University of Massachusetts Occasional Papers in Linguistics

Volume 12 *University of Massachusetts Occasional Papers in Linguistics - Special Issue: Psycholinguistics- Volume 9*

Article 3

1986

The Role of Context in Resolving Syntactic Ambiguity

Fernanda Ferreira
University of Massachusetts at Amherst

THE ROLE OF CONTEXT IN RESOLVING SYNTACTIC AMBIGUITY

FERNANDA FERREIRA

UNIVERSITY OF MASSACHUSETTS AMHERST

A great deal of current research on sentence comprehension is concerned with the issue of how different categories of information potentially available to the language processor are used. A central question is whether the human sentence parsing mechanism (the parser) can be influenced by nonsyntactic information sources, such as semantic, pragmatic, and discourse-contextual information, or whether the parser assigns syntactic structure independently. The research discussed in this paper is based on the position that the language processor consists of a number of subprocessors, each with its own distinct properties, and that these subprocessors interact in a constrained fashion. From this perspective, the fact that people typically arrive at the most plausible reading of a sentence is considered the outcome of the language processor's highly structured use of different information sources.

The Syntactic Processor

The operation of the sentence comprehension mechanism can be revealed by observing how it behaves when it is confronted with syntactic ambiguity, i.e., with more than one potential analysis of a portion of a sentence. Frazier (1978) outlines a number of ways the parser might cope with ambiguity. The parser might compute all possible analyses in parallel (e.g., Crain & Steedman, 1984), delay making a decision about the analysis until disambiguating information has been received (e.g., Marcus, 1980), or select only one analysis.

The first two possibilities are computationally costly,and the last possibility would mean that the parser might choose the wrong analysis and have to reanalyze the ambiguous material.

There is substantial evidence that the parser computes only one analysis, and that it computes the first analysis available (Frazier, 1978; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983). Frazier (1978) proposed that the parser follows the minimal attachment strategy: Attach incoming material into the phrase-marker being constructed using the fewest syntactic nodes consistent with the well-formedness rules of the language. This strategy is efficient in terms of computational and memory load: Only one analysis at a time is constructed, and all incoming material is structured as it is received. Frazier and Fodor (1978) argue that a minimal attachment analysis is more efficient than a nonminimal attachment analysis because the minimal attachment analysis requires the accessing of fewer phrase structure rules. Under the reasonable assumption that accessing rules takes time, the minimal attachment analysis would be constructed first. The parser then stops once the minimal attachment analysis is complete. If the analysis turns out to be incorrect and the parser is led down the garden-path, then the parser will have to reanalyze the misanalyzed material. Since it takes extra work to do so, the minimal attachment strategy predicts that people will take more time and have more difficulty processing nonminimal attachment sentences.

Frazier (1978) tested this prediction by presenting either minimal attachment or nonminimal attachment sentences to subjects and asking them to judge whether the sentences were grammatical. She found that subjects took less time to make their decisions for the minimal attachment than nonminimal attachment sentences. Frazier and Rayner (1982), using an eye movement recording technique, found that subjects took more time to read nonminimal attachment sentences, and their fixation durations were longer in the region of the sentence that disambiguated the analysis. Subjects also made regressive eye movements to earlier portions of the nonminimal attachment sentences, consistent with the prediction that people misanalyze nonminimal attachment sentences and then go back and revise them.

Autonomy and Interaction in Sentence Processing

Although it has been shown that the behavior of the parser can be characterized in terms of the minimal attachment strategy, it is possible that the strategy is not obligatory, and that it can be overridden by nonsyntactic information. The issue here is whether the syntactic processor operates independently and autonomously of other information sources, or whether different information sources interact and communicate freely. The former alternative is consistent with proposals made by Fodor (1983) concerning the modularity of mind. Fodor argues that the mind consists of a number of input systems or

modules that provide information to the central processor. Input modules are informationally encapsulated: Their operation cannot be affected by information at relatively high levels of representation. The syntactic processor can be viewed as a module, since it uses syntactic information to compute a syntactic representation, and it does not have access to information at higher levels, such as semantic constraints and world knowledge. To use Pylyshyn's (1980) terminology, the operation of the syntactic processor is cognitively impenetrable.

Forster (1979) has proposed a model of the language processor in accord with this modular approach. The language processor is organized into three microprocessors: a lexical processor, a syntactic processor, and a message processor. Each level operates autonomously; no microprocessor can receive information from a microprocessor at a higher level. The syntactic processor cannot be affected by the message processor; it can only send its output up to the message processor. The task of the message processor is to use semantic information to convert the representation into a conceptual structure. At this point, the language processor no longer operates, and the general problem solving system takes over.

In contrast to this model is Marslen-Wilson's fully interactive model (Marslen-Wilson & Tyler, 1980; Tyler & Marslen-Wilson, 1977). According to this model, different information sources (lexical, syntactic, and semantic-pragmatic) interact and communicate freely during sentence comprehension. The output of these different sources is potentially available to any other. Marslen-Wilson argues that such a flexible and unconstrained system results in the most efficient processing.

A more extreme view of the use of semantic information during sentence processing is Schank's Conceptual Dependency Theory (Schank 1972; Riesbeck & Schank, 1978), which states essentially that a syntactic representation is not computed at all, except in unusual situations (i.e., when lexical constraints are not sufficient to constrain the analysis of a sentence). This view is not unusual in the psychological literature (c.f., Clark & Clark, 1977). However, there is substantial evidence that people do compute a syntactic representation (e.g., Marslen-Wilson & Tyler, 1980; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983), and there is some evidence (Frazier, 1983) that they will not assign a pragmatically tempting reading to a sentence if it is not grammatically licensed (e.g. people will not assign the sensible reading to John claimed the gangster will die yesterday at the shootout). Therefore, any theory of language processing which does not attribute psychological reality to a level of syntactic representation cannot be adequate.

Information Sources in Sentence Comprehension

It is important to distinguish among the different sources of information that could potentially interact with syntactic information. One source of information is the output of the parser itself, i.e., syntactic information. Frazier, Taft, Roeper, Clifton, and Ehrlich (1984) presented sentences consisting of two clauses to subjects. The clauses were either parallel (both minimal attachment or both nonminimal attachment) or nonparallel. They found that the second segment of parallel form sentences was read faster than the second segment of nonparallel structures. A large facilitation effect was found when the nonminimal attachment form in the second clause followed a nonminimal attachment form in the first clause, but no facilitation was found when the minimal attachment form was preceded by the minimal attachment form.

This result is interesting because although it demonstrates that the behavior of the parser can violate the minimal attachment strategy, it does not conflict with the autonomy of syntactic processing assumption. The facilitation is due to processes occurring within a level, not from a higher level process feeding down to a lower one. Facilitation within a single level of the language processor does not violate the autonomy assumption. This argument is analogous to the argument made by Forster (1979) concerning facilitation within the level of the lexical processor. Forster argues that the standard DOCTOR-NURSE priming effect does not violate the autonomy of lexical processing assumption because the effect is intralexical (i.e., the effect is due to associations that exist within the lexicon). Thus, as long as facilitation occurs within a level of the language processor, the autonomy of levels assumption is not violated.

A second source of information that could potentially interact with the syntactic processor is semantic information. For the purposes of this paper, "semantic" information will be considered both readers' linguistic knowledge of the meaning of the expressions of their language and some aspects of their knowledge of real world properties of objects and events (Rayner, Carlson, & Frazier, 1983, footnote 1). In addition, semantic information will be (arbitrarily) defined in terms of within sentence plausibility. Since our intuitions suggest that semantic information can strongly affect how we interpret sentences, it has been argued by proponents of the interactive model that semantic information can completely override any syntactic bias. In contrast, the autonomous model would predict that semantic information cannot affect the parser's initial syntactic decisions, but could be used to aid reanalysis of a syntactically misanalyzed string.

Rayner, Carlson, and Frazier (1983) conducted a study using the eye-movement recording technique designed to evaluate the claims made by the interactive and autonomous models. They proposed that if the interactive model is correct, then they would expect subjects

initially to adopt the semantically preferred reading of the sentences and therefore not to be garden pathed in sentences semantically biased towards the nonminimal attachment reading. If, on the other hand, the autonomous model is correct, then they would still expect garden paths in the nonminimal attachment sentences.

Their experiment contrasted sentences such as the following:

(1) The performer sent the flowers was very pleased. (Reduced plausible NMA)
(2) The florist sent the flowers was very pleased. (Reduced implausible MA)
(3) The performer who was sent the flowers was very pleased. (Unreduced plausible unambiguous)
(4) The performer sent the flowers and was very pleased with herself. (active implausible MA)

Rayner et al. predicted that, if readers initially compute the minimal attachment reading regardless of semantic plausibility, then readers should be garden-pathed in (1) and (2) but not in (3) and (4). Consistent with these hypotheses, they found that the reduced sentences took longer to read than the other sentences forms. Further, subjects took much longer to read the reduced relative sentences in the disambiguating region (was very) for both first pass and total reading times. Analysis of the number of regressive eye movements revealed that there were more regressions in the reduced relative sentences than in the others. This pattern of data indicates that semantic plausibility did not determine readers' initial structural analysis of the sentences, providing evidence for the autonomy of syntax hypothesis.

A third and final sources of information that needs to be considered is contextual and discourse knowledge, which will be defined as semantic plausibility factors operating across sentences and integrative processes occurring during text processing. Very little research has been done on the question of how syntactic and discourse-contextual information are integrated. This topic is important, however, and for at least two reasons. First, preceding context might provide the syntactic parser with information about the pragmatically preferable analysis before the parser actually commits itself to a structural analysis of an ambiguous string. Thus, this information source would be especially useful to a parser that could be influenced by nonsyntactic information. Second, studying sentences in context might provide information about language processing in a more natural setting. Sentences rarely appear in isolation, and so it is possible that results from single sentence experiments are based on tasks that are unnatural to the subject and force him to adopt an unusual or atypical strategy.

This second point can be considered in terms of one of the main functions of language, namely communication. A speaker uses language to convey information to his listener(s). Some sentence structures are better suited to conveying certain kinds of messages than others; different structures are appropriate to different discourse functions. For example, the cleft construction It is the girl that is petting the cat is appropriate to a situation in which the listener already knows that someone is petting the cat, but doesn't know who that someone is; it would be inappropriate to use this sentence if the listener knew the girl was petting something, but didn't know what it was (Clark & Haviland, 1977). Further, this sentence may be difficult to process in isolation compared to a simple active sentence, but in an appropriate context, the difficulty of the sentence might disappear.

Davison and Lutz (1984) explored this issue by presenting sentences of different syntactic forms to subjects together with a preceding context sentence that served to establish a semantic relation between the two sentences. They hypothesized that if the preceding context sentence mentioned the sentence topic of the next (target) sentence, the target sentence would be easier to understand than if it were presented in a neutral context. The transformed sentence forms studied were Passive, There-insertion, Adverb preposing, Raising to subject position, and Raising to object position. Each structure introduces a different entity as sentence topic. Corresponding untransformed sentences were also presented. Subjects were told to read each sentence and press a button when they felt they had understood the sentence. From their data, Davison and Lutz conclude that (1) the transformed version is harder to process than the untransformed version in a neutral context; (2) reading time decreases if the context mentions the topic of the next sentence; and (3) reading times increase (somewhat) if the context and target sentence are incongruent.

These conclusions, however, are not supported by their results. First, the difference between the transformed and untransformed sentences in the neutral contexts is quite small (2395 msec vs. 2360, respectively, on one analysis; 2410 vs. 2350 on an analysis excluding the Raising to subject data). Second, the conclusions imply that there was a significant interaction between sentence form (transformed vs. untransformed) and context type (neutral, biased-untransformed, and biased-transformed), but no such interaction was found. Third, for each sentence form, reading times for the neutral context were longer than for either of the biased contexts, and there was virtually no difference between congruent and incongruent conditions. Even when there was a mismatch between context type and sentence form, there was little increase in the reading times, and subjects were faster than in the neutral conditions. Thus, if context had any effect at all on the processing of the sentences (and the absence of a significant interaction suggests that it did not), it seems that the effect is due

to the presence of context, regardless of how the context relates to the next sentence. Thus, although some of the obtained effects are in the predicted direction, the effects are small and not statistically significant. The claim that context can affect ease of syntactic processing--although intuitively plausible--has not been supported by Davison and Lutz' obtained results.

Crain and Steedman (1984) also studied the effects of context on syntactic processing. They proposed that syntactic ambiguities are resolved by semantics and specific conversational context, rather than by structural-syntactic mechanisms. To clarify this proposal, they distinguish between weak and strong interaction. In a language processing system with weak interaction between syntactic and nonsyntactic information, the semantic processor can decide whether to terminate or continue a given syntactic analysis. In a system with strong interaction, the semantic processor decides which syntactic structure(s) get proposed in the first place. Crain and Steedman argue for weak interaction, but the interaction takes place on a word-by-word basis (consistent with Marslen-Wilson and Tyler's (1980) data on the availability of semantic information). The parser proposes syntactic analyses in parallel, and these analyses are then compared in terms of contextual information. Therefore, there should be no difficulty associated with a syntactic structure if it is placed in an appropriate context. This prediction contrasts with the results obtained by Rayner, Carlson, and Frazier (1983): When the preferred syntactic analysis and the semantic bias were in conflict, subjects had more trouble understanding the sentence than when the two information sources were consistent. Nevertheless, Crain and Steedman might argue that within-sentence plausibility effects studied by Rayner et al. are not nearly as influential as plausibility effects due to contextual information.

The strong prediction made by the Crain and Steedman model is that all structural garden paths can be eliminated by manipulation of semantics and context. In the right context, the difficulty of a syntactically complex sentence will be eliminated. In particular, it is crucial that context satisfy the presuppositions of a structurally complex sentence. For example, complex structures such as reduced relative clause constructions (The horse raced past the barn fell) contain a number of presuppositions: first, that there is a set of horses in focus; second, that the relative clause applies to a horse in the set; and third, that there is a single horse identified by the relative clause. Because of the presuppositions, this sentence is almost impossible to understand out of context or in a context that does not satisfy the presuppositions. This argument has implications for the neutrality of "neutral" contexts used in experiments on the processing difficulty of sentences. If a reduced relative is placed in a neutral context that does not satisfy the presuppositions, the simpler analysis will be preferred on pragmatic grounds, not on syntactic grounds.

Crain and Steedman conducted three experiments to test these proposals. In the first experiment, subjects were presented reduced relative sentences such as 5a and 5b, and they were asked to judge whether the sentences were grammatical or ungrammatical.

5(a). The teachers taught by the Berlitz method passed the test. (Implausible)
5(b). The children taught by the Berlitz method passed the test. (Plausible)

The implausible sentences were judged ungrammatical 55% of the time, and the plausible sentences were judged ungrammatical 35% of the time. No response time data were presented. The results of this experiment (as well as Experiments 2 and 3) cannot be taken as evidence for the model Crain and Steedman offer, because the measure is not sensitive to online processes during comprehension. This experiment does not provide convincing evidence that context can influence the processing of structural garden path sentences.

In their second experiment, Crain and Steedman attempted to create garden paths in sentences which syntactically should not garden path the parser. They constructed contexts which set up the reader to expect either a relative or complement reading of the ambiguous portion of a target sentence such as (6) (the complement reading is the simpler construction).

6(a) The psychologist told the wife that he was having trouble with her husband. (Complement)
6(b) The psychologist told the wife that he was having trouble with to leave her husband. (Relative)

The contexts either picked out one member of a couple (Complement-inducing context) or picked out one of the two couples (Relative-inducing context). Crain and Steedman predicted that the subjects would be garden pathed by the simpler construction (which structurally should not garden path the parser) if it were preceded by the relative-inducing context. Consistent with this prediction, they found that a complement target preceded by a relative context was judged ungrammatical 54% of the time (compared to 12% ungrammatical when preceded by an appropriate context); a relative target preceded by a complement context was judged ungrammatical 50% of the time (compared to 22% when preceded by an appropriate context).The task, however, was not a strict grammaticality judgment. The subjects were told to try to integrate the context sentence with the target sentence, and to decide whether there was anything peculiar about the context-target pairs. The pattern of data they obtained, then, is not surprising. In addition, as Crain and Steedman acknowledge, this experiment does not permit evaluation of the claim that the parser first selects a minimal attachment analysis and then revises it; the judgment task measures the product of the integration of many

different information sources but does not measure how the information sources are integrated.

Their third experiment was designed to get at the question whether a minimal attachment analysis is initially proposed by the parser. Crain and Steedman suggested that if a context permits both a minimal attachment and nonminimal attachment analysis equally well, any residual preference for one analysis or the other would have to be attributed to structural biases. An example of such a neutral context is the following.

> Three new players joined the baseball team. Several older players picked one of the new players to join them for the first practice. The other new players were upset by this.

The target sentence either appeared in complement (minimal attachment) or relative (nonminimal attachment) form, as in (7):

> 7(a). The coach convinced the player that the old pros chose to practice by themselves for the rest of spring training. (Complement)
> 7(b). The coach convinced the player that the old pros chose to practice by himself for the rest of spring training. (Relative)

Subjects rated both the relative and complement sentences as grammatical 61% of the time. From this result, Crain and Steedman argue that analyses are not proposed serially on the basis of structural criteria, as argued by Frazier. However, again, this experiment cannot be taken as conclusive evidence against the Frazier model because of the kind of measure employed; Crain and Steedman did not tap online processes occurring during sentence processing. As well, it is difficult to draw strong conclusions from a finding of no difference between two conditions. It is quite possible that the experimental technique was not sensitive enough to pick up actual differences. In sum, neither the Crain and Steedman nor the Davison and Lutz study provide evidence that the behavior of the parser can be directly affected by contextual information.

The question how contextual information affects syntactic processing is clearly an important one for evaluating autonomous and interactive models of language processing in a more natural situation. Contextual information is potentially available to the parser before the parser actually begins a structural analysis. Therefore, in a biasing context, it might be possible to influence the parser's initial structural decision and cause the parser to select the less preferred syntactic analysis (nonminimal attachment).

Alternatively, contextual information may act just like semantic information in the Rayner, Carlson, and Frazier (1983) study: The parser initially selects the first analysis available (minimal attachment) and only later revises the analysis if it turns out to be incorrect or anomalous (or both). To distinguish these possibilities, neutral contexts must be carefully constructed in accordance with the points raised by Crain and Steedman: Their neutrality should consist not in their permitting neither the minimal nor the nonminimal analysis, but rather in permitting both analyses with close to equal plausibility. Unless the neutral context is constructed so as to permit both readings, the target sentence would be hard to process on purely pragmatic grounds. The speaker would be violating the cooperative principle if the complex sentence did not appear in a reasonably appropriate context.

EXPERIMENT

The purpose of the experiment was to examine the role of context in parsing strategies. This topic can be divided into two issues: (1) how structural analyses are proposed, and (2) how contextual information interacts with syntactic information. The first issue refers to how the parser copes with syntactic ambiguity. The parser could compute all analyses in parallel, delay making a syntactic decision until disambiguating information is received, or compute only one analysis at a time, based on structural or semantic criteria. The second issue concerns whether contextual information can be used to influence the syntactic analyses that get proposed in the first place, or whether contextual information is used to reject an analysis (or analyses) or to propose alternatives to analyses independently computed by the syntactic processor.

On the basis of these issues, it is possible to identify four major classes of models. First, according to a parallel model, all syntactic analyses are computed in parallel (e.g., Bever, 1970), and context is used to select among the alternatives. This model implies that there are no purely structural garden paths; syntactic structures are difficult to process if they do not appear in an appropriate context (Crain & Steedman, 1984). As long as the neutral contexts satisfy the presuppositions of the target sentence, neutral contexts should qualify as appropriate as well as biased contexts.

Second, when confronted with an ambiguous string, the parser could delay making a syntactic decision until disambiguating information is received (e.g., Marcus, 1980). This model also predicts that there should be no garden paths. Furthermore, context should have no effect since the parser simply waits for clearly disambiguating information.

Third, the parser could compute only one analysis, and only the semantically preferred analysis. This model is assumed by

semantically-driven language processors and detective-style parsers (Fodor, Bever, & Garrett, 1974). No garden paths should occur if the ambiguous string appears in an appropriately biasing context, and sentence processing should be easier in the presence of a biasing context than in a neutral context. The latter prediction follows since a parser of this sort gathers information in favor of a syntactic analysis; more information should facilitate construction of the analysis.

Finally, the parser might compute only one analysis, as in the third model, but always the syntactically preferred (minimal attachment) analysis. Contextual information would be used to aid reanalysis if it turns out that the minimal attachment analysis is incorrect. According to this model, garden paths should be observed for nonminimal attachment sentences even in strongly biasing contexts. Reanalysis may occur more quickly or more easily if the nonminimal attachment sentence appears in a biased rather than a neutral context. Once the parser realizes that it has been led down the garden path and it must therefore search for alternative analyses, the information in the biased context is likely to be useful.

The experiment was designed to assess how the parser operates when it has available contextual information that biases the interpretation of a syntactically ambiguous string. Target sentences appeared in either minimal attachment or nonminimal attachment form, as in (8).

8(a). The horse raced past the barn / and fell in a puddle. (MA)
8(b). The horse raced past the barn / fell in a puddle. (NMA)

Notice that the two sentences are identical except for the presence of the word and in 8(a). The sentence is ambiguous up to the slash, and is disambiguated by the next one or two words (and fell vs. fell).

Target sentences such as these were placed in either a biasing or neutral context. The biasing contexts contained information that strongly selected for a particular reading of the target sentence. A biasing minimal attachment context contextually biased the interpretation of (for example) the horse raced past the barn towards a minimal attachment reading by mentioning and describing a particular horse that could be taken as the referent of the horse that raced past the barn. A biasing nonminimal attachment context biased the string towards a nonminimal attachment reading by explicitly mentioning two horses (a mare and a stallion), one of which was raced past a barn, the other of which was raced to the boundary of the next farm. The next sentence, the target sentence, referred to the horse raced past the barn. This context should strongly bias the reading of the target sentence, since the reduced relative structure is used to distinguish between two referents previously mentioned which could possibly be confused.

The neutral contexts contained information that permits both analyses, but does not strongly select for one or the other. The neutral contexts were constructed in accordance with the points made by Crain and Steedman (1984) concerning the presuppositions of complex syntactic structures. Continuing with the horse example, more than one horse was mentioned, but in contrast to the nonminimal attachment biasing contexts, the horses were only mentioned (but not given distinguishing characteristics). In addition, each passage had one sentence following the target sentence, so that subjects did not become sensitive to the critical sentences.

Context-target passages appeared in four different forms: minimal attachment (MA) context-MA target, nonminimal attachment (NMA) context-NMA target, Neutral context-MA target, and Neutral context-NMA target. (Context bias and target sentence bias were always congruent rather than crossed. Crossing of the contexts and target sentences would most likely simply have produced anomalous passages.)

Predictions

The four conditions in the experiment are: (1) MA-context-MA sentence (MA-MA), (2) NMA-NMA, (3) Neutral-MA (N-MA), and (4) Neutral-NMA (N-NMA). In addition, collapsing across sentence form, there are two conditions of context: (1) biased (including MA-MA and NMA-NMA) and (2) neutral (including N-MA and N-NMA).

The parallel model would predict no difference between MA-MA and NMA-NMA conditions, no difference between N-MA and N-NMA conditions, and no difference between the neutral and biased conditions. These predictions follow since both the neutral and biased contexts satisfy the presuppositions of the nonminimal attachment sentence; above and beyond the satisfaction of presuppositions, the parallel model does not specify the role of context. The delay model also predicts no differences between any of the conditions, since the parser simply waits until disambiguating information is received and then computes a syntactic analysis.

According to the third model--which states that only one analysis is computed at a time, and only the semantically-pragmatically preferred analysis--the MA-MA and NMA-NMA conditions should be the same, and the N-MA condition should be faster than the N-NMA condition (since in the absence of a biasing context, the simpler structure should be computed more quickly). This model also predicts that the biased condition should be faster than the neutral conditions, since the presence of potentially useful information should facilitate construction of the analysis of the target sentence.

According to the fourth model--which states that only one analysis at a time is computed and always the minimal attachment--the

NMA-NMA condition should be slower than the MA-MA condition, and this difference in reading time should appear only in and after the disambiguating region of the sentence, not before. The N-NMA condition should be slower than the N-MA analysis. The biased condition may also be faster than the neutral condition after the disambiguating region of the sentence if contextual information is used quickly to revise an incorrect syntactic analysis.

Method

Subjects. Thirty-two students from the University of Massachusetts Psychology Department human subjects pool participated in this experiment.

Materials. Target sentences such as (8a or b) were placed in either a congruent or neutral context, resulting in the four different versions, as shown in Table 1.

Table 1

Example of a Passage in the Four Experimental Conditions

MA-MA
Sally's job was to prepare horses for racing competitions. Probably the most important aspect of her work was selecting the horses that were strong and healthy so that her training regimen would be most effective. Her qualification test involved taking the horses down to a nearby farm. She would let each horse run freely through the fields and by the farmer's old barn, and then to the boundary of the next farm. Sally put one horse to the test. The horse raced past the barn and fell in a puddle. Right then and there, Sally knew that horse did not have the right stuff.

NMA-NMA
Sally's job was to prepare horses for racing competitions. Probably the most important aspect of her work was selecting the horses that were strong and healthy so that her training regimen would be most effective. Her qualification test involved taking the horses down to a nearby farm. She and her assistant would each take a horse and ride it as hard as it would be ridden in a race. Sally rode a mare through the field and by the farmer's old barn; her assistant rode a stallion right to the boundary of the next farm. The horse raced past the barn fell in a puddle. Right then and there, Sally knew that horse did not have the right stuff.

N-MA
Sally's job was to prepare horses for racing competitions.

Probably the most important aspect of her work was selecting the horses that were strong and healthy so that her training regimen would be most effective. Her qualification test involved taking the horses down to a nearby farm. She and her assistant would each take a horse and ride it as hard as it would be ridden in a race, through the fields and by the old barn. If a horse seemed good, they'd let it run freely. The horse raced past the barn and fell in a puddle. Right then and there, Sally knew that horse did not have the right stuff.

N-NMA
Sally's job .... The horse raced past the barn fell in a puddle. Right then and there, Sally knew that horse did not have the right stuff.

Sixteen different passages such as those presented in Table 1 were constructed. The target sentences from each passage are presented in Appendix 1. Subjects saw only one version of each passage, and they saw four passages in each of the four conditions. The experimental stories were presented together with sixteen filler stories, and the presentation of experimental and filler stories was randomized individually for each subject.

Each story was divided into segments or regions so that segment-by-segment reading times could be obtained. The segments ranged between one and five words, and corresponded roughly to phrases. Each target sentence was divided into regions as follows: Noun phrase, ambiguous segment (verb phrase vs. reduced relative clause), disambiguating segment (conjunction plus verb vs. verb), and phrase after the disambiguating segment. For example (c represents the critical or disambiguating region):

c-2	c-1	c	c+1
The horse	raced past the barn	(and) fell	in a puddle.

Each story (experimental and filler) was followed by three questions designed to assess the subjects' comprehension of the story. The second question asked them about their interpretation of the target sentence (e.g., Which horse fell in a puddle?), and was written so that a different answer was correct for the minimal attachment and nonminimal attachment sentences. If the subject gave an ambiguous answer, he was probed for more information until he either gave an unambiguous answer or indicated that he didn't know the correct answer. The other question-answering data were not scored, and were included simply to prevent the subjects from becoming aware of the purpose of the experiment.

Procedure. Each of the thirty-two different passages was presented to subjects on a computer screen. The initial display for a passage presented a dash in place of each letter, but preserved

spaces. The subjects hit a button to bring up the first phrase or segment of the story. When they had read and understood the segment, they pressed the button again to bring up the next phrase, and the previous phrase was replaced with dashes. The subjects continued in this manner until all 32 stories were read. The subjects' reading times for each phrase were recorded.

Results

The mean reading times were computed for each segment of each target sentence. The data are presented in Table 2.

Table 2

Mean Reading Time for Each Region in Each Condition and Percent Correct for Each Condition

	Region				
Condition	c-2	c-1	c	c+1	% correct
NMA-NMA	637	756	847	880	81
MA-MA	582	715	682	715	76
Neutral-NMA	595	729	775	855	55
Neutral-MA	650	762	742	746	63

Two analyses of variance were performed on the reading time data, one treating subjects as the random variable, and the other treating sentences as the random variable. The fixed effects variables were segment (four levels), sentence form (MA vs. NMA), and context bias (biased vs. neutral).

Subjects took longer to read the nonminimal attachment than the minimal attachment sentences. The effect was significant on the subjects analysis ($F_1(1,31)=9.97$, $p<.004$), but not on the items analysis ($F_2(1,15)=2.92$, $p<.11$). Table 2 shows that the difference between the sentence forms appeared in and after the critical segment. The interaction between form and segment was significant on the subjects analysis ($F_1(3,93)=3.61, p<.02$) and approached significance on the items analysis ($F_2(3,45)=2.61, p<.07$). Reading time also varied among the different segments of the sentence ($F_1(3,93)=17.77$, $p<.0001$, $F_2(3,45)=7.43$, $p<.0004$, but this effect is uninterpretable, because the segments differed in length. The interaction between context bias and sentence form was significant on the subjects analysis ($F_1(1,31)=6.00$, $p<.02$) but not on the items

analysis ($F_2(1,15)=2.40, p<.14$).

The differences in mean reading times between the minimal and nonminimal sentences were compared using a Bonferroni t-test ($p<.05$). Considering the biased condition first, subjects took longer to read the nonminimal attachment than the minimal attachment sentences in the critical region and in the region following the critical region. These effects were significant on both the subjects and items analysis. The minimal attachment and nonminimal attachment reading times for the other regions were not significantly different. In the neutral condition, subjects took longer to read the nonminimal attachment sentences in the region following the critical region; this effect was significant in the subjects analysis but not in the items analysis. The minimal attachment and nonminimal attachment reading times for the other regions were not significantly different.

For the question data, the percentage correct in each of the four conditions was computed. The subjects understood the target sentence more often in the biased than in the neutral conditions, 81% correct for the nonminimal attachment and 76% correct for the minimal attachment sentences in the biased condition, and 55% correct for the nonminimal attachment and 63% correct for the minimal attachment sentences in the neutral condition. This effect of context bias was significant in the analysis of variance treating items as a random variable, $F(1,15)=12.30$, $p<.003$, but the effect of sentence form ($F(1,15)=.02, p<.89$) and the interaction between the two variables ($F(1,15)=1.42, p<.26$) were not.

Discussion

The main results of the experiment are the following: Subjects took longer to read the nonminimal attachment sentences than the minimal attachment sentences, as has been found in previous studies (Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983). This difference held up even in the presence of biasing context; i.e., context did not override the syntactic bias of the parser. Second, there is some evidence that context facilitated reanalysis of the nonminimal attachment sentences. In the biased context condition, subjects apparently started their reanalysis of the misanalyzed string one segment earlier than subjects in the neutral condition. Finally, the question data suggest that context allowed subjects eventually to get the right reading of the reduced relatives: Subjects were more accurate answering questions about the nonminimal attachment sentences in the biased than in the neutral conditions.

The goal of the experiment was to examine the role of context on parsing strategies, and this goal was divided into two issues: (1) how are syntactic analyses proposed when a string is syntactically ambiguous (in parallel, serially, or delayed), and (2) how is contextual information integrated with syntactic information. The

results support the position that the parser proposes only one analysis at a time, the first analysis available (minimal attachment), arguing against both the parallel and delay models. The role of context is not to guide the parser's initial syntactic decisions, but to provide information to aid reanalysis of a misanalyzed string. These findings are consistent with the Rayner, Carlson, and Frazier (1983) study, which showed that within-sentence plausibility information did not affect the parser's initial analysis. Similarly, the experiment has demonstrated that contextual across-sentence plausibility information does not affect the parser's initial analysis. These results argue against the Crain and Steedman model, which predicted that all garden paths could be eliminated if a sentence appeared in an appropriate context (namely, a context that satisfies the presuppositions of the utterance). Not only were people garden-pathed in the neutral context (which satisfied the presuppositions of the reduced relatives), they were garden-pathed in the biased contexts which both satisfied the presuppositions of the sentence and strongly selected for a reduced relative reading of the target sentence.

What kind of a mechanism can account for both the initial analysis and the reanalysis results? First I will consider a simple model, in which reanalysis is guided just by syntactic information. When the syntactic processor parses the nonminimal attachment sentence in the biasing context, it initially assigns a minimal attachment structure to the temporarily ambiguous string The horse raced past the barn. When the parser comes to the word fell, it gets a syntactic error signal. At this point the parser must reanalyze the string, and so it goes back and computes a nonminimal attachment structure. The parser could do this in a few different ways. One possibility is that, rather than terminating the analysis once it arrives at a minimal attachment structure, it continues with the other analysis until the nonminimal attachment structure is also completed. Another possibility is that the minimal attachment structure analysis is not computed at all during reanalysis, and the parser computes just the nonminimal attachment structure. Once the parser has computed the reduced relative reading, it finds a referent for the complex noun phrase in the developing discourse representation, and the parser takes fell to be the main verb of the sentence.

This sort of a model can handle almost all of the results of the experiment; the only finding it has some trouble explaining is the earlier use of context in the biased condition. The model, however, cannot explain nonsyntactically signalled reanalysis. In the Rayner, Carlson, and Frazier (1983) study, subjects reanalyzed sentences such as John saw the cop with a revolver but the cop didn't see him from a minimal attachment to a nonminimal attachment reading. The reanalysis occurred not because there was anything syntactically wrong with the analysis, but because it is unlikely that anyone would use a revolver as an instrument for seeing. To account for this, Frazier proposed a

thematic processor that selects the most plausible frame from alternative lexical frames listed for the head of a phrase. The syntactic and the thematic processor operate independently, but their outputs are quickly integrated so that the most sensible reading of a sentence is eventually computed. This finding, in addition to the earlier use of context found in the present experiment, suggests the need for a more complex model than the one outlined above.

A more complex model would allow the text processor to propose the referent of the noun phrase The horse raced past the barn to the parser before the parser reanalyzes the string. Using the information that there is a referent in the discourse model that may match an alternative analysis to guide its operation, the parser reanalyzes and computes the nonminimal attachment structure. This sort of model can explain the earlier use of context in the biased condition. In the neutral condition, a referent is not immediately available in the discourse representation, so the subject goes on to see if the items after the disambiguating region (fell) can salvage the initial syntactic analysis. When nothing does, the parser tries to find a referent in the discourse representation again. Because the referent has to be inferred, sometimes a referent is identified and sometimes not, accounting for the long reading times in the segment following the critical segment. This sort of mechanism can also explain why subjects were less accurate in the neutral than in the biased condition.

Although the results of this experiment provide quite convincing evidence that the syntactic processor operates independently and autonomously of other information sources, a few improvements could be made to eliminate some minor problems. The first problem is the lack of significance on the statistical analyses treating items as a random variable. Some of the effects that were significant on the subjects analysis were close to significant on the items analysis, but it would be much better if the effects were significant on both analyses. This pattern indicates a problem with some of the stories used in the experiment; in particular, it is possible that some of the first verbs of the target sentences permitted a reduced relative reading more easily than others. A second problem is the technique used. Although segment-by-segment reading times do tap online processes occurring during language comprehension, the technique is a little crude. A far more sensitive technique is eye-movement recording, which not only measures first-pass reading times and fixations but also picks up regressive eye movements (which one would expect to see for garden-path sentences).

To follow up this study, two more experiments using the eye tracker will be conducted. The first experiment is designed simply to replicate this experiment using a more sensitive technique. For the second experiment, ambiguous sentences such as John threw the ball in the garage will be placed in an appropriately biasing (minimal or

nonminimal attachment) or neutral context. This experiment will be done to assess the generality of the results of the reduced relative experiment by looking at a different and less complex structure (in particular, a structure that does not involve any deletions of optional elements).

References

Clark, H., & Clark, E. Psychology and language: An Introduction to Psycholinguistics. New York: Harcourt, Brace, Jovanovich.

Clark, H. & Haviland, S. (1977) Comprehension and the given-new contract. In R. Freedle (Ed.), Discourse comprehension and production. Norwood, N.J.: Ablex Publishers.

Crain, S., & Steedman, M. (1984) On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen and A. Zwicky, Natural Language Parsing, Cambridge University Press.

Davison, A., & Lutz, R. Measurement of syntactic complexity relative to context. In D. Dowty, L. Karttunen and A. Zwicky, Natural Language Parsing, Cambridge University Press.

Fodor, J., Bever, T., & Garrett, M. (1974)The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar.

Fodor, J. Modularity of Mind. Cambridge, Mass.: MIT Press.

Forster, K. (1979) Levels of processing and the structure of the language processor. In W. E. Cooper and E. C. T. Walker (Eds.), Sentence Processing. Hillsdale, N.J.: Erlbaum Press, 27-86.

Frazier, L. (1978) On comprehending sentences: Syntactic parsing strategies. Bloomington, Ind.: Indiana University Linguistics Club.

Frazier, L. (1983) Processing sentence structure. In K. Rayner (Ed.), Eye Movements in Reading: Perceptual and Language Processes. New York: Academic Press, 215-236.

Frazier, L., & Fodor, J.D. (1978) The sausage machine: A new two-stage parsing model. Cognition, 6, 291-326.

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.

Marcus, M. (1980)A theory of syntactic recognition for natural language. Cambridge, Mass.: MIT Press.

Marslen-Wilson, W., & Tyler, L. (1980) The temporal structure of spoken language understanding. Cognition, 8, 1-72.

Pylyshyn, Z. (1980) Computation and cognition: Issues in the foundation of cognitive science. Behavioral and Brain Sciences, 3, 111-169.

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

Riesbeck, C., & Schank, R. (1978) Comprension by computer: Expectation-based analysis of sentences in context. In W.J.M. Levelt and G.B. Flores d'Arcais (Eds.), Studies in the perception of language. New York: Wiley, 247-294.

Schank, R. (1972) Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 3, 552-631.

Tyler, L. & Marslen-Wilson, L. (1977) The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior, 16, 683-692.

Appendix 1

Target Sentences and Questions Used in the Experiment

The man expected to die (but) would not give up easily.
Who expected the situation was terminal?
MA--patient; NMA--doctor.

The editor played the tape (and) agreed the story was a big one.
Who played the tape?
MA--editor; NMA--the reporter.

The man taught the new method (but) thought the standard one might be superior.
Who thought the standard method might be superior?
MA--teacher; NMA--student.

The troll brought the princess (and) thought she looked good enough to eat.
Who thought the princess looked good?
MA--troll who went out; NMA--troll who stayed at home.

The horse raced past the barn (and) fell in a puddle.
What fell in a puddle?
MA--horse running freely; NMA--horse being ridden.

The woman told the joke (but) didn't think it was funny.
Who told the joke?
MA--the woman; NMA--the man.

The union sued for damages (but) didn't expect the settlement to be large.
Who got sued for damages?
MA--company; NMA--union.

The companies mailed the information (but) decided to discontinue certain services immediately.
Who mailed out the information?
MA--companies; NMA--students.

The man sold the Vega (but) knew he wasn't getting a very good deal.
Who didn't think the deal was very good?
MA--salesman; NMA--buyer.

The woman served the caviar (and then) almost fell into the pool.

Who almost fell?
MA--maid; NMA--the guest.

The man ordered the drink (but) refused to drink it.
Who ordered the drink?
MA--one of two men; NMA--another guy at the bar.

The woman paid (and) left the store immediately.
Who left the store immediately?
MA--customer; NMA--clerk.

The man read the story (and) said it just confirmed his suspicions.
Who read the story?
MA--a coworker; NMA--Paul.

The man asked for directions (and) pulled out a map of New England.
Who pulled out the map?
MA--the skier; NMA--old man on the steps.

The woman delivered the letter (and) suddenly got very upset.
Who got upset?
MA--secretary; NMA--manager.

The company awarded the contract (and) was anxious for the project to get started.
Who was anxious for the project to get started?
MA--computer company; NMA--construction company.