vs. direct object). Reading times were significantly shorter for the disambiguating region ("started to prepare") when the sentence interpretation was in line with the verb preference.

These results have to be taken with caution, though. The phrasal presentation suggests segmentation in a meaningful way. It is more likely that readers take each display as a phrase than to place a clause boundary in the middle of the display. In the materials used in Mitchell and Holmes (1985), this is exactly what was required in the region before the test segment (e.g., "arrived his wife"). An analysis of the reading times in this crucial segment could clarify if the segmentation induced processing difficulties for sentences of type 1). Unfortunately, Mitchell and Holmes do not report this analysis.

Although they acknowledge the possible influence of the segmentation patterns on the reading times, their second and third experiments cannot resolve this issue. In the second experiment, processing times for entire sentences were measured; and in the third experiment, the sentences were disambiguated using punctuation or different word orders. While the global reading times replicated the verb effects, the disambiguation eliminated them.

In a follow-up experiment, Mitchell (1987) showed that the influence of segmentation is in fact very strong. In this study, similar sentences were segmented in two different ways, as the following example illustrates ( / indicates segmentation):

- 1) First segmentation:  
After the child had sneezed / the doctor prescribed a course of injections.
- 2) Second segmentation:  
After the child had sneezed the doctor / prescribed a course of injections.

Garden-pathing was eliminated if the segmentation was at the clause boundary, as in 1). These results, however, do not provide information about the on-line processes involved. The segments were rather long, and the reading times per segment were about three seconds. An on-line task which allows to assess the location of the processing differences would be necessary.

Ferreira and Henderson (1990) investigated if syntactic attachment preferences could be overridden by verb subcategorization information. In addition, they attempted to systematically compare three on-line paradigms. In three experiments, they used locally ambiguous sentences of the form

- 1) direct-object bias (non-minimal attachment verb):  
He wished Pam needed a ride with him.
- 2) direct-object or complement (minimal attachment verb):  
He forgot Pam needed a ride with him.

In both cases, "Pam" can be interpreted as the direct object of the verb until the disambiguating information is encountered. This interpretation is the preferred one as specified by the principle of minimal attachment. The verbs differed in their subcategorization preference. While "wish" only rarely takes a direct object, "forget" can either be followed by a direct object or by a sentence complement. Each pair of verbs was inserted in four different sentence frames. As a control, two matched unambiguous sentences containing the complementizer "that" were constructed. If verb information guides parsing decisions initially, no processing difficulties should be observed for sentence 1). If, on the other hand, the minimal attachment principle always determines initial parsing decisions, no differences between the two sentences are expected.

In the first experiment, first pass eye movement data showed no effects of verb type. As predicted, the presence of the complementizer led to faster reading times in the disambiguating region

(e.g., "needed"), but there was no interaction of complementizer presence with verb type. For the total fixation times, there was an indication of verb bias effects in the postdisambiguating region (e.g., "a ride") with shorter reading times for sentences containing a non-minimal attachment verb. However, this effect was only significant by subjects but not by items. The lack of a verb effect can be interpreted in two ways. Consistent with the predictions, it might be that the parsing mechanism is not influenced initially by verb information. Another possibility is that the selection of the materials was not adequate. The verbs were chosen in part on the basis of normative data on subcategorization preference (Connine, Ferreira, Jones, Clifton, & Frazier, 1984), and in part on the intuitions of the authors, but no pretest was conducted to confirm the effectiveness of the bias.

To rule out this second explanation, Ferreira and Henderson conducted a second experiment using a word-by-word self-paced non-cumulative reading paradigm. Besides the strong effect of complementizer presence on reading times of the disambiguating region, there was an interaction between complementizer presence and verb type in the postdisambiguating region. Only when the complementizer was absent reading times were longer for the sentences in which the verb allowed a minimal attachment reading. An analysis of sentence reading times yielded a similar pattern, showing a verb bias effect only for the sentences without the complementizer. Ferreira and Henderson take these results as proof that the independent variable verb type was effectively manipulated. Moreover, they conclude that verb information is not taken into account immediately but at a later stage in processing.

The third experiment is interesting from a methodological point of view. Using the same materials, the data from a cumulative self-paced reading task were not as differentiated as the data from the non-cumulative presentation. In the disambiguating region, only the effect of complementizer presence was significant, while the total reading times were affected only by verb type.

The results of these experiments are taken as strong evidence for a garden-path model of parsing in which verb frame information does not enter the analysis of a sentence immediately. The syntactic principle of minimal attachment could not be overridden by lexical verb subcategorization preferences. However, this conclusion is drawn mainly based on the null effects of verb bias in Experiment 1. Thus, it seems like the appropriateness of the selection of verbs and experimental sentences has to be confirmed more convincingly than only with the interaction results of Experiment 2. Inspection of the sentence frames shows that about half of the sentences are not pragmatically plausible if interpreted using minimal attachment. E.g., the sentence fragments "Sue wrote Iowa...", "Jan warned the fire...", "Ed recalled wine..." use arguments which are pragmatically inadequate as direct objects of the verb, since they violate thematic constraints. If besides lexical information pragmatic information is also used in the parsing process, this might be a crucial factor (see Chapter 5).

More importantly, although the non-minimal attachment verbs might have a preference against a direct object, some of them do allow a direct object, and in many cases the preference does not seem to be very strong. For example, the sentence fragments "Jack revealed news...", or "She asserted ideas..." have perfectly plausible continuations in which the second noun is the direct object of the verb (e.g., "He wished Pam good luck," "Jack revealed news about his supervisor," "She asserted ideas which were not proven"). A more thorough analysis of the subcategorization preferences of the verbs would have been useful. Without a pretest of the sentences used in the study, it is not safe to conclude that lexical information is not immediately available.

In the comparison of the three paradigms, the outcome is also not absolutely clear. The eye-movement data and the data from the non-cumulative reading task did not yield converging results. The theoretically crucial interaction between verb type and complementizer presence in the postdisambiguating region was only significant in the self-paced reading task, but was not found in the analyses of eye fixations. Nevertheless, the authors argue that eye movement monitoring is the preferred paradigm. They explain the differences with a task specific strategy the subjects use. Since in the word-by-word presentation reanalysis is not possible, subjects might take advantage of the verb information more than under natural reading conditions. In the comparison between the

cumulative and non-cumulative presentation modes, on the other hand, the conclusions were different. The lack of significant effects in the cumulative presentation (which allows reanalysis) led Ferreira and Henderson to conclude that this paradigm is not sensitive enough to capture syntactic processing differences.

However, Holmes (1987) found verb effects using a cumulative self-paced reading paradigm and similar materials. In contrast to Ferreira and Henderson, she confirmed the selection of the verbs using a sentence completion task. Only 16 verbs with a very strong preference for either a sentential complement or a direct object were chosen from a large pool. Verbs biased towards taking a direct object led to large reading time increases on the disambiguating word if the complementizer was absent, and this effect was much smaller for verbs biased towards a sentential complement (i.e., an effect corresponding to an interaction between the variables complementizer presence and verb bias was found). Thus, it does not seem conclusive to dismiss the cumulative presentation paradigm based on the null effects obtained in Ferreira and Henderson (1990).

Holmes, Stowe, and Cupples (1989) also contrasted what they call complement verbs, which can take either a direct object or a clausal complement, with verbs which are strongly biased in favor of a direct object continuation. In addition, they systematically manipulated pragmatic plausibility of the noun as a direct object of the verb. A prior sentence continuation experiment confirmed the selection of verbs in the two classes. An example of a sentence pair used in the study is:

- 1) complement bias:  
The historian suspected the manuscript of his book had been lost.
- 2) direct-object bias:  
The historian read the manuscript of his book had been lost.

Control sentences containing the complementizer "that" were also included. Moreover, the plausibility of the second noun as a direct object of the verb was varied in order to bias towards a sentential complement interpretation. For example, since a sentence such as "The candidate doubted his champagne." is highly implausible, it is likely that a sentential continuation follows (e.g., "The candidate doubted his champagne would be appreciated."). If pragmatic knowledge and verb information are taken into account immediately, processing difficulties for the complement clauses are only expected for the plausible sentences containing direct-object bias verbs, but not for the other three sentence types.

In Experiment 1 reading times for each word were measured in a cumulative grammaticality judgment task. The displays crucial to the experimental manipulations were the locally ambiguous second noun phrase of the sentences and the first (usually an auxiliary verb) and second word of the subsequent disambiguating phrase. The pattern of results for the direct-object verbs revealed that plausibility had an impact only on the reading times of the noun, but not on the reading times of the disambiguating information. Verb bias influenced the magnitude of the processing cost increase on the auxiliary (the first disambiguating word). For sentences containing a direct-object verb, subjects needed 400 msec more to read the auxiliary in the reduced relative forms than in the unreduced forms. For sentences containing complement verbs, this difference was only 110 msec.

Similar evidence for verb effects was found in a second experiment employing a cumulative reading task with the instruction to memorize the sentences (instead of word-by-word grammaticality judgments). In this second experiment, however, the processing difficulties for the direct-object bias sentences manifested themselves not on the first, but on the second word of the disambiguating phrase, and there was no effect of plausibility.

Finally, in a third experiment, Holmes et al. (1989) investigated whether the length of the noun phrase is related to the observation of garden path effects (plausibility was not considered in this experiment). The sentences were presented word-by-word, non-cumulatively, and subjects' comprehension was assessed using simple yes/no questions. The second noun phrase was either short (e.g., "the story"), or long (e.g., "the story about the accident"). For direct-object bias

sentences, there were no effects of noun phrase length. For the complement bias sentences, the auxiliary after the long noun phrase in the reduced relative sentences was read longer than in the other three sentence types. Although the interaction between length of noun phrase and complementizer presence was not significant, pairwise comparisons showed that a significant effect of complementizer presence was found only after long noun phrases but not after short ones. This result was not expected, and thus deserves some comment. It seems as if the hypothesis of the noun phrase being the subject of a complement was abandoned during processing of the long noun phrases in favor of direct object reading. If this interpretation is correct, subjects should have been led down the garden-path twice. Longer reading times, thus, should have been observed also for the region before the disambiguating information in the reduced relative sentences as compared to the unreduced sentences. Unfortunately, Holmes, Stowe, and Cupples (1989) do not report analyses of the reading times for these three words.

For the word after the auxiliary, longer reading times were observed for reduced sentences than for unreduced sentences, and longer reading times were observed for sentences containing long noun phrases than for sentences containing short noun phrases. To summarize, the results for the short noun phrase sentences in Experiment 3 replicate those of the two previous experiments, even though different presentation paradigms were employed. Cumulative presentation with either grammaticality judgment or memorization instructions, and non-cumulative presentation with subsequent comprehension questions yielded similar lexical effects. However, an effect of plausibility was only observed in the grammaticality judgment task.

The studies on subcategorization preference yielded different results. While studies using a dual-task paradigm (Clifton, Frazier & Connine, 1984) and self-paced reading paradigms (e.g., Mitchell & Holmes, 1984) confirmed that lexical information is used immediately in parsing, the only study which employed an eye movement monitoring paradigm did not (Ferreira & Henderson, 1990). Because of the methodological problems in this study, further research is necessary to investigate whether effects of subcategorization preferences are induced by task specific strategies. If lexical information plays an important role in syntactic processing, its effects should also manifest themselves in eye fixation patterns.

### 3. Argument preference

Assuming that subcategorization information or argument structures are used to guide parsing decisions, the postulation of a subcategorization frame should lead the parsing mechanism to attempt to fill its empty slots. For instance, if a verb is associated with a subcategorization frame of the form [ \_ NP PP ], a prepositional phrase should be expected after reading the subject and the verb of the sentence. Ford (1989) formulated this mechanism as the Lexical Preference Principle, and tested it in several experiments.

In Experiment 1, Ford (1989) contrasted unambiguous sentences of the form

- 1) The secretary owned the paintings in the office.
- 2) The secretary stuck the paintings in the office.

The verb "owned" in sentence 1) has a simple structure which postulates only a subject and a direct object, whereas "stuck" is associated with a direct object and a locative prepositional phrase. If the lexical structure facilitates processing, sentence 1) is predicted to be more difficult than sentence 2).

Using a cumulative self-paced reading task, in which subjects were asked to judge the grammaticality of the sentence after each word, Ford (1989) confirmed this prediction. Processing times increased by about 100 msec on both the preposition ("in") and the noun in the prepositional phrase ("office") in sentence 1) as compared to sentence 2). In particular the ease of processing the



preposition in sentences of type 2) leads to the conclusion that the structural expectations influenced syntactic processing.

In a second experiment (using the same paradigm), this conclusion was tested using sentences in which the expectations for a prepositional phrase were violated. Adding a relative clause to modify the object noun phrase yielded sentences of the following types:

- 1) The secretary owned the paintings that we hated.
- 2) The secretary stuck the paintings that we hated under the table.

As predicted, the relative pronoun "that" was more difficult to process in sentence 2) than in sentence 1). The need of the parsing mechanism to fill the prepositional phrase slot in the postulated subcategorization frame led to increased reading times for the unexpected continuation.

Clifton, Speer and Abney (1990) also tested the lexical preference principle, although their theoretical background is slightly different. According to Abney's parsing model (Abney, 1989), there is a preference ordering for three syntactic principles. The most important principle is the preference of arguments over adjuncts. This principle holds that it is easier to interpret a noun phrase as an argument of an already encountered lexical item, than to interpret it as an adjunct. Thus, lexical information about the argument structure of the verb is used in the application of this principle. The second principle corresponds to the minimal attachment principle, and states that verb phrase attachments of prepositional phrases are preferred over noun phrase attachments. These two distinctions were crossed in a study by Clifton, Speer and Abney (1991). Sentences were constructed in which the syntactically ambiguous prepositional phrase attachment was disambiguated using pragmatic information, resulting in the following four conditions:

- 1) verb attachment, argument:  
The saleswoman tried to interest the man in a wallet during the storewide sale at Steigers.
- 2) verb attachment, adjunct:  
The man expressed his interest in a hurry during the storewide sale at Steigers.
- 3) noun attachment, argument:  
The man expressed his interest in a wallet during the storewide sale at Steigers.
- 4) noun attachment, adjunct:  
The saleswoman tried to interest the man in his fifties during the storewide sale at Steigers.

According to the principle of minimal attachment, processing difficulties should arise only for sentences 3 and 4 in which a non-minimal attachment is required. If the argument preference principle precedes minimal attachment, in contrast, the predictions are that processing difficulties should be observed for sentences 2 and 4, in which the prepositional phrase is interpreted as an adjunct.

The results of a phrasal self-paced reading task supported the predictions of the phrase-structure model in the region of the ambiguous prepositional phrase (e.g., "in a wallet"). Reading times per character were longer for the non-minimal attachment sentences. However, for the phrase after the prepositional phrase ("during the storewide sale"), the results supported the predictions of a verb frame based model, because the reading times were longer for the sentences containing an adjunct than for the sentences in which the prepositional phrase was an argument. For none of the regions an interaction between attachment type and argument type was found.

In order to confirm that these results were not task-specific, the same materials were used in an eye movement monitoring paradigm. Replicating the early influence of minimal attachment found in the first experiment, the first pass eye fixations indicated an advantage for minimal attachment sentences in the prepositional phrase. Replicating the later influence of argument status, an advantage of arguments over adjuncts was found in the next region. The total reading times for either region were only affected by the argument vs. adjunct distinction. An unexpected result was

found in the analysis of regressions. Contrary to the predictions of garden-path models, subjects were more likely to regress out of the prepositional phrase when it was minimally attached.

The results of both experiments provide evidence against parsing models which propose that lexical information is immediately used in sentence processing. Although strong effects of both attachment and argument status were observed, the lexical verb frame information was taken into account later than the purely syntactic principle. Therefore, the results did not confirm the preference ordering of the principles as suggested by Abney (1989).

Two issues have to be discussed before these results can be interpreted. The first is that the verbs used in the study were not analyzed with respect to their subcategorization (or argument) frames. The sentences used in the study contained verb arguments in the form of prepositional phrases. Thus, the subcategorization frame which has to be used is of the form [\_ NP PP]. It seems that the argument preference principle should be affected by the subcategorization possibilities. If a verb is also subcategorized for an alternative structure, the principle might not be as effective. Of course, this reasoning rests on the assumption that subcategorization information is indeed taken into account immediately, and that the argument preference principle serves the purpose of filling slots in a postulated predicate-argument frame (Ford, 1989).

A more important issue, because it does not rest on a theoretical argument, is that the sentences used by Clifton, Speer and Abney were not syntactically identical. The wording of sentences 2) and 3) before the prepositional phrase differed from that of sentences 1) and 4). In general, it cannot be evaluated how the differences in sentence context affected processing. Specifically, the lexical information necessary to obtain the correct interpretation of a thematic argument was associated with the verb ("interest") in sentences 1) and 4), but with the noun ("interest") in sentences 2) and 3). It is far from obvious that lexical information associated with nouns yields similar effects on syntactic processes as information associated with verbs. In fact, in the discussion of their experiments, Clifton, Speer and Abney (1990) propose a modification of Abney's model which makes an explicit distinction concerning the processing of verbs and nouns.

Nevertheless, the experiments by Ford (1989) and Clifton, Speer and Abney (1990) show that after postulating a subcategorization frame, the parsing mechanism forms expectations consistent with this structure. Obtaining converging results in different paradigms, these studies also provide further evidence for the use of lexical information in syntactic processing. However, the location of the influence of lexical information is not agreed upon.

#### 4. Selection restrictions

The second type of lexical information which the syntactic processing mechanism can only use after the postulation of a subcategorization frame are selection restrictions. Instead of merely attempting to fill empty slots in the frame, the parsing mechanism should attempt to find appropriate fillers. Selection restrictions have also been interpreted as semantic, since they concern category membership of concepts. However, independent of the classification of this type of information, evidence for its immediate use cannot be explained within a rule-based parsing model.

Ferreira and Clifton (1986) used an animacy constraint and contrasted the processing of sentences such as

- 1) animate subject:  
The defendant examined by the lawyer turned out to be unreliable.
- 2) inanimate subject:  
The evidence examined by the lawyer turned out to be unreliable.

The syntactic ambiguity of the verb "examined" as either a past tense main verb, or the past participle in a reduced relative clause, is disambiguated by the phrase "by the lawyer." However, if selection restrictions are taken into account immediately, only sentence 1) should lead to the incorrect, preferred reading of "examined" as the verb of the main clause.

Ferreira and Clifton (1986) used an eye movement monitoring paradigm to test whether this information affected parsing. The comparison of sentence 1) and 2) with unambiguous sentences containing a relative pronoun showed that the ambiguity led to processing difficulties. However, for both first-pass and second-pass reading times, the inanimacy of the subject in sentence 2) did not eliminate the garden-path. Although reading times for the reduced, inanimate condition were longer on the verb ("examined"), this did not lead to easier processing of the following disambiguating region. Thus, in this experiment, selection restrictions did not seem to play a role in syntactic processing.

In a similar experiment, Stowe (1989) did find effects of selection restrictions. In sentences with ambiguous clause boundaries, such as used by Mitchell and Holmes (1984), the animacy of the first noun phrase was varied. Only if this noun phrase is animate, a direct object reading of the second noun phrase is felicitous. In addition to resulting sentences of the form

- 1) animate subject:  
When the police stopped the driver was already getting nervous.
- 2) inanimate subject:  
When the truck stopped the driver was already getting nervous.

unambiguous control sentences were used. In a cumulative, self-paced grammaticality judgment task, there were no differences in reading times for any of the words except for the disambiguating second verb ("was"). The ambiguous, animate condition elevated the reading times for this verb by about 600 msec compared to the reading times in the other three conditions. The inanimacy of the subject was used successfully to facilitate processing of the ambiguous sentences.

Just and Carpenter (1992; MacDonald, Just & Carpenter, 1991; King & Just, 1991) try to account for these seemingly contradictory findings with their capacity theory of parsing. In this theory, the working memory capacity for language is considered a crucial explanatory factor for reading skills (see Crain & Shankweiler, 1990; McRoy & Hirst, 1990; for other approaches which consider working memory effects in parsing). Assuming that processing costs are associated with considering multiple sources of information and with keeping multiple interpretations activated, non-syntactic effects in sentence processing should only be observable in readers with a large memory span. As in previous work (Daneman & Carpenter, 1980; Carpenter & Daneman, 1981), working memory capacity for reading is assessed using the reading span task. This task requires subjects to hold the last word of each of several presented sentences in working memory, while comprehending subsequent sentences. The number of last words recalled is defined as the reading span. Good readers have a reading span of about five to six words, whereas the reading span of bad readers is only about two words.

Evidence for the impact of reading skills on the use of selection restrictions in sentence processing comes from an experiment reported in Just and Carpenter (1992). Before monitoring eye movements during reading of the sentences used in Ferreira and Clifton (1986, Experiment 1), Just and Carpenter assessed subjects' reading spans. As predicted by the capacity theory, only the low span subjects could not take into account the animacy of the sentence subject. For these subjects, the Ferreira and Clifton results were replicated, in that no differences between the two sentence types were found in the first-pass reading times. The high span subjects' fixations on the disambiguating phrase, on the other hand, revealed a different pattern. While the fixation times for the sentences which started with an animate noun were similar to those of the low span subjects, the high span subjects had a considerable advantage on the sentences starting with an inanimate noun. For the total reading times, the interaction between animacy and reading span was no longer significant. Only the high span subjects could make use of the selection restriction immediately, whereas all subjects took

this information into account eventually. Thus, the results for the high span subjects indicated on-line influences of the lexical information.

These results cannot be accounted for in a strictly rule-based parsing model. While the experiment of Stowe (1989) could be criticized for the use of a grammaticality judgment task, the findings of Just and Carpenter (1992) were obtained in an eye movement paradigm. Thus, selection restrictions seem to have an immediate effect in sentence processing.

## 5. Summary

The studies reviewed in this section were concerned with the psychological validity of lexical information in sentence processing. In particular, the influence of subcategorization information associated with verbs was addressed. Although the notion of syntactic verb complexity as one determinant of syntactic processing difficulty yielded some support, it came only from experiments on spoken language comprehension. The preference of a particular subcategorization frame, on the other hand, could only be shown to influence parsing in self-paced reading tasks. In both domains, eye movement paradigms (which are often considered to yield the most sensitive data) could not confirm the positive results.

In contrast, eye-movement studies on argument preferences and selection restrictions did yield evidence for lexical effects. For argument preferences, the unresolved issue is the exact location of the use of lexical information. For selection restrictions, seemingly contradictory findings could be reconciled if subjects' reading skills were taken into account. Since employing these two types of lexical information presumes the prior postulation of a particular subcategorization frame, these studies provide additional, indirect evidence for the importance of subcategorization information.

## CHAPTER 5 DISCOURSE EFFECTS ON SYNTACTIC PROCESSING

The previous chapter showed that lexical information can influence syntactic processing. However, the differences induced by lexical information can also be explained using the conceptual relationships between the words in an utterance. The thematic grid of a verb, for instance, contains information about the roles with which its arguments are associated, and selection restrictions provide constraints about the type of filler which can appear in the argument slots. Depending on the underlying grammar theory, and on the assumptions about the representation of semantic and pragmatic knowledge, this type of information can be interpreted as either lexical or conceptual. The boundary between pragmatic, semantic, and syntactic information is therefore not well-defined (see Fillmore, 1984).

In this chapter, studies are reviewed which investigate the influence of clearly non-syntactic information on parsing processes. The experiments have in common that the contextual manipulations cannot be accounted for in a purely lexical framework. In the first section, I consider studies which are concerned with the role of pragmatic plausibility, and more importantly, of sentence context. The second section contains a review of psycholinguistic research on the effects of discourse context.

### 1. Pragmatic knowledge and sentence context

Rayner, Carlson & Frazier (1983) varied the pragmatic plausibility of the filler for the agent role of verbs, in both syntactically ambiguous and unambiguous sentences. To each of twelve verbs, nouns were selected which were typical fillers for the agent or for the recipient role. In the pragmatically implausible condition, the action (e.g., "pay") was performed by the typical recipient (e.g., "clerk"), instead of by the typical agent (e.g., "customer"). The locally ambiguous sentences were reduced relative clauses, whereas the unambiguous sentences were unreduced relatives and active sentences. Examples for the resulting four types of experimental sentences are (p. 361):

- 1) reduced implausible:  
The florist sent the flowers was very pleased.
- 2) reduced plausible:  
The performer sent the flowers was very pleased.
- 3) unreduced plausible:  
The performer who was sent the flowers was very pleased.
- 4) active implausible:  
The performer sent the flowers and was very pleased.

A sentence completion experiment confirmed that the pragmatic plausibility determined the frequency of assigning a reduced relative reading to the sentence fragment "sent the flowers." However, in the analyses of eye fixations there were no effects of plausibility. For both versions of the reduced relative sentences, reading times for the disambiguating region were considerably longer than in the unambiguous sentences.

Unfortunately, the unambiguous sentences 3) and 4) did not have the same syntactic structure, and "plausibility" and "ambiguity" were not treated as two independent variables. Thus, the statistical comparisons of the four sentence types are not as informative as they could be. First, the active sentence 4) was read more easily than the three sentences involving a relative clause. Second, the high number of regressions for sentence 3), which contained an unambiguous relative clause, suggested that processing a relative clause is difficult even without garden-pathing. The

experimental design would have been more convincing if the sentences in the unambiguous, implausible condition had been sentences of the form

4) unreduced, implausible:

The florist who was sent the flowers was very pleased.

In this case, the processing costs associated with the implausibility of the agent could have been measured in an unambiguous syntactic structure. Since a direct comparison of plausible and implausible unambiguous sentences is not available, the lack of a plausibility effect in the reduced relative clauses is not interpretable.

An alternative conclusion is that plausibility differences as used by Rayner, Carlson, and Frazier (1983) are indeed not taken into consideration during syntactic processing. The degree of the implausibility induced by the choice of an unexpected agent might not be sufficient to affect the syntactic analysis. In fact, the implausible sentences used in this study do make sense, although they are slightly odd.

In a similar study, Marslen-Wilson, Brown and Tyler (1988) used pragmatic violations which seem more severe. In addition to plausibility, they investigated processing difficulties caused by subcategorization violations and selection violations. In short paragraphs, such as

1) The crowd was waiting eagerly. The young man \_\_\_\_\_ the guitar ... .

four types of verbs were inserted, so that the resulting sentence was a) normal ("grabbed"), b) pragmatically implausible, but grammatical ("buried"), c) violating selection restrictions ("drank"), or d) violating subcategorization restrictions ("slept"). Naming latencies to the target argument ("guitar" in the example) in a word monitoring task showed that comprehension processes reflected this order of violations. The reaction times were slowest when subcategorization information was violated, but there was also a significant cost of pragmatic implausibility. These results indicate once more that lexical information is available in the immediate vicinity of the verb. More importantly, they also suggest that pragmatic knowledge is used early in syntactic processing.

As reviewed in the previous chapter, Stowe (1989) showed that selection restrictions could eliminate the incorrect assignment of a clause boundary in sentences such as

1) animate, plausible:

When the police stopped the driver became very frightened.

2) inanimate, plausible:

When the truck stopped the driver became very frightened.

In a second experiment, she additionally varied the pragmatic plausibility of the noun phrase to render a direct-object interpretation infelicitous. For instance, the sentence "The police stopped the driver." is more plausible than the sentence "The police stopped the silence." Examples for the resulting experimental sentences are:

3) animate, implausible:

When the police stopped the silence became very frightening.

4) inanimate, implausible:

When the truck stopped the silence became very frightening.

In contrast to the distinction between animate and inanimate subjects, the distinction between plausible and implausible objects is not considered a lexical constraint. As expected, the reading times for the disambiguating second verb (again from a cumulative, word-by-word grammaticality judgment task) were longest in the animate, plausible sentences. The reading times for this word in the inanimate, implausible sentences were shorter, replicating the findings from Experiment 1. However, for both implausible conditions, reading times were similar and fell between these two

extremes. Thus, although the effects of selection restriction suggest that the inanimate sentences should be processed more easily, even in the case of an implausible noun, this was not the case. A strong interaction between plausibility and animacy showed that animacy only had an effect in the presence of a plausible noun.

Stowe (1989) takes these results as suggesting a non-syntactic reanalysis process. Even in the absence of a syntactic misanalysis, a thematic plausibility check is postulated which evaluates the relationships between neighboring pairs of content words. In fact, the noun "silence" used in both implausible sentences is not associatively related to any of the preceding words. In contrast, in the plausible sentences, the sentence topic of "driving" facilitates the thematic integration of the subsequent noun "driver." Thus, the prior sentence context might have influenced processing times.

In one of the first studies addressing such a more global influence of sentence context on on-line comprehension processes, Tyler and Marslen-Wilson (1977) used sentences containing ambiguities such as "flying airplanes" in which the syntactic class of the first word is initially not determined. "Flying" can be interpreted as either a deverbal adjective modifying the noun "airplanes," or it can be understood as the gerundive form of the verb "to fly." Tyler and Marslen-Wilson wrote sentences which biased towards one of these two interpretations. Consider the examples

- 1) adjective:  
If you walk too near the runway, landing planes are ... .
- 2) gerundive:  
If you have been trained as a pilot, landing planes is ... .

An appropriate continuation of sentence fragment 1) requires the plural form of a verb, whereas only a singular verb is appropriate in sentence fragment 2). A prior study confirmed that the sentence contexts biased sufficiently towards one or the other interpretation.

These sentences were used in a cross-modal priming paradigm. The sentences were presented orally, and naming probes (singular or plural auxiliary verbs) were presented immediately after the ambiguous words. The naming latencies supported the hypothesis that the context bias influenced lexical class assignment. The appropriate verb form (i.e., "are" in 1), "is" in 2)) was more easily accessible than the context inappropriate continuation. Tyler and Marslen-Wilson concluded that the syntactic processor which assigns lexical class to incoming words makes use of the sentence context quickly and efficiently.

Cowart and Cairns (1987) (see also Townsend & Bever, 1982) pointed out that these results could be accounted for with a purely syntactic explanation. Most of the sentences biasing towards the gerundive interpretation contained a plural pronoun (e.g., "Because they are so awful in the kitchen, cutting boards..."). In order to assign a referent to this pronoun, the ambiguous words had to be interpreted as a plural noun phrase. To confirm that this confound was indeed crucial, Cowart and Cairns (1987) conducted a study in which, instead of the sentence bias, the presence of a plural verb was manipulated. Thus, the processing of sentence frames like

- 1) noun phrase:  
While the boxes usually come with several internal partitions, packing cases... .
- 2) plural pronoun:  
While they usually come with several internal partitions, packing cases... .

were contrasted in the Tyler and Marslen-Wilson design. The results were as expected. While there were no measurable differences between the naming latencies for the two targets in the noun phrase sentences 1), it took subjects longer to name the inappropriate "is" after the pronoun context than to name "are." Unfortunately, Cowart and Cairns (1987) did not attempt to write materials in which this confound was not present, i.e., to embed the ambiguous words in sentences which biased towards the noun phrase interpretations without containing a plural pronoun. Thus, while their experiment provided evidence for the strong tendency of the human parsing processes to find

referents for unassigned pronouns, it did not rule out the effectiveness of a sentence context bias for disambiguation of lexical class information.

However, Tyler (1989) reports an experiment (Marslen-Wilson & Young, unpublished) in which the confound was avoided. An example for sentence fragments which biased towards one interpretation only through appropriate pragmatic knowledge, without containing plural pronouns, is the following:

- 1) gerundive bias:  
If you want a cheap holiday, visiting relatives ... .
- 2) adjective bias:  
If you have a spare bedroom, visiting relatives ... .

Using these materials yielded the same results as in Tyler and Marslen-Wilson; naming latencies to inappropriate probes were longer than naming latencies to appropriate probes. However, since this experiment is not published, it is not possible to inspect the full set of materials, and evaluate the generality or the magnitude of this effect.

In a different framework, Frazier and Rayner (1987) exploited the same type of ambiguity. They attempted to provide evidence for the so-called "delay strategy hypothesis" which states that the syntactic processor activates all possible lexical classes of a word in a parallel fashion and delays the decision to select one of them until disambiguating information is encountered. (Note that the authors claim that this strategy is only valid for the assignment of lexical class, whereas structural decisions are proposed to be made without delay; Frazier, 1989) Two ambiguous forms of sentences, in which the two target words were either resolved as a plural noun phrase or a gerundive nominal, were paired with two unambiguous forms. For example, the targets "ringing bells" were embedded in four sentences of the following form (Experiment 3):

- 1) ambiguous, adjective:  
Without a doubt, ringing bells are disturbing to everyone.
- 2) ambiguous, gerundive:  
Without a doubt, ringing bells is disturbing to everyone.
- 3) unambiguous, adjective:  
Without a doubt, loud ringing bells are disturbing to everyone.
- 4) unambiguous, gerundive:  
Without a doubt, ringing loud bells is disturbing to everyone.

Based on results of a prior sentence continuation study, half of the sentences were classified as adjective biasing, and half of the sentences as gerundive biasing. Eye fixations were measured, and fixation times per character were calculated for each of the two target words, as well as the remainder of the sentence. The only effects found were that the target words were read slower in the unambiguous sentence forms, and, as expected, that the disambiguating region was read slower in the ambiguous sentence forms. None of the other variables (sentence bias, structural form) had an impact on fixation times.

This experiment is not directly comparable to the Tyler and Marslen-Wilson and Cowart and Cairns efforts. Whereas Frazier and Rayner pay more attention to the reading times for the ambiguous materials, the previously described studies were mainly concerned with the question which of the alternative interpretations is more accessible immediately after encountering the ambiguous phrase; and even more importantly, if this accessibility can be influenced by providing biasing sentential contexts. Although Frazier and Rayner also report results for the disambiguating region, the size of the analyzed region is rather large (including about seven words). Instead of reporting fixation times for the one or two words following the ambiguous words (e.g., "is/are disturbing"), they calculate reading times per character for the entire end of the sentence. Although the data suggest that the disambiguating regions in sentences in which structure and preference corresponded were read more easily (about 37 msec per character) than in the inconsistent conditions



(about 41 msec), this trend was not significant. The only difference in these regions, however, is if the verb is singular or plural. Thus, it is not surprising that eventual small differences in the fixation times for the disambiguating words are hidden in the reading time analyses of the long regions.

Taraban and McClelland (1988) argue that the principle of minimal attachment, which was first demonstrated by Rayner, Carlson and Frazier (1983), can be overridden by content based expectations. Using a sentence completion task and expectancy ratings, Taraban and McClelland (1988) confirmed that the sentence contexts in the materials used by Rayner et al. pragmatically biased towards verb phrase attachment. Taraban and McClelland constructed a second set of materials in which the sentence contexts suggested a noun phrase attachment. For each sentence context, two continuations were used which disambiguated the prepositional phrase attachment. Examples of the four resulting sentence types are:

- 1) non-minimal attachment, minimal attachment bias:  
The spy saw the cop with a revolver.
- 2) non-minimal attachment, non-minimal attachment bias:  
The reporter exposed corruption in the government.
- 3) minimal attachment, minimal attachment bias:  
The spy saw the cop with binoculars.
- 4) minimal attachment, non-minimal attachment bias:  
The reporter exposed corruption in the article.

The analysis of reading times from a word-by-word, non-cumulative self-paced reading task replicated the processing difficulties for Rayner et al.'s non-minimal attachment sentences. Thus, the self-paced reading task was as sensitive to the processing differences as eye-movement monitoring. However, confirming the importance of the sentence context, the results were exactly opposite for the new sentences. Thus, for the noun in the prepositional phrase reading times were longer in the inconsistent conditions (1 and 4) than in the consistent conditions (2 and 3). Averaged across the two sets of materials, no effect of attachment was found.

Similar results were obtained in a second experiment in which the sentences were continued after the prepositional phrase. There was no effect of attachment. However, the location of the processing differences shifted. For the noun in the prepositional phrase, the reading times were indistinguishable in the four conditions. For the three words following the noun, on the other hand, the strong interaction between attachment and sentence context was replicated.

Taraban and McClelland discuss several explanations for their context effects. Instead of the violation of an expected prepositional phrase attachment, the processing difficulties could have been due to violations of the expectations for a thematic role, or, even more specifically, they could have been due to violations of the expectations for a particular noun filler. In sentence 1) and 3), for instance, an instrument might be the expected thematic role for the prepositional phrase, whereas in sentence 2) and 4) the specification of an instrument is less likely.

Since these factors were confounded in Experiment 1, Taraban and McClelland (1988) conducted a second experiment designed to test for these effects differentially. Sentence frames were written which could be continued with nouns in one of four categories. The noun was either fully consistent with the expectations, it was an unexpected filler for the expected role, it was a filler for an unexpected role, or it required the unexpected attachment. In the sentence fragment

- 5) The janitor cleaned the room with the \_\_\_\_ because of many complaints.

for instance, "broom" is fully consistent as a typical instrument for cleaning rooms, "solvent" is a less expected instrument, "manager" violates the role expectation but still requires minimal attachment, and finally, "odor" requires a non-minimal noun-phrase attachment. A variety of thematic roles was used, and an equal number of sentence frames biased towards verb phrase attachment and noun phrase attachment was written. In a cloze task, the sentences were pretested to

ensure that the constraint for attachment preferences and role expectations were met. In addition, expectation ratings and plausibility ratings were collected for all the experimental materials.

Reading times were analyzed for the noun filler and the three words immediately following it. Of particular interest, besides overall effects of filler type and word position, were pairwise comparisons to pinpoint the influence of violated expectations. The reading times for the noun filler did not vary across conditions. Processing of an unexpected filler for the expected role was somewhat harder for the word after the noun. Violations of the role expectation led to increased reading times for the two words after the noun. Finally, small additional costs of processing an unexpected attachment over and above the costs of processing an unexpected role appeared only in the reading times for the third word after the noun (and this effect was only significant by subjects but not by items).

To summarize, this experiment provided evidence for the importance of thematic expectations in sentence processing. It was shown that violation of filler or role expectations increased reading times in the vicinity of the noun. In contrast, the reading time patterns for the sentences in which syntactic attachment expectations were violated and the sentences in which thematic role expectations were violated were almost indistinguishable. Moreover, as in Experiment 1, there was no overall advantage of minimal attachment sentences over non-minimal attachment sentences.

While the sentence context did indeed account for processing difficulties in this study, no distinction between lexical information and sentence context was attempted. Since the expectations for a specific argument (e.g., instrument), or no argument, could be guided by the thematic frame of the verb, a more thorough analysis of the sentences would be desirable. Lexical information has been shown to influence syntactic processing (Chapter 4), so that a convincing argument for contextual penetrability can only be made if the sentence contexts are controlled for lexical factors. Thus, an evaluation of the thematic argument structures of the verbs used in Taraban and McClelland's materials is necessary. A further study in which lexical and contextual factors are independently manipulated could provide additional information on which of these factors affects syntactic processing.

## 2. Discourse context

The best way to test the effects of non-syntactic information is to assess the syntactic processes involved in comprehending a sentence which is embedded in a discourse context. Although studies on sentence context and pragmatic knowledge address this question as well, the sentences used in these studies differ across conditions. Thus, it is difficult to establish exactly which factor contributed to eventual processing differences. Moreover, as seen in the previous chapter, the distinction between lexical information and contextual or pragmatic information is not always clear. What could be interpreted as an effect of the reader's general world knowledge, might, given the appropriate representation within the lexicon, also be interpreted as a syntactic effect. In addition, processing sentences in isolation is not the usual comprehension task. In natural language understanding contextual information plays an important role.

Given these considerations, it is surprising that only few studies have been conducted which are concerned with the effects of discourse context on parsing. One reason for this lack of activity might be that there have been no specific proposals for criteria to establish the effectiveness of a given context to bias towards a particular syntactic reading. Although most theories of parsing include a mechanism which takes into account the plausibility of a syntactic interpretation, it is not clear which features of the context rule out a given syntactic analysis. Even if one interpretation is more plausible in a particular context, an alternative interpretation might still be acceptable.

An example which illustrates this point is a study by Holmes (1984). This study was one of the first to explicitly test the validity of the Minimal Attachment Principle within biasing contexts.

Holmes (1984) used a verification task to investigate which of the two readings of an ambiguous sentence was more prevalent. One of the structural ambiguities used in the study was the attachment of a prepositional phrase. According to Frazier's principle of minimal attachment, the prepositional phrase in sentences such as

- 1) The parents discussed the problem with the mathematician.

tends to be attached to the verb phrase (i.e., the parents talked with the mathematician about the problem), instead of to the noun phrase (i.e., the problem with the mathematician was the topic of the parents' discussion). Holmes asked subjects to indicate if target sentences of type 1) were consistent with a context paragraph presented before. The last sentence of these paragraphs biased the reader towards one of the two possible interpretations of the target sentence. The contexts used for sentence 1) were:

- 1) minimal attachment supporting (i.e., VP attachment):  
The parents discussed with the math teacher what could be done about the difficult exam, but no solution was reached.
- 2) non-minimal attachment supporting (i.e., NP attachment):  
The parents discussed among themselves what could be done about the math teacher, but no general agreement was reached.

If the context biased towards the preferred reading, as in 1), subjects found the target sentence consistent with the preceding paragraph in 79% of the cases, compared with only 44% for the contexts biasing towards the non-preferred reading, as in 2) (i.e., the non-minimal attachment). Thus, many subjects failed to consider the non-preferred reading even when it seemed to correspond closely to the previously established situation model. Holmes concluded that this difference provided evidence for the psychological reality of the principle of minimal attachment and against contextual effects.

Even using a global measure, (the verification of the consistency of a sentence with the preceding paragraph), Holmes could not find a very strong context effect. Although the contexts seemed to bias towards one of the two attachments, they did not produce a reversal of the attachment preferences. Thus, it is necessary to establish the effectiveness of the intended bias before studying the on-line effects of discourse context on parsing. In addition, in order to successfully construct biasing contexts it is necessary to formulate criteria specifying exactly which features of a discourse influence the parsing process.

Crain and Steedman (1985) were the first to attempt to formulate criteria for the effectiveness of context bias. Their principles of parsimony and referential support (see Chapter 2) provide guidelines for how to construct discourse paragraphs which facilitate the interpretation of a syntactic ambiguity. Crain (1980; reported in Crain & Steedman, 1985) used a word-by-word grammaticality judgment task to show that these principles could indeed induce syntactic processing differences. Sentences such as

- 1) complement:  
The psychologist told the woman that he was having trouble with her husband.
- 2) relative clause:  
The psychologist told the woman that he was having trouble with to visit him again.

were preceded by one of the following context paragraphs:

- 1) complement-supporting:  
A psychologist was counselling a man and a woman. He was worried about one of them but not about the other.

2) relative-supporting:

A psychologist was counselling two women. He was worried about one of them but not about the other.

The only difference between these paragraphs is that in 1) only one woman is introduced into the discourse model, whereas in 2) two women are referred to. Consequently, the reference to "the woman" in the target sentences is unambiguous in context 1), but an additional specification is needed in context 2). Using a word-by-word grammaticality judgment task, Crain (1980; cited in Crain & Steedman, 1985) showed that in the two conditions in which contexts and targets were consistent, the target sentences were judged grammatical significantly more often than in the inconsistent conditions. Since this study used a constant presentation rate for each word, and since only the judgments, but not reaction times were measured, this experiment does not shed light on the on-line processes which gave rise to the context effect.

Altmann & Steedman (1988) conducted two follow-up experiments. For these experiments, they used similar materials as Crain (1980). The context paragraphs were minimally different, i.e., they only differed with respect to the number of possible referents to the crucial noun phrase in the target sentences. An example of the two context types is the following (*italics added*):

1) minimal attachment supporting (i.e., VP-supporting):

A burglar broke into a bank carrying some *dynamite*. He planned to blow open a safe. Once inside he saw that there was a safe with a *new lock* and a *strongbox* with an old lock.

2) non-minimal attachment supporting (i.e., NP-supporting):

A burglar broke into a bank carrying some *dynamite*. He planned to blow open a safe. Once inside he saw that there was a safe with a *new lock* and a *safe* with an old lock.

Target sentences using either minimal attachment or non-minimal attachment were written in which pragmatic information allowed to disambiguate the syntactic structure. Examples are:

1) Minimal attachment (VP attachment, instrument):

The burglar blew open the safe with the *dynamite* and made off with the loot.

2) Non-minimal attachment (NP attachment, object-modifier):

The burglar blew open the safe with the *new lock* and made off with the loot.

Consequently, four experimental conditions were tested, crossing the two context types with the two target sentence types (there was an additional factor, number of instruments, which did not have an impact on the results, and which is therefore omitted in this review). In the first experiment, reading times for the entire target sentences were measured in a self-paced reading task, in which the sentences were presented one at a time. The results confirmed the effectiveness of the context manipulation. In particular, there was a strong interaction between the two factors context type and target type, indicating that for both target types the reading times in the supporting context were shorter than in the infelicitous context. Surprisingly, the reading times were shorter overall for the non-minimally attached targets than for the minimally attached targets. This experiment confirmed the effectiveness of the context bias, but since a global measure was employed, no information was obtained about the location at which the context exerted its influence.

To be able to point to the location of the processing differences, Altmann and Steedman (1988) conducted a second experiment. The same materials were used in a cumulative self-paced phrasal reading task. An example of the segmentation of the sentences into short phrases is the following:

## 1) The burglar / blew open / the safe / with the dynamite / and made off / with the loot.

The results of this second experiment confirmed that the reading time patterns for the prepositional phrase ("with the dynamite/new lock") mirrored those for the entire sentences.

Therefore, the processing differences occur early in the sentence, at the disambiguating phrase, and are not due to reanalysis at the sentence boundary.

It is important to note again that there was an unexpected main effect of target type, caused by *longer* reading times for the syntactically preferred minimal attachment sentences. There is a confound in the materials, however, which might explain this counterintuitive result. The instrument of the target sentence verb is always mentioned in the first context sentence, whereas the object(s) is (are) mentioned in the last context sentence (*italicized* in the example). The relative ease of establishing coherence in the latter case, compared to the first, might have lead to faster reading times. This confound in Altmann and Steedman's materials was first pointed out by Clifton and Ferreira (1989). In an attempt to control for this recency effect, Clifton and Ferreira (1989) conducted a similar experiment in which the context paragraphs were slightly altered. In their materials, recency of the critical antecedent favored the minimal attachment sentences. The results were uninterpretable, however. The condition which is predicted to be most difficult (non-minimal attachment in a context which biases towards minimal attachment) yielded the shortest reading times. Moreover, the comprehension questions only tapped information contained in the target sentences, so that subjects seemed to have adopted the strategy of ignoring the context altogether (see Steedman & Altmann, 1989, for a more extensive discussion of this point). Thus, because of methodological problems, this experiment does not allow to settle the issue of the cause of the context effects documented by Altmann and Steedman.

It is important to point out, however, that even if recency of the antecedent was a crucial factor for processing of the target sentences in Altmann and Steedman (1988), this effect is a discourse effect as well. Although either the principle of referential support, or the recency of mention might have been responsible for the context effects, both variables are inherent to the discourse paragraphs. Syntactic principles alone cannot account for the context effects found by Altmann and Steedman (1988). However, since these effects were found in a cumulative reading task, they need to be replicated with another on-line paradigm.

Ferreira and Clifton (1986) used an eye-movement monitoring paradigm to provide evidence against context effects in parsing. In their second experiment, they used globally ambiguous sentences with respect to the attachment of a locative prepositional phrase, for example:

- 1) Sam loaded the boxes on the cart.

The interpretation of this sentence using a minimally attached reading of the prepositional phrase leads to a situation model in which Sam moves boxes from their current location onto the cart, whereas according to the non-minimally attached reading Sam moves the boxes which are on the cart to an unspecified location. In order to force subjects to decide on one reading, two sentence continuations were written which disambiguated the prepositional attachment:

- 1) minimal attachment:  
Sam loaded the boxes on the cart before the coffee break.
- 2) non-minimal attachment:  
Sam loaded the boxes on the cart onto the van.

In Ferreira and Clifton's opinion, these two continuations disambiguated the sentence sufficiently to guide readers to one of the aforementioned interpretations. Since the verb "load" subcategorizes for a direct object and a locative prepositional phrase, the temporal continuation in 1) forces the interpretation of "on the cart" as the locative. To confirm that the disambiguations were sufficient, questions were constructed to verify that subjects indeed interpreted the target sentences in the intended way.

In addition, Ferreira and Clifton (1986) wrote context paragraphs which were either neutral with respect to the attachment choice, or supportive of the appropriate reading. Examples for the two types of contexts for the previously cited sentences are:

- 1) Minimal-attachment supporting:  
Sam wanted to go for his coffee break, but his boss said Sam had to fill up one more cart before he could go. The boss knew some guys from another department needed the cart.
- 2) Neutral:  
Sam wanted to go for his coffee break, but his boss said he had a little more work to do. He wanted Sam to free up a cart for some guys in another department.

Whereas the felicitous paragraphs contained explicit reference to the objects as needed in the suggested attachment, the neutral context of the provided example does not seem to equally support both attachments. Before the target sentence, the phrase "free up a cart" is used, which seems to imply that something is to be removed from a cart. If this inference is made, however, it biases towards the non-minimal reading.

The contexts were not crossed with targets. Instead, the target sentences were either preceded by the neutral context, or by the felicitous context. Analyses of first pass fixation times provided no evidence for a context effect, while the non-minimal targets were more difficult to process in the last, disambiguating region (e.g., "before the coffee break" vs. "onto the van"). For the second pass fixation times, the triple interaction between region, attachment, and context was marginally significant by subjects, indicating that it was somewhat easier to recover from the garden path in the felicitous context than in the neutral context.

Similar results were obtained using reduced relative clauses in both an eye-movement paradigm and in a phrasal self-paced reading paradigm. Context did not facilitate the processing of the non-minimal attachment interpretation.

Two explanations have been brought forward to explain this study's failure to obtain a context effect. The first explanation rests on an inspection of the argument structures of the verbs used in the study (Britt et al., 1992). The verbs subcategorized two obligatory arguments (e.g., "load" takes a direct object and a locative). Context influences might not be obtained as easily as long as the syntactic processor has not filled the obligatory argument slots. Evidence for this interpretation is also found in a study by Rayner, Garrod and Perfetti (1992). However, this explanation requires the assumption that lexical information takes precedence over contextual or pragmatic information. The validity of this claim has to be tested more thoroughly.

The second, and more important, objection concerns the effectiveness of the context bias. The paragraphs were not pretested to confirm that their selection served the intended purpose. Moreover, the neutral contexts did not render both attachment alternatives equally plausible. The contexts did not only differ with respect to the number of antecedent referents for the crucial noun phrase, but also in other ways which might have biased towards one reading. In fact, as Altmann and Steedman (1988) point out, the neutral paragraph preceding the sentence "The editor played the tape agreed the story was a big one" did indeed bias towards minimal attachment. The last sentences of these three context paragraphs were:

- 1) non-minimal attachment bias:  
He ran a tape for one of his editors, and he showed some photos to the other.
- 2) minimal attachment bias:  
He gave a tape to his editor and told him to listen to it.
- 3) neutral:  
He gave a tape to one of his editors and told him to listen to it.

Thus, although the preceding context contained several editors, in 2) and in 3) the last sentence before the target sentence explicitly foregrounded one of them. For the minimally attached sentences, therefore, the contexts cannot induce processing differences. For the non-minimally attached sentences, on the other hand, the longer processing times in the felicitous context could have been due to a variety of factors besides syntactic attachment. Since two editors were in the focus of discourse, the processing increase could have reflected the difficulty to establish this more

complex discourse model. In addition, sentences containing reduced relative clauses are more difficult to process than simple active sentences. These factors can only be controlled for if the biasing contexts are crossed with target types.

Following these arguments, Mitchell, Corley & Garnham (1992) used a design in which three target sentence types were fully crossed with two context types, but they also failed to obtain context effects. They used local ambiguities of the form

- 1) The headmaster told the boy that ... .

which can be resolved as either a sentential complement or a relative clause. In addition, the relative clause continuation can modify the subject or the object. Examples for these types of resolutions are:

- 1) complement:  
The headmaster told the boy that he had been watching him.
- 2) object-relative:  
The headmaster told the boy that he had been watching to go and wait outside.
- 3) subject-relative:  
The headmaster told the boy that had been standing there to go and wait outside.

Sentences of these types, as well as unambiguous control sentences (in which the verb did not allow a sentential complement, such as "call," or in which the verb was replaced by "and") were presented to subjects in a phrasal self-paced reading task. The crucial display in all sentences was "had been." These two words disambiguate the sentences 3) as subject-relative clauses. In the absence of an ambiguity it took slightly longer to read this display when it was part of an object-relative clause or a complement, than when it was part of a subject-relative clause. The opposite pattern was observed in the ambiguous cases. Reading times for this target display were almost 200 msec longer if it ruled out an expected sentential complement interpretation. In the unambiguous cases, there was no such effect of sentence type (note that this result provides additional evidence for the use of subcategorization information at an early point in parsing, since one of the control conditions contained verbs which do not take a complement).

Since several parsing models can account for these data, a second experiment was conducted in which short context paragraphs were added. The relative-supporting contexts introduced two referents for the object in the target sentence, so that a relative clause reading was more likely than a complement reading. In the complement-supporting contexts, on the other hand, only one referent was introduced, so that a relative clause interpretation was functionally less plausible. These context paragraphs, therefore, were written according to the principle of referential support.

The three target sentence types used in Experiment 1 were fully crossed with the two context types. Four displays were used to present the target sentences in a phrasal reading task. An example of the segmentation, indicated by /, is:

- 1) The politician told the woman that he / had been / meeting / that he was going to see the minister.

All parsing theories predict that the context bias facilitates the felicitous target sentence at sentence wrap-up; however, the interaction between context type and target type was reliable neither for total reading times nor for the last display. An analysis restricted to the two sentence types in which the last display contained disambiguating information (complement and object-relative), in contrast, yielded a marginally significant interaction between context type and sentence type.

The theoretically more interesting analyses of the reading times for the second display replicated the results of the first experiment. The inclusion of the pronoun in the first display (in complement and object-relative sentences) made it easier to process the "had been" display. However, this effect did not vary with context type.

Unfortunately, no analysis of the reading time patterns for the third display, containing either only the word "meeting" or the words "meeting him" (for the subject-relative sentences), was reported. In the relative-supporting contexts, reading times for this third display in the subject-relative sentences were 140 msec longer than in the complement-supporting contexts. Although it is not clear if this difference was significant, it could be explained by the functional redundancy of the modifier in the complement-supporting context. In the relative-supporting context the information "meeting him" is necessary to assign the correct referent to the noun phrase (i.e., one of the two women introduced in the paragraph).

Although the context paragraphs in this study were constructed following the principle of referential support, and the context types were crossed with target types, several other methodological problems can be identified. First, the materials were not pretested using a global measure. The context manipulation is successful only if the likelihood of a relative clause continuation increases in the relative-supporting paragraphs. Although all parsing models predict the eventual use of discourse information, in Mitchell et al.'s study neither the total reading times, nor the reading times for the last display were affected by context. This finding suggests that the context bias was not sufficient.

Second, inspection of the example materials given in Mitchell et al. (1992) shows that the predicate in all target sentences was the verb "tell." Since only 26 filler paragraphs were used with 24 experimental paragraphs, this repetition could have led the subjects to adopt a strategy specific to the experiment. Moreover, according to the norms of Connine et al. (1984), "tell" has a subcategorization preference for a sentential complement, which is more pronounced in conjunction with the function word "that." This extremely strong bias towards a complement reading was indeed confirmed in the first experiment. Therefore, the subcategorization information could have had more influence on the processing times than the contextual information. Since a similar explanation for the failure to find context effects has been suggested in the interpretation of the study by Ferreira and Clifton (1986), this argument has to be taken seriously. Further research is necessary to evaluate whether lexical information has precedence over discourse information.

Third, in the phrasal reading task, the segmentation of the materials has a large impact on the results (see Chapter 3). Mitchell et al. divided the sentences in a way that the displays which are crucial for assessing early context effects were much shorter (one or two words) than the first and last display (about 8 words). Again, this could have prompted subjects to adopt a particular reading strategy in reaction to the experimental procedure. It seems that the length of the display could have been used as a cue for the importance of the information.

Fourth, the use of three sentence types makes the results more difficult to interpret, and it seems that the authors focussed on one effect while neglecting others. Mitchell et al. concentrated on the reading times of the words "had been." These words are the disambiguating material only in the subject-relative sentences. Processing difficulties observed in this region can therefore not be attributed to the erroneous postulation of a sentential complement. The processing difficulties could also be due to the need to give up an object-relative interpretation. Since the context paragraph bias was only concerned with the distinction relative clause versus complement, the disambiguation between the two *types* of relative clause could not be affected by the context.

The disambiguating information needed to distinguish between a complement sentence and an object-relative sentence, on the other hand, was not presented before the last display. Recall that there was indeed a marginally significant interaction between context and sentence type for this last display. Mitchell et al. interpret this finding as evidence for late effects of context. That is, they conclude that contextual information is not immediately taken into account. However, the segmentation of the materials did not allow to distinguish early and late effects in this study. As mentioned earlier, the last display contained about eight words. These eight words consisted of a full sentential complement in the object-relative sentences, and of a noun phrase as well as a prepositional phrase in the complement sentences. Thus, the reading times for the last display are affected by sentence wrap-up effects, the length of the display, and the complexity of the information



contained in it. The disambiguating information, on the other hand, consisted only of the first one or two words of the display. Consequently, to be able to distinguish between early and late context effects in this case, it would have been necessary to measure processing times for the disambiguating information only. Since the context manipulation specifically addressed the distinction between complement sentences and relative sentences, an early context effect could only have been observed at this location. Because processing times for this location were not measured by Mitchell et al., the claim that their study provides evidence for late context effects only cannot be evaluated.

Britt, Perfetti, Garrod, & Rayner (1992) find immediate discourse effects on the processing of prepositional phrases using a more fine-grained level of analysis. The targets were preceded by either a neutral context paragraph, or by a context which supported the appropriate attachment. Following the principle of parsimony, the felicitous contexts explicitly introduced two instances of the crucial concept. For instance, for the target sentences

- 1) Minimal attachment (instrument, VP):  
The doctor examined the child with the needle instead of another technique.
- 2) Non-minimal attachment (object, NP):  
The doctor examined the child with the doll instead of another child.

the supporting context for 1) contained the doctor's attempt to select an appropriate technique for the examination, whereas the supporting context for 2) contained reference to two children, one of whom held a doll. Furthermore, in both contexts, the object ("needle" or "doll") belonged unambiguously to either the child or the doctor. The neutral contexts, on the other hand, did not bias towards either attachment. In order to confirm that the context paragraphs written did indeed follow these criteria, a group of independent subjects indicated their confidence to questions such as "Who has the needle?".

Self-paced reading tasks with word-by-word as well as phrasal presentations were used, and the results for the two presentation types were similar. Analyses of reading times per character indicated that the context bias was effective. Whereas in neutral contexts subjects took longer to read the final disambiguating region (e.g., "instead of the other technique/child") for non-minimally attached targets than for the preferred minimally attached targets, there was no such difference in the supportive contexts. However, no comparable results were reported for the region which contained the word crucial for the correct attachment of the prepositional phrase ("needle" vs. "doll"). Moreover, the analyses conducted consisted only of t-tests. No interactions between the factors were reported.

Two interpretations for this pattern of results are possible. The first is that the immediate processing of the disambiguating region is not influenced by context, but that the context only exerts its effect at the end of the sentence (this interpretation would be in line with garden path models). The second explanation is that the longer reading times in the last region are caused by a spillover effect from the previous region. The results are consistent with Britt's (1988), but not with Ferreira & Clifton's (1986).

One possible reason is that subject paced reading tasks might not be sensitive enough to capture subtle differences. Britt et al. (1992), thus, conducted a similar experiment employing an eye monitoring technique. For each of 12 sentences, three versions were written. The first version contained a minimally attached prepositional phrase. Correct non-minimal attachment in the second sentence required the use of semantic information, whereas in the third sentence pragmatic information was needed. Examples of these three sentence types are:

- 1) minimal attachment:  
Peter read the books on the chair instead of lying in bed.
- 2) non-minimal attachment (semantic):  
Peter read the books on the chair instead of the other books.

3) non-minimal attachment (pragmatic):

Peter read the books on the war instead of the other books.

Context paragraphs were used which biased towards the correct interpretation as in Experiment 1. Examples for a context biasing towards sentence 2), for instance, included reference to two piles of books, one of which was on the chair.

The sentences were presented in isolation or in these supportive contexts (no neutral contexts were used). In isolation, processing difficulties (as reflected by first-pass fixation times per character) were found for the disambiguating regions in the non-minimal targets. In particular, compared to the preferred minimal attachment sentence 1), the reading times were longer for the phrase "on the war" for sentence 3), and for the phrase "instead of the other books" for sentence 2). These differences were eliminated in context. Analyses of the total reading times yielded analogous context effects.

The results of this second experiment show that syntactic processing is immediately influenced by discourse information. Fixation times for the disambiguating regions confirmed that felicitous contexts can alleviate processing difficulties observed when the same sentences are presented in isolation.

The explanation Britt et al. (1992) provide for the inconsistency of their results with those of Ferreira and Clifton (1986) rests on the notion of verb argument structure. In Ferreira and Clifton's study, the second argument of the verbs is often obligatory, whereas the verbs Britt et al. used have an optional second argument. The need to fill an obligatory argument slot might override other syntactic processes (cf. Chapter 4).

In their third experiment, Britt et al. (1992) test the hypothesis that contextual override of syntactic preferences is only possible for some syntactic phenomena, but not for others. While they demonstrated that context effects could be obtained for prepositional phrase attachment, they hypothesized that for reduced relative clauses no contextual override would be observed. This prediction is based on the Cooperative Language Processor, a model proposed by Perfetti (1990) (for a discussion, see Britt et al., 1992). Sentences with local ambiguities such as

1) main clause:

The woman rushed to the hospital without taking her laundry.

2) reduced relative clause:

The woman rushed to the hospital had given birth today.

were considered. In 1) the phrase "rushed to the hospital" is the active verb phrase of the sentence, whereas in 2) this phrase is a reduced relative clause modifying the sentence subject. A relative clause modification of a noun is needed in discourse if there are more than one possible antecedents (see the principle of referential support). It is also useful if the antecedent is not any more in the focus of the discourse. Following the latter explanation, Britt et al. wrote context paragraphs which manipulated if the antecedent for the target sentence subject was foregrounded in the discourse or not. Since these context paragraphs are crucial for evaluation of the results, two examples of target sentence within their appropriate contexts are given in full (Britt et al., 1992, p. 307):

1) active bias:

Suzanne had spent a fortune decorating the living room and fitting an expensive new rug. So the night before her friends were invited over, she had a dream that someone spilled coffee all over her new rug. The next day, her friends came over and she really enjoyed showing off her new living room to everyone. Imagine her embarrassment when she tripped over the table with a tray in her hand.

*The coffee spilled on the rug and even marked the new wallpaper.*

2) relative clause bias:

Suzanne didn't know what to say to her parents when they came home. She had thrown

a party and her friends had spilled coffee on the rug and scratched her mother's new table. She tried to do her best to clean the rug and table a bit. But she knew it was too late when she heard her parents at the front door. As they entered she tried to distract their attention.

*The coffee spilled on the rug was difficult for her to conceal.*

Although intuitively it is not clear that the paragraphs indeed biased towards the intended interpretation, a pretest showed that their selection facilitated the correct interpretation. A group of subjects rated answers to questions after reading the target sentences either in isolation or embedded in the felicitous paragraphs. For the example sentences above, for instance, the question was "How did the woman get to the hospital?", and the answers to be rated were "She rushed herself to the hospital," or "Someone else rushed her to the hospital." Whereas the ratings for the sentences presented in isolation indicated almost random choices between the two interpretations, the ratings in context corresponded strongly to the intended interpretation.

Using a word-by-word self-paced reading task, the results did not provide evidence for context effects. As expected, for isolated sentences, reading times for the disambiguating region of the reduced relative sentences were longer than for the corresponding region of the active sentences. However, this was also the case when the target sentences were presented within their appropriate contexts. Britt et al. take these results as support for their hypothesis that only the interpretations of particular grammatical structures can be directly influenced by discourse information.

This conclusion can be criticized on several grounds. First, the statistical analyses were not conducted as carefully as desired. As pointed out frequently, reading times collected in button-press tasks should not be normalized for length of the region simply by dividing by the number of characters. This procedure rests on the erroneous assumption that the reaction time for a display of 0 characters is 0. More sensitive analyses can be carried out if, for each subject, the residuals from regressions of the reaction times on the length of the region are used as the dependent variable. Moreover, even under the assumption that the pattern of results is comparable using the uncorrected measure reading time per character, more thorough analyses would be helpful. The argument that context effects were not obtained rests on the result that the disambiguating information was significantly more difficult to process in the reduced relative sentences as compared to the active sentences. No comparisons between the sentence types for the other two regions, or comparisons of the last region with the previous two are reported. In order to evaluate the conclusion, it is crucial, however, to know if there were differences between the locally ambiguous part of the sentence and the disambiguating information. In addition, processing of sentences with an embedded relative clause, even if unreduced, is more complex than processing of sentences with a simple active structure (Fodor, Bever, & Garrett, 1974). Since the relative clause structure is not unambiguously revealed before the last display, the increases in reading times in this region need not be due to garden-path effects but could also be due to the processing complexity. To answer this question it would have been necessary to include unreduced relative control sentences (e.g., "The woman who had been rushed to the hospital had given birth safely.").

Finally, and most importantly, the effectiveness of the discourse bias was not established convincingly. In the pretest, the high ratings for the appropriate interpretation could have been due to facilitated comprehension, independent of expectations for a particular syntactic structure. It has been shown repeatedly (and is replicated in this experiment) that reading times are shorter when sentences are presented in a context which allows to integrate the information into an existing situation model. Bias towards the correct syntactic structure could only be proven if crossing context paragraphs with target sentence types gave rise to reversed rating patterns.

A second way to test the context bias rests on the assumption that discourse foregrounding influences the expectations for a noun phrase modified by a relative clause. A pretest could have been conducted to establish if the antecedent of the target noun phrase was indeed in the focus of one context paragraph but not in the focus of the other. Intuitive inspection of the context paragraphs presented by Britt et al. (1992) suggests that the context bias was not sufficiently strong. While the

contexts for the first two studies made use of the principle of referential support, the context paragraphs in the third experiment did not. The alternative hypothesis that discourse foregrounding of an antecedent plays a role in syntactic processing was not convincingly confirmed. In order to establish that the processing of reduced relative clauses cannot be facilitated by embedding them in an appropriate context, it would have been more promising to rely on the principle of referential support once more.

A further argument against Britt's et al. conclusion that reduced relative structures are not susceptible to context effects comes from a study by Trueswell and Tanenhaus (1991). Instead of basing the discourse bias on the principle of referential support or discourse foregrounding, they chose to focus on how temporal information influences the situation model. The tense of verbs in a discourse helps to establish the temporal order of the described events relative to the time of the utterance. For instance, in the short paragraphs

- 1) Past tense:  
John went to the hospital. He twisted his ankle on a patch of ice.
- 2) Future tense:  
John will go to the hospital. He twisted his ankle on a patch of ice.

the tense of the verb in the first sentence determines how the event sequence is constructed. In the future tense example 2), John twisted his ankle prior to the time of speech, and he will go to the hospital at some time after the time of speech. In the past tense example 1), on the other hand, John twisted his ankle while going to the hospital, and both events precede the time of speech. If the causally implied order of twisting the ankle before going to the hospital was to be preserved in the past tense context, the past perfect "had twisted" would have to be used.

The dependency of the correct temporal interpretation of a sentence on the preceding discourse can be used to bias towards one interpretation of a locally ambiguous syntactic structure, as the following example illustrates:

- 1) Future tense context:  
Several students will be sitting together taking an exam tomorrow. A proctor will come up and notice one of the students' cheating.
- 2) Past tense context:  
Several students were sitting together taking an exam yesterday. A proctor came up and noticed one of the students' cheating.

If the sentence fragment "The student spotted..." is preceded by 1), the interpretation of "spotted" as a past tense verb leads to inconsistency in the discourse model, whereas its interpretation as the past participle in a reduced relative clause leaves the temporal relationships between the events intact. On the other hand, if the discourse is written in past tense as well, as in 2), both interpretations are felicitous.

To confirm that the tense of the preceding discourse indeed biases towards one of these interpretations, Trueswell and Tanenhaus conducted a sentence completion study. Twenty sets of materials similar to the one shown above were prepared, and subjects were asked to complete the ambiguous sentence fragments. In half of the cases the sentence fragment was preceded by the word "only." This condition was intended to induce a relative clause interpretation, because the phrase "only the student" implies reference to one particular student among others. In the past tense contexts, only 21% of the completions used a relative clause, compared to 69% in the future contexts. For the target fragments "Only the student spotted..." the same preference pattern was found, but it was not as pronounced (58% vs. 77%).

In two further experiments the on-line effects of the tense manipulation were studied. The sentence fragments were completed to full sentences, such as

- 1) past tense contexts:  
The student spotted by the proctor received a warning.
- 2) future tense contexts:  
The student spotted by the proctor will receive a warning.

and similar sentences containing disambiguated unreduced relative clauses were added as controls. Reading times were measured in a moving window paradigm in which two words were presented at a time. The analyses showed that there was an effect of context. In the future contexts, which rendered an active main clause reading of the verb "spotted" infelicitous, reading times were shorter for the reduced relative clause than in the past tense contexts. Moreover, there was no difference between reduced and unreduced relative sentences in the future contexts. Thus the temporal information about the discourse events was taken into account early and facilitated the processing of the locally ambiguous phrase.

Trueswell and Tanenhaus used the same procedure and similar materials in a third experiment. In this experiment, the verb used in the target sentence's relative clause was identical to the verb used in the preceding context paragraph. In the example provided above, for instance, "The proctor noticed one of the students..." was replaced with "The proctor spotted one of the students....." This minor change eliminates the need to make the inference that both sentences refer to the same event. If one of the reasons for increased processing difficulties on the phrase "spotted by" is the processing cost associated with making this inference, the differences should now disappear. Contrary to the hypothesis, however, the opposite effect was observed. In fact, the target segment "spotted by" caused more difficulties in the reduced relative clauses as compared to the unreduced relative clauses. Trueswell and Tanenhaus provide a plausible post-hoc explanation for this unexpected result. Similar to effects observed in the lexical access of ambiguous words, it could be the case that all interpretations of the verb are activated simultaneously, and that the most appropriate one is selected later on. Since the correct interpretation in the context sentence is the active past tense, the activation of the past participle might subsequently be inhibited. More important than this unexpected result, however, is the finding that the context effects were replicated once more.

This study demonstrates nicely that the situation model can indeed influence syntactic decisions early in the parsing process. Using context paragraphs which only differed in tense, but were identical with respect to content, reading times for the locally ambiguous segment could be effectively manipulated. This work illustrates a promising approach to study discourse context effects which do not depend on the principle of referential support. Further research is needed to determine which other types of situational information, besides tense, are accessed immediately during parsing.

### 3. Summary

The studies examining pragmatic and contextual effects on parsing do not yet provide a clear picture of the language processing mechanism. While there seems to be evidence for the influence of these factors (for plausibility and sentence context, Stowe, 1989; Tyler & Marslen-Wilson, 1977; Taraban & McClelland, 1988; for discourse context Crain & Steedman, 1985; Altmann & Steedman, 1988; Britt et al., 1992; Trueswell & Tanenhaus, 1991), in other studies pragmatic and contextual information did not affect syntactic processing (Rayner, Carlson, & Frazier, 1983; Ferreira & Clifton, 1986; Mitchell, Corley, & Garnham, 1992). The most obvious difference between these studies is the experimental paradigms employed. Studies using self-paced reading tasks tend to obtain context effects, whereas the analysis of eye movement data suggests contextual impenetrability of the syntactic processor (Rayner, Flores d'Arcais, & Balota, 1990).

There are a few exceptions to this apparent task dependency of the results. Mitchell, Corley, & Garnham (1992) used a phrasal self-paced reading task and failed to obtain context effects. The methodological problems of this study have been extensively discussed. Britt, et al. (1992), in

contrast, did provide evidence for context effects using both reading time data and eye fixation data. Moreover, Taraban and McClelland (1988) replicated results from an earlier eye movement study using a self-paced reading paradigm.

The conclusion that the task demands are indeed responsible for the differences in findings is not warranted, though. Since only a few studies used identical materials to compare various on-line paradigms directly, the different findings could also have been caused by the syntactic phenomenon studied, or by the selection of particular materials. Thus, the question if and under what circumstances syntactic processing is facilitated by non-syntactic information must be studied more extensively. Further research is necessary which more accurately separates task specific effects from effects specific to the materials.

## CHAPTER 6 CONCLUSIONS

The present review of recent psycholinguistic studies on lexical effects and context effects in syntactic processing leaves many questions unanswered. As pointed out in the previous chapters, the results are often contradictory and cannot easily be reconciled. In this last chapter I attempt to evaluate the results with respect to the issues discussed in the beginning of the paper. First, conclusions regarding the experimental methodology are drawn, and suggestions for further research are made. Second, the results of the studies are used to evaluate the psychological validity of models of the human parsing mechanism. Finally, I summarize two general models of text comprehension which contain a parsing component, but whose scope goes beyond syntactic processing. An evaluation of these models in light of the present experimental data is sketched.

### 1. Methodological implications

Effects of syntactic processing are not easy to document. As described in Chapter 3, only a few on-line paradigms are appropriate to study the time course of comprehension. Besides syntactic factors, word frequency, semantic associations, and pragmatic considerations affect the ease of processing an utterance. Thus, task specific demands have to be carefully analyzed in order to be able to interpret experimental results.

#### 1. Comparison of experimental paradigms:

In both the literature on lexical effects, and the literature on context effects in parsing, it is apparent that the type of results depends on the paradigms employed. Using an eye movement paradigm usually provides little evidence for immediate lexical or non-syntactic effects, while results from self-paced reading tasks tend to confirm non-syntactic effects. Exceptions are the studies by Just and Carpenter (1992), Britt, Perfetti, Garrod and Rayner (1992), and Clifton, Speer, and Abney (1991). The verb effects found in this latter study, however, did not manifest themselves immediately.

Although the apparent task dependency of the results on parsing has been widely acknowledged (Altmann & Steedman, 1988; Rayner, Flores d'Arcais, & Balota, 1990; Rayner, Garrod & Perfetti, 1992), a direct comparison of the tasks was only attempted in Ferreira and Henderson (1990), Clifton, Speer and Abney (1991), and Taraban and McClelland (1988). Only in these studies identical materials were used in different paradigms. In the latter two studies, the results converged. Reading times in a non-cumulative self-paced reading task, and eye movement data yielded comparable descriptions of the comprehension process. In the former study, the cumulative self-paced reading task and eye movement monitoring converged, while the non-cumulative reading task yielded a verb effect not obtained using the other two paradigms.

The reasons for this discrepancy could lie in the task specific strategies subjects develop. In a cumulative reading task, the reading times for a given word or phrase do not necessarily reflect processing of this unit. Evidence for this interpretation comes from two studies concerned with the effects of the absence of the complementizer "that." Holmes, Kennedy, and Murray (1987) concluded that both complement sentences with and without a complementizer were equally difficult to process. Rayner and Frazier (1987), in contrast, showed that eye movement data did reflect the increased difficulty of the complement sentences when the complementizer was absent. In a non-cumulative self-paced reading task, on the other hand, the fact that it is not possible to regress to previously read materials is likely to induce more thorough processing of the current word or phrase. Evidence for this interpretation comes from the fact that reading times for similar materials are

usually longer in button-press tasks than the corresponding eye fixations (Rayner & Frazier, 1987). This hypothesis could be tested by recording subjects' eye movements during a memorization task, instead of a comprehension task. This variation of the task instructions might lead to more complete processing comparable to that found in button-press tasks. More research is needed to be able to rule out the possibility that non-syntactic effects are caused by the unnaturalness of the reading conditions in self-paced tasks. In particular, if the goal is to provide evidence for context effects, it is necessary to supplement existing evidence with eye movement data, or to show the equivalence of the tasks by replicating previous eye movement results.

Similarly, the comparison between lexical decision and naming tasks on one hand, and eye movement data on the other hand leaves some questions open. Results on verb complexity obtained using a dual-task paradigm (Shapiro, Zurif, & Grimshaw, 1987, 1989; Shapiro, Brookins, Gordon, & Nagel, 1991) could not be replicated using an eye movement paradigm (Schmauder, 1991). Lexical class effects studied by Tyler and Marslen-Wilson (1977) in a priming task could not be directly compared to eye movement results obtained by Rayner, Carlson, & Frazier (1983), because the location of the analyses differed.

A direct comparison of these types of task is aggravated because different modalities are employed. Priming tasks are usually conducted in a cross-modal fashion in which the sentences are presented aurally, whereas eye movement data are obtained in reading tasks. Once more, further studies which compare the tasks using identical materials are needed to be able to evaluate if naming and lexical decision latencies provide a more sensitive measure of comprehension processes than eye fixations.

Moreover, the development of other methods to study parsing processes might be fruitful. Tanenhaus, Garnsey, & Boland (1990), for instance, report studies on gap-filling which use evoked brain potentials as the response measure. While the subject is reading or listening, his or her brain potentials are recorded. This measure is unintrusive and can be taken during normal comprehension. It has been shown that the resulting waveforms show a negative peak about 400 msec after the onset of an anomalous word (e.g., Kutas & Hillyard, 1983). Tanenhaus, et al. (1990) demonstrated that the N400 was sufficiently sensitive to capture the processing differences between plausible and implausible objects, and that this effect could be used to study gap-filling processes.

## **2. Selection of materials:**

Another reason for the differences between studies could lie in the selection of materials. In empirical work on parsing, the independent variables are manipulated through the use of different materials. For example, various context paragraphs, sentence frames in which different verbs are inserted, ambiguous vs. disambiguated forms of the sentences are the basis for comparisons of processing. Since all theories on parsing agree that lexical and pragmatic information plays a role during processing eventually, an effective manipulation must be observable in global measures. Thus, many studies use cloze tasks, sentence completion tasks, rating tasks, or paraphrasing tasks to pretest the materials. The preferred subcategorization frame of a verb, for instance, can be found in the most frequent continuation of a sentence containing the verb as the main predicate. The appropriateness of a context for biasing towards one particular interpretation of a syntactically ambiguous structure can be found using a cloze task or a sentence completion task. Paraphrasing can provide information on the preferred interpretation of a globally ambiguous sentence.

In the studies which argued most strongly against non-syntactic effects the materials were not pretested, or they were pretested in an insufficient way. For instance, the subcategorization preferences in Ferreira and Henderson (1990) were assessed intuitively, and lexical effects were not obtained. In contrast, the selection of strongly biased verbs allowed Holmes (1987) to document lexical effects. The bias of the context paragraphs in the studies by Ferreira and Clifton (1986), and Mitchell, Corley, & Garnham (1990) was not confirmed in an independent experiment. The pretest of the paragraphs in Britt et al.'s third experiment (which did not yield context effects) did establish



facilitated overall processing, but did not convincingly confirm the intended bias. Taraban and McClelland (1988) and Trueswell and Tanenhaus (1990), on the other hand, did provide on-line data as evidence for context effects, but only after the appropriateness of the materials had been confirmed using sentence completion and rating tasks.

A test of the materials, therefore, seems to be crucial. The failure to obtain immediate lexical or context effects in an on-line experiment can only be taken as evidence for a rule-based parsing model if global measures validate the selection of the materials.

### **3. Further research directions:**

Besides an extension of the present research using more thorough comparisons of different on-line tasks, and more careful selections of the materials, several other research directions are feasible. The experiments on lexical effects are mostly concerned with verb information, while the lexical entries of other lexical items have not been studied. The psycholinguistic work on discourse context effects is so far restricted to a few syntactic phenomena. Prepositional phrase attachments, reduced relative clauses, and reduced complement clauses have been the only grammatical structures considered in this literature. It is desirable to enhance the generality of the results by including other syntactic structures. For instance, the investigation of lexical class assignment (Tyler & Marslen-Wilson, 1977; Cowart & Cairns, 1987; Frazier & Rayner, 1987) could be extended to include discourse context in addition to (or instead of) a biasing sentence context. In the research on lexical effects, studies have also been conducted on gap-filling and long-distance dependencies (which are not reviewed in this paper), as well as studies employing ambiguous phrase boundaries, clause boundaries, or adverbial phrase attachments. Eventual context effects in processing of these syntactic phenomena could shed light on the question if contextual override can only take place for certain syntactic structures but not for others (Perfetti, 1990; Britt et al., 1992).

For the selection of context paragraphs, a similar extension is proposed. Most of the studies made use of the principle of referential support in order to construct facilitating context paragraphs. Exceptions are the experiments by Britt et al. (1992; also Rayner, Garrod, & Perfetti, 1992), in which discourse focus was manipulated, and the experiment by Trueswell and Tanenhaus (1991), in which tense information was varied. The formation of a situation model or discourse model depends on many factors, though. Identification of some of these factors could be used to provide the starting point for further research. For instance, the domain knowledge of subjects is an important factor in comprehension. If pragmatic knowledge is necessary for successful disambiguation, the availability of this knowledge is predicted to determine the ease of processing.

A final proposal for an extension of the empirical research concerns the relationship between lexical information and pragmatic information. The interaction between these two types of knowledge has not been extensively studied. Using materials in which both lexical preference and discourse bias are varied could provide additional information on how and when these two types of knowledge are used in syntactic processing.

## **2. Implications for parsing theories**

Given the inconsistency of the experimental results, and the possibility that some of the obtained effects might be induced by task specific strategies, it is difficult to evaluate the theoretical proposals for the human parsing system. Nevertheless, leaving aside methodological considerations, I will try to draw some conclusions about the architecture of the syntactic processor. Furthermore, I will outline research questions concerning the architecture of the processing system which have not been sufficiently investigated.

## **1. Conclusions on the architecture of the syntactic processor:**

The issue of parallelism in syntactic processing has not been directly addressed in this paper. However, the findings that verb complexity increases processing load (Shapiro, et al., 1991) can only be accounted for if multiple representations are considered concurrently. Moreover, MacDonald, Just, and Carpenter (1991) provided evidence for parallel consideration of several structural alternatives. In contrast, the results consistent with the garden-path model of parsing, suggest a serial, single-representation model (e.g., Ferreira & Clifton, 1986; Mitchell, Corley, & Garnham, 1992). In an attempt to reconcile these contradictory results, Just and Carpenter (1992) suggested to include the reading skills (or the working memory capacity) of the reader. Only readers with good reading skills can maintain multiple representations. If this account can resolve the question of parallelism in syntactic processing has yet to be seen.

The issues of lexical versus rule-based syntactic processing, and of informational encapsulation were the topic of the reviews in the previous two chapters. As discussed there and in the beginning of this chapter, the inconsistency of the results and methodological questions do not allow to come to definite conclusions. However, it seems that the evidence for lexical effects is strong enough to argue that lexical information plays a very important role in the parsing process. Even in a model which postulates that lexical information is only taken into account after an initial rule based interpretation has been completed, the mechanisms which make use of the lexical information have to be spelled out explicitly.

Similarly, there is considerable support for effects of pragmatic knowledge and context. The use of these knowledge sources is not postponed until the clause or sentence boundary. Thus, while pragmatic knowledge might not be available immediately upon encountering a word to suggest an appropriate analysis, it is used early in processing. A purely rule based model, in which syntactic principles guide the structural interpretations, cannot account for these facts (but see Frazier, 1990b, for an attempt to extend the model). Once more, a model of the parsing process has to specify explicitly when and how non-syntactic information is used to aid in the syntactic analysis.

## **2. Alternative research directions:**

It is not surprising that so many questions remain open, since psycholinguistic research on the architecture of the syntactic processor is still in its infancy. Although there are many suggestions on how the empirical research can be extended, it is not likely that the type of psycholinguistic studies reviewed here will suffice to describe the parsing mechanisms comprehensively. Thus, converging evidence from other areas of language comprehension research should be considered as well. Two of the most interesting approaches are the study of other subject populations, (besides adults with good comprehension skills), and cross-linguistic studies.

To answer questions on the architectural constraints of the syntactic processing mechanisms, neuropsychologists and neurolinguists study the effects of brain damage on language comprehension (for reviews see Linebarger, 1990a, 1990b; Zurif, Swinney, & Garrett, 1990). If the syntactic processor is a modular input system, language pathology should provide evidence for the dissociability of linguistic processes and other cognitive processes. If there is modularity within the language processing system, selective disruptions of syntactic, lexical, or semantic processes should be observable in certain aphasic patients. In fact, patients have been documented who seem to be able to analyze a sentence grammatically without understanding its meaning.

The second area of research to be considered here is cross-linguistic research. Many linguistic theories of syntax claim to be universal. With respect to the parsing theories discussed here, for instance, the principles of minimal attachment and late closure are assumed to be applicable in all languages. Thus, in order to test the validity of the principles, psycholinguistic studies which compare processing of similar syntactic structures in different languages are needed. Cuetos & Mitchell (1988; Mitchell, Cuetos, & Zagar, 1990), for example, used a set of English sentences whose preferred interpretation was obtained by applying the principle of late closure. For the

Spanish translations, however, whose syntactic structures mirrored those of the English sentences, the preferences were different.

Furthermore, cross-linguistic studies can provide information going beyond the data obtained in parsing studies in English, because differences in the syntactic structures of the languages can be used to make more focussed predictions (e.g., Frazier, 1987; MacWhinney, Bates, & Kliegl, 1984). For instance, Strube, Hemforth, & Wrobel (1990) report experiments which make use of the fact that in German past perfect sentences, the main verb appears at the end of the sentence. For past tense sentences, on the other hand, the word order is the same as in English. Processing times for these sentences indicated that the availability of the verb before the object noun phrase influenced the preferred attachment of a prepositional phrase. Results like this demonstrate that studying other languages, besides English, could provide useful insights into syntactic processing (Flores d'Arcais, 1990).

Hopefully, bringing together psycholinguistic research with these (and other) areas will lead to a more detailed and more accurate description of the language processing mechanisms, and allow to distinguish between the theoretical proposals.

### 3. Implications for text comprehension theories

Although most psychological models of parsing contain assumptions about the use of non-syntactic information, the goal of the models is to find a grammatically correct representation of the utterance. The question of how comprehenders use the resulting representation to derive the meaning of an utterance lie outside the scope of parsing models. Text comprehension theories, in contrast, are developed in order to account for phenomena such as recall, recognition, inferences, or eventual distortions of the textual input. Thus, the syntactic processor is considered one component, among others, which enables subjects to encode the meaning of sentences. Pragmatic knowledge and context might be used to facilitate syntactic processing, but in addition, they might also be used to form a representation of the gist of the utterance. A complete model of the comprehension process must therefore include representations on different levels. Syntactic and non-syntactic mechanisms as well as specific proposals about their interactions are needed to explain how readers or listeners construct these representations.

Jurafsky (1991) put forward three criteria for the evaluation of a parsing component in a larger framework of comprehension. First, the model has to be functionally adequate, that is, the representations computed by the model have to be "rich and complete enough to function as an interpretation of the sentence in a larger model of language understanding." Second, the model has to be representationally adequate, that is, it "must include a declarative and linguistically motivated representation of linguistic knowledge." And finally, the model must be cognitively adequate, that is, it "must meet standards of psycholinguistic and general cognitive validity."

As an illustration of the difficulty of evaluating whether a theory of language comprehension fulfills these criteria, (in particular the criterion of cognitive adequacy), I will briefly describe two models which make very different assumptions about the architecture of the language processing system. Both models postulate representations of the input on a text based and a situational level, and both models are specified enough to be implemented in a computer simulation.

#### 1. NL-SOAR

NL-SOAR (Lehman, Lewis, & Newell, 1991a; 1991b) is an implementation of a language processing system within the universal cognitive architecture SOAR (Newell, 1990; Laird, Newell, & Rosenbloom, 1987). This production system is explicitly based on the assumption that all cognitive processes can be explained within a common framework. Thus, language comprehension is not considered having a special status in the cognitive system. Like other cognitive processes,

language comprehension is formalized as the execution of production rules (cf. Anderson, Kline, & Lewis, 1977; Thibadeau, Just, & Carpenter, 1982). Through a process called chunking, the system can acquire new productions which make the comprehension process increasingly efficient. Initial deliberate comprehension, which resembles problem solving, is thus gradually replaced by later recognitional comprehension, which resembles adult reading.

To illustrate the type of productions and the knowledge representation used, consider the following example. This production finds the correct interpretation of "John" as the subject in the sentence "John knows":

```

IF the problem space is Comprehension and
   the Comprehension-state has node p2 on the edge and
   p2's profile has
       number = 3s
       subclass = proper
       base-level class = person
       receives = subject and
   the Comprehension-state has an incoming node n1 and
   n1 corresponds to the word knows and
   the comprehend operator is being applied to n1
THEN add a change record to the comprehend operator that links p2 to n1 as subject and
   mark p2's profile as
       number = 3s
       tense = present
       assigns = subject & object & clausal-object
etc.

```

The production encodes a constraint on number agreement, uses a word order cue (subject - verb), and contains subcategorization information which can be used in subsequent processing.

Syntactic relationships are represented in the so-called utterance model which is a network consisting of word nodes and links denoting simple syntactic relationships (e.g., determiner, subject, complement). In addition, the model builds a situation model of the utterance which encodes the conceptual relationships between nodes. The two representations are linked via their common elements (each word in the utterance model is connected to corresponding object in the situation model). Although there is no mechanism at present which specifies how discourse context is taken into account, the situation model can also contain elements from previously encountered sentences.

NL-SOAR is a strongly interactive model of parsing in which all knowledge sources are immediately brought to bear; the two types of representation are built concurrently. Consider, for example, the following production which SOAR computed to disambiguate the adjective "rich" (Lehman et al., 1991b, p. 464):

```

IF comprehending rich,
   and there is a verb (...),
   and the verb has an unfilled pred-complement role,
   and the subject of the verb is a person,
   and the subject's referent isn't already wealthy,
THEN the sense of rich is wealthy,
   assign rich to be the pred-complement of the verb,
   assert the wealthy property of the subject's referent.

```

[lexical]  
[syntactic]  
[syntactic]  
[semantic]  
[pragmatic]  
[lexical]  
[syntactic]  
[pragmatic]

In both the condition and the action sides, several levels of analysis are utilized. The condition labeled "pragmatic" refers to a feature of the situation model (rather than the general world knowledge), the condition labeled "semantic" (about the animacy of the subject) concerns selection

restrictions, and the second "syntactic" condition specifies that an open slot in the postulated verb frame has to be filled. Thus, sentence context, lexical properties, and syntactic considerations are accessed in parallel, in a single processing step.

In the case of syntactic ambiguity, however, the model selects only a single interpretation. If this interpretation turns out to be incorrect (this can occur either in a syntactic garden path sentence in which a subsequent word cannot be integrated in the structure, or if semantic and pragmatic constraints are violated), then the model can apply a so-called snip operator to undo a previously made link. Another syntactic attachment is tried and a lookahead facility evaluates if the snip leads to a more appropriate interpretation. Since the snip operator can only cut links in the active edge set (i.e., links made recently), the model predicts that comprehension cannot succeed for garden-path sentences in which an earlier attachment decision has to be revised.

The model in its present form does not make specific predictions about the time course of processing. Since it is embedded in SOAR, however, reaction time predictions can be obtained by counting the number of productions executed. Given the relatively sparse examples provided, it is not possible to evaluate if the resulting predictions on the reading times correspond to empirical data. Moreover, it is not yet clear how the productions built through chunking generalize to other utterances. The examples given seem to be very context specific. For instance, the production for the disambiguation of the adjective "rich" would not be applicable to the sentence "The rich woman bought an airplane". Similarly, it is not clear if the successful disambiguation of a syntactic structure would lead to a general production for parsing other, syntactically similar sentences.

NL-SOAR is functionally and representationally adequate. However, the psychological validity of the model can only be thoroughly evaluated if the system is developed further. Since the focus of the model lies on explaining the transition from deliberate comprehension to recognitional comprehension, the knowledge which is explicitly represented in the system is only sparse. Only after exposing the chunking mechanism to many examples of utterances can an analysis of the resulting productions be attempted. Moreover, empirical data on processing difficulties and the time course of processing can only be compared to predictions of NL-SOAR after the model has developed a large degree of recognitional comprehension.

To summarize, NL-SOAR is a strongly interactive, single-representation model of language comprehension. The model is a production system within the general architecture SOAR. In the following section, a weakly interactive, multiple-representation model is described.

## **2. The construction-integration model of text comprehension:**

Language comprehension has also been modeled in the connectionist framework. Purely connectionist systems have been shown to account for phenomena on various levels (e.g., Elman, 1990; Gallant, 1991; Jain, 1991; McClelland & Kawamoto, 1986; Sharkey, 1990; St. John, 1992; St. John & McClelland, 1990; Waltz & Pollack, 1985). However, the scope of the models is usually very small. They are concerned with word recognition, syntactic processes, thematic role assignment, the use of schemata, and many other issues, but there have been few attempts to combine the models into a larger framework. One answer to the question of whether connectionist systems can efficiently represent symbolic information (see Fodor & Pylyshyn, 1988; Pinker & Prince, 1988) is the development of hybrid models (e.g., Just & Carpenter, 1992; Holyoak, 1990).

The construction-integration model of text comprehension (Kintsch, 1988; Kintsch & Welsch, 1990) combines a propositional, symbolic representation with a connectionist spreading activation process. In the construction phase, the text input is used to form an associative network. The nodes correspond to concepts and propositions, and the links are formed according to argument overlap. For instance, the sentence "Yesterday, John called Sharyn." is represented in the two connected propositions:

P1 call [John, Sharyn]

P2 time-of [P1, yesterday].

An important feature of this construction process is that the rules guiding it are "sloppy" and general. Several hypotheses about the correct interpretation of a sentence can be represented in the propositional network at once. If there are inconsistent hypotheses, the link strengths between the corresponding propositions are negative. For the temporarily ambiguous sentence "Mary knows the guest was late," for instance, the propositions entered into the network include

P1 know [Mary, guest]  
 P2 know [Mary, P3]  
 P3 is-late [guest],

with a negative link between the inconsistent propositions P1 and P2.

The goal of the model is to explain the use of general world knowledge and discourse context in comprehension. Thus, associations to the propositions are retrieved from long-term memory which is also assumed to be represented in the form of an associative network. In the case of comprehending an utterance in context, a few nodes from the previous processing step are carried over to the current one. The resulting network contains information about the sentence context, word associations, and the prior discourse, as well as the current input.

The second stage, initiated after words, short phrases, or clauses, consists of the integration of the network using a spreading activation process. The context inappropriate nodes are deactivated, while the central propositions of the sentence (often the proposition encoding the main clause) remain at a high activation level. The activation pattern after the network settled is considered the comprehenders' representation of the utterance. Simulations using this representation have been shown to account for recall and verification data, for the disambiguation of homonyms, for pronoun assignment, and many other phenomena in language comprehension (Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Kintsch, 1992).

The parsing component in this model is not concerned with the interpretation of syntactic ambiguities. The model is a multiple representation model, in which all alternative structures are activated in parallel. Linguistic knowledge is therefore only needed to translate the input into the propositional format used in the spreading activation stage. This component of the model is at present not explicitly formulated. However, since the propositional representation so closely resembles subcategorization frames or verb argument frames, it seems that besides simple syntactic rules, lexical information should play an important part in guiding the propositionalization. Non-syntactic knowledge, such as the discourse context, or pragmatic knowledge is put into the constructed network immediately, but it is not used to propose syntactic analyses. The propositional representation is elaborated by the non-syntactic information, but it is not altered in changing contexts. The non-syntactic knowledge exerts its effect only in the integration stage, in which the connections between text propositions and elaborations help to establish the consistency of the input with pragmatic context. Thus, the model is weakly interactive. Non-syntactic information is used after the parallel activation of syntactic alternatives, and it affects parsing decisions on-line.

The selection of the appropriate interpretation is accomplished through the constraint-satisfaction process in the integration phase. Only few examples are available to show the effectiveness of this process, and to illustrate which types of knowledge are used. Consider the example "The janitor cleaned the room with the broom/windows." used by Taraban and McClelland (1988). In this sentence, the prepositional phrase attachment is disambiguated using pragmatic knowledge. As a simulation using the construction-integration model shows, associative connections between "broom" and "cleaning" and between "room" and "windows" are sufficient to select the appropriate attachment. It is not necessary to encode thematic roles or selection restrictions explicitly in the model. In the garden-path sentence "The horse raced past the barn fell," only connections encoding the syntactic inconsistency between the two interpretations of "raced" are needed to select the correct interpretation. Further experimentation is needed to explore the generality of these early results.

As for NL-SOAR, several important components have not yet been addressed in the construction-integration theory. First, the model does not make predictions about the time course of processing. Reading times, or backtracking behavior cannot be predicted in the present model. However, an extension of the model is feasible which takes into account the processing costs associated with creating new propositions, and the processing costs associated with integrating the network. Second, the work on parsing has not yet simulated discourse context effects. Since the theory has been successfully applied to model contextual phenomena, however, the extension of this work to parsing effects is straightforward.

A thorough comparison of the two models is beyond the scope of this paper. It is important to note, however, that both models are representationally and functionally adequate, and both models are considered to be psychologically valid. Both models assume the representation of lexical and rule-based syntactic knowledge. However, NL-SOAR is strongly interactive, whereas the construction-integration model is weakly interactive; NL-SOAR is a single-representation model, CI is a multiple representation model. NL-SOAR disambiguates by choosing the interpretation which fulfills constraints encoded in production rules, whereas CI disambiguates using a relaxation procedure in an associative network.

#### 4. Summary

The literature on lexical effects and context effects in parsing allows to draw conclusions concerning the methodology of on-line comprehension experiments. In particular, a direct comparison of different on-line paradigms is needed to resolve inconsistencies in the empirical results. Moreover, the selection of materials must be confirmed by global measures. Only if a lexical or pragmatic bias is effective on a global level, can on-line effects be detected.

In the absence of more psycholinguistic experiments to reconcile contradictory results, it is difficult to evaluate the psychological validity of the theories of parsing. Nevertheless, a purely rule-based, serial model does not seem to account for the data as well as a model which takes into account non-syntactic information and lexical information. In all models of parsing, a more detailed account of the specific interactions between the different types of knowledge is needed.

Embedding a parsing component in a larger model of text comprehension can provide the means for defining these interactions. Two psychological theories of text comprehension were described which use a larger architecture successfully to model how general world knowledge and discourse context interact with syntactic knowledge. However, these models make diametrically opposite assumptions about the architecture of the parser.

These apparent differences and contradictions, as well as the inconsistent empirical results should not discourage, however. The study of parsing mechanisms seems to be an active research area in which many questions are unanswered. Theoretical work and empirical research can supplement each other in the attempt to come to a better understanding of the human language comprehension system.

## References

- Aaronson, D., & Ferres, S. (1984). The word-by-word reading paradigm: An experimental and theoretical approach. In D. E. Kieras & M. A. Just (Eds.), New methods in reading comprehension research (pp. 31-68). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Abney, S. P. (1989). A computational model of human parsing. Journal of Psycholinguistic Research, *18*, 129-144.
- Altmann, G. T. M. (1988). Ambiguity, parsing strategies, and computational models. Language and Cognitive Processes, *3*, 73-97.
- Altmann, G. T. M. (1990). Cognitive models of speech processing: An introduction. In G. T. M. Altmann (Ed.), Cognitive models of speech processing (pp. 1-23). Cambridge, MA: MIT Press.
- Altmann, G. T. M. (Ed.) (1990). Cognitive models of speech processing. Cambridge, MA: MIT Press.
- Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. Cognition, *30*, 191-238.
- Anderson, J. R., Kline, P. J., & Lewis, C. H. (1977). A production system model of language processing. In M. A. Just & P. A. Carpenter (Eds.), Cognitive Processes in Comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baddeley, A. D. (1986). Working memory. New York: Oxford University Press.
- Balota, D. A., Flores d'Arcais, G. B., & Rayner, K. (Eds.) (1990). Comprehension processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Beech, J. R., & Colley, A. M. (Eds.) (1987). Cognitive approaches to reading. Chichester: Wiley.
- Berwick, R. C., Abney, S. P., & Tenny, C. (Eds.) (1991). Principle-based parsing: Computation and psycholinguistics. Dordrecht: Kluwer Academic Publishers.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language. New York: John Wiley and Sons.
- Bock, J. K., & Kroch, A. S. (1989). The isolability of syntactic processing. In G. N. Carlson, & M. K. Tanenhaus (Eds.), Linguistic structure in language processing (pp. 157-196). Dordrecht: Kluwer Academic.
- Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. Journal of Memory and Language, *29*, 413-432.
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, *2*, 331-350.
- Bresnan, J., & Kaplan, R. M. (1982). Introduction: Grammars as mental representations of language. In J. Bresnan (Ed.), The mental representation of grammatical relations. Cambridge, MA: MIT Press.
- Britt, M. A., Perfetti C. A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: Context effects and their limits. Journal of Memory and Language, *31*, 293-314.
- Carlson, G. N., & Tanenhaus, M. K. (Eds.) (1989). Linguistic structure in language processing. Dordrecht: Kluwer Academic.
- Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. Journal of Verbal Learning and Verbal Behavior, *20*, 137-160.
- Carroll, P. J. & Slowiaczek, M. L. (1987). Modes and modules: Multiple paths to the language processor. In J. L. Garfield (Ed.), Modularity in knowledge representation and natural-language understanding (pp. 221-248). Cambridge, MA: MIT Press.
- Chomsky, N. (1957). Syntactic structures. The Hague: Mouton.
- Chomsky, N. (1981). Lectures on government and binding. Dordrecht: Foris.
- Clark, H. H., & Clark, E. V. (1977). The psychology of language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.
- Clifford, J. (1990). Grammar and the activation of referents in working memory: The case of particle movement. Unpublished manuscript, University of Colorado.



- Clifton, C., Jr., & Ferreira, F. (1987). Discourse structure and anaphora: Some experimental results. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 619-634). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Clifton, C., Jr., & Ferreira, F. (1989). Ambiguity in context. Language and Cognitive Processes, 4, 91-104.
- Clifton, C., Jr., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. N. Carlson, & M. K. Tanenhaus (Eds.), Linguistic structure in language processing. Dordrecht: Kluwer Academic.
- Clifton, C., Jr., Frazier, L., & Connine, C. (1984). Lexical expectations in sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 23, 696-708.
- Clifton, C., Jr., Speer, S., & Abney, S. P. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. Journal of Memory and Language, 30, 251-271.
- Coltheart, M. (Ed.) (1987). Attention and performance XII: The psychology of reading. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Connine, C. M., Ferreira, F., Jones, C., Clifton, C. E., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. Journal of Psycholinguistic Research, 13, 307-319.
- Cowart, W. (1982). Autonomy and interaction in the language processing system: A reply to Marslen-Wilson and Tyler. Cognition, 12, 109-117.
- Cowart, W., & Cairns, H. (1987). Evidence for an anaphoric mechanism within syntactic processing: Some reference relations defy semantic and pragmatic constraints. Memory and Cognition, 15, 318-331.
- Crain, S., & Fodor, J. (1985). How can grammars help parsers? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), Natural language parsing: Psychological, computational, and theoretical perspectives (pp. 320-358). Cambridge: Cambridge University Press.
- Crain, S., & Shankweiler, D. (1990). Explaining failures in spoken language comprehension by children with reading disability. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 539-556). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax parser. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), Natural language parsing: Psychological, computational, and theoretical perspectives (pp. 320-358). Cambridge: Cambridge University Press.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. Cognition, 30, 73-105.
- Daneman, M., & Carpenter, M. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-466.
- Dowty, D. R., Karttunen, L., & Zwicky, A. M. (Eds.) (1985). Natural language parsing: Psychological, computational, and theoretical perspectives. Cambridge: Cambridge University Press.
- Elman, J. L. (1990). Representation and structure in connectionist models. In G. T. M. Altmann (Ed.), Cognitive models of speech processing (pp. 345-383). Cambridge, MA: MIT Press.
- Ericsson, K. A., & Kintsch, W. (1991). Memory in comprehension and problem solving: A long-term working memory (Tech. Rep. No. 91-13). Boulder: University of Colorado, Institute of Cognitive Science.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. Journal of Memory and Language, 25, 348-368.
- Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 555-568.
- Fillmore, C. J. (1968). The case for the case. In E. Bach, & R. Harms (Eds.), Universals in linguistic theory. New York, NY: Holt, Reinhart and Winston.
- Fillmore, C. J. (1984). Some thoughts on the boundaries and components of linguistics. In T. G. Bever, J. M. Carroll, & L. A. Miller (Eds.), Talking minds: The study of language in the cognitive sciences (pp. 73-108). Cambridge, MA: MIT Press.

- Flores d'Arcais, G. B. (1987). Syntactic processing during reading for comprehension. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 619-634). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Flores d'Arcais, G. B. (1990). Parsing principles and language comprehension during reading. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 345-357). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fodor, J. A. (1983). The modularity of mind. Cambridge, MA: MIT Press.
- Fodor, J. A. (1987). Modules, frames, fridges, sleeping dogs, and the music of the spheres. In J. L. Garfield (Ed.), Modularity in knowledge representation and natural-language understanding (pp. 25-36). Cambridge, MA: MIT Press.
- Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). The psychology of language: An introduction to psycholinguistics and generative grammar. New York: McGraw-Hill.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28, 3-71.
- Fodor, J. D. (1988). On modularity in syntactic processing. Journal of Psycholinguistic Research, 17, 125-168.
- Fodor, J. D., & Frazier, L. (1980). Is the human sentence parsing mechanism an ATN? Cognition, 8, 417-459.
- Ford, M. (1989). Parsing complexity and a theory of parsing. In G. N. Carlson, & M. K. Tanenhaus (Eds.), Linguistic structure in language processing. Dordrecht: Kluwer Academic.
- Ford, M., Bresnan, J., & Kaplan, R. M. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), The mental representation of grammatical relations (pp. 727-796). Cambridge, MA: MIT Press.
- Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), Sentence processing: Psycholinguistic studies presented to Merrill Garrett (pp. 27-85). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Foss, D. J. (1988). Experimental psycholinguistics. Annual Review of Psychology, 39, 301-348.
- Frazier, L. (1985). Syntactic complexity. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), Natural language parsing: Psychological, computational, and theoretical perspectives (pp. 320-358). Cambridge: Cambridge University Press.
- Frazier, L. (1987a). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 554-586). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frazier, L. (1987b). Theories of sentence processing. In J. L. Garfield (Ed.), Modularity in knowledge representation and natural-language understanding (pp. 291-307). Cambridge, MA: MIT Press.
- Frazier, L. (1987c). Syntactic processing: Evidence from Dutch. Natural Language and Linguistic Theory, 5, 519-559.
- Frazier, L. (1989). Against lexical generation of syntax. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 505-528). Cambridge, MA: MIT Press.
- Frazier, L. (1990a). Exploring the architecture of the language-processing system. In G. T. M. Altmann (Ed.), Cognitive models of speech processing (pp. 409-433). Cambridge, MA: MIT Press.
- Frazier, L. (1990b). Parsing modifiers: Special purpose routines in the human sentence processing mechanism? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 303-330). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. Cognition, 6, 291-325.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology, 14, 178-210.
- Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. Journal of Memory and Language, 26, 505-526.
- Gallant, S. (1991). A practical approach for the representation of context and for performing word sense disambiguation using neural networks. Neural computation, 3, 293-309.

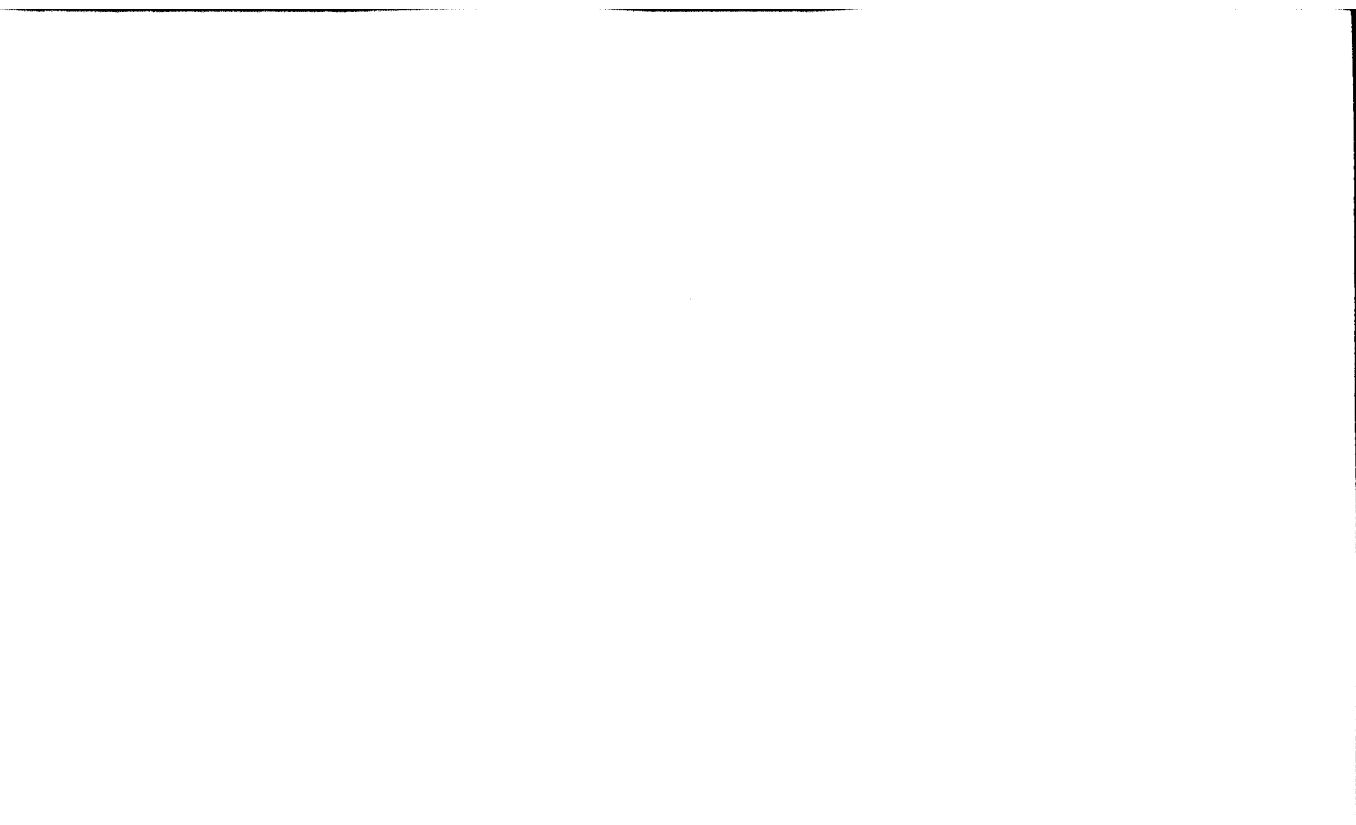
- Garfield, J. (Ed.) (1987). Modularity in knowledge representation and natural language understanding. Cambridge, MA: MIT Press.
- Garnham, A. (1989). Integrating information in text comprehension: The interpretation of anaphoric noun phrases. In G. N. Carlson, & M. K. Tanenhaus (Eds.). Linguistic structure in language processing. Dordrecht: Kluwer Academic.
- Gernsbacher, M. A., & Shroyer, S. (1989). The Cacioppo use of the indefinite *this* in spoken narratives. Memory and Cognition, 17, 536-540.
- Givon, T. (1992). The grammar of referential coherence as mental processing instructions. Linguistics, 30, 5-55.
- Gorrell, P. (1989). Establishing the loci of serial and parallel effects in syntactic processing. Journal of Psycholinguistic Research, 18, 61-73.
- Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 266-283.
- Healy, A. F., Oliver, W. L., & McNamara, T. P. (1987). Detecting letters in continuous text: Effects of display size. Journal of Experimental Psychology: Human Perception and Performance, 13, 279-290.
- Holmes, V. M. (1984). Parsing strategies and discourse context. Journal of Psycholinguistic Research, 13, 237-257.
- Holmes, V. M. (1987). Syntactic parsing: In search of the garden path. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 587-600). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Holmes, V. M., Kennedy, A., & Murray, W. S. (1987). Syntactic structure and the garden path. The Quarterly Journal of Experimental Psychology, 39A, 277-293.
- Holmes, V. M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. Journal of Memory and Language, 28, 668-689.
- Jackendoff, R. S. (1972). Semantic interpretations in generative grammar. Cambridge MA: MIT Press.
- Jain, A. N. (1991). Parsing complex sentences with structured connectionist networks. Neural computation, 3, 110-120.
- Johnson-Laird, P. N. (1983). Mental models. Cambridge, MA: Harvard University Press.
- Jurafsky, D. (1991). An on-line model of human sentence interpretation. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society (pp. 449-454). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.
- Just, M. A., & Carpenter, P. A. (1984). Using eye fixations to study reading comprehension. In D. E. Kieras & M. A. Just (Eds.), New methods in reading comprehension research (pp. 151-182). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122-149.
- Kaplan, R. M., & Bresnan, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), The mental representation of grammatical relations (pp. 173-281). Cambridge, MA: MIT Press.
- Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing complements: Comments on the generality of the principle of minimal attachment. Language and Cognitive Processes, 4, S151-76.
- Kieras, D. E., & Just, M. A. (Eds.) (1984). New methods in reading comprehension research. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. Cognition, 2, 15-47.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. Journal of Memory and Language, 30, 580-602.
- King, M. (Ed.) (1983). Parsing natural language. New York, NY: Academic Press.
- Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Kintsch, W. (1984). Approaches to the study of the psychology of language. In T. G. Bever, J. M. Carroll, & L. A. Miller (Eds.), Talking minds: The study of language in the cognitive sciences (pp. 111-146). Cambridge, MA: MIT Press.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95, 163-182.
- Kintsch, W. (1992). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), From learning processes to cognitive processes: Essays in honor of William K. Estes. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W., & vanDijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 85, 363-394.
- Kintsch, W., & Welsch, D. (1991). The construction-integration model: A framework for studying memory for text. In W. E. Hockley, & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. Journal of Memory and Language, 29, 133-159.
- Konieczny, L., Hemforth, B., & Strube, G. (1991). Psychologisch fundierte Prinzipien der Satzverarbeitung jenseits von Minimal Attachment. Kognitionswissenschaft, 1, 58-70.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. Memory and Cognition, 11, 539-550.
- Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. Artificial Intelligence, 33, 1-64.
- Lehman, J. F., Lewis, R. L., & Newell, A. (1991a). Natural language comprehension in SOAR (Tech. Rep. No. CMU-CS-91-117). Pittsburgh, PA: Carnegie Mellon University, School of Computer Science.
- Lehman, J. F., Lewis, R. L., & Newell, A. (1991b). Integrating knowledge sources in language comprehension. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society (pp. 461-466). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Linebarger, M. C. (1989). Neuropsychological evidence for linguistic modularity. In G. N. Carlson, & M. K. Tanenhaus (Eds.), Linguistic structure in language processing. Dordrecht: Kluwer Academic.
- Linebarger, M. C. (1990). Neuropsychology of sentence parsing. In A. Caramazza (Ed.), Cognitive neuropsychology and neurolinguistics (pp. 55-122). Hillsdale, NJ: Lawrence Erlbaum Associates.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. Cognitive Psychology, 24, 56-98.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. Journal of Verbal Learning and Verbal Behavior, 23, 127-150.
- Marcus, M. P. (1980). A theory of syntactic recognition for natural language. Cambridge, MA: MIT Press.
- Marcus, M. P. (1984). Some inadequate theories of human language processing. In T. G. Bever, J. M. Carroll, & L. A. Miller (Eds.), Talking minds: The study of language in the cognitive sciences (pp. 253-278). Cambridge, MA: MIT Press.
- Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. Science, 189, 226-228.
- Marslen-Wilson, W. D. (Ed.) (1989). Lexical representation and process. Cambridge, MA: MIT Press.
- Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. Cognition, 8, 1-71.
- Marslen-Wilson, W. D., & Tyler, L. K. (1983). Reply to Cowart. Cognition, 15, 227-235.
- Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. L. Garfield (Ed.), Modularity in knowledge representation and natural-language understanding (pp. 37-62). Cambridge, MA: MIT Press.

- McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 3-38). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McClelland, J. L., & Kawamoto, A. H. (1986). Mechanisms of sentence processing: Assigning roles to constituents of sentences. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), Parallel distributed processing: Explorations in the micro-structure of cognition: Vol. 2. Psychological and biological models (pp. 272-325). Cambridge, MA: MIT Press.
- McKoon, G., & Ratcliff, R. (1984). Priming and on-line text comprehension. In D. E. Kieras & M. A. Just (Eds.), New methods in reading comprehension research (pp. 119-128). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McRoy, S. W., & Hirst, G. (1990). Race-based parsing and syntactic disambiguation. Cognitive Science, *14*, 313-353.
- Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras, & M. A. Just (Eds.), New methods in reading comprehension research (pp. 69-89). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mitchell, D. C. (1987a). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 601-618). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mitchell, D. C. (1987b). Reading and syntactic analysis. In J. R. Beech & A. M. Colley (Eds.), Cognitive approaches to reading (pp. 87-112). Chichester: Wiley.
- Mitchell, D. C. (1989). Verb-guidance and other lexical effects in parsing. Language and Cognitive Processes, *4*, S1123-154.
- Mitchell, D. C., Corley, M. M. B., & Garnham, A. (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. Journal of Experimental Psychology: Learning, Memory, and Cognition, *18*, 69-88.
- Mitchell, D. C., Cuetos, F., & Zagar, D. (1990). Reading in different languages: Is there a universal mechanism for parsing sentences? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 285-302). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. Journal of Memory and Language, *24*, 542-559.
- Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.
- Norris, D. (1987). Syntax, semantics, and garden paths. In A. Ellis (Ed.), Progress in the psychology of language: Vol. 3 (pp. 233-252). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Norris, D. (1990). Connectionism: A case for modularity. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 331-343). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perfetti, C. A. (1990). The cooperative language processors: Semantic influences in an autonomous syntax. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 205-230). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. Cognition, *28*, 73-193.
- Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. Language, *64*, 539-576.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. Journal of Verbal Learning and Verbal Behavior, *22*, 358-374.
- Rayner, K., & Carroll, P. J. (1984). Eye movements and reading comprehension. In D. E. Kieras & M. A. Just (Eds.), New methods in reading comprehension research (pp. 129-150). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. Memory and Cognition, *14*, 191-201.
- Rayner, K., Flores d'Arcais, G. B., & Balota, D. A. (1990). Comprehension processes in reading: Final thoughts. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 631-638). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. The Quarterly Journal of Experimental Psychology, *39A*, 657-673.
- Rayner, K., Garrod, S., & Perfetti, C. A. (1992). During parsing discourse influences are delayed. Cognition, *45*, 109-139.
- Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 327-362). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. Language and Cognitive Processes, *4*, S121-50.
- Riesbeck, C., & Schank, R. (1978). Comprehension by computer: Expectation-based analysis of sentences in context. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), Studies in the perception of language (pp. 247-294). New York: Wiley.
- Schmauder, R. (1991). Argument structure frames: A lexical complexity metric? Journal of Experimental Psychology: Learning, Memory and Cognition, *17*, 49-65.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowsky, M.A. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, *14*, 489-537.
- Sells, P. (1985). Lectures on contemporary syntactic theories: An introduction to government-binding theory, generalized phrase-structure grammar, and lexical-functional grammar. Stanford, CA: Center for the Study of Language and Information.
- Shapiro, L. P., Brookins, B., Gordon, B., & Nagel, N. (1991). Verb effects during sentence processing. Journal of Experimental Psychology: Learning, Memory and Cognition, *17*, 983-996.
- Shapiro, L. P., Zurif, E., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. Cognition, *27*, 219-246.
- Shapiro, L. P., Zurif, E., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. Journal of Psycholinguistic Research, *18*, 223-243.
- Sharkey, N. E. (1990). A connectionist model of text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 487-514). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Small, S., & Rieger, C. (1982). Parsing and comprehending with word experts: A theory and its realization. In M. D. Ringle, & W. Lenhart (Eds.), Strategies for natural language processing. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Speas, M. J. (1990). Phrase structure in natural language. Dordrecht: Kluwer Academic Publishers.
- St. John, M. F. (1992). The story gestalt: A model of knowledge-intensive processes in text comprehension. Cognitive Science, *16*, 271-306.
- St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. Artificial Intelligence, *46*, 217-257.
- Steedman, M. J. (1989). Grammar, interpretations, and processing from the lexicon. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 463-504). Cambridge, MA: MIT Press.
- Steedman, M. J., & Altmann, G. T. M. (1989). Ambiguity in context: A reply. Language and Cognitive Processes, *4*, S1105-122.
- Stowe, L. A. (1989). Thematic structures and sentence comprehension. In G. N. Carlson, & M. K. Tanenhaus (Eds.), Linguistic structure in language processing. Dordrecht: Kluwer Academic.

- Strube, G., Hemforth, B., & Wrobel, H. (1990). Auf dem Weg zu psychologisch fundierten Modellen menschlicher Sprachverarbeitung: Echtzeitanalysen des Satzverstehens. In S. W. Felix, S. Kanngießer, & G. Rickheit (Eds.), Sprache und Wissen: Studien zur kognitiven Linguistik (pp. 115-133). Opladen, Germany: Westdeutscher Verlag.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: Reconsideration of some context effects. Journal of Verbal Learning and Verbal Behavior, *18*, 645-659.
- Tanenhaus, M. K., & Carlson, G. N. (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 529-561). Cambridge, MA: MIT Press.
- Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. S. (1985). Do listeners compute linguistic representations? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), Natural language parsing: Psychological, computational, and theoretical perspectives (pp. 359-408). Cambridge: Cambridge University Press.
- Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. Language and Cognitive Processes, *4*, SI211-234.
- Tanenhaus, M. K., Garnsey, S. M., & Boland, J. (1990). Combinatory lexical information and language comprehension. In G. T. M. Altmann (Ed.), Cognitive models of speech processing (pp. 383-408). Cambridge, MA: MIT Press.
- Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. Journal of Memory and Language, *27*, 597-632.
- Taraban, R., & McClelland, J. L. (1990). Parsing and comprehension: A multiple-constraint view. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 231-263). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. Cognitive Science, *6*, 157-203.
- Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. Memory and Cognition, *16*, 283-298.
- Townsend, D. J., & Bever, T. G. (1982). Natural units interact during language comprehension. Journal of Verbal Learning and Verbal Behavior, *28*, 681-703.
- Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. Language and Cognitive Processes, *6*, 303-338.
- Tyler, L. K. (1989). The role of lexical representations in language comprehension. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 439-462). Cambridge, MA: MIT Press.
- Tyler, L. K., & Marslen-Wilson, W. M. (1977). The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior, *16*, 683-692.
- vanDijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.
- Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. Cognitive Science, *9*, 51-74.
- Wanner, E. (1980). The ATN and the sausage machine: Which one is baloney? Cognition, *8*, 209-225.
- West, R. F., & Stanovich, K. E. (1986). Robust effects of syntactic structure on visual word processing. Memory and Cognition, *14*, 104-112.
- Winograd, T. (1983). Language as a cognitive process, Vol I: Syntax. Reading, MA: Addison-Wesley.
- Wright, B., & Garrett, M. (1984). Lexical decision in sentences: Effects of syntactic structure. Memory and Cognition, *12*, 31-45.
- Zurif, E., Swinney, D., & Garrett, M. (1990). Lexical processing and sentence comprehension in aphasia. In A. Caramazza (Ed.), Cognitive neuropsychology and neurolinguistics (pp. 123-136). Hillsdale, NJ: Lawrence Erlbaum Associates.



-

-