Ambiguity Resolution and Constraints on Syntactic Processing

by

Neal Jay Pearlmutter

A.B., Amherst College (1989)

Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1993

Author .....
Department of Brain and Cognitive Sciences
August 11, 1993

Certified by .....
Mary C. Potter
Professor, Department of Brain and Cognitive Sciences
Thesis Supervisor

Accepted by .....
Emilio Bizzi
Chair, Department of Brain and Cognitive Sciences

Ambiguity Resolution
and Constraints on Syntactic Processing

by

Neal Jay Pearlmutter

Submitted to the Department of Brain and Cognitive Sciences
on August 11, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Ambiguity resolution studies in human sentence processing have focused on the applicability of the two heuristics proposed in the Garden Path theory of Frazier (1979, 1987): Minimal Attachment and Late Closure. The first of these forces the parser to construct a phrase structure representation of an input using the fewest grammatically permitted nodes. The second heuristic specifies that new material should be attached to more-recently-processed structure in favor of less-recent structure, applying when two permitted representations involve the same number of nodes, so that Minimal Attachment does not arbitrate between them. This thesis investigates the generality of these two processing strategies and suggests several reformulations in the context of a new theory.

Chapter 1 describes a set of experiments investigating the effects of contextual information on the application of Minimal Attachment. When context supports the eventually correct interpretation of an ambiguity (in opposition to Minimal Attachment), the ambiguity is no more difficult than an unambiguous control. When context does not support the correct interpretation (instead either providing no useful information or supporting the incorrect interpretation), the ambiguity produces significant difficulty. Furthermore, the strength of contextual bias predicts degree of difficulty across items: The stronger the contextual support for the appropriate interpretation, the smaller the difficulty. Thus context seems to affect the processing of syntactic ambiguity rapidly, overriding any Minimal Attachment-based preference.

Chapter 2 provides a new theory of ambiguity resolution which accounts for the effects in Chapter 1, as well as a wide range of other related results in the literature. The theory unifies lexical and syntactic ambiguity resolution by providing an enriched theory of lexical representation and processing, showing how syntactic ambiguities which had been handled under Minimal Attachment in fact derive from ambiguities over lexical representations. Thus the same mechanisms which are used to resolve lexical (meaning) ambiguity can be used to resolve syntactic ambiguity. In such a system, three types of constraints can conceivably affect processing: relative frequency of alternatives in the lexicon, contextual biases, and grammatical

knowledge. This chapter focuses on the effects of the first two types (frequency and context) and their interaction, but it shows how grammatical constraints will apply in the same system, so that the process of understanding a sentence is a matter of identifying mutually compatible alternatives at all levels of representation in the language processing system.

Chapter 3 turns to the second Garden Path theory heuristic, Late Closure, and shows first that Late Closure must be extended, and, second, that an additional principle is required to account for data from several different syntactic ambiguities in English and Spanish. Three experiments show that while a preference for attachment to recent (low in a phrase structure tree) constituents is uniformly present, an additional preference for attachment to non-recent (high in a structural tree) constituents in a sequence also exists, for certain syntactic ambiguities, but not uniformly. Furthermore, the relative strength of these preferences must be parameterized across languages.

Thesis Supervisor: Mary C. Potter
Title: Professor, Department of Brain and Cognitive Sciences

Acknowledgments

Maryellen MacDonald has been a superb advisor throughout my graduate career, as well as a good friend. There is no question that without her, this dissertation would not exist, and she deserves a large proportion of the credit for many of the ideas discussed in it. Conversations with her have never failed to stimulate my interest, and she has always managed to provide insight, valuable advice, and plenty of support, regardless of whether we happened to be on the same coast or not. I look forward to continued collaboration with her in the future.

Ted Gibson and Mark Seidenberg also deserve a great deal of credit for the development of this thesis. Both kept me on my intellectual toes at all times, forcing me to refine my ideas continually and to work through the details I was always trying to ignore. This is the foundation of any productive research environment, and Ted and Mark have each done a superb job of creating such an environment, again, notably, regardless of whether we happened to be on the same coast or not. Enriqueta Canseco-Gonzalez and Greg Hickok deserve similar credit for helping in the development of the work reported in the third chapter, and I also greatly appreciate a number of discussions that I have had about this work with Stephen Crain.

Gary Marcus and I have been officemates for all four of our years here, and, among other things, he has provided me with the most continuously challenging discussions I have ever encountered during episodes of frisbee. The indoor conversations were not too shabby, either: Gary helped to make the graduate student lounge one of the most interesting places to be in E10, particularly in the Spring of 1992.

I am also indebted to many other lounge-dwellers, particularly Sandeep Prasada, John Kim, Kevin Broihier, Suzie Johnson, David Poeppel, and Fei Xu, all of whom kept life in E10 interesting while simultaneously insuring that my dart-throwing skills improved. I thank the rest of the E10 graduate students and post-docs as well, for their constant support and friendliness, and for help on plenty of specific questions, particularly over the past few months.

Bob Berwick, Molly Potter, and Ken Wexler each deserve thanks for providing

many helpful comments, suggestions, and criticisms, both early on in this process, at my orals, and, later, during the writing of the dissertation. Molly, Jan Ellertsen, and Debbie Reynolds also deserve thanks for shepherding me through the administrative detail of thesis preparation and defense.

Finally, I must thank my parents and brother for their support and understanding throughout this process. Despite my uncertainty until very nearly the last possible moment about when and, indeed, whether anything would be finished, they were always supportive and confident that I would do just fine. I can only hope that I have.

Preface

This dissertation is composed of three papers, each of which was originally written as an independent work. The first two are nevertheless quite closely related, in obvious ways, but the third begins with a slightly different perspective and is in part concerned with a different range of issues. It can be reconciled with the second paper in particular, however, if the conceptualization of syntax in the lexicon in the second paper is taken seriously, and if the cost functions of the third paper are inverted and treated as activation functions. I leave working out the details of this reconciliation to future work.

Despite their apparent differences, all three papers draw on the same literature, and thus all references are listed together at the end, followed by the appendices to the first and third papers. All figures and tables are placed appropriately within the text.

Contents

List of Figures

List of Tables

Chapter 1

Plausibility Effects in Syntactic Ambiguity Resolution[1]

1.1 Introduction

An important question in sentence processing research concerns the type of information that the sentence processor or *parser* can use in deciding which interpretation(s) to pursue for a syntactic ambiguity. The most clearly articulated position on this issue has been that the human syntactic processor consults only syntactic information during the initial construction of an interpretation for a syntactic ambiguity, so that it operates as an autonomous module in Fodor's (1983) sense (Frazier, 1979; Rayner, Carlson & Frazier, 1983; Ferreira & Clifton, 1986; Perfetti, 1990). This position is often accompanied by two other claims: 1) that the parser must choose only one interpretation for an ambiguity and cannot pursue multiple alternatives simultaneously, and 2) that the parser must make its initial choice immediately, as soon as the ambiguity is encountered. A modular parser that obeys these two additional constraints could make many costly mistakes if it were to choose the wrong interpretation when forced to make immediate decisions without the benefit of potentially helpful semantic or discourse-level information. In order to increase the chances of making correct

[1]This work was conducted in collaboration with Maryellen MacDonald.

choices, the parser could rely on strategies, and a number of these have been proposed. The best known is Minimal Attachment (Frazier, 1979): The parser should choose the "simplest" interpretation so that it can build a phrase structure representation with the fewest nodes. According to Frazier and colleagues' *garden path* model (Ferreira & Clifton, 1986; Frazier & Rayner, 1982; Rayner et al., 1983), the one interpretation initially adopted for an ambiguity is guided by only three things: 1) knowledge of the allowable syntactic structures in the language in the form of stored phrase structure rules, 2) category information in the input (e.g., whether the current word is a noun, a verb, etc.), and 3) parsing heuristics such as Minimal Attachment. On this view, semantic and/or discourse information may affect the parse only after an initial structure is built, possibly through the operation of a separate thematic processor working in parallel with the syntactic parser (Rayner et al., 1983; Ferreira & Clifton, 1986). Rather than influencing initial parsing decisions, the thematic processor instead triggers backtracking in the parser by signaling that the initial interpretation of an ambiguity is implausible.

In this paper, we argue that the parser is not as limited in the resources available to guide decisions about ambiguities as the garden path model suggests. We will present evidence that some kinds of non-syntactic information can guide the parser in the initial interpretation of syntactic ambiguities, so that heuristics like Minimal Attachment are not the only guidance for the parser. Specifically, we claim that lexical or discourse information can influence decisions about *thematic roles* that are assigned to noun phrases in a sentence, and that these role assignments can in turn affect the choices that are made by the parser when it is confronted with a syntactic ambiguity.

A thematic role is an abstract semantic relation between a verb and one of its noun phrase arguments; these roles include the *Agent* of the action described by the verb, the *Theme* (sometimes called *Patient*) of the action, the *Instrument* of the action, and others. Many verbs can optionally omit or include some of these arguments. For example, the verb *cut* may appear with three arguments, Agent, Theme, and Instrument, as in *Mary cut the rope with a knife*, and it can also omit the Instrument role, as

in *Mary cut the rope.* Tanenhaus and Carlson (1989) have proposed that the various argument structure options for each verb are stored with that verb's lexical entry, so that the entry for *cut* would contain at least two different argument structures, <Agent, Theme, Instrument> and <Agent, Theme>. They have argued that the argument structures play an important role in parsing: When a verb is encountered in the input, all of its argument structures are activated in parallel, and tentative thematic role assignments are attempted for the nouns that have been encountered to that point. Possible assignments are evaluated by comparing information in the noun's lexical entry with the verb's argument structure and the semantic information in its lexical entry. For example, Agents must be animate, and so an inanimate noun will not be assigned an Agent thematic role. If one verb argument structure produces compatible assignments for the noun phrases encountered to that point and other argument structures do not, then this argument structure will be adopted by the parser, and the parser will adopt a syntactic (phrase structure) representation of the input that is compatible with the chosen argument structure. At this point it is unclear whether the thematic assignments could be influenced only by information in the lexicon or if they could also draw on real-world knowledge and information in the discourse representation. We will side-step this question for the time being and will refer to these effects as "plausibility" effects of thematic role assignment without regard to their lexical or discourse origin. Regardless of the outcome of the lexical/discourse issue, this account of parsing differs from Frazier's garden path model, in which thematic role assignments are seen as part of the semantic/discourse components and have no influence on initial parsing decisions.

A number of studies have provided evidence for each of these positions in recent years. The studies we will focus on involve the main verb/reduced relative (MV/RR) ambiguity, though the effects of thematic role assignments on some other ambiguous structures have been explored (Taraban & McClelland, 1988). The MV/RR ambiguity can arise with verbs that have identical past tense and past participle forms, such as *use.* For example, in the fragment *The actor used...*, used might be the past tense main verb of the sentence, as in *The actor used a prop in the commercial*, or

it might be a past participle introducing a reduced relative clause, as in *The actor used in the commercial had a British accent.* Comprehenders often have great difficulty parsing the reduced relative interpretation (Bever, 1970; Frazier, 1979; Rayner et al., 1983), which is exactly what the Minimal Attachment strategy would predict. The phrase structure representation of the reduced relative contains more nodes than does the main verb structure, and so application of the Minimal Attachment strategy would result in the main verb interpretation being chosen for this ambiguity, rather than the reduced relative. It is important to note that the thematic role assigned to *actor* will also be different in each interpretation. In the main verb interpretation, *actor* is the Agent of *used*, whereas in the reduced relative, it is the Theme of *used.* The hypothesis we will evaluate is whether tentative assignments of thematic roles to the subject of the sentence, in this case *actor*, could influence how the MV/RR ambiguity is resolved, or whether Minimal Attachment always determines the choice of interpretations.

Rayner, Carlson and Frazier (1983) investigated this question using sentences such as *The florist/performer sent the flowers was very pleased.* In this sentence, the verb *sent* would initially be analyzed as the past tense main verb of the sentence, using the Minimal Attachment strategy, but this interpretation turns out to be incorrect, because *sent* instead introduces a reduced relative clause, equivalent to *the florist* (or *performer*) *who was sent...* However, if the parser had plausibility information, for example that florists typically send flowers to other people whereas performers typically receive flowers, then it might be able to avoid misanalyzing the *performer* sentence. When *sent* is encountered, any of the three thematic roles could be assigned to either *florist* or *performer*: Either noun could be the Agent of *sent*, the Theme (so that the performer or florist was sent somewhere), or the Goal (or Recipient), in which something was sent to the florist or performer. However, when the parser encounters *flowers*, plausibility information could become available to help the parser. It is implausible that *performer* is the Agent of *sent*, since performers do not send flowers, so the parser might instead assign *performer* the Goal role and adopt the cor-

responding reduced relative clause structure.[2] Rayner et al. (1983) found that this was not the case; subjects had just as much difficulty comprehending the sentences with plausibility cues (*performer*) as those without helpful cues (*florist*). These results were interpreted as indicating that the parser cannot use plausibility information in constructing its initial analysis, and that it instead chooses the Minimal Attachment interpretation in all cases.

Some researchers have criticized the Rayner et al. (1983) study as having only extremely weak cues in its stimuli (e.g., that it is not terribly implausible that a performer would send flowers, McClelland, 1987). A related issue is that the plausibility information is available only after the ambiguity is introduced at *sent*, and thus the decision concerning which interpretation to pursue might already have been made at *sent*, before the potentially helpful information was encountered. These criticisms suggest that the Rayner et al. study was not the strongest test of the hypothesis that non-syntactic information can influence the initial parse, but the study nonetheless demonstrates that there are limits to the use of plausibility information by the parser.

A stronger test of the hypothesis has been provided by Ferreira and Clifton (1986), who also examined the MV/RR ambiguity. Their stimulus items differed from the Rayner et al. (1983) stimuli in two ways. First, the plausibility information was available when the ambiguous verb was encountered, and second, Ferreira and Clifton manipulated the animacy of the subject noun phrase, providing a stronger manipulation than Rayner et al. had used. For example, in *The defendant/evidence examined by the lawyer was uninteresting*, there is semantic information at the point of the ambiguous verb *examined*: *Defendant*, being animate, may be the Agent of *examined*, whereas *evidence* is inanimate and cannot be the Agent of *examined*. If the syntactic parser can make use of this non-syntactic information, then it should be better equipped to parse the version of the sentence in which *evidence* is the subject than when *de-*

[2]The Theme assignment to *florist* or *performer* is ruled out with the appearance of *the flowers* not because of plausibility but because the parser would have had to encounter a prepositional phrase after the verb in order to assign a Theme role to the first NP (e.g. *The florist sent into the greenhouse...*). Because the parser encounters an NP (*the flowers*) instead of a PP, the Theme assignment to *florist/performer* is no longer possible.

fendant is the subject. The parser would reject assigning the Agent thematic role to *evidence*, and instead assign the Theme role, which is compatible with a reduced relative structure, but it would presumably assign the Agent role to *defendant* and therefore adopt the (incorrect) main verb structure. If this non-syntactic information is ignored, however, and if the Minimal Attachment strategy is uniformly applied, subjects should choose the main verb interpretation for both versions of the sentence and should therefore have difficulty comprehending both versions. Ferreira & Clifton found that the predictions of the garden path model were supported: Subjects read more slowly in the disambiguating region (*by the lawyer*) in both the *evidence* and the *defendant* versions of the sentence, compared to unambiguous control sentences. They concluded that Minimal Attachment had been applied to both sentence versions and that the animacy information was not used by the parser at the time it was making its initial decisions about which syntactic structure to adopt for the ambiguity.

Recently, however, Trueswell, Tanenhaus and Garnsey (in press) have provided evidence that may call into question Ferreira and Clifton's results. Trueswell et al. noted that in roughly half of Ferreira and Clifton's stimuli, the ambiguous verbs allowed an inanimate subject in the main verb interpretation of the ambiguity (See Burgess, 1990, for norming data supporting this point). An example of such a sentence is *The car towed from the parking lot was parked illegally.* Here, *car* is inanimate, but it could receive an Instrument thematic role from *tow* in an active sentence, as in *The car towed the trailer.* Trueswell et al. modified the original Ferreira and Clifton stimulus set to remove items which allowed this additional possibility and found different results than Ferreira and Clifton had obtained. Subjects in the Trueswell et al. study did have difficulty with animate-subject ambiguous sentences, as in *The defendant examined...*, compared to unambiguous controls, but they did not read ambiguous sentences with inanimate subjects (e.g., *The evidence examined...*) more slowly than unambiguous controls. These results suggest that non-syntactic information (knowledge of whether a noun is animate or inanimate) can affect thematic role assignments, which can in turn influence the initial decisions of the parser.

The Trueswell et al. results call into question the garden path model of parsing

in which the Minimal Attachment structure is always the first interpretation adopted for the MV/RR ambiguity. Given an alternative model that incorporates early use of thematic roles, an obvious question is how the parser evaluates the thematic role assignments that it attempts as it is constructing a syntactic representation of the input. The Trueswell et al. data show that the parser uses the animacy information of the noun in assigning thematic roles and guiding the construction of a syntactic representation, but the Rayner et al. data show that weaker plausibility information, such as who is likely to send or receive flowers, did not appear to influence parsing operations. One reason for this difference might be that the noun animacy distinction could be represented in the lexical entries of nouns, independent of context: *Evidence* is inherently inanimate and the Agent role for *any* verb must be assigned to an animate noun.. If only the most basic lexical features become available during word recognition, such as noun animacy and verb argument structure, then the parser might easily use such features to evaluate thematic role assignments. In contrast, there should be no core lexical information in the individual entries for *performer*, *sent* and *flowers* that would influence thematic role assignments; *performer* is an acceptable Agent for *sent*, *flowers* is an acceptable Theme for *sent*, and there is no restriction in the lexical entries that could prevent *performer* and *flowers* from appearing in the same clause. Instead, the plausibility information in the Rayner et al. stimuli would either have to stem from a discourse representation incorporating real-world knowledge about floral deliveries, or it could perhaps arise from the combined activations of *performer*, *sent*, and *flowers*, given an extremely rich lexical representation (Tanenhaus, Garnsey & Boland, 1990).

One conclusion from these results might be that the most basic lexical semantic information, such as the feature ±Animate, can influence thematic role assignment, but plausibility information, such as the probability of a performer sending flowers, cannot. As we have noted, however, the Rayner et al. (1983) stimuli differed from the Trueswell et al. (in press) stimuli not only in the plausibility versus animacy nature of the information, but also in the timing of the constraints and the ambiguity. The Rayner et al. constraints appear only when *flowers* is encountered after the

ambiguous verb, but the Trueswell et al. constraints occur at the ambiguous verb (e.g. the conjunction of *evidence* and *examined*). Therefore, if we are to examine further the role of non-syntactic information in thematic role assignment, these factors must be unconfounded so that a plausibility cue is available to the parser at the point of ambiguity. The experiments below manipulate features of the subject noun phrase, as in the Ferreira and Clifton (1986) and Trueswell et al. (in press) studies, but instead of investigating dichotomous lexical features like ±Animacy, we manipulate *plausibility* of a particular thematic role assignment. Experiment 1 varies plausibility using inanimate subject noun phrases, and Experiment 2 manipulates plausibility in animate noun phrases. A replication of the Trueswell et al. effects with these manipulations would suggest that plausibility information can influence thematic role assignment, but if the plausibility effects in the studies below do not guide the parser, this would be evidence that only basic lexical information like animacy can influence thematic role assignment, but that plausibility information cannot.

1.2 Experiment 1A

This experiment investigates plausibility effects on thematic roles assignments within the class of inanimate nouns. It is divided into two parts: Experiment 1A examines reading times on MV/RR ambiguous sentences that vary in plausibility, and Experiment 1B collects off-line data about the strength of the cues in each stimulus item used in Experiment 1A, in order to determine whether the parser is responding to the graded nature of the cues. That is, if the parser can use plausibility information in evaluating thematic role assignments and not merely dichotomous lexical information like animacy, then subjects should show less garden pathing (as measured by reading time differences in ambiguous and unambiguous conditions) for those items in which the helpful plausibility cues are particularly strong.

Because we want to manipulate compatibility of an inanimate noun phrase with a thematic role, we must use a set of verbs that allow inanimate subjects in both main verb and reduced relative constructions. One such set is the class of *spray/load*

verbs, named after two of its members, which has received extensive discussion in the linguistics literature (Levin & Rappaport, 1986; Rappaport, Laughren & Levin, 1987). These verbs can assign up to three thematic roles: an Agent role, a Material role, and a Location role (Levin & Rappaport, 1986). These latter two roles may appear in either order in the verb phrase, as shown in the examples in (1), in which the thematic roles assigned to each noun phrase are noted in subscripts.

(1)

a. [Chris]$_{Agent}$ scattered [the marbles]$_{Material}$ across [the floor]$_{Location}$.

b. [Chris]$_{Agent}$ scattered [the floor]$_{Location}$ with [the marbles]$_{Material}$.

In addition to these constructions with an Agent, some spray/load verbs permit an *ergative* reading, which is an active sentence in which the noun phrase receiving the Material role is the subject, as in (2). Notably, though, a similar version in which the subject noun phrase receives the Location role is not grammatical, as shown by (3).

(2) [The marbles]$_{Material}$ scattered across [the floor]$_{Location}$.

(3) *[The floor]$_{Location}$ scattered with [the marbles]$_{Material}$.

The sensitivity of the ergative structure to thematic roles could be extremely important for parsing the MV/RR ambiguity. Note that the reduced relative sentences in (4) and (5) begin with exactly the same words as the ergative sentences in (2) and (3).

(4) [The marbles]$_{Material}$ scattered across [the floor]$_{Location}$ made the room look messy.

(5) [The floor]$_{Location}$ scattered with [the marbles]$_{Material}$ made the room look messy.

If the parser encounters a string like *The marbles scattered...*, it must cope with a temporary MV/RR ambiguity between the main verb interpretation in (2) and the reduced relative in (4). However, the situation may be different if the parser encounters the string *The floor scattered...* and attempts to assign a thematic role

to *floor*. The Agent role should be ruled out for both *The floor scattered* and *The marbles scattered* versions because, as in the Trueswell et al. experiments, *floor* and *marbles* are inanimate and thus cannot be Agents. The parser has two roles left, Material and Location. If it had access to plausibility information, then it would assign the Material role to *marbles* and the Location role to *floor*, because marbles make better Materials than Locations for the verb *scatter* (marbles more often scatter across something than have something scattered on them), whereas floors make better Locations than Materials for *scatter*.

These assignments of Material and Location roles could affect the process of choosing the main verb or reduced relative interpretation for the MV/RR ambiguity. If the Material role is assigned to the subject noun phrase, as in *The marbles scattered...*, then the sentence remains ambiguous between an active, ergative construction, as in (2), and a reduced relative clause, as in (4). If the Location thematic role is assigned to the subject noun phrase, however, the sentence cannot have an active ergative structure, as the ungrammaticality of (3) demonstrates; the only option is a reduced relative clause, as in (5). Therefore, if Material and Location thematic role assignments can be made on the basis of plausibility information and can influence the choice of interpretations for an ambiguity, then sentence (5), which contains the Location thematic role in subject position, should be easier to process than (4), which contains the Material thematic role in subject position.

In what way does compatibility of the noun phrase with the Material role differ from compatibility of the noun phrase with the Agent role in Trueswell et al. (in press)? The difference is that the Trueswell et al. study used noun phrases (e.g., *the evidence*) that were inherently bad Agents because they were inanimate — *evidence* can never receive the Agent role from any verb. In contrast, it is not the case that *marbles* is an inherently good Material or a bad Location. For example, *marbles* is assigned the Location role from the spray/load verb *smear* in *Timmy smeared Play-doh on the marbles.* In other words, the compatibility of the assignment of the Material or Location thematic roles emerges from the combination of features of the noun and the verb; there is no dichotomous lexical feature such as animacy to aid

the parser. This is the same situation as in Rayner et al.'s (1983) stimuli like *The performer sent the flowers* sentences, except that in the case of the Rayner et al. stimuli, the thematic assignment cues emerged from knowledge about performers, sending, and flowers, whereas in the *marbles scattered* sentence, only the subject and verb enter into the computation.

Method

Subjects

Forty-one Massachusetts Institute of Technology students were paid for their participation. One subject's data were removed when a post-experiment conversation indicated that the subject had been using a specific strategy rather than reading normally. The subjects in this and all subsequent experiments were native English speakers.

Materials

Ten verbs which permit an ergative reading were chosen from Levin's (1989) list of spray/load verbs. All of the verbs had identical past tense and past participle forms, so that they could trigger a MV/RR ambiguity. For each of these verbs, two pairs of noun phrases were constructed; within each pair, one NP was, in the experimenters' judgment, a good Material for the verb and a poor Location (the "material" noun phrase), and the other was a good Location and a poor Material (the "location" NP). Twenty sentence sets were constructed, each using one NP pair and one spray/load verb that introduced a reduced relative clause. The two noun phrases could appear in either order in the sentence. In the material-first versions, the material NP was the subject of the sentence and the location NP appeared within a prepositional phrase immediately following the spray/load verb; the location-first versions reversed this order. Unambiguous control sentences were constructed by replacing the ambiguous spray-load verb with a verb that had a past participle form distinct from its past tense form, such as *strewn*, *blown*, etc. Each sentence could therefore appear in four versions

Table 1.1: Experiment 1 Example Sentences

Introduction	The children could pick out the spots where
Material-first	
Ambiguous	the green peas scattered in the tiny garden were doing very well.
Unambiguous	the green peas sown in the tiny garden were doing very well.
Location-first	
Ambiguous	the tiny garden scattered with the green peas was doing very well.
Unambiguous	the tiny garden sown with the green peas was doing very well.

resulting from the crossing of two independent variables: ambiguity (ambiguous vs. unambiguous) and thematic role (material-first vs. location-first). A sample sentence set is shown in Table 1.1; all experimental items are contained in Appendix A.1.

Four lists were created, each containing one sentence from each of the twenty experimental sentence sets, along with forty filler sentences. The lists were constructed so that no list contained the same NP or verb in more than one item, and so that each list contained five experimental items in each condition.

A yes/no comprehension question was prepared for each sentence. Questions for the experimental items did not focus on the ambiguity, although most of them did ask about one or both of the Material and Location NPs. Questions mentioning one or the other of the NPs occurred equally often.

Procedure

Subjects read the sentences on a CRT in a Moving Window display (Just, Carpenter & Woolley, 1982) in which they pressed a key to see each word of the sentence. A trial began with a display of the sentence to be read, modified so that all non-space characters were replaced by hyphens. When the subject pressed the space bar on a computer terminal, the first word appeared, replacing the corresponding hyphens. With each successive press of the space bar, the visible word reverted to hyphens and the next word appeared. When the subject read the last word of the sentence

and pressed the space bar, the hyphen display was replaced by a question about the sentence. The subject responded by pressing a key labeled either "Yes" or "No" above the space bar. Subjects did not receive feedback regarding their answers, and the next trial began as soon as the question was answered.

Sentences were double-spaced on the screen, and experimental sentences were displayed so that the head noun prior to the ambiguous verb, the ambiguous verb itself, the following PP, and at least the next three words (the final disambiguation) were all on the same screen line.

Following 5 practice items, the 20 experimental items and 40 fillers were presented in a pseudo-random order, so that at least one filler trial always intervened between presentations of experimental trials. The order of items was different for each subject. Subjects completed the experiment without a break in one 20 minute session.

Results

Only trials on which subjects answered the comprehension question correctly were included in reading time analyses. The data from one item (2b in Appendix A.1) were dropped because subjects answered its comprehension question correctly only 70% of the time while answering the other questions correctly an average of 94% of the time.

Reading Time Analyses

The reading time data for the experimental items were trimmed separately for each word position for each subject. Data points more than three standard deviations from the relevant mean were replaced by the three standard deviation cutoff value. This manipulation affected less than 2% of the data.

The reading time data were grouped into regions for analysis. Region 1 consisted of the Material or Location NP, Region 2 contained only the ambiguous or unambiguous verb, Region 3 contained the prepositional phrase, and Region 4 contained the first two words of the main VP of the clause (the disambiguation). Reading times for these regions can be seen in Figure 1-1.

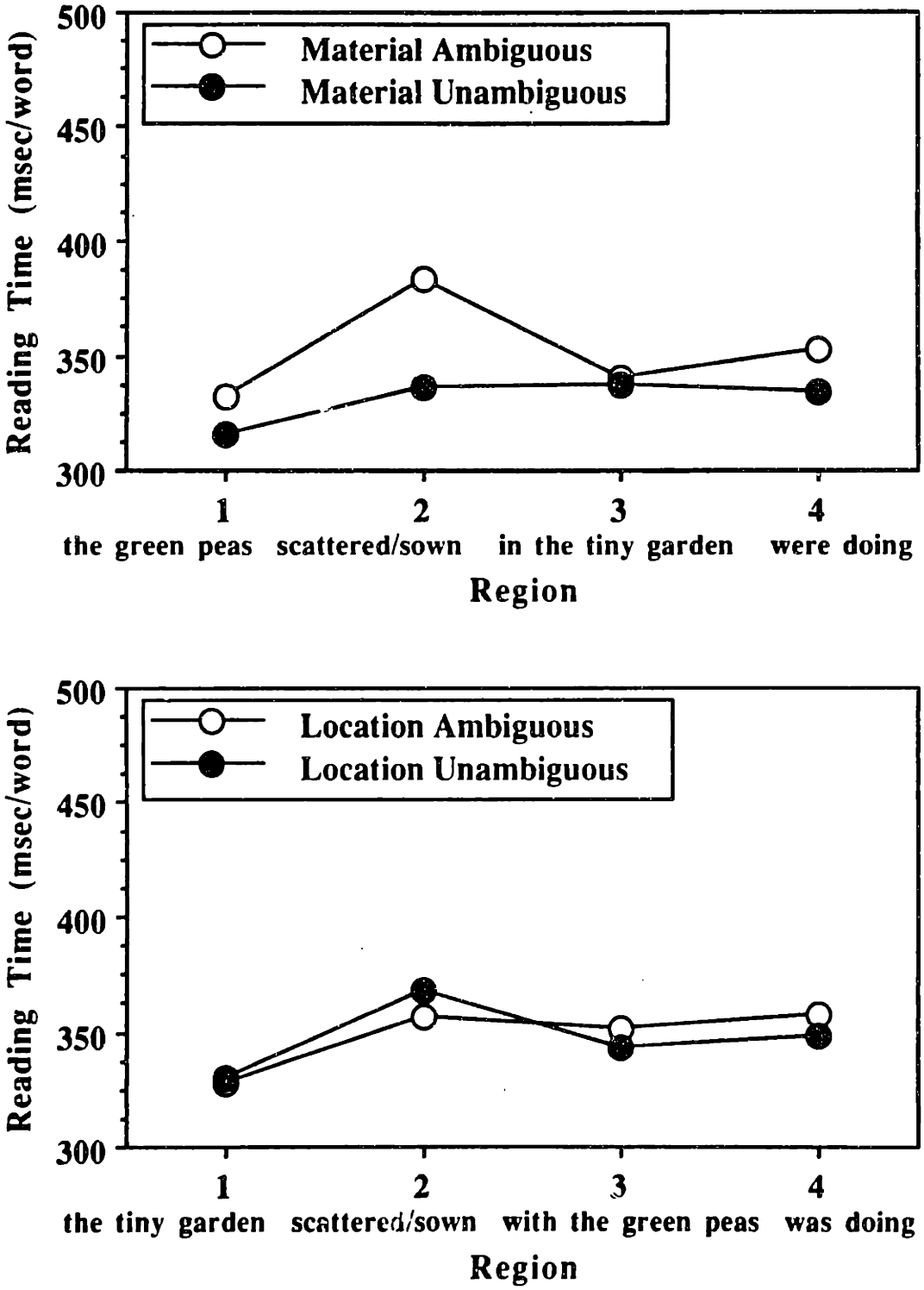


Figure 1-1: Reading time in ambiguous and unambiguous conditions for material-first (top) and location-first (bottom) sentences.

Analyses of the reading times in Figure 1-1 revealed several interesting effects. First, there was no main effect of ambiguity ($F_1 < 1$; $F_2(1,18) = 3.42$, $p < .10$, $MS_e = 1609.45$). This result is surprising because the MV/RR ambiguity is generally found to be extremely difficult, producing much longer reading times in the ambiguous than in the unambiguous conditions. There was, however, a significant Thematic Role $\times$ Ambiguity $\times$ Region interaction, indicating that the thematic role manipulation was modulating an ambiguity effect across regions, ($F_1(3,117) = 4.01$, $p < .01$, $MS_e = 3072.98$; $F_2(3,54) = 4.20$, $p = .01$, $MS_e = 1202.87$). In addition, there was a trend for material-first sentences to be read more quickly than location-first sentences, independent of ambiguity ($F_1(1,39) = 7.00$, $p < .05$, $MS_e = 1585.90$; $F_2(1,18) = 1.16$, $p > .25$). The only other effect approaching significance was that of region ($F_1(3,117) = 2.64$, $p = .053$, $MS_e = 7196.71$; $F_2(3,54) = 4.79$, $p < .01$, $MS_e = 1584.50$).

Separate analyses were conducted at each region to examine these effects. In Region 1, the material or location NP, the Role $\times$ Ambiguity interaction was significant by subjects, though not by items ($F_1(1,39) = 4.79$, $p < .05$, $MS_e = 1008.59$; $F_2(1,18) = 2.30$, $p > .10$). Because the stimulus materials are identical at each level of ambiguity at this region, we must conclude that this marginal result is spurious.

For Region 2, the ambiguous or unambiguous verb, a marginal Role $\times$ Ambiguity interaction was found ($F_1(1,39) = 3.86$, $p < .10$, $MS_e = 9264.56$; $F_2(1,18) = 3.74$, $p < .10$, $MS_e = 4245.21$). This effect was due to the fact that reading time increased with ambiguity in the material-first sentences (384 ms in the ambiguous condition vs. 336 ms in the unambiguous condition, $t_1(39) = 2.26$, $p < .05$; $t_2(18) = 2.24$, $p < .05$), but not in the location-first sentences ($ts < 1$). These and all t-tests reported here are 2-tailed.

The only effect at Region 3 (the prepositional phrase) was a trend for an effect of thematic role, so that the material-first versions were read somewhat more quickly than location-first versions ($F_1(1,39) = 5.59$, $p < .05$, $MS_e = 678.64$; $F_2(1,18) = 3.48$, $p < .10$, $MS_e = 421.98$). This result is difficult to interpret, as different words appear in this region at each level of the role factor. That is, it may be that the

material-first construction is simply easier than the location-first construction at this point in the sentence, or the difference may stem from the characteristics of the different words in Region 3 at each level of role.

The analysis of the disambiguating region (4) revealed marginal effects of thematic role ($F_1(1,39) = 5.00$, $p < .05$, $MS_e = 1260.12$; $F_2(1,18) = 1.74$, $p = .203$, $MS_e = 1146.38$) and ambiguity ($F_1(1,39) = 1.11$, $p > .20$, $MS_e = 2510.82$; $F_2(1,18) = 7.47$, $p < .05$, $MS_e = 484.83$). As in Region 2, ambiguity affected reading times in the material-first condition but not in the location-first condition. In the material-first condition, reading times were slow in the ambiguous condition (352 ms) compared to the unambiguous condition (334 ms), a difference that was significant by items ($t_1(39) = 1.07$, $p > .20$; $t_2(18) = 3.16$, $p < .01$). The effect of ambiguity in the location-first condition (358 ms vs. 348 ms for ambiguous and unambiguous conditions, respectively) was not at all robust ($t_1 < 1$; $t_2(18) = 1.22$, $p > .20$). Analyses of the remainder of the sentence (which varied in length across items) revealed no significant or marginal effects.

Comprehension Questions

Responses to the comprehension questions were uniformly high: 95% and 94.5% in the material-first ambiguous and unambiguous conditions, respectively, and 93.2% and 90% in the location-first ambiguous and unambiguous conditions. There were no reliable differences among these conditions (all ps $> .20$).

Discussion

One notable finding of this experiment is the lack of a strong ambiguity effect, which is surprising in light of previous findings of garden pathing on this construction. We will return to this point in the General Discussion. The second notable finding is that these small ambiguity effects are concentrated entirely in the material-first conditions, which is what is predicted if the parser can use tentative thematic role assignments to assist in interpreting ambiguities. In the case of the location-first items, assignment of the Location role to the first NP rules out an ergative interpretation, so that the

sentence is essentially disambiguated in favor of the reduced relative interpretation well before the main verb of the sentence is reached at Region 4. For the material-first items, however, Material role assignments do not aid in disambiguation, as a Material role assigned to the first NP still admits both an ergative and a reduced relative structure, and the disambiguation arrives only with the main verb at Region 4. The fact that the location-first items do not demonstrate the same processing difficulty that is observed with the material-first items does not support the garden path model, which predicts equal garden pathing on both the material-first and location-first items.

The results of Experiment 1A suggest that the plausibility of alternate thematic role assignments is computed early enough to influence the choice of interpretations pursued for a syntactic ambiguity. Because plausibility is a continuous variable that varies from sentence to sentence, the division of the stimulus set into material-first and location-first conditions is probably only a rough approximation of the plausibility differences across items. It should be the case that the more plausible an NP and spray/load verb are as the introduction of a relative clause, the easier the reduced relative should be to process. In other words, the scenario set up by the first few words of the relative clause should correlate with reading times later in the sentence: The more plausible a relative clause appears to be, the smaller the differences should be between ambiguous and unambiguous reading times for sentences that do contain a relative clause. Such an effect would lend strong support to the view that rapid plausibility computations affect thematic role assignment and ambiguity resolution during parsing. This effect is investigated by measuring plausibility in the stimulus items in off-line norms in Experiment 1B and then correlating these values with reading times from Experiment 1A.

1.3 Experiment 1B

The purpose of this experiment is to assess the relative plausibility of the material and location NPs and the spray/load verbs in relative clause structures. Several different types of norms could be collected. For example, we could present fragments such

as *The green peas scattered...* and ask subjects to complete the sentence (Burgess, 1990), or we could ask subjects to rate the plausibility of unambiguous main verb and relative clause structures containing the NPs and verbs from Experiment 1A. We chose the rating option over the sentence completion option, as the ratings seemed to be a closer reflection of how the language processing system perceived the items, and less a reflection of subjects' completion biases (e.g. to write the shortest possible sentence).

Given a rating methodology, it is still not clear exactly what subjects should rate. For example, it might be the case that the plausibility of a relative clause interpretation should be assessed — the more plausible a relative clause is, the faster subjects read ambiguous sentences that are resolved with a reduced relative clause interpretation. On the other hand, perhaps the main verb interpretation should be assessed, in which case we would expect that the *less* plausible the main verb interpretation is, the faster subjects will read ambiguous sentences that are resolved with a relative clause interpretation. Because of the exploratory nature of these ratings, we conducted ratings to assess plausibility of both main verb and relative clause interpretations, and we conducted a third set of ratings on the unambiguous items to assess their plausibility as well. Because the nouns and verbs from Experiment 1A were the items of interest, we presented only small sentences and sentence fragments to the norming subjects.

Method

Subjects

Eighty Massachusetts Institute of Technology students not included in the reading portions of Experiments 1 and 2 participated as unpaid volunteers.

Materials

The ambiguous items were presented in two different forms, in order to obtain the two different ratings. To obtain a "subject" rating (how plausible an NP was as the subject

of its ambiguous verb in a main verb interpretation), complete ergative sentences were prepared from the NP and verb, as in *The green peas scattered.* In order to rate the plausibility of the NP as object of its ambiguous verb in a relative clause structure (the "object" rating), sentence fragments containing an unreduced relative clause were prepared, as in *The green peas that were scattered across the...* The "unambiguous" ratings also used fragments, but were reduced, because the unambiguous verb alone was enough to disambiguate: *The glass marbles strewn across the...*

Eight twenty-item lists were constructed so that no NP or verb appeared more than once on a single list. Each item from Experiment 1A was therefore rated on only one of the three ratings on each list. Each item appeared with a 7-point scale next to it, with "Good" next to the "1" and "Bad" next to the "7". Each list contained unambiguous rating items and either subject or object rating items, randomly intermixed. Lists were printed with 10 items on each page. The order of the two pages was randomized across subjects.

Procedure

Subjects received instructions to rate the items on two criteria: "whether they are grammatical in English" and "whether they describe something that could happen in real life." The instructions provided examples and explained the rating scale. Subjects were instructed not to rate the fragments as being bad just because they were not complete sentences. After reading the instruction sheet, subjects circled a rating for each item. Each list was rated by 10 subjects.

Results & Discussion

The mean subject, object, and unambiguous ratings for each NP-verb combination are shown for each item in Appendix A.1. The item excluded from Experiment 1A because subjects answered its comprehension question very poorly was also excluded from all of the following analyses.

Initial t-tests comparing the NPs coded as Materials to the NPs coded as Locations in Experiment 1A showed that, as expected, Materials were rated significantly better

than Locations as subjects of ergatives (from the subject rating data) ($t(36) = 9.62$, $p < .001$), but that, in the object rating data, Materials were better objects than the Locations were ($t(36) = 3.90$, $p < .001$). The Materials were also somewhat better objects than the Locations in the unambiguous ratings ($t(36) = 1.614, p < .12$). These results on the object and unambiguous ratings suggest that some of our Location NPs may have been somewhat awkward in the relative clause construction, which might account for the tendency of location-first versions to be read more slowly than material-first versions in Experiment 1A, independent of ambiguity. However, it is unlikely that any overall increased difficulty in the location-first sentences would have been covering up ambiguity effects; if anything, ambiguity effects appear to be exacerbated in more difficult sentences (Stowe, 1990). This question may be addressed directly by examining the correlations between the ratings and reading times, to which we now turn.

Each of the ratings was correlated with the ambiguous and unambiguous reading times from Experiment 1A at each of the regions that were shown in Figure 1-1. A number of different reading-rating correlations were explored, using the subject ratings, the object ratings, the unambiguous ratings, and both subject and object ratings in regressions to predict reading times. Only the object ratings were good predictors of reading times; adding the subject ratings (which were positively correlated with the object ratings, $r = .54$, $p < .001$) to the regression equation had little effect. We therefore report results from only the object ratings in Table 1.2, though all three ratings are included for each item in Appendix A.1. Table 1.2 presents correlations between the object ratings and ambiguous reading times, unambiguous reading times, and the difference between ambiguous and unambiguous reading times.

As can be seen from Table 1.2, the ratings did not reliably correlate with reading times in Region 1 (the material or location NP), Region 2 (the ambiguous verb), or Region 4 (the disambiguation) (all $ps > .40$). However, in Region 3 (the prepositional phrase), the object ratings did predict ambiguous reading times and the difference between ambiguous and unambiguous reading times: The more plausible the initial NP was as an object of the verb, the faster the ambiguous reading times became on

Table 1.2: Correlation of Object Ratings and Reading Times, Experiment 1B

	Region			
Condition	1	2	3	4
Ambiguous	0.01	-0.13	0.33*	0.01
Unambiguous	-0.07	-0.13	-0.12	0.04
Ambig − Unambig Difference	0.06	-0.01	0.36*	-0.03

Note: Ratings were scaled from 1 = good to 7 = bad; positive values therefore indicate that faster reading times or smaller reading time differences correspond to better object ratings.
*$p < .05$.

the next region, the prepositional phrase, so that the difference between unambiguous and ambiguous reading times decreased.

Why would an effect appear at the prepositional phrase in Region 3, and not at the disambiguating main verb in Region 4? One answer to this question can be found by considering the material-first and location-first items separately. If we are correct that assignment of a Location role to the location-first items rules out the ergative interpretation, then in fact the disambiguation occurs much earlier for these items, namely at that point when thematic role assignments have been computed. It is not possible to say exactly how long these assignments take, but it is reasonable to assume that encountering the preposition *with* in the next region would further encourage the Location assignment, as *with* is more compatible with this assignment than the Material assignment (*with* in the Material assignment forces an awkward "accompaniment" reading, as in *The peas scattered with the beans across the garden*). Therefore, subjects should be reading Region 3, the prepositional phrase, as they are disambiguating the Location first items. In contrast, the prepositional phrase in Region 3 is not the point of syntactic disambiguation for the material-first items — even when the Material role has been assigned, the structure is still ambiguous between an ergative and a reduced relative, and the prepositional phrase does not

bear on the ambiguity. Therefore, if the location-first and material-first items are examined separately, it should be the case that the object rating should correlate with reading times in Region 3 for the location-first items, as this is the region being read when the reduced relative is being disambiguated, but it should not correlate with reading times in the material-first items. This is exactly the pattern that was found: Differences between ambiguous and unambiguous reading times decreased as object ratings improved for the location-first items ($r = .51$, $p < .05$), but there was no such effect for the material-first items ($p > .75$). The fact that correlations between the ratings and the reading times appeared within one of the thematic role conditions is quite impressive, considering the conservative nature of this type of procedure (Knight, 1984) and the small number of stimulus items entering into the analysis.

Summary

The object ratings were significantly correlated with reading times in Region 3 for the location-first items, so that items that were rated as good objects produced small ambiguity effects, but poor objects produced large ambiguity effects. This result is not predicted by the garden path model (Frazier, 1987), in which plausibility information is ignored when the initial structure interpretation is built. The garden path model could account for these data only by claiming that the reading time effects and correlations found in Experiments 1A–B were produced by garden pathing in both material-first and location-first conditions, but that there was quick recovery in the location-first condition, aided by the thematic processor's use of plausibility information. This claim is difficult to maintain in light of the complete lack of evidence for any garden pathing in the location-first condition; reading times in the ambiguous sentences do not differ from the unambiguous control sentences at any point. Moreover, the garden path model cannot easily account for the reading time-rating correlation for the location-first items in the prepositional phrase region. The garden path model predicts that the (incorrect) main verb interpretation is being pursued while subjects are reading this region, and that it is only at the disambiguating verb

in the following region that the parser finds its error and backtracks with the aid of the thematic processor. The results are readily interpreted by a model in which plausibility information could influence thematic role assignment, which in turn allowed the early disambiguation of the location-first ambiguous sentences. The better the location NP was rated as an object, the more certain the choice of the reduced relative clause structure (which permits the NP to be an object of the following verb), so that reading times became more similar to those of the unambiguous relative clauses when the NP was a good object.

The object ratings were not successful in predicting ambiguous – unambiguous reading time differences for the material-first items in any region. The ratings may have been less successful for the material-first items because many of these items were rated as being both good objects and good subjects. The ratings thus may not have assessed the relative goodness between the main verb and reduced relative interpretations for this subset of the stimulus set.

The relationship between the object ratings and the ambiguity effect for the location-first items in Experiment 1 is encouraging, but given the small effects of ambiguity overall and the lack of a relationship between the ratings and the material-first items, a conclusion that plausibility affects syntactic ambiguity resolution would be premature. Experiment 2 therefore examines the same issues using a new stimulus set with animate noun phrases that can receive Agent or Theme roles from the ambiguous verb. A second difference from Experiment 1 is that we conducted subject and object ratings on a pool of stimulus items before selecting the items for the reading time study, to guarantee that subject ratings and object ratings were negatively correlated, instead of positively correlated, as they had been for the items in Experiment 1. Experiment 2 should therefore provide a stronger test of the hypothesis that plausibility of thematic role assignments influences ambiguity resolution.

1.4 Experiment 2

Trueswell et al. (in press) found that animacy was an important cue for assignment of an Agent thematic role, so that sentences with inanimate nouns and verbs requiring an Agent (e.g. *The evidence examined...*) did not produce garden paths. However, if plausibility information can also influence the assignment of thematic roles, then for a given verb, certain animate nouns will make better Agents than others. For example, consider the sentences in (6) from Crain and Steedman (1985), who explored a very similar hypothesis.

(6)

a. The teachers taught by the Berlitz method passed the test.

b. The students taught by the Berlitz method passed the test.

Here, *teachers* is a good Agent for the verb *taught*, but *students* is not. This plausibility information is the same sort that was manipulated in Experiment 1 and in Rayner et al.'s (1983) studies; it is not that *students* is inherently a poor Agent, it is merely a poor Agent for this particular verb. If the parser could evaluate the plausibility of thematic role assignments, then the parser should have less difficulty parsing the temporary ambiguity in (6b) compared to (6a). In (6b), *students* is a poor Agent but a good Theme for *teach*, and the assignment of the Theme role to *students* could be accomplished if the sentence had the reduced relative clause structure. By contrast, *teachers* in (6a) makes a good Agent and a poor Theme for *teach*; if *teachers* is assigned the Agent role, then the compatible structure is one in which *taught* is the main verb of the sentence. When the parser later encounters the disambiguating word *passed* in (6), it should easily continue parsing (6b), because it has already posited a reduced relative structure for this sentence, but (6a) should be more difficult, as the main verb structure was initially chosen for this sentence.

Crain and Steedman (1985) report only slight effects in a grammaticality judgment experiment that manipulated plausibility. However, an examination of their stimuli reveals a heterogeneous set of items: Some have the plausibility cues in the subject noun phrase and the verb, as in (6) above, some use the animacy cue, as in Trueswell

et al. (in press), and some have no plausibility information available at the point of the ambiguous verb — the information comes only later, as in the Rayner et al. (1983) stimuli. As a result, it is not clear from the Crain and Steedman data whether plausibility information can influence assignment of Agent and Theme thematic roles. Experiment 2 is designed to answer this question. As in the previous experiment, we conducted a reading time study and also correlated reading times with norming data on the stimulus items.

Method

Ratings

Preliminary ratings similar to those in Experiment 1B were conducted on a large pool of NP-verb combinations in which the NP was judged by the experimenters to be strongly biased toward receiving either an Agent or a Theme thematic role. We obtained ratings of the NP as both a subject (Agent role assigned, main verb reading) and as an object (Theme role assigned, relative clause reading) of a particular verb. We also obtained a comparative preference rating between these two readings.

Because raters were directly comparing the two readings, both readings were conveyed with sentence fragments, as in *The doctor treated the...* and *The doctor who was treated by the...*, for the subject and object rating respectively. As in Experiment 1B, the fragments were rated individually on a scale from 1=good to 7=bad. The comparative rating was also performed on a 7-point scale, with 1 corresponding to a very strong subject preference and 7 a very strong object preference.

Four lists were constructed, each with 30 items and 18 similar items from an unrelated experiment. An item consisted of a subject rating fragment on the left side of the page, the corresponding object fragment on the right, the 7-point comparative scale in between them, and a blank under each version in which the individual rating (1–7) was written. Eight items were placed on each page, and the order of the six pages was randomized across subjects.

Forty-five Massachusetts Institute of Technology students participated as unpaid

volunteers (11 in 3 lists, 12 in the fourth). The rating criteria were described as in Experiment 1B, and subjects were instructed to complete the comparative rating first for each pair of fragments, then the individual ratings.

Materials

Using the pool of rated fragments, 24 stimulus items were constructed. Each item was constructed from two pairs of rated fragments that had contained the same ambiguous verb. The NP in one fragment pair had received a very low comparative rating (such that the NP was strongly preferred as an Agent, e.g., *The hunter captured the...* was preferred to *The hunter that was captured by the...*), and the NP in the other pair had received a very high comparative rating (such that this NP was strongly preferred as a Theme, e.g. *The prisoner that was captured by the...* was preferred over *The prisoner captured the...*). Introductory phrases and sentence completions were constructed for each item. Unambiguous control sentences were prepared, so that each stimulus item could appear in four versions by crossing two independent variables: Ambiguity (ambiguous vs. unambiguous) and Thematic Role Preference (Agent vs. Theme). A sample sentence set is shown in Table 1.3; all items and their ratings are contained in Appendix A.2.

The unambiguous items in this experiment contained unreduced relative clauses created by the insertion of *that was* or *that were* before the ambiguous verb. This disambiguation was employed because very few English verbs with distinct past participles can be used with animate subjects. Trueswell et al. (in press) pointed out that the addition of these extra words before the ambiguity could create differences between the ambiguous and unambiguous conditions that are not due to ambiguity. They suggested that the addition of the extra words might help the parser prepare for a relative clause, which may add to the complexity of a sentence, independent of any ambiguity. Two additional *word control* conditions were included to examine the effect of adding the two words (*that was/were*) to the unambiguous conditions. Both of these conditions contained relative clauses with unambiguous past participles like *strewn* and *blown*. In the unreduced unambiguous condition, the unambiguous verb

Table 1.3: Experiment 2 Example Sentences

Introduction	Major Brown quietly told the shocked debriefing officers about how
Agent-NP Preference	
Ambiguous	the hunter captured by the rebels begged for freedom for weeks with no success.
Unambiguous	the hunter that was captured by the rebels begged for freedom for weeks with no success.
Theme-NP Preference	
Ambiguous	the prisoner captured by the rebels begged for freedom for weeks with no success.
Unambiguous	the prisoner that was captured by the rebels begged for freedom for weeks with no success.
Word Control Items	
Unreduced	The children could pick out the spots where the green peas that were sown in the tiny garden were doing very well.
Reduced	The children could pick out the spots where the green peas sown in the tiny garden were doing very well.

was preceded by *that was/were*, as in *The seeds that were strewn...* In the reduced unambiguous condition, the relative clause was introduced only by the verb. A sample pair of sentences is at the bottom of Table 1.3. Ten material-first unambiguous items from Experiment 1 were used for the word control conditions, five each in the reduced and unreduced unambiguous conditions, counterbalanced across subjects. These 10 items are noted in Appendix A.1.

Fifty-one filler items with a variety of syntactic structures were prepared, as were yes/no comprehension questions for all items. In the case of the experimental items, the comprehension question always asked about whether the NP in the prepositional phrase (the by-phrase) of the relative clause had or had not performed the action described in the disambiguation of the sentence. On the relative clause reading of these sentences, the NP in the by-phrase does not perform the action, but if a sentence is misread with the main verb interpretation, then the by-phrase NP would seem to be the most likely Agent for the action in the disambiguation. Thus an incorrect answer suggests a misparse in which the main verb interpretation was adopted. The correct answer was "yes" for half of these questions.

Subjects

Fifty-two MIT students were paid for their participation in the reading experiment. An additional four subjects were tested but were removed from the data set because of comprehension question error rates of more than 30% for the experimental items. Due to a programming error, three experimental items were presented in an incorrect condition for 13 subjects. Three of these subjects were randomly selected for replacement with three additional subjects, who saw the corrected stimulus list. This procedure was employed so that those three experimental items would be seen in all conditions across subjects and could thereby be included in item analyses.

Procedure

The procedure was identical to that of Experiment 1A. Subjects saw five practice items, followed by the 24 experimental items, the 10 word control items, and 51

fillers. The order of items was pseudo-randomized so that at least one filler preceded any experimental or word control item.

Experimental items were positioned on the screen so that the Agent or Theme NP, the entire relative clause (reduced or unreduced), and the following four words were all on the same screen line. The word control sentences were positioned so that the head noun prior to the relative clause, the relative clause itself, and at least the next three words were all on the same line.

Results & Discussion

The reading time and comprehension question data will be discussed separately for the experimental and word control items. The analyses reported here for the experimental items were also conducted without the three replacement subjects. The power of the items analyses was reduced with the removal of these three subjects (because only 21 items were seen in all conditions without inclusion of the replacement subjects), but the nature of the effects were unchanged.

The reading time data were trimmed as in the previous experiment, affecting less than 3% of the data. The experimental and word control items were trimmed separately.

Reading Time on Experimental Items

The experimental sentences were broken into four regions for the reading time analysis, as shown in Figure 1-2. Region 1 consisted of the Agent or Theme NP. Region 2 contained the ambiguous verb, as well as *that was/were* in the unambiguous conditions. Region 3 consisted of the by-phrase, and Region 4 contained the first two words of the main verb phrase (the disambiguation). The remainder of the sentence was separated into two additional regions which were analyzed separately. Region 4 is considered the disambiguation because the by-phrase in Region 3 does not necessarily disambiguate the sentence. The preposition *by* is ambiguous between an agentive phrase (which would disambiguate) and a locative (e.g. *by the sea*, which would not disambiguate if the preceding verb had an intransitive interpretation, as some Ex-

periment 2 stimuli do have; see MacDonald, in press; Pritchett, 1989). Only the last word of Region 3 suggests the correct interpretation, and a locative interpretation is never entirely ruled out.

As can be seen in Figure 1-2, there was a marked effect of thematic role on the extent to which ambiguity increased reading times over the unambiguous conditions. An initial omnibus ANOVA revealed a significant Thematic Role Preference × Ambiguity × Region interaction ($F_1(3,153) = 5.71$, $p = .001$, $MS_e = 5402.07$, $F_2(3,69) = 2.87$, $p < .05$, $MS_e = 3757.33$). The Role × Region interaction was significant ($F_1(3,153) = 13.96$, $p < .001$, $MS_e = 4092.99$; $F_2(3,69) = 13.63$, $p < .001$, $MS_e = 2242.69$), as was the Ambiguity × Region interaction ($F_1(3,153) = 2.92$, $p < .05$, $MS_e = 6865.93$; $F_2(3,69) = 2.72$, $p = .051$, $MS_e = 2205.69$). There was also a marginal main effect of role, such that sentences with the Theme NP produced faster reading times than sentences with the Agent NP ($F_1(1,51) = 3.28$, $p < .10$, $MS_e = 7704.80$; $F_2(1,23) < 1$), and a main effect of region ($F_1(3,153) = 20.70$, $p < .001$, $MS_e = 8372.13$; $F_2(3,69) = 12.09$, $p < .001$, $MS_e = 7274.48$).

As expected, there were no significant effects in Region 1, before the ambiguity was introduced. In Region 2, however, analyses indicated that Theme sentences tended to be read more slowly than Agent sentences ($F_1(1,51) = 2.79$, $p = .10$, $MS_e = 7370.09$; $F_2(1,23) = 7.05$, $p < .05$, $MS_e = 3532.56$). A marginal Role × Ambiguity interaction was also present by subjects ($F_1(1,51) = 3.63$, $p < .10$, $MS_e = 11845.67$) but not by items ($F_2(1,23) = 1.83$, $p > .15$). Reading times were identical in the two unambiguous conditions for this region (399 ms and 401 ms for the Agent and Theme conditions, respectively). By comparison, Agent Ambiguous reading times were relatively fast (380 ms), while Theme Ambiguous times were relatively slow (432 ms), a significant difference ($t_1(51) = -2.58$, $p < .05$; $t_2(23) = -3.14$, $p < .005$). This result suggests that subjects appear to be choosing a structure at this early point: When subjects had plausibility information that the NP was a good Agent for an ambiguous verb, suggesting a simple main verb interpretation, they read this verb quickly, but when they had information that the NP was a good Theme for the verb, implying a reduced relative clause analysis, they read the verb more slowly. An

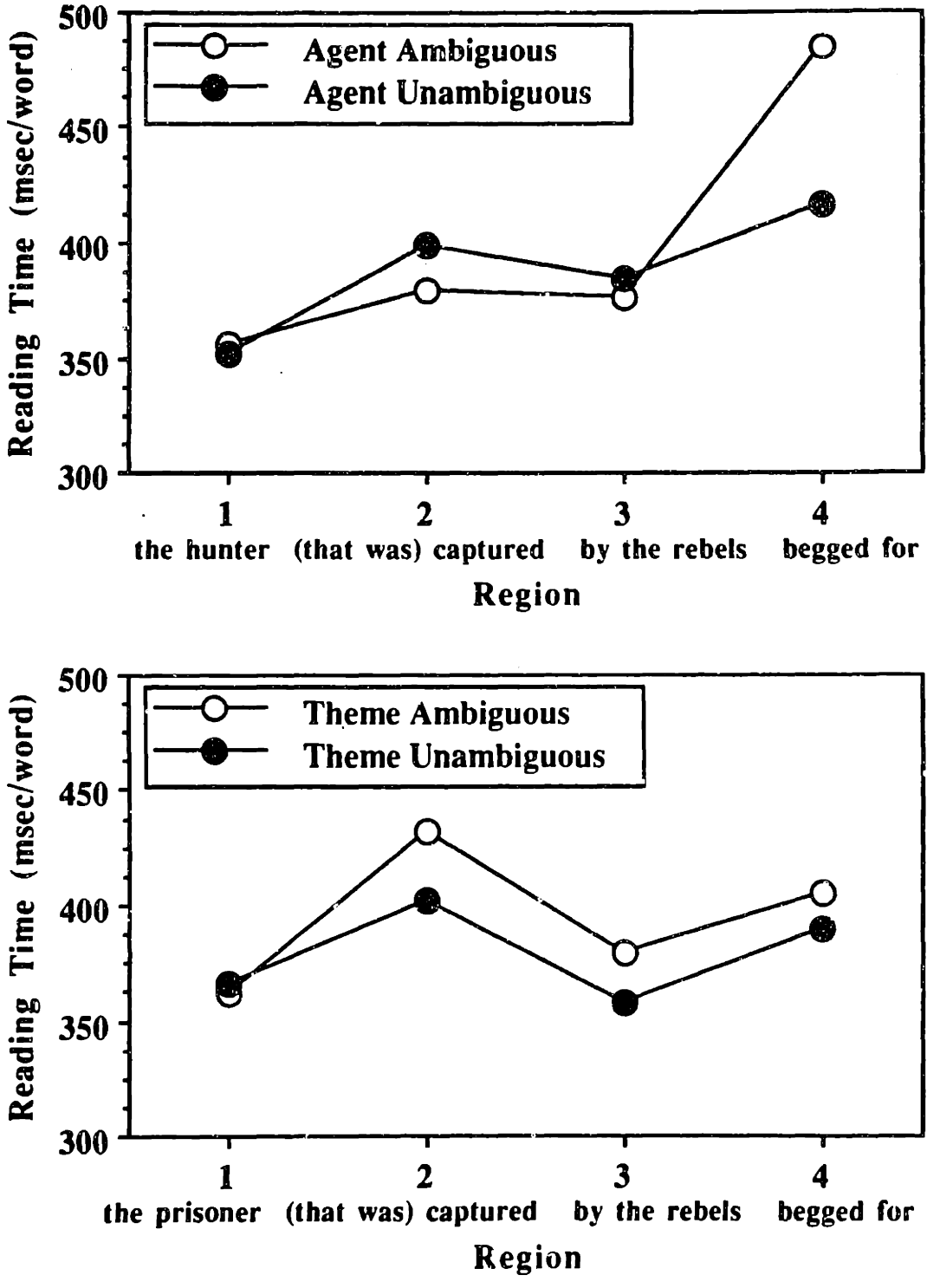


Figure 1-2: Reading time in ambiguous and unambiguous conditions for Agent-NP Preference (top) and Theme-NP Preference (bottom) sentences.

alternative interpretation that would be favored by the garden path model is that the parser is choosing a main verb interpretation in both cases and is slowing down in the Theme Ambiguous condition because of the implausibility of the NP being an Agent in this structure. This interpretation would predict that there should be some effect on the unambiguous sentences at this point as well, that subjects would be noting the implausibility of a good Agent receiving a Theme role in the unambiguous conditions and would have slow reading times in the Agent Unambiguous condition compared to the Theme Unambiguous condition. The finding of identical reading times in the unambiguous conditions does not support this alternative interpretation.

Analyses in Region 3, which contained the by-phrase, revealed a Role × Ambiguity interaction ($F_1(1,51) = 4.28$, $p < .05$, $MS_e = 9650.09$; $F_2(1,23) = 4.15$, $p = .053$, $MS_e = 5909.89$). The source of this interaction was that the Theme Unambiguous condition (358 ms) was at least marginally faster than the other three conditions: Theme Ambiguous (379 ms) ($t_1(51) = 1.79$, $p < .10$; $t_2(23) = 2.00$, $p < .10$), Agent Unambiguous (384 ms) ($t_1(51) = 2.81$, $p < .01$; $t_2(23) = 1.70$, $p < .15$), and Agent Ambiguous (377 ms) ($t_1(51) = 2.76$, $p < .01$; $t_2 < 1$), but these latter three conditions did not differ from one another (all ps $> .30$). There was also a marginal main effect of thematic role, so that Agent sentences were read more slowly than Theme sentences, ($F_1(1,51) = 4.80$, $p < .05$, $MS_e = 8305.15$; $F_2 < 1$).

The analysis for Region 4, which contained the syntactic disambiguation, also revealed a Role × Ambiguity interaction ($F_1(1,51) = 4.68$, $p < .05$, $MS_e = 16481.39$; $F_2(1,23) = 2.22$, $p = .15$, $MS_e = 7017.01$). Further analysis revealed that in the Agent conditions, reading times in Region 4 were significantly slower in the ambiguous sentences (485 ms) than in the unambiguous controls (416 ms) ($t_1(51) = 2.81$, $p < .01$; $t_2(23) = 2.58$, $p < .05$), but there was no effect of ambiguity in the Theme conditions (405 vs. 390 ms in ambiguous and unambiguous conditions, ts < 1). In other words, subjects were not garden pathed in the presence of plausibility information favoring a Theme assignment for the first NP, but they were garden pathed when an Agent role was favored. In addition, Theme Unambiguous reading times were marginally faster than Agent Unambiguous reading times ($t_1(51) = 1.90$, $p < .10$; $t_2(23) = 2.10$,

$p < .05$), indicating an effect of plausibility even in the absence of ambiguity.

To determine whether any effects of ambiguity or role continued after the disambiguation, additional analyses were conducted on the final six words in the sentence. An ANOVA on the next five words of the sentence showed only a slight trend for Agent sentences to be slower than Theme sentences (375 vs. 363 ms) ($F_1(1,51) = 3.96$, $p < .10$, $MS_e = 14201.01$; $F_2(1,23) = 1.97$, $p > .15$, $MS_e = 8880.11$). A similar pattern appeared at the final word: Agent sentences (554 ms) were significantly slower than Theme sentences (503 ms) ($F_1(1,51) = 5.92$, $p < .05$, $MS_e = 31230.80$; $F_2(1,23) = 4.92$, $p < .05$, $MS_e = 11876.09$). Neither the effect of ambiguity nor its interaction with role approached significance in either region (all $ps > .30$).

In sum, the reading time data show a clear difference between the ambiguity effects for the Agent and Theme conditions. The Agent conditions demonstrate the classic garden path effect: Reading times at the ambiguous verb were faster than in the unambiguous condition, but they increased over unambiguous times at the disambiguation. It appears that subjects parsed these sentences exactly as the garden path model would predict, initially interpreting the ambiguous verb as the past tense main verb of the sentence, and then showing surprise at the disambiguation. The garden path model predicts the same pattern in the Theme conditions, but the data do not support this prediction. Reading times for the ambiguous verb in the Theme Ambiguous condition are slower than in the unambiguous condition, suggesting that subjects are already considering a relative clause interpretation at this point, and they show no significant increase in reading times over the unambiguous condition at the disambiguation. This reading pattern strongly suggests that subjects are using the plausibility information to choose a relative clause interpretation of the ambiguity well before they reach the disambiguation. An alternative interpretation of the slowed reading at the ambiguous verb, that subjects are using Minimal Attachment to choose the main verb interpretation but are slowed by the implausibility, is not supported by the pattern of reading times across regions. This hypothesis would predict that subjects would also show large increases in reading time at the disambiguating region, when the parser gets information that the main verb interpretation

is incorrect. The data do not support this view; subjects show no surprise at the syntactic disambiguation in the Theme Ambiguous condition.

Ratings

The reading experiment demonstrated that the sentences in the Agent conditions did not show the same ambiguity effects as those in the Theme conditions; the ratings will investigate whether the plausibility constraints operate along a continuum. Ambiguous and unambiguous reading times and ambiguous − unambiguous difference scores in the four regions were correlated with the subject and object ratings we obtained in our initial norming procedures, and the combination of the subject and object ratings were evaluated as predictors of reading times in regression analyses. As in Experiment 1B, the object ratings again produced the clearest pattern; the subject ratings produced consistent but weaker results. Therefore, only the object rating correlations are reported here unless otherwise noted.

Because the stimulus items were selected on the basis of the comparative ratings (whether subjects thought the Agent or Theme role for a noun was more plausible, given a particular verb), it is important to ask whether the distribution of the object ratings is a perfect reflection of the Agent/Theme dichotomy in the reading experiment. If so, then examining the correlations between reading time and the ratings would add little new information, because the ratings would simply recreate the assignment to conditions. An examination of the object ratings listed with each item in Appendix A.2 indicates that this is not the case, however. There was a small range of object ratings in the Theme condition (1.07–2.27) and a wide range of object rating values within the Agent condition (1.95–5.27), so that the distributions of ratings in the two conditions actually overlap. Given this range of ratings, correlations between object ratings and reading times should be informative. Moreover, it may also be useful to examine the correlations within the Agent and Theme items separately, though it should be recognized that only the very strongest relationships could be identified, because the range of ratings and statistical power are reduced when only half of the items enter into the analysis.

Table 1.4: Correlation of Object Ratings and Reading Times, Experiment 2

	Region			
Condition	1	2	3	4
Ambiguous	-0.10	-0.27*	-0.05	0.49***
Unambiguous	-0.22	0.07	0.34**	0.28**
Ambig − Unambig Difference	0.07	-0.23	-0.29**	0.26

Note: Ratings were scaled from 1 = good to 7 = bad; positive values therefore indicate that faster reading times or smaller reading time differences correspond to better object ratings.
*$p < .10$. **$p < .05$. ***$p = .0005$.

The object rating correlation results for the entire stimulus set are shown in Table 1.4. Correlations for Region 1 (the Agent or Theme NP) were all non-significant (all $ps > .10$), but in Region 2 (the ambiguous verb), the negative correlation with ambiguous reading times was marginally significant, indicating that ambiguous verbs following good object NPs tended to be read more slowly than when the NP was a poor object. In other words, subjects slowed their reading of the ambiguous verb when plausibility information gave the processor some reason to expect a relative clause. As mentioned above, finding this effect in this region is important, as it indicates that the plausibility of alternate interpretations is affecting the *initial* interpretation of the sentence. If plausibility influenced only backtracking operations, as the garden path model suggests, then there would be no such effect on reading times in this early region. Importantly, this correlation was also present even when the Theme stimulus items were examined individually ($r = -.36$, $p < .10$), though the effect was not robust for the Agent items ($p > .40$).

In Region 3 (the by-phrase), ambiguous reading times were unrelated to the ratings, but unambiguous reading times decreased as the quality of the initial NP as an object increased. The correlation between the ratings and the ambiguous − unambiguous reading time differences reflected this combination of effects. These results

show that within a few words after encountering the Region 2 verb, the language processor is sensitive to the anomaly of assigning a Theme role to a good Agent in a structurally unambiguous relative clause. These effects on unambiguous reading times were also present, though smaller, within the Agent items ($r = .26$, $p < .25$) and within the Theme items ($r = .38$, $p < .10$).

In Region 4 (the disambiguation), both the ambiguous and unambiguous reading times correlated well with the object ratings, though the effect was much stronger for the ambiguous items. This result suggests that plausibility accounts for some effects in all sentences regardless of ambiguity, but that plausibility had an additional effect on the ambiguous items. The reading time difference correlation indicates that as the initial NP becomes a better potential object, the ambiguity effect decreases at the point of disambiguation: Subjects are surprised to find the reduced relative structure only when plausibility information favors the main verb structure, and not when it promotes a relative clause structure. When the Agent and Theme items were considered separately in this region, the correlation between ratings and ambiguous reading times was present for the Agent items ($r = .036$, $p < .05$) but was not robust for the Theme items ($p > .40$). The correlation between the ratings and ambiguous − unambiguous difference scores was not robust for either group of materials when only the object ratings were considered ($ps > .20$). However, when the subject ratings and object ratings were combined in a multiple regression, the combined ratings were a good predictor of reading times for the Agent items in the ambiguous condition ($r = .48$, $p < .10$), in the unambiguous condition ($r = .45$, $p < .10$), and most importantly, the combined subject and object ratings strongly correlated with the ambiguous − unambiguous difference score ($r = .60$, $p = .01$).

In sum, the correlational analyses confirmed and extended the major effects in the reading experiment: As the Region 1 NPs became poorer objects, reading time on the ambiguous verb in Region 2 became faster, and reading time in the disambiguating Region 4 became slower, indicating that subjects were garden pathed at that point. As the NP became a good object, however, subjects were slow to read the ambiguous verb in Region 2 and showed no slowing in the disambiguating region. As mentioned above

in the reading time results, it is difficult to ascribe these effects to backtracking and revision following garden pathing, as the garden path model would predict. Instead, it appears that the early plausibility information supporting a relative clause encouraged comprehenders to adopt that structure well before the definitive disambiguation was reached. Moreover, the fact that meaningful correlations can be found within the Agent and Theme items considered separately demonstrates that these effects were not due to a few items at the extremes of the ratings scale. Instead, these results reaffirm the graded nature of the plausibility cues investigated here and emphasize the value of combining factorial and correlational methods in ambiguity research.

Reading Time on Word Control Items

These items were broken into four regions for the reading time analyses, as shown in Figure 1-3. The effect here is similar to that in Trueswell et al. (in press), with reduced unambiguous items being read more slowly than their unreduced counterparts ($F_1(1,51) = 4.99$, $p < .05$, $MS_e = 3878.38$; $F_2(1,9) = 5.28$, $p < .05$, $MS_e = 731.52$), but it is not clear enough to be localizable to a particular region: There was no Region $\times$ Reduction interaction (Fs < 1). The main effect of Region was also marginal ($F_1(3,153) = 4.92$, $p < .005$, $MS_e = 4277.64$; $F_2(3,27) = 2.59$, $p < .10$, $MS_e = 1316.48$). Because the effects appear to be small and inconsistent across experiments, the addition of extra words in the unambiguous experimental conditions probably can be safely ignored here.

Comprehension Questions

The comprehension question accuracy for the four experimental conditions ranged from 86.1% to 87%, and produced no significant effects (Fs < 1). For the word control items, accuracy was 95.4% and 96.9% for the reduced and unreduced unambiguous items, respectively. This difference was also nonsignificant (Fs < 1). As discussed above, the questions for the experimental items were designed to be somewhat harder than those in the previous experiment (and in the word control items, which were drawn from the previous experiment), and the reduced accuracy reflects

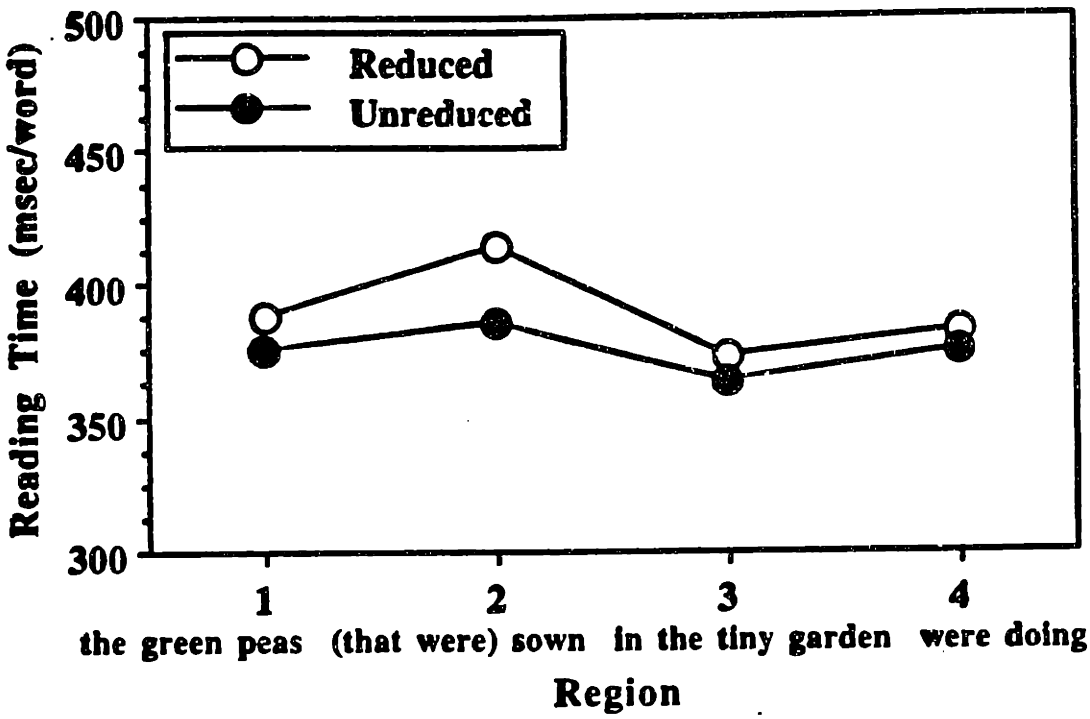


Figure 1-3: Reading time for reduced and unreduced unambiguous word control items.

this manipulation.

1.5 General Discussion

Both Experiments 1 and 2 have provided evidence that plausibility information can help the parser choose an initial interpretation for main verb/reduced relative clause ambiguities. There was no ambiguity effect in either experiment in the two conditions (location-first and Theme conditions) that promoted a reduced relative clause interpretation, but there was an ambiguity effect in both experiments for those conditions in which the plausibility cues encouraged a main verb interpretation (material-first and Agent conditions). In addition, reading times prior to and including the disambiguation region in both experiments correlated with the plausibility ratings collected in off-line norms. These results argue strongly for an important role for the effects of plausibility on thematic role assignment and syntactic ambiguity resolution. Dichotomous features like noun animacy are not the only influences on thematic role assignment and ambiguity resolution; much more subtle plausibility manipulations

across combinations of lexical items also influence the initial parse adopted for an ambiguity.

Although both experiments produced an ambiguity effect in the absence of helpful plausibility information, the effects were much stronger in Experiment 2's Agent sentences than in Experiment 1's material-first items. This difference in the magnitude of ambiguity effects in the two "difficult" conditions may be linked to the animacy of the subject NPs in the two experiments. Apparently when the subject NP was inanimate in Experiment 1, the reduced relative interpretation was readily available so that the ambiguity effect was surprisingly small in the Material-first condition. In Experiment 2, the animate NPs appeared to work against the relative clause interpretation, so that there was a large ambiguity effect in the Agent condition. This pattern suggests that both animacy and plausibility information are available to the processor as it selects the initial interpretation of an ambiguity. This hypothesis, which is based on comparisons across two different experiments, must of course remain tentative; a study that manipulates both kinds of information is needed.

The results of the present experiments contrast with the Rayner et al. (1983) results that suggested that plausibility information could not help to determine the course of first-pass parsing. The differences across studies suggest that sufficiently strong plausibility cues that are present at the onset of an ambiguity can aid in initial syntactic parsing, as in the present experiments, but that the weaker and late-occurring plausibility cues in the Rayner et al. study could not redirect the course of processing once initial decisions had been made, so that subjects were still garden pathed in those sentences when they arrived at the disambiguation.

Our results further suggest that "sufficiently strong" cues do not aid parsing in an all-or-nothing manner, but rather that the ease with which the parse is handled is directly related to the strength of the plausibility cue. This conclusion implies that a model of human syntactic processing which relies solely on general purpose strategies such as Minimal Attachment in the initial parsing of structural ambiguities is not correct, as there is no way to accommodate the graded effects of the plausibility cues within these general purpose strategies. Our results are instead consistent

with the class of models discussed by Tanenhaus and Carlson (1989), in which the syntactic parser builds its syntactic interpretation based on the best-supported of the available verb argument structures. Trueswell et al. (in press) showed that the "best-supported" argument structure could be at least in part determined by relevant lexical semantic information such as noun animacy; the current experiments indicate that plausibility information arising from the combination of noun and verb meanings can also help to make this determination. Our results are also compatible with Altmann and Steedman's (1988) "weakly interactive" model of parsing in which the parser proposes multiple structural analyses in parallel, and pragmatic factors determine which interpretation will be pursued by the parser. What is not yet clear from any of this research is the level of representation that is mediating the plausibility effects: it may be from information extracted from a discourse representation of the sentence, incorporating real world knowledge, or it may be that activation of a rich lexical/semantic representation of *prisoners* and *captured* is sufficient to guide thematic role assignments. Continued investigation of this issue will have important implications for both theories of syntactic parsing and for theories of the lexicon.

Modularity and Non-Syntactic Effects on Parsing

An important claim that is frequently made in comparisons of alternative models of ambiguity resolution is that the Frazier garden path model is modular, whereas models in which the initial parse is influenced by plausibility information (whether from lexical, discourse, or real world knowledge) are not modular (Ferreira & Clifton, 1986; Frazier, 1987; but see Altmann & Steedman, 1988, and Clifton & Ferreira, 1989, for an alternative view). If so, it would seem that the evidence provided here for plausibility effects on syntactic ambiguity resolution would argue against a modular architecture in the human language processor. Before adopting this conclusion, however, it is necessary to explore exactly what constitutes modularity in the language processor, as there are several different definitions that might be adopted.

Fodor (1983) has argued that the most important criterion for modularity is *informational encapsulation*, the isolation of a lower level processing system from "central

processes" and knowledge, including the plausibility information we have been considering. Fodor offers two cogent arguments for why modules ought to be informationally encapsulated. First, any subsystem that has access to all the information and beliefs of an organism cannot possibly function rapidly, as modules do, because the subsystem would have to determine for each bit of prior knowledge whether or not it is relevant to the current processing; and second, if beliefs and/or expectations are allowed to affect the processing of new input, then an organism will not be able to correctly handle unexpected or surprising stimuli.

The specific claim about modularity in the garden path model of parsing is that the syntactic parser is a module within the language processing system, in that it constructs a syntactic phrase structure representation before any non-syntactic information is consulted (Frazier, 1987; Rayner et al., 1983). The garden path model is not the only one that can claim modularity on these criteria, however. In the Tanenhaus and Carlson model, the syntactic parser receives input that has been affected by thematic role information, but operations of the syntactic processor on its input are themselves unaffected. Syntactic operations are thus just as fast and automatic as those of the parser in the garden path model. Similarly, Altmann and Steedman (1988) have argued that rapid use of pragmatic information does not prevent the parser from being modular. They suggest a "weakly interactive" model in which alternative parses are generated in parallel by the parser and then filtered by discourse/pragmatic information (see Clifton & Ferreira, 1989; Steedman & Altmann, 1989; for further discussion). The syntactic component in all these models may thus be equivalently modular on these criteria.

Another level of modularity that is important to this discussion is the modularity of the language processor as a whole, independent of whether the syntactic component is itself a sub-module. At this level, it seems that *no* current model of parsing may be considered modular if the strict criteria of informational encapsulation are maintained. Any model, such as the Tanenhaus and Carlson model or the Altmann and Steedman model, that allows use of plausibility and pragmatic information violates informational encapsulation. Frazier's garden path model fares no better: The

syntactic component is modular, but another important component, the thematic processor, does not appear to be. The thematic processor operates in parallel with the syntactic processor and makes use of plausible thematic role assignments and discourse information to suggest alternative analyses to the syntactic parser (Clifton & Ferreira, 1989; Ferreira & Clifton, 1986; Frazier, 1987; Rayner et al., 1983). If the thematic processor or any component that uses real world knowledge is placed within the language processor, then informational encapsulation is violated. Removing the thematic and semantic/discourse components from the language processing module would keep it shielded from "central processes," but this move does not seem to be an optimal solution, for it violates our intuitions that thematic role assignments are part of the operations of the language processor and are no less automatic than assignment of phrase structure.

Considering the complexity of the lexical/syntactic/discourse interactions postulated in current models, and considering the number of architectural assumptions underlying each model, it thus seems that the dichotomy between "modular" and "non-modular" is not particularly useful. In contrast, the important questions raised by the modularity debate remain. These questions concern the nature of the influence of non-syntactic information on the syntactic parser. Fodor claims that the two main reasons for encapsulating modules are to ensure that they are fast and that they recognize unexpected input, but, at least for the semantic and discourse information that we and others (Altmann & Steedman, 1988; Crain & Steedman, 1985; Taraban & McClelland, 1988; Trueswell et al., in press) have considered, the empirical evidence suggests that the use of such information does not have these harmful effects. Relevant semantic and discourse information has been shown to affect processing times almost immediately at the point of its access, so a module accessing such information does not appear to suffer in terms of speed. Moreover, there is no evidence that the contextual information can actually override the decisions of the syntactic component, so that the parser still would recognize unexpected input. None of these models allows non-syntactic information to do more than constrain the range of syntactic options, none propose that non-syntactic information can actually force the parser

to build any ungrammatical structure. Indeed, Altmann and Steedman (1988) have argued that any model in which discourse or pragmatic factors do any more than constrain the syntax is unworkable.

This "filtering" role for contextual information has of course always been a claim of the garden path model (Frazier, 1987), and the currently available data do not challenge this view. Altmann and Steedman's (1988) "weakly interactive" model and Tanenhaus and Carlson's (1989) thematic role assignment model also claim that context filters syntactically permissible alternatives. The differences lie in the proposed mechanisms of the filtering: The garden path model has the parser adopt only the syntactically simplest parse on the first pass, and the thematic processor later proposes alternatives; in the more interactive models, plausibility information biases the choice of syntactic alternatives but could not force the adoption of an ungrammatical structure. Given this position, that plausibility influences are present but restricted, a major challenge for psycholinguistic research remains the investigation of the limits of these influences. Focusing on the graded nature of the cues, as in the correlation studies here, should help to test these limits. However, given that there do appear to be limits on the extent of contextual influences on parsing, considering cue strength alone, without taking into account different levels of linguistic representation and process, is unlikely to account for all of the complex interactions between different types of information during comprehension.

Chapter 2

Lexical Mechanisms of Syntactic Ambiguity Resolution[1]

2.1 Introduction

The Ambiguity Resolution Problem

One of the principal goals of a theory of language comprehension is to explain how the reader or listener copes with a pervasive ambiguity problem. Languages are structured at multiple levels simultaneously, including the lexical, phonological, morphological, and syntactic levels, as well as the discourse or message level. At any given point in a sentence, the information that is available can be consistent with multiple alternatives at each level. To take a simple example, the word *watch* is ambiguous between alternative meanings ("a time-piece," "to observe"). It is therefore also ambiguous with regard to grammatical category (whether it is a noun or verb). The verb sense of *watch* creates further ambiguity because it can participate in several different syntactic structures, including transitive (*John watched Mary.*) and intransitive (*John watched intently.*). Comprehending a sentence involves resolving several types of ambiguities so as to converge on a single interpretation (usually the one intended by the speaker or writer).

[1]This work was conducted in collaboration with Maryellen MacDonald and Mark Seidenberg.

The complexity of the ambiguity resolution problem is greatly increased by a fundamental fact about the nature of linguistic stimuli: The information in a sentence becomes available over time. Because of the nature of the articulatory system, a spoken sentence consists of a sequence of articulatory gestures. In contrast to the printed stimulus in reading, the parts of a spoken utterance are fast-fading and persist only briefly (Hockett, 1963). In reading, the distal stimulus is unchanging, but severe limitations on the perceptual span (Rayner, 1978) necessitate multiple fixations on each sentence, creating a situation in which sentences in this modality also become available to the perceiver over time. Thus, whereas languages exhibit multiple levels of *structure* simultaneously, sentences are necessarily *processed* sequentially. This fundamental tension creates a difficult computational problem. The information relevant to resolving a particular type of ambiguity often will not be available immediately, and the amount of time before it will become available is indeterminate. Yet the ambiguities must be resolved in order for sentences to be understood. Moreover, despite its computational complexity, language comprehension occurs rapidly, unconsciously, and typically without error. The purpose of this article is to develop the basic principles underlying a theory of sentence comprehension that explains how this outcome is achieved, given the types of information that are available and the conditions under which sentences are processed.

There has been a large amount of research on ambiguity resolution in language comprehension; it is one of the central topics in psycholinguistics. Research has focused on two general classes of ambiguities: lexical and syntactic. Lexical ambiguity research has addressed how the reader or listener determines the contextually-appropriate meaning of a word with multiple senses (e.g. *watch*, as noted above). Lexical ambiguities pervade natural language. Words exhibit different types and degrees of ambiguity. *Homonymy* refers to the case in which two words have the same sound and/or spelling but different meanings. The words can be spelled and pronounced the same (e.g., the alternative senses of *tire*); spelled the same but pronounced differently (e.g., *bass* or *wind*); or spelled differently but pronounced the same (e.g., *team*, *teem*). The extent to which the alternative meanings are semantically re-

lated varies. In extreme cases such as *tire*, the alternative meanings are unrelated; it is a historical accident of English that the same orthographic/phonological form is used to encode two unrelated senses (other examples are *watch* and *rose*). The term *polysemy* refers to instances in which alternative meanings are semantically related; for example, there are noun and verb senses of *take* that are clearly semantically related ("to take something," "the take from a casino"), but others senses that are less related (*someone's take on an issue*). Almost all words in the English lexicon exhibit a nonzero degree of homonymy or polysemy, some acutely so. For example, the American Heritage Dictionary lists 41 separate meanings of *take*, several of which have multiple polysemous senses.

Research on lexical ambiguity has addressed two main issues: First, how aspects of lexical structure such as the relative frequencies of alternative meanings affect ambiguity resolution; and second, how different types of information provided by the linguistic and nonlinguistic context are used, in conjunction with the perceiver's knowledge of the world, in resolving lexical ambiguities (see Simpson, 1981, for a review).

Other research has addressed analogous questions that arise regarding the comprehension of syntactic ambiguities. These issues have been extensively studied with respect to several syntactic ambiguities. One construction, the main verb/reduced relative (MV/RR) ambiguity illustrated in (7), has been the object of attention for nearly 25 years (dating from Bever, 1970). (7a) is a simple sentence in which *raced* is the past tense main verb of the sentence. In (7b) *raced* is a past participle introducing a reduced relative clause construction, meaning "the horse that was raced past the barn," and *fell* is the main verb of the sentence. This ambiguity is permitted in English because certain words introducing a relative clause (e.g., *that was*) can be optionally omitted. This omission can create a problem for the human parser because a verb such as *raced* is ambiguous between past tense and past participle forms. If *raced* is treated as a past tense, (7a) will be analyzed correctly but (7b) will not. If *raced* is treated as a past participle, (7b) will be correctly analyzed but not (7a). Numerous studies have attempted to identify the types of knowledge and processes that are

used in analyzing such structures (Britt, Perfetti, Garrod & Rayner, 1992; Ferreira & Clifton, 1986; MacDonald, in press; Pearlmutter & MacDonald, 1992; Rayner, Carlson & Frazier, 1983; Rayner, Garrod & Perfetti, 1992; Spivey-Knowlton, Trueswell & Tanenhaus, 1993; Trueswell, Tanenhaus, & Garnsey, in press; Trueswell, Tanenhaus, & Kello, 1993). A hint about the complexity of the problem is provided by the observation that sentences with identical syntactic structures can differ greatly in terms of ease of comprehension. Thus (7b), the classic sentence discussed by Bever (1970) and in many subsequent papers, is very hard for naive subjects to comprehend, whereas (7c), which has essentially the same syntactic structure, is quite easy. The theory that we develop in this article addresses how such structures are analyzed and explains the differences in processing difficulty that have been identified in behavioral studies.

(7)

a. The horse raced past the barn.

b. The horse raced past the barn fell.

c. The prisoner captured by the cannibals escaped.

Like lexical ambiguity, syntactic ambiguity is omnipresent. Sentences such as (7b) are famous among psycholinguists because they illustrate an ambiguous structure in dramatic fashion: In such *garden-path* sentences, a particularly vicious ambiguity makes it very difficult to derive a meaningful interpretation of the sentence. The difficulty of (7b) has sometimes left the impression that such ambiguities must be extremely rare, but in fact the MV/RR ambiguity, resolved with the reduced relative interpretation, occurs throughout natural discourse and typically goes unnoticed. To illustrate, Table 2.1 presents examples of reduced relative constructions taken from several recent papers on syntactic ambiguity resolution.[2]

Theories of Lexical and Syntactic Ambiguity Resolution

The theories that have been dominant over the past 15 years suggest that lexical and syntactic ambiguities are resolved by very different mechanisms. Theorizing about

[2]This sentence also contains a reduced relative.

Table 2.1: Examples of Sentences Containing Reduced Relatives

Example (Reduced Relatives *Italicized*)	Source
The ambiguity *associated with reduced relative clauses* ...provides another testing ground...	Tanenhaus, Carlson & Trueswell (1989, p. SI217)
Models *formulated in the classical framework* face many serious problems...	McClelland, St. John & Taraban (1989, p. SI295)
...recent research *reported by Ferreira and Clifton (1986)* has demonstrated that syntactic processing is quite independent and that the initial syntactic analysis *assigned to a sentence* is little influenced by the semantic information *already analyzed.*	Frazier & Rayner (1987, p. 521)
...the pattern of results *obtained with global reading times* is reflected by differences *located on the critical disambiguating phrase...*	Altmann & Steedman (1988, p. 226)
A person *presented with "The evidence examined by the lawyer"* may interpret "the evidence" as a theme, not an agent, since "examined" requires an animate agent.	Ferreira & Clifton (1986, p. 350)
In all cases, the examples *cited here* were not the only reduced relatives in the paper.	Note to this table.

Note: All of these examples contain a reduced relative modifying the subject noun phrase. For clarity, italicized text in the original has been set off with quotation marks. Additional reduced relatives in the same sentence are italicized where appropriate. In all cases, the examples cited here were not the only reduced relatives in the paper.

lexical ambiguity resolution has been heavily influenced by the finding (first reported by Conrad, 1974) that subjects briefly activate multiple senses of ambiguous words even in clearly disambiguating contexts. Many of these studies have employed what Swinney (1979) termed a "cross-modal priming" paradigm. The subject hears a sentence such as (8a), followed by a visually-presented target word related to one of the senses of the ambiguous word (e.g., *watch*: "observe," "time-piece"). Latencies to name the target or make a lexical decision to it are compared to those that occur when the target follows an unambiguous, unrelated control sentence such as (8b).

(8)

a. John began to watch.

b. John began to eat.

The pattern obtained in numerous studies is that when targets are presented immediately after the ambiguous word, there is significant facilitation in responding to targets related to either alternative meaning. If targets are presented somewhat later (e.g., after a delay of 200 msec or after several intervening words), there is priming only for the target related to the contextually-appropriate meaning. These studies (e.g., Swinney, 1979; Tanenhaus, Leiman & Seidenberg, 1979) suggested a model in which multiple alternative senses of ambiguous words are briefly activated, followed by the use of contextual information to select the appropriate meaning. The counterintuitive aspect of these results was that people are only consciously aware of the contextually-appropriate meaning that has been chosen, not the initial activation of multiple senses. (This is true except in degenerate cases such as puns, in which the punner manipulates the ambiguity resolution process with the intent of causing the listener to become aware of an alternative, humorously related word sense.) This phenomenon is particularly striking when activation of multiple alternatives occurs in contexts such as (8a), in which the syntax of the sentence clearly dictates one and only one meaning of the ambiguous word *watch*.

Table 2.2 lists the characteristics commonly attributed to the process of lexical ambiguity resolution. The alternative meanings of words are thought to be stored in

Table 2.2: Properties of Ambiguity Resolution Systems

Lexical Ambiguity	Syntactic Ambiguity
multiple alternatives initially considered	single alternative initially considered
parallel processing	serial processing
capacity free	capacity limited
context used to select appropriate meaning	context used to confirm analysis or guide reanalysis
meanings are stored and accessed	syntactic structures constructed by rule

memory and "accessed" in processing. In the initial stage of processing an ambiguous word, multiple meanings are accessed. Swinney (1979) proposed that access of meaning is "exhaustive" (i.e., that all meanings are accessed), whereas Seidenberg, Tanenhaus, Leiman, and Bienkowski (1982) proposed that access was limited to the most common meanings. Both models held that multiple meanings are considered in parallel, with contextual information used shortly afterward to select the relevant one and suppress all alternatives. It was assumed that multiple meanings could be accessed in parallel because this process is automatic and capacity-free (in the sense of Posner & Snyder, 1975, and others).

Syntactic ambiguity resolution has been viewed quite differently (see Table 2.2). The most fully developed theory of syntactic ambiguity resolution is the Garden Path theory of Frazier and her colleagues (Frazier, 1979, 1987; Frazier & Rayner, 1982; Rayner et al., 1983; Ferreira & Clifton, 1986). The theory proposes that the first stage in the comprehension process is the computation of the basic syntax (phrase structure) of the input. This analysis is the responsibility of a syntactic processor (or "parser") thought to have the capacity to compute only a single analysis at a time. It receives as input only the grammatical categories of the words being processed

(e.g. Determiner, Adjective, Noun, Verb) and does not have access to their identities or meanings. When the sequence of categories that the parser is analyzing is compatible with multiple phrase structure analyses, the path it follows is determined by a processing strategy called *Minimal Attachment*. According to this strategy, the preferred syntactic interpretation is the syntactically simpler one, where simplicity is indexed by the number of nodes in a phrase structure tree. Minimal Attachment directs the parser to pursue the main verb interpretation of the MV/RR structure illustrated in (7). These initial decisions are thought to occur independently of other sentence processing operations, including analyses of word meanings. Other processing subsystems subsequently either confirm or reject the initial syntactic analysis. Frazier introduced a second parsing principle, Late Closure, that prevails in cases where Minimal Attachment does not adjudicate between two alternative analyses. The Late Closure principle also insures that only a single analysis will be pursued.

Many of the differences between the lexical and syntactic ambiguity resolution mechanisms summarized in Table 2.2 derive from assumptions about the types of knowledge involved in each domain (Frazier, 1989). Lexical ambiguity is thought to involve meanings that are stored in the lexicon. Processing involves *accessing* this information, which is assumed to be accomplished automatically and in parallel. Syntactic structures, in contrast, are thought to be *constructed* on the basis of grammatical rules rather than stored directly in memory. This computation is assumed to place demands on working memory and attentional resources that are limited in capacity (Frazier, 1979; Gibson, 1991; MacDonald, Just & Carpenter, 1992). Because of these limitations, the parser can pursue only a single analysis at a time. Thus, lexical ambiguity resolution is thought to routinely involve the temporary access of multiple alternatives, with other processing subsystems coming along later to select the contextually-appropriate word meaning. Syntactic ambiguity resolution, in contrast, is thought to be strictly serial, with only one parse computed at a time. Other processing subsystems come along later to confirm or disconfirm this analysis, initiating reanalyses as necessary.

It is interesting to note that although the lexical and syntactic ambiguity resolu-

tion mechanisms differ in fundamental ways according to these theories, both have been explained in terms of Fodor's (1983) concept of *modularity*. The language processing system is thought to be composed of a small number of components or modules that are responsible for analyzing different types of information. The mental lexicon and the parser are two of these components. One of the core claims of Fodor's theory is that the modules of the mind are *informationally encapsulated*. That is, they operate autonomously, responding only to the type of information to which they are attuned. In the case of lexical ambiguity, the lexical processor responsible for accessing the meanings of words is isolated from the processors responsible for computing other types of information. The consequence is that even when there is nonlexical information in the prior context indicating which meaning of an ambiguous word is appropriate, the lexical processor cannot utilize it. Hence, multiple meanings of ambiguous words are accessed. Other modules operate on the output of the lexical processor, determining which meaning is appropriate. In the case of syntactic ambiguity, the parser is also isolated from other components of the comprehension system. Here the consequence is that even when there is nonsyntactic information in the prior context favoring a more complex syntactic analysis, the parser cannot utilize it. Instead, the single, simplest parse is pursued. Other modules operate on the output of the parser, determining whether the proposed analysis is consistent with the other information to which they have access.

Challenges for the Standard Theories

Both empirical and theoretical considerations have called into question the accounts of ambiguity resolution summarized in Table 2.2. First, there have been challenges to the standard assumptions about context effects. Both of the standard theories predict that manipulations of contextual bias should have little effect on the initial processing of either type of ambiguity, a prediction that has generated many empirical studies. In both domains there have been replications of the core findings on which the theories were originally based, but also important deviations from predicted patterns of behavior. Second, there are questions about the role of frequency

information in modulating ambiguity resolution. Third, there are questions about the types of knowledge representations and processes involved in the two domains that have undermined the distinction between "accessing" a meaning and "constructing" a syntactic representation. We will briefly review how these three issues have arisen in each domain.

Lexical Ambiguity

1. Context effects. Even though theorizing about lexical ambiguity resolution focused on the multiple access pattern, several studies failed to yield this outcome (e.g., Swinney & Hakes, 1976; Seidenberg et al., 1982, Experiment 2; Tabossi, 1988; Tabossi, Colombo and Job, 1987; Simpson & Krueger, 1991). These results were initially attributed to methodological limitations. For example, Swinney (1979) suggested that some of the response measures used in earlier studies reflected the outcome of the ambiguity resolution process but were insensitive to the initial activation of multiple word senses. This argument has been weakened by the finding of selective access using the same methods as in Swinney (1979); see, for example, Tabossi (1988) and Tabossi and Zardon (1993). In other cases, the "modular" theory of lexical access was preserved by attributing the contextual effects to a process termed *intra-lexical* priming (Forster, 1979; Seidenberg et al., 1982). It was known that the identification of a word is facilitated in the context of a semantically or associatively related priming word (Meyer & Schvaneveldt, 1971). Studies such as Seidenberg et al.'s (1982) Experiment 2 were interpreted as showing that a word in the sentence context could prime a meaning of an ambiguous word, resulting in selective access. These priming effects were attributed to spreading activation within the lexical module. As such, they did not necessarily represent the penetration of the lexical module by nonlexical types of information. This interpretation of contextual effects on the access of meaning is unsatisfying, however, because there is no independent theory of the scope of such "intra-lexical" effects. It has not been clear, for example, which aspects of semantic relatedness contribute to priming effects (see McRae, de Sa, & Seidenberg, 1993; and Tabossi, 1988; for attempts to address this theoretical gap). Therefore, the

conditions under which a context word could prime one meaning of an ambiguous word could not be independently established.

The priming explanation of context effects has also been called into question by studies in which selective access is observed in sentence contexts that do not contain words that are strongly related to the ambiguity (e.g., Morris, in press). Foss (1982) and J. N. Williams (1988) provided other evidence calling into question whether associative or semantic priming play a significant role in sentence processing.[3] Thus a number of studies have yielded selective access of meaning in biasing contexts. However, the reasons why contexts apparently differ in their effects on activation of meaning remain obscure.

2. Frequency. The meanings of an ambiguous word vary in frequency of use. Frequency has a large impact on lexical processing, a fact acknowledged by all models of the lexicon dating from Morton (1969). It is therefore important to consider how frequency affects the processing of ambiguous words. Studies such as Swinney (1979) and Tanenhaus et al. (1979) examined ambiguous words such as *tire* or *straw* that have two primary meanings that are roughly equal in frequency. Swinney (1979; Onifer & Swinney, 1981) interpreted the fact that multiple access was observed with such "equibiased" ambiguous words as evidence that access of meaning is exhaustive (i.e., that lexical access makes available all the meanings of a word, regardless of context). Most ambiguous words have alternative meanings that differ considerably in frequency, however. For example, in the Francis and Kučera (1982) corpus of American English, the noun sense of the word *barge* occurs twice as often as the verb sense. Similarly, *bluff* occurs 11 times more often in this sample as a noun than as a verb. *Brief* occurs 64 times in the corpus as an adjective, 10 times as a noun,

[3]Glucksberg, Kreuz and Rho (1986) raised a somewhat different question concerning the role of priming in ambiguity resolution. Their hypothesis was that the multiple access pattern is itself a methodological artifact due to backwards priming from a semantically related target to the preceding ambiguous word. Thus, the contextually-inappropriate meaning of *tire* would show activation in a sentence such as *The car was missing one tire* because it is primed by the target *sleep*, not because access of meaning is independent of context. Burgess, Tanenhaus and Seidenberg (1989), however, provided evidence calling into question whether backward priming could be the source of the multiple access pattern.

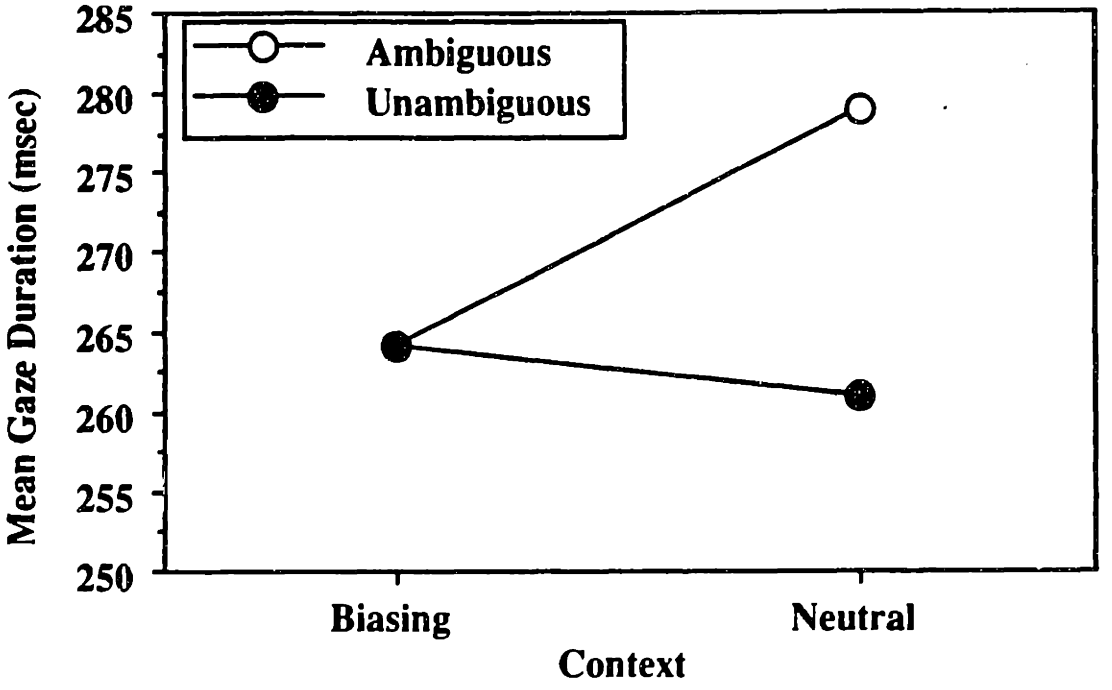
and 3 times as a verb. *Tire*, in contrast, is more equally balanced: It occurs 31 times as a noun and 46 times as a verb. Nevertheless, frequency asymmetries appear to be common: In the Francis and Kučera corpus, 1366 words occur as both nouns and verbs. These include both words whose noun and verb meanings are semantically unrelated (such as *tire* and *bluff*) and words with semantically related meanings (such as *cap* and *coin*). For 73% of these noun-verb ambiguities, one grammatical category is used at least twice as often as the other. Analyses of the words that are ambiguous between noun and adjective meanings yield a similar picture. There are 155 such words in the Francis and Kučera corpus, including *nude*, *special*, and *current*. The ratio of the frequencies is 2:1 or better for 81% of these words. Thus, it appears that many ambiguous words exhibit asymmetries with regard to the frequencies of their alternative meanings.

Simpson (1981) reviewed a large number of studies suggesting that frequency affects the order in which the meanings of ambiguous words are accessed. An early model incorporating this idea was proposed by Hogaboam and Perfetti (1975), and Seidenberg et al. (1982) also suggested that frequency modulates the activation of meaning, with the multiple access pattern the particular case in which the two most common meanings of a word happen to be similar in frequency. Important studies by Rayner and Duffy (1986) and Duffy, Morris, and Rayner (1988) have provided additional evidence concerning the frequency factor. Rayner and Duffy (1986) established that equibiased ambiguous words (such as *pitcher*) produce longer fixation durations (using an eyetracking paradigm) than matched unambiguous words in neutral contexts (i.e., contexts that did not favor either meaning). Thus, multiple access of meaning was associated with longer fixation durations. Fixation durations did not differ for nonequibiased words (such as *port*) compared to controls, indicating that for these words, in which one meaning is much more frequent than others, only the dominant meaning is accessed and integrated with the context. These effects can be seen in the top panel of Figure 2-1. Duffy et al. (1988) then examined how the processing of these words is affected by biasing contextual information. Contexts were constructed so as to favor one meaning of the equibiased ambiguous word and the

subordinate (less frequent) meaning of the nonequibiased ambiguous word. Contextual bias was achieved without including any words that were highly semantically or associatively related to the ambiguous word, eliminating lexical priming as a possible source of contextual effects. As the bottom panel of Figure 2-1 demonstrates, latencies for equibiased words in a biasing context were equal to those for unambiguous controls. For non-equibiased ambiguous words, however, a context supporting the subordinate meaning resulted in longer fixations than for unambiguous controls.

These results indicate that equibiased ambiguous words in *biasing* contexts act like biased words in *neutral* contexts. In both cases, a single meaning is accessed on line and integrated with the context, and no effect of ambiguity is observed. Thus, context can promote one meaning of the equibiased word, resulting in the "selective access" pattern. Duffy et al. therefore interpreted the data as evidence for a model in which context can "reorder" the access of meaning. The nonequibiased words in biasing contexts favoring the subordinate meaning acted like equibiased words in neutral contexts. In both cases, latencies were longer in the ambiguous condition than in the unambiguous control condition, indicating that multiple meanings had been accessed. Thus, contexts that promote the less frequent meaning of a biased word favor that reading, another instance of re-ordered access. However, the effect of contextual information is limited in this case: Whereas the context increases the activation of the lower frequency meaning, the higher frequency meaning is still activated. Thus, context can increase the salience of a lower frequency meaning but not inhibit the higher frequency meaning.

There are several important findings here. First, the studies provide clear evidence that frequency of meaning has an impact on processing, consistent with earlier studies. Second, contextual information can result in activation of only a single meaning of an ambiguous word. Third, the two factors interact: Biasing contexts produce selective access when the alternative meanings are similar in frequency and when they are consistent with the higher frequency meanings of nonequibiased words, but not when they favor lower frequency meanings. This interaction between frequency and degree of contextual constraint will play an important role in the theory presented below.



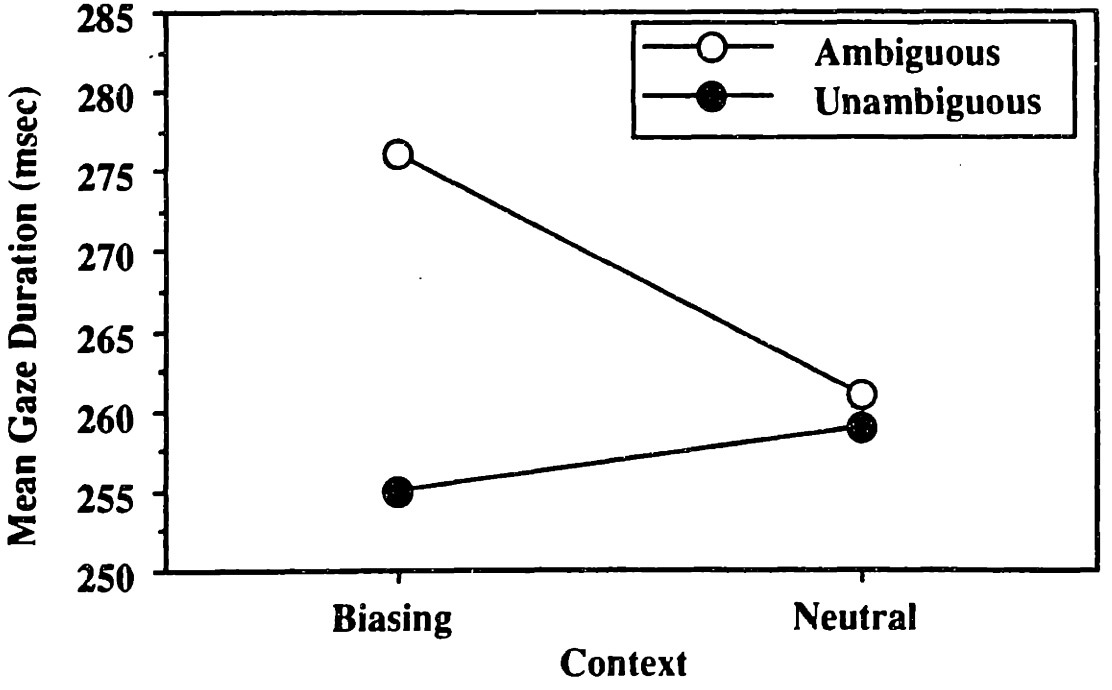


Figure 2-1: Mean gaze duration for equibiased (top) and non-equibiased (bottom) words in Duffy, Morris, and Rayner (1988).

3. Representation: Are meanings "accessed?" Lexical ambiguity resolution concerns the processing of words. Most models of word recognition developed over the past 20 years have construed the process as one involving the "access" of lexical information (see, for example, Allport & Funnell, 1981; Coltheart, 1987; Forster, 1976; Monsell, 1985). The words in a person's vocabulary are thought to be encoded as entries in a mental lexicon; recognizing a word involves accessing its entry. For ambiguous words, the issue was whether readers access a single meaning or multiple meanings of a word in context.

The metaphor of a dictionary-like mental lexicon with entries that are searched and accessed has played an important part in thinking about lexical processing. In recent years, however, the assumptions about knowledge representation and processing that are presupposed by the concept of "lexical access" have begun to be critically evaluated. For example, the idea that meanings are fixed entries that can be "accessed" has been strongly questioned on both theoretical and empirical grounds. Current theories of the representation of meaning assume that the meanings of words are computed as part of the recognition process (Barsalou, 1987). The meaning that is computed may differ depending on the context in which it occurs. An elegant experiment by Merrill, Sperber, and MacCauley (1981) illustrated this phenomenon clearly. They examined the processing of simple, minimally ambiguous words such as *cat*. On the lexical access view, recognizing a word provides "access" to a meaning stored in memory. This suggests that the same information should become available in all contexts in which the word occurs. Merrill et al. showed that even for simple words such as *cat*, different information became available in different contexts. The study employed a priming methodology very similar to the one used in studies of lexical ambiguity. A sentence ending with the word *cat* was followed by a target word such as *fur* or *claws*, to which the subject had to respond. For the sentence *John petted the cat*, latencies to *fur*, but not *claws*, differed from latencies to an unrelated control word. For the sentence *John was scratched by the cat*, the opposite pattern was observed: *Claws*, but not *fur*, differed from the control. The results are problematic for the view that meanings are "accessed." Rather, they suggest that

the pattern of semantic information that becomes available when a word is processed depends on the context in which it occurred. Hence, the meaning of a word must be computed based on knowledge of the word and knowledge of the context, rather than "accessed." Barsalou (1987) provides other demonstrations of the constructive nature of meaning computation.

These results have important implications for the concept of "lexical access." The view that recognizing a word involves "accessing" its entry in memory implies that access is an event that occurs at a specific point in time and that meanings are distinct, countable entities. It therefore makes sense within this framework to investigate questions such as how frequency affects the latency of lexical access, or how many meanings of an ambiguous word are accessed in context. However, the Merrill et al. and Barsalou studies suggest that even the meanings of relatively unambiguous words are computed anew each time they are processed, in a highly context-dependent manner. It may therefore be misleading to think of lexical access as an event that is achieved at a distinct moment in time. Balota (1990) refers to this as the "magic moment" conception of lexical access and summarizes a broad range of phenomena that are inconsistent with it.

These phenomena seem to demand a different notion than that of accessing meanings from memory (Seidenberg, 1990). An alternative is provided by the distributed representations employed in connectionist networks. The meaning of a word can be represented as a pattern of activation over a set of units encoding semantic primitives (Hinton, 1986); each unit participates in the meaning of many words. In a system employing this type of knowledge structure, the meanings of words are not stored as separate entries and therefore cannot be "accessed;" rather, they are computed, with different patterns activated in different contexts. Moreover, in networks that incorporate attractors (e.g., Hinton & Shallice, 1991; Plaut & Shallice, in press), the network settles into a pattern of activation over time, rather than computing it in a single step, providing a specific alternative to the metaphor of "accessing" a meaning at a fixed latency. The notion of distributed representation can be extended to the other codes for words as well. Thus, in Seidenberg and McClelland's (1989) model of

word recognition, the spellings and pronunciations of words are also represented as patterns of activation over relevant sets of units. A word's spelling and pronunciation are computed in the same manner as meaning. Hence their model completely eschews the notion of accessing lexical entries.

Many of these issues and their implications for lexical ambiguity resolution were first addressed by Kawamoto (1988, 1993). In Kawamoto's model, each code for a word (spelling, meaning, phonology, grammatical category) is represented by a pattern of activation over a set of units encoding that type of information. The model computes patterns of activation corresponding to these codes. The model was developed using a greatly simplified domain consisting of a small number of words and contexts. The vocabulary items included homographs such as *wind* and *bass*, which have semantically-unrelated meanings that are associated with different pronunciations. The architecture of the system allowed the computation of word meaning to be influenced by contextual information. The model computed multiple meanings of homographs even when there was contextual information biased toward one meaning. The basic reason for this outcome is very simple. Multiple meanings are computed for words such as *wind* because the association between each meaning and its sensory form (orthography or phonology) is much stronger than the association between a component meaning and the information provided by the biasing context. Thus, the model exhibits the property Marslen-Wilson (1987) termed *bottom-up priority*.

Kawamoto's model is important for two reasons. First, it shows that the multiple access pattern that had been taken as diagnostic of a modular lexicon could arise from other sources. It placed the explanation for this effect with the nature of contextual constraints, not the architecture of the processing system. It was simply because the contextual constraints were too weak that multiple access occurred in Kawamoto's model, not because the lexical processor was an autonomous module (see McClelland, 1987, for discussion). The model does not show that the modular account is incorrect; rather it shows that the data are compatible with very different theoretical conclusions.[4] Second, the model was an early demonstration of how

[4]There are now several demonstrations of this kind in the literature: connectionist models showing

lexical information could be represented in a system without entries for individual words. This idea has now been incorporated in several models of lexical processing (e.g., Seidenberg & McClelland, 1989; Hinton & Shallice, 1991). In models employing distributed representations of lexical codes, the strong distinction between "accessing" and "constructing" representations is not preserved. In Kawamoto's model, for example, the meaning of a word is computed each time it is recognized. Thus, the model does not maintain the central property of lexical representation that had been used to differentiate it from syntax.

Syntactic Ambiguity

Exactly the same issues — the scope of contextual effects, the role of frequency, and the types of knowledge representations and processes involved — have arisen with regard to syntactic ambiguity resolution.

1. Context effects. Frazier and colleagues' Garden Path theory has provided the focus for much of the research on sentence processing over the past 15 years. Whereas lexical ambiguity research focused on whether contextual information could override the multiple access pattern, research on syntactic ambiguity addressed whether contextual information could override the minimal attachment pattern, so that a syntactically more complex interpretation of an ambiguity would be favored in some context. So-called "interactive" or "constraint-based" models of sentence processing (Bates & MacWhinney, 1989; McClelland, 1987; Marslen-Wilson, 1975) emphasize how the processing of words in sentences depends on the availability of different types of contextual information. Thus, whether minimal attachment is observed or not will also depend on contextual information.

Research on context effects in syntactic ambiguity resolution has focused on the MV/RR ambiguity illustrated in (7) (repeated below). The interpretation that is *not* favored by the minimal attachment algorithm is the one in which the ambigu-

---

that inferences that have been drawn from behavioral data about the architecture of one or another cognitive system are equally compatible with connectionist models with very different properties. See Seidenberg (1993) for discussion.

ous phrase modifies the preceding noun (e.g. *raced past the barn* modifies *horse* in (7b), and *captured by the cannibals* modifies *prisoner* in (7c)). Two different types of contextual manipulations favoring such noun-modification interpretations have been investigated. The first of these is the plausibility of a noun-modification interpretation relative to the alternatives (e.g. how plausible is it for a prisoner to capture something vs. to be captured). If the surrounding context makes the noun modification interpretation more plausible than the minimal attachment interpretation, it might be pursued by the parser. The second context manipulation is a pragmatic one — whether or not it is felicitous to modify a noun at a particular point in the discourse. Crain and Steedman (1985) argued that noun modification was not necessary or felicitous when the discourse had established only one possible referent for a noun (e.g., only one horse in the discourse), but modification was felicitous and necessary in a discourse in which several potential referents had been established. Thus the noun modification interpretation in (7b) might be preferred in a context that established several horses in the discourse (one of which was known to have been raced past a barn, another of which was standing in a pasture), because this modification would serve to indicate exactly which horse the speaker was discussing.

(7a) The horse raced past the barn.

(7b) The horse raced past the barn fell.

(7c) The prisoner captured by the cannibals escaped.

According to the garden path model, neither the plausibility nor the pragmatic context manipulations should affect the initial interpretation of syntactic ambiguities, because initial parsing decisions are guided only by minimal attachment, not by context. The theory predicts that the use of contextual information should be delayed: It can be used only to identify and possibly correct a phrase structure analysis that turns out to be inappropriate. The effects of plausibility and pragmatic context have been investigated with these kinds of sentences using a variety of methodologies, including cross-modal priming and measurements of reading times. Many studies yielded findings consistent with the Garden Path theory's claim for delayed use of context, but at the same time there are some notable exceptions to this pattern.

In a paradigmatic study of this type, Ferreira and Clifton (1986) examined the effect of plausibility in the MV/RR ambiguity by varying the animacy of the noun phrase in subject position. They examined reading times for sentences such as (9) and (10) compared to unambiguous control sentences (containing *that was* between the subject noun and *examined*). The plausibility manipulation relies on the nature of the subject noun phrase and the ambiguous verb. *Examine* is a verb that normally takes an animate subject. *Witness*, being animate, is consistent with the simple past tense interpretation of *examined* (in which the witness examined something), which turns out to be incorrect. *Evidence*, however, is incongruent with the simple past tense interpretation because inanimate things such as evidence cannot examine something. Hence it favors the reduced relative alternative, which turns out to be correct. If subjects can use this information to guide the initial parse, (10) should cause less of a garden path effect than (9). Ferreira and Clifton (1986) found that subjects had equal difficulty reading the two types of sentences, however. They concluded that such cues were used only after the initial parse was assigned on the basis of minimal attachment.

(9) The witness examined by the lawyer was useless.
(10) The evidence examined by the lawyer was useless.

In later research, Trueswell, Tanenhaus, and Garnsey (in press) raised several questions about the validity of the Ferreira and Clifton results. First, they questioned whether the contextual manipulation was adequate. The manipulation presumes that verbs such as *examine* cannot take inanimate nouns with the main verb interpretation of the ambiguity. However, some of Ferreira and Clifton's sentences included verbs such as *towed* and *felt*, for which the presence of an inanimate subject noun phrase yields a plausible main verb interpretation (e.g. *The car towed the boat*, *The skin felt soft.*). Thus these items did not succeed in strongly promoting the reduced relative interpretation. Second, Trueswell et al. suggested that the nature of the stimulus display in the Ferreira and Clifton experiments (which was only 40 characters per line and introduced line breaks in some critical regions) affected subjects' abilities to read the sentences naturally. Third, Trueswell et al. argued that Ferreira and Clifton's

method of analyzing their eyetracking data obscured some context effects. When Trueswell et al. modified the stimuli, the display, and the data analyses, they found a robust context effect, such that helpful contexts removed all difficulty associated with the reduced relative interpretation of the ambiguity. This result favors an account of syntactic ambiguity resolution in which both syntactic and contextual information interact in directing the comprehender to an interpretation of a syntactic ambiguity, and it argues against a first-stage parser that makes initial parsing decisions without regard to context.

As in the research on lexical ambiguity, there have been attempts to attribute these kinds of conflicting results to methodological differences between studies. Clifton and Ferreira (1989), responding to an early presentation of some of the Trueswell et al. (in press) data, noted that Trueswell et al.'s subjects read somewhat faster than their own and suggested that the eyetracking measure they used may not have been able to detect subtle garden path effects prior to the influences of context. This is reminiscent of Swinney's (1979) suggestion that some response measures were not able to detect the activation of multiple meanings of ambiguous words prior to the influences of context. In the sentence processing literature, this is a pattern that has been repeated several times: First there are studies suggesting a lack of context effects (e.g., Britt et al., 1992; Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Mitchell, Corley & Garnham, 1992; Rayner et al., 1983; 1992). Then there are criticisms that context effects were obscured by poor stimulus construction or other methodological problems (e.g., McClelland, 1987; Taraban & McClelland, 1988; Trueswell et al., in press; Trueswell et al., 1993), accompanied by replication studies demonstrating early effects of context. These studies are in turn questioned for using methods that might be insensitive to the initial operation of the parser (e.g., Clifton & Ferreira, 1989; Frazier, 1987).

The studies showing early effects of context nonetheless raise questions concerning the generality of the Minimal Attachment strategy and the ways in which contextual information does and does not influence the comprehension process. Under at least some conditions, readers' behavior does not conform with predictions of the Garden

Path theory. However, the kinds of information that contribute to "contextual bias" and the mechanisms by which this information is utilized in sentence processing are poorly understood. These questions are addressed by the theory presented below. This theory suggests that the conflicting results of existing studies are not in fact due to methodological differences, but rather to the fact that factors that are relevant to the ambiguity resolution process varied across stimulus materials.

2. Frequency. A second issue concerns the role of frequency in syntactic ambiguity resolution. It is necessary at the outset to consider what kinds of frequency information might be relevant in this domain. This question is more difficult to answer than in the case of lexical ambiguity, where the simple idea that people keep track of the frequencies of meanings of words has proven useful. In syntactic ambiguity resolution it is less clear what kinds of frequency information are implicated. Until recently, frequency information has not been considered very relevant to sentence comprehension, partly because of the influence of Noam Chomsky, whose remarks on the issue in *Syntactic Structures* (1957) may have diverted attention away from it (although perhaps more thoroughly than he intended). On pp. 16–17, Chomsky developed the famous conception of grammar as a theory of the well-formedness of utterances. He argued that this notion of grammar cannot be equated with statistical properties of language, such as the frequencies with which words co-occur. He noted that a sentence may be grammatical but have a low frequency of occurrence (his examples included *Colorless green ideas sleep furiously* and *I saw a fragile whale*) and that ungrammatical utterances may occur with relatively high frequency. He said, "I think that we are forced to conclude that ... probabilistic models give no particular insight into some of the basic problems of syntactic structure." (p. 17) Most theoretical linguists have accepted this argument and shown little interest in issues concerning statistical properties of language such as the relative frequencies of syntactic structures. It should be noted, however, that Chomsky (1957) then made a point of acknowledging that such statistical information might prove quite relevant to understanding language *use*. Continuing on p. 17 of *Syntactic Structures*, he wrote,

"Given the grammar of a language, one can study the use of the language statistically in various ways; and the development of probabilistic models for the use of language (as distinct from the syntactic structure of language) can be quite rewarding."

Although Chomsky's observations were about the irrelevance of statistical properties of language to his conception of the *grammar*, the Frazier et al. theory assumes that statistical properties are also irrelevant to the initial syntactic *processing* of sentences. Minimal attachment is a strategy based on the relative simplicity of syntactic structures, not their relative frequencies. The studies that were designed to test this theory examined the ease of processing different structures, ignoring facts about their relative frequencies of occurrence. Thus Ferreira and Clifton (1986) considered whether subjects would prefer main verb or reduced relative interpretations of phrases such as *The witness/evidence examined by the lawyer...* However, a person interested in the statistical properties of language might ask, by analogy to lexical ambiguity, whether the verb *examined* is used more often in simple transitive sentences (as a main verb) or reduced relatives, whether there are verbs in which the relative frequencies are reversed, and whether the comprehension process exploits such frequency information. With few exceptions, (e.g., Ford, Bresnan & Kaplan, 1982), there has been little investigation of frequency factors in parsing. From the perspective of the Garden Path model, no effects of frequency are expected during the initial analysis of a syntactic ambiguity, because the parser operates only on strings of grammatical categories and has no access to any information about frequency. Over the past few years, however, there has been a revival of interest in questions concerning the statistical properties of natural language (see, for example, papers in Zernik, 1991). If the human sentence processing mechanism exploits such properties of language, it will be necessary to reassess earlier studies in which they were not controlled.

3. Representation. Finally, there are questions as to whether the data interpreted as evidence for the Garden Path theory necessarily implicate a modular parser. As in the case of lexical ambiguity, there is a signature finding — the garden path effect in sentences such as *The horse raced past the barn fell* — taken as evidence for an

autonomous module in the comprehension system. Kawamoto's model alerts us to the possibility that this kind of outcome might result from a parser with a very different type of architecture. The claim that some types of biasing information are merely too weak to overcome Minimal Attachment is certainly reminiscent of the claim, embodied by Kawamoto's model, that multiple access of meaning occurs in contexts that are too weakly constraining.

The distinction between lexical and syntactic forms of ambiguity has been further weakened by subsequent developments in linguistic theory. Although current linguistic theories still rely on the notion of computing some sort of independent syntactic structure (a phrase structure tree, for example), recent work has emphasized how the form of this representation is constrained by properties of the lexical items included in it. Early work in generative grammar (up to and including the Standard Theory of Chomsky, 1965) relied on the mechanism of phrase structure rules to generate basic syntactic structures. Nothing about these rules necessarily associated them with particular lexical items, and thus the syntactic component of the linguistic system could be fairly easily isolated and viewed as an independent module with its own idiosyncratic representations and processes. More recently, however, collections of phrase structure rules have been abandoned in favor of formulations in which lexical items are directly associated with the information used to construct a complete syntactic representation (e.g., X-bar structures in Government-Binding Theory (Chomsky, 1981) and predicate-argument structures in Lexical-Functional Grammar (Bresnan, 1982)). On this view, the lexicon and syntax are very tightly linked, and to the extent that information required by the syntactic component is stored with individual lexical items, it will be difficult to find a boundary between the two systems. Thus, whereas Kawamoto's model provided an alternative to the concept of accessing a lexical entry, syntactic theory has provided, as an alternative to the concept of constructing a phrase structure representation by rule, the idea that syntactic information is just part of a word's representation in memory. These independent developments suggest the possibility that the computation of both "lexical" and "syntactic" information in sentence comprehension are governed by common lexical processing mechanisms. We

explore just this possibility below.

2.2 Toward an Integrated Theory of Lexical and Syntactic Ambiguity Resolution

In this section, we develop our principal theoretical claim, that both lexical and syntactic ambiguity are governed by the same types of knowledge representations and processing mechanisms. Several considerations motivate this approach. We have just seen that very similar empirical, theoretical and methodological issues have arisen in both the lexical and syntactic domains. These issues include: the role of frequency information, the types of information involved in contextual constraints, the relative degree to which contextual information constrains the interpretation of ambiguities, whether the processing system is modular or interactive, and the sensitivity of different measures to the time course of processing. Moreover, we have seen that recent types of theorizing eliminate the strong distinction between accessing a meaning and computing a syntactic representation, which was central to previous accounts. We suggest that these parallels between the domains are not coincidental; they reflect common underlying processes and types of knowledge representations. The parallels derive from the fact that the syntactic ambiguities in question are based on ambiguities at the lexical level. The same ambiguity resolution mechanisms apply in both domains because both involve ambiguities over various types of lexical representations.

Knowledge of words is standardly assumed to be encoded in the mental lexicon, a part of long-term memory. Research on what has been called "lexical" ambiguity resolution has focused on one aspect of this knowledge, word meaning. The research attempted to determine how the reader or listener identified which of the meanings of an ambiguous word is appropriate in a given context. More is known about words than just their meanings, however. For example, each meaning of a word is associated with a grammatical category such as noun or verb; pronouns in English have associated genders (e.g., *she*, *him*); and verbs have associated tenses (*run*, *ran*; *shows*, *shown*).

As we describe in some detail below, current linguistic theories assume that each word's lexical entry consists of many types of information which partially specify the grammatical (and other) structures in which the word participates. These types of lexical information create additional ambiguities because a single word is often associated with multiple alternatives of a given type. To illustrate, a word can be ambiguous between different grammatical categories (e.g., noun or verb); verbs can be ambiguous between different tenses (e.g., past tense or past participle); and nouns can be ambiguous between different numbers (e.g., singular or plural). The link between the mechanisms of "lexical" and "syntactic" ambiguity resolution is provided by the fact that syntactic ambiguities such as the MV/RR construction derive from ambiguities over these other types of lexical information. In *The horse raced past the barn fell*, for example, the ambiguity turns on the lexical fact that *raced* can be either a past tense or a past participle, and that *the horse* can be either Agent (*The horse raced past the barn.*) or Theme (*The horse (that was) raced ... fell.*).

Thus "lexical" ambiguity research has been concerned with one aspect of lexical representation, meaning. "Syntactic" ambiguity research has been concerned with ambiguities arising from other aspects of lexical representation, those determining the syntactic structures in which words participate. Our integrated view of ambiguity resolution assumes a model in which both types of information are part of the lexicon, the organization of which we consider in the next section. A corollary of this view is that similar processing mechanisms should underlie the resolution of the different types of ambiguities. These mechanisms are described in the subsequent section. We then show that this account of ambiguity resolution allows us to explain a broad range of findings in the experimental literature, including some that were previously seen as conflicting. We conclude with a discussion of some of the unresolved issues and directions for future research suggested by this theory.

Levels of Representation

We assume that comprehension is the process of concurrently deriving a number of linked representations of different types for a given input sentence or sequence

of sentences. A single appropriate representation must be computed at each level in the course of processing (although a complete record of all of this information will not necessarily have to be maintained throughout the processing of a sentence). Sentences are represented at three major levels: (1) lexical, the level that is the focus of this article, (2) syntactic, and (3) discourse. We will briefly describe each level and the links between them before turning to a detailed consideration of the lexical component.

The structure of the mental lexicon and the process of word recognition have been widely studied. The standard view of the lexicon in psycholinguistics is that it is a mental dictionary containing an entry for each word in a person's vocabulary, specifying its semantics, phonology, and orthography. The lexical representations that we postulate include these three types of information, but also several other types of information that are relevant to syntactic processing (cf. Boland & Tanenhaus, 1991, for a closely-related proposal). Our claims for a richer lexical representation gain support from recent work within linguistic theory, which has become increasingly focused on issues concerning the structure of the lexicon and the relationships between different types of information (e.g., syntactic and semantic) within the lexicon. We describe the structure of the lexicon in more detail below. For now, it is sufficient to note that lexical processing involves activating different types of information associated with a word form and then using this information to generate representations at other levels.

The second major level of representation is syntax. Syntactic structures encode configurational information (linear order and immediate dominance) that is relevant to a wide variety of grammatical constraints, including the enforcement of feature agreement (number, person, gender, etc.) and the computation of potential coreference relationships (as in *Sue saw herself*, in which *herself* must refer to the same person as *Sue*; versus *Sue saw her*, in which *her* must not refer to the same person as *Sue*). The traditional representation of structural relationships is a *phrase structure tree*, which preserves the order of the input words (linear order) and adds hierarchical information (specified in terms of immediate dominance) about which words are more closely associated with which other words (constituency). Figures 2-2 and 2-3 show

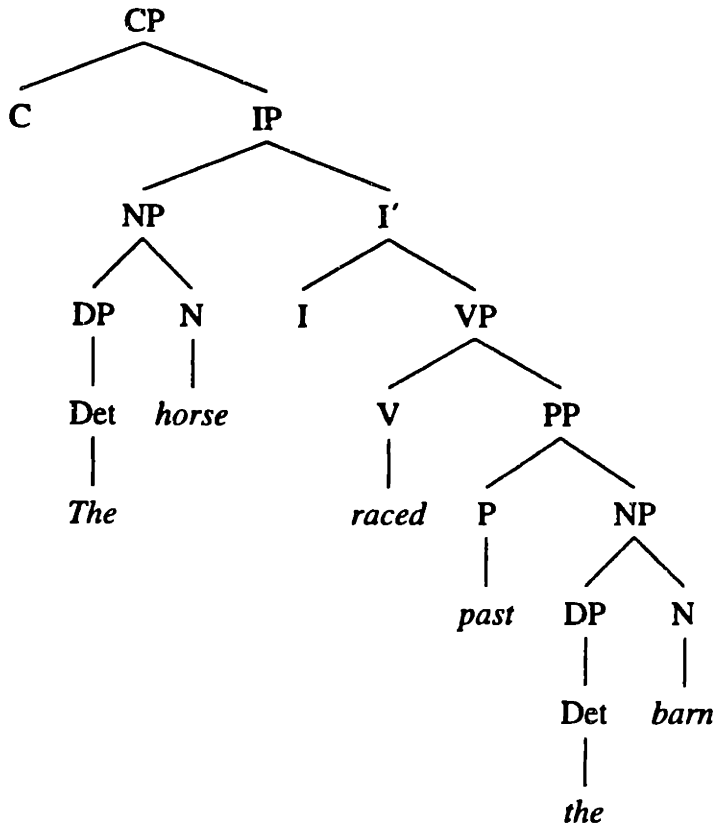


Figure 2-2: Phrase structure tree for *The horse raced past the barn.*

two such trees (to be discussed more extensively below), one for the sentence *The horse raced past the barn* (Figure 2-2), and the other for *The horse raced past the barn fell* (Figure 2-3).

Work in Government-Binding Theory (GB; Chomsky, 1981, 1986a, 1986b; Chomsky & Lasnik, 1991) and Generalized Phrase Structure Grammar (GPSG; Gazdar, 1981; Gazdar, Klein, Pullum, & Sag, 1985) has relied on the idea that complete syntactic trees are built from sub-components called X-bar structures (often written as X′) (Chomsky, 1970; Jackendoff, 1977).[5] X′-structures are generic pieces of phrase structure consisting of a *head* (X, a grammatical category such as Noun or Verb associated with a lexical item) and various *projections*, as shown in Figure 2-4. The head

[5]The terms of the following discussion are primarily those of GB, and in general, our analyses are closest to those of GB throughout, but everything we discuss will have an analysis in other syntactic frameworks.

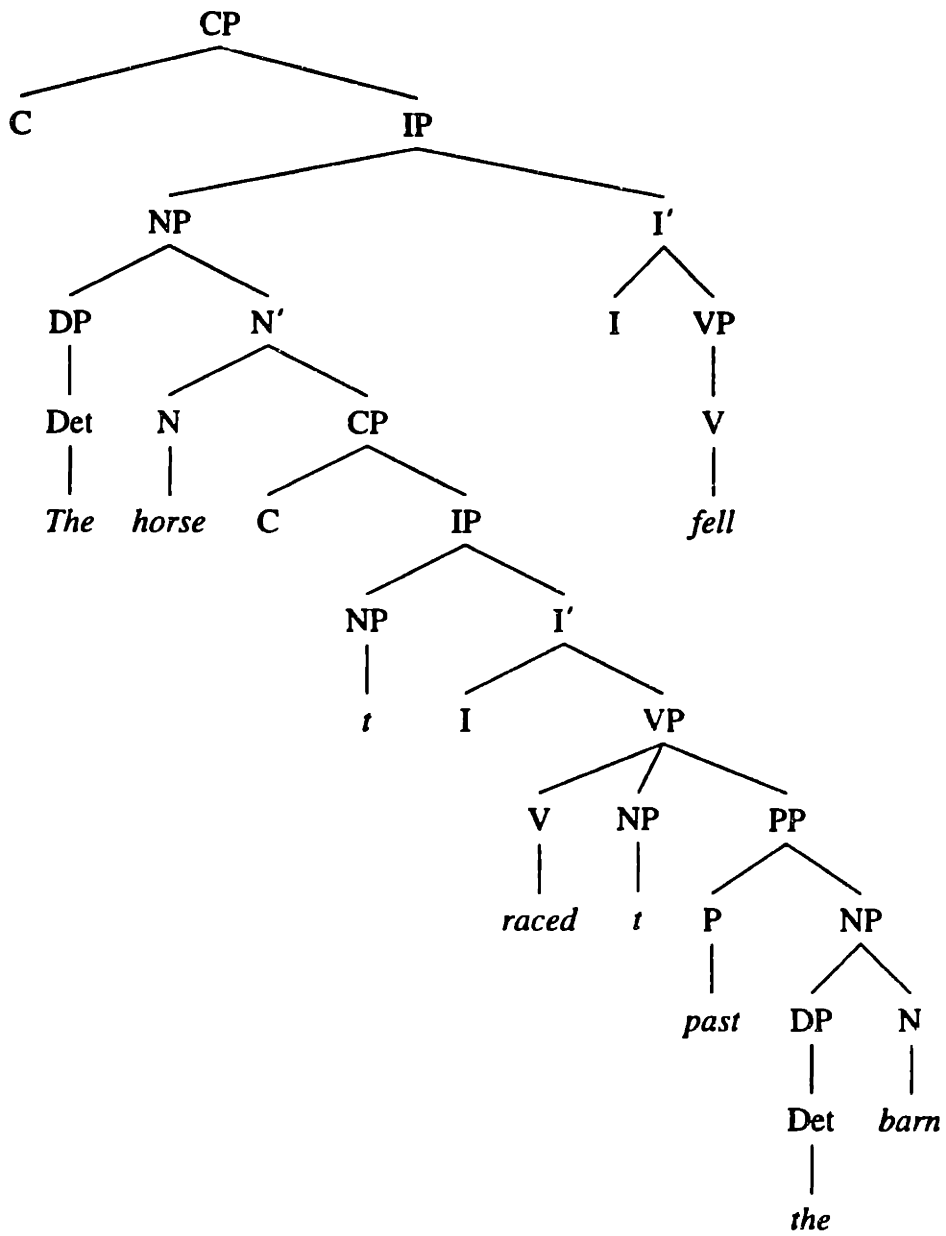


Figure 2-3: Phrase structure tree for *The horse raced past the barn fell.*

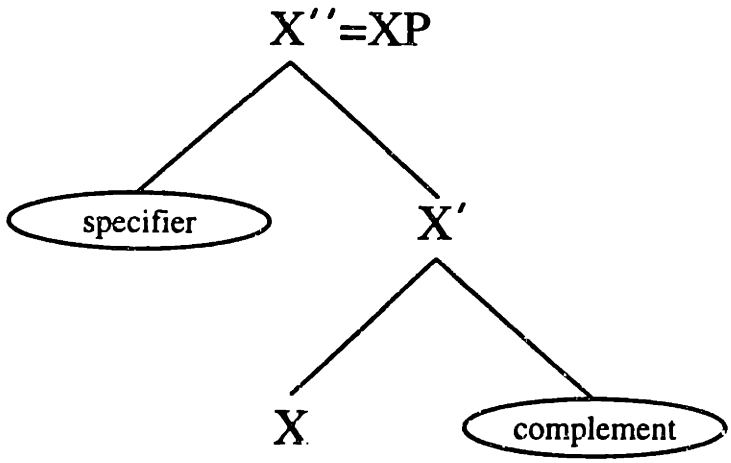


Figure 2-4: A generic X-bar structure for English.

is said to have a first-order projection (an X′ node), which in turn has a second-order projection (an X″ node) that is usually referred to as the head's phrasal or maximal projection (XP). Thus a noun head projects an N′ node and an N″ or NP (noun phrase) node, while a verb projects a V′ and a VP node.

X′-structures are attached to each other to create larger phrase structure trees by placing the XP node of one X′-structure at either the *specifier* or *complement* position (Figure 2-4) of the other X′-structure. The order of the specifier and complement positions with respect to the head varies parametrically across languages; Figure 2-4 shows the order for English. Thus in Figure 2-2, the noun *horse* projects an NP, the determiner *the* projects a DP, and the two are connected: The DP is attached as the specifier of the NP.[6] On the right side of the tree, *raced* projects a VP and has a complement PP headed by the preposition *past*. The PP also has a complement: the NP *the barn*.

The other two X′-structures in Figure 2-2, at the top of the tree, are "functional" projections associated specifically with clauses (Chomsky, 1986a). The higher X′-

[6]Both *the* and *horse* also project X′ nodes (D′ and N′, respectively), but the X′ node (as opposed to either the maximal projection (XP) or the head) is typically ignored when either a specifier or a complement (or both) is not present. Note that in Figure 2-3, the N′ node projected by *horse* is present, because the X′-structure for *horse* has both a specifier (the DP) and a complement (the relative clause).

structure, the CP, is headed by C (or COMP, for complementizer), which is phonologically and orthographically empty for matrix clauses in English. In embedded clauses, however, C will often be associated with *that*, as in *I wish that he would leave*; or *for*, as in *I wished for him to leave.* Similarly, the head I (INFL) of the IP (inflectional phrase) appears to be empty. However, INFL contains grammatical feature information required for subject-verb agreement, tense marking, etc. and is also the position of modal and other auxiliary verbs (*can*, *will*, *should*, *have*, *be*, etc.) when they are present, as in *The horse had raced past the barn.* The importance of these two functional projections can also be seen in the embedded relative clause (*raced past the barn*) of Figure 2-3: As in Figure 2-2, C and I are not associated with overt forms, but as noted above, the reduced relative of Figure 2-3 has a corresponding unreduced form (*that was raced past the barn*), and in this version, C would be associated with *that* and I would be associated with *was*.

Figure 2-3 also contains two NP *traces* (labeled *t*), one in the specifier position of the embedded IP (the position where the subject usually occurs in a clause (cf. the position of *the horse* in Figure 2-2), and the other attached at the position corresponding to the direct object of *raced.* Traces, often referred to as *gaps* or *empty categories*, function as place-holders for other NPs in a syntactic structure.[7] In Figure 2-3, both traces are place-holders for the NP *the horse*: The trace in direct object position (in the VP) indicates that *the horse* behaves in certain ways like the direct object of *raced*, as a result of the clause being passive, as in *The horse was raced past the barn (by someone).* The subject position trace is present because all relative clauses are "missing" one of their phrases, and the missing phrase corresponds to the one to which the relative clause is attached. This is more apparent in sentences like *The man that we saw t was hurt* and *The man that t saw us was hurt*, where *t* indicates the position of the missing phrase corresponding to *the man.*

Thus syntactic structures contain configurational information about both the relative order and the hierarchical grouping into phrases (constituency, defined in terms of dominance) of words in a sentence. This information is used in a number of language

[7]For evidence that gaps are detected in language comprehension, see MacDonald (1989).

comprehension processes, and much of the organization of this structure depends on properties of its lexical items. We will see in more detail below how lexical entries constrain syntactic structure, but first we briefly describe the other major level of representation in the processing system.

The third level of representation that the parser maintains is the discourse model (see Grosz, Pollack & Sidner, 1989, for discussion). This representation contains information about events and entities which have been discussed or are currently under discussion. Unlike the preceding representations, the discourse model includes information from earlier sentences as well as from the current one. It includes at least a specification of which entities/events are in the foreground, what the current topic is, what is known or presupposed about objects in the discourse, and the temporal and logical relationships between objects in the discourse. In addition to serving as the ongoing representation of what is being talked about, the discourse model will also provide information potentially useful to the parser's other processes, such as identifying the antecedents of pronouns (involving consideration of discourse focus and topic, etc.). The discourse model, then, is the level of representation which is most directly relevant to connecting language with other non-linguistic systems, and it has access to both lexical and syntactic information about sentences.

The Lexicon

Psycholinguistic research on the mental lexicon has emphasized the orthographic, phonological, and semantic codes of words and the associative links among them (e.g., Forster, 1976; Seidenberg & McClelland, 1989). Two other types of information are especially relevant to ambiguity resolution: what current syntactic theories such as GB call *argument structures* (e.g., Grimshaw, 1990; Hale & Keyser, 1986, 1987; Levin & Rappaport, 1986; Marantz, 1984; Rappaport & Levin, 1988; Stowell, 1981; E. Williams, 1981; see also Bresnan, 1982; Jackendoff, 1983, 1990; Pinker, 1989) and the pieces of syntactic structure specified by X-bar theory.

Argument Structure

The argument structures associated with a word serve two critical functions. First, they encode the relationships between the word and the phrases that typically occur with it (the word's arguments). Second, argument structures capture important facts about correlations between syntactic and semantic information. As an interface between these two types of information, they play a central role in language comprehension and production. The concept of argument structure is similar to the earlier notion of verb *subcategorization frames*, which indicate the kinds of syntactic phrases that optionally or obligatorally occur with a verb in a sentence (Chomsky, 1965). For example, the subcategorization frame for the verb *put* indicates that it must occur with both a direct object noun phrase (NP) and a prepositional phrase (PP). An argument structure representation specifies the same information (indirectly, when it interacts with syntax), but in addition it provides a limited amount of semantic information about the relationship between the word and each of its associated arguments (including, for a verb, its subject, which is typically excluded from its subcategorization frame(s)). For the verb *put*, its argument structure information captures an English speaker's knowledge that not just any combination of a subject NP, an object NP, and a PP is acceptable for this verb — the subject NP must take the role of Agent (the thing doing the putting), the object NP must be the Theme (the thing being put), and the PP must specify a Location (where the Theme is put). This coarse-grained semantic information about a phrase, such as the requirement that the PP mark a location of some sort, is called the *thematic role* assigned to the phrase (Chomsky, 1981; Fillmore, 1968; Gruber, 1976; Jackendoff, 1972; see also discussion in Dowty, 1991). Argument structures thus specify two things about each phrase that occurs with a word: the grammatical category of the phrase (whether it is an NP, a PP, etc.), and the thematic role assigned to the phrase. This conjunction of syntactic and semantic information in one representation captures interdependencies between the syntactic and semantic information that are important for language processing and ambiguity resolution (e.g., Boland & Tanenhaus, 1991; Carlson & Tanenhaus, 1988; Stowe, 1989; Tanenhaus & Carlson, 1989).

The number of discrete thematic roles is assumed to be finite and relatively small, although we do not attempt to provide an exhaustive list. In addition to Agent, Theme, and Location, some important thematic roles are listed in Table 2.3. According to some theories, thematic roles are not themselves primitives in linguistic representation; rather they reflect combinations of finer-grained semantic features such as animacy, sentience, humanness, volition, causation, and movement (Dowty, 1991). Agents, for example, are nearly always animate and capable of causing an action, and they are typically volitional. Themes tend to be at the other end of a continuum from Agents with respect to most of these features: They typically undergo movement and are the objects of causation, rather than the causers of any particular action. Although sentience and humanness are not necessarily required of Agents, such features are likely to be more strongly associated with Agents than with Themes. Instruments will tend to be close to Agents in terms of causation, but not in terms of animacy and volition. Similar featural relationships hold between other sets of roles as well, and thus thematic roles can at least be taken to reflect generalizations about co-occurring clusters of more primitive semantic features.

Argument structures in the lexicon. We assume that the lexical entry of (at least) each verb, preposition, noun, and adjective includes a representation of argument structure information. (11a) shows the argument structure for the verb *put*, and (11b) shows a corresponding sentence, in which *Mary* is the Agent, *the book* is the Theme, and *on the desk* is the Location.

(11)

a. put: <Agent, Theme> <Location>

b. Mary put the book on the table.

The argument structure for *put* contains two thematic grids, indicated in angle brackets.[8] The "core" thematic grid is on the left; the phrases that take the thematic

[8]Our thematic grid notation differs slightly from related forms in the literature (cf. Levin & Rappaport, 1986; Marantz, 1984; E. Williams, 1981) in its specification of core versus non-core grids. However, the notion *thematic grid* is at best shorthand for a more complex argument structure

Table 2.3: Some Major Thematic Roles

Role	Function	Examples (Illustrated Role *Italicized*)
Agent	Performer of action	*Mary* ran. *John* cut the cake.
Theme[a]	Thing acted on or perceived	John cut *the cake.* Mary saw *the race.*
Instrument	Device used in action	John cut the cake *with a knife.*
Location	Location	John put the cake *on the plate.*
Goal	Endpoint of action	John gave the cake *to Mary.* John gave *Mary* the cake.
Source	Starting point of action	Mary took the cake *from John.*
Beneficiary	Beneficiary of action	John baked the cake *for Mary.*
Manner	Description of action	Mary took the cake *with a smile.*
Attribute	Description of a thing or event	Mary took the cake *with green frosting.*
Experiencer	Perceiver or experiencer of mental state	Mary frightened *Susan.* *Susan* feared Mary.
Proposition	Statement, question	Mary knew *the race would be difficult.* John wondered *who frightened Susan.*

[a]Some descriptions of thematic roles use the term *Patient* for *Theme*; others use both terms to distinguish percepts from things affected by actions. Similar subdivisions are possible for some other roles such as Proposition, but we will not rely on any of these distinctions.

roles specified in the core grid are licensed by the verb without any additional support, so they appear in the sentence as NPs (*Mary* and *the book*, in (11b)), except in the case of the role Proposition, which appears as a complete embedded clause (CP). Thematic roles in non-core grids (such as the <Location> grid for *put*) can only be assigned indirectly, with the help of prepositions, so they will be assigned to PPs (such as *on the table*). The underlined position in a core grid corresponds to an "external" argument (external to the verb phrase), which is usually the verb's subject (*Mary* in (11b)). Non-underlined positions are internal arguments and are either direct, if they are part of a core grid, as for the Theme *the book* in (11b); or indirect, if they are part of a non-core grid, as for *on the desk*. This same notation applies to the argument structure of words other than verbs. Nouns, for example, cannot license NPs (either as external or internal arguments), so their core grids are often empty, but some can take clausal complements (CPs), as in (12). The thematic grid in (12a) is *report*'s core grid, and in (12b), its role is assigned to the direct internal argument *that Bill died*. Similarly, prepositions take direct internal arguments (direct objects, usually, because they are usually NPs), just as verbs do: *With* in (13) assigns the Instrument role to its direct internal argument (the NP *the hammer* in (13b)).

(12)

a. report: <Proposition>

b. The report that Bill died was false.

(13)

a. with: <Theme>

b. Carol hit the nail with the hammer.

Thus, argument structure representations specify information about the category of associated phrases (e.g., NP vs. PP vs. CP), the thematic roles of the phrases,

representation, in which the core versus non-core distinction, the external versus internal distinction, the direct versus indirect distinction, and, probably, the identity of assigned thematic roles, are all derivable from more basic properties of the representation, possibly in interaction with the syntactic system. See, for example, Hale and Keyser's (1992) lexical relational structures, Rappaport and Levin's (1988) predicate argument structures, and Pinker's (1989) lexicosemantic structures.

and their syntactic position (e.g., subject, direct object). We will assume that this argument structure information is simply listed as part of the lexical representation of a word.

Because many words can appear with more than one configuration of arguments, they will be associated with several different argument structures. The examples in (14) illustrate three common English core thematic grids, for the verb *cook*.

(14)

a. cook:	<Agent, Theme>	John cooked the stew.
b.	<Agent>	John cooked.
c.	<Theme>	The stew cooked.

In (14a), *John* is the Agent and *the stew* is the Theme. We refer to argument structures like that shown in (14a), in which the core grid contains both an external and an internal argument, as *transitive* argument structures, and corresponding sentences with both a subject and direct object (as in (14a)) as transitive sentences. (14b) and (14c) are *intransitive* argument structures because either an internal (14b) or external (14c) argument is not present, and the corresponding sentences are intransitive as well.[9] Note that (14b), in which *John* is an Agent, indicates that John cooked something, whereas (14c), in which *the stew* is the Theme, indicates that the stew is itself being cooked. Although the Theme role is assigned to the direct internal argument, the NP receiving the role appears as a subject in the corresponding sentence, because no external argument is present.

The argument structure representations in (14) are part of the lexical representation of *cook*. The fact that the verb is associated with more than one argument structure creates an ambiguity that must be resolved in the comprehension process. When the comprehender encounters a string such as *John cooked...*, an argument

[9]For historical reasons, intransitives with only an external argument (14b) are often termed *unergative*, while those with only an internal argument (14c) are called *unaccusative*. The two types behave differently in many languages on a number of tests: Subjects of unaccusatives often behave like direct objects of transitives, while subjects of unergatives typically behave like subjects of transitives. The origin and appropriate characterization of these differences has been extensively debated, and while our notation obviously captures the distinction, we will have nothing to say about its origin. See Levin and Rappaport Hovav (1992) and references cited therein for discussion.

structure for *cooked* must be selected and one of its thematic roles must be assigned to *John*: Either *John* is the Agent (the cooker) or the Theme (the entity cooked). This type of ambiguity illustrates a crucial difference between our proposal and other characterizations of parsing. Previous theories have emphasized syntactic phrase structure ambiguity; in this example, the syntactic ambiguity concerns the sort of verb phrase to construct for *cooked*. These structures were assumed to be built on the basis of phrase structure rules, and proposals about the architecture of the parser described strategies for dealing with ambiguities defined at this level. In the Frazier et al. model, for example, the parser builds only the single simplest structure consistent with the rules of the grammar; the assignment of thematic roles and other aspects of comprehension are secondary and temporally distinct. By contrast, we are claiming that this ambiguity has a fundamentally different character: It is lexical insofar as it derives from the fact that the lexical representation of a verb such as *cook(ed)* contains multiple argument structures. The ambiguity is resolved not by strategies governing the application of structure-generating rules, but rather by the processes that govern the resolution of other types of lexical ambiguities.

Lexical Syntax

As we saw above, X-bar syntax is used in current linguistic theory to construct syntactic phrase structure trees from coherent smaller pieces, each associated with an individual lexical item. From our perspective, then, X-bar structures are just another type of information that is part of a lexical representation. Just as a word has associated argument structure(s) in its lexical entry, it will also have associated X′-structures: A verb will have an associated V head, which will be to linked to (project) a V′ node, which will in turn be linked to a VP node; a noun will have associated N, N′, and NP nodes, etc. Words which are ambiguous between multiple grammatical categories will then have multiple X′-structures as well, each associated with the appropriate grammatical category. Furthermore, with X′-structures in the lexicon, it is possible to treat the entire process of constructing a syntactic tree (syntactic processing) as a matter of connecting X′-structures in the lexicon to each other in

the appropriate (syntactically-required) manner. Thus under this proposal, we could conceivably do away with the idea of an independent syntactic level of representation, with properties completely unrelated to (and unconstrained by) those of other parts of the processing system.

Associating X′-structures directly with their lexical entries has another advantage as well, in that it provides a way of directly linking syntactic knowledge with other aspects of lexical representation. As we saw above, argument structures do much of the work of constraining what the phrases appearing with a word are like, and these constraints are enforced by syntactic principles (e.g., the Theta Criterion and the Projection Principle in GB (Chomsky, 1981)) in combination with a set of rules linking argument positions to positions in syntactic structures.[10] With X′-structures linked directly to individual lexical items, the rules linking arguments to syntactic positions can be specified by directly linking positions in a lexical entry's argument structure(s) to the corresponding positions in the lexical entry's X′-structure(s). We will see below in more detail how the argument structure and syntactic structure representations constrain each other, and how linking them tightly as we have proposed provides an appropriate mechanism for the use of lexical information in the resolution of syntactic ambiguity.

Other Lexical Representations

In addition to argument structures and X-bar structures, lexical representations contain several other types of information. As mentioned earlier, there is information about the grammatical category of lexical items (noun, verb, etc.); there is also other grammatically-relevant featural information, specifying tense (past, present, future), finiteness (finite or infinitive), voice (active or passive), number (singular or plural), person (first, second, third), and gender (masculine, feminine, neuter; in English, this is only realized in the pronominal system: *him*, *herself*, *itself*, etc.). The possible

[10]Specifying the linking rules is beyond the scope of this article. Their formulation depends both on a number of theory-internal syntactic details and on a more finely articulated theory of argument structure. See, for example, Pinker (1989), as well as most of the other references cited above in the discussion of argument structure.

thematic roles which a noun can take will also be listed, so that the lexical entry of *desk*, for example, would include the roles Theme, Location, perhaps Instrument, etc. Morphological information is available as well, though English morphology is incomplete in a number of respects and is often another source of ambiguity. For example, only about 50 English verbs have different past tense and past participle forms (e.g. *broke/broken*, *drank/drunk*); for the rest, the two forms are identical (*raced*, *walked*, *sat*, etc.). The latter verbs are therefore associated with a morphological ambiguity.

Putting all these different pieces of information together, a lexical representation for some word will therefore include a representation of the word's phonological form, orthographic form, semantics, grammatical features (including grammatical category and allowed thematic roles), morphology (at least inflectional), argument structure, and syntactic structure, as illustrated for the word *cooked* in Table 2.4. Words that are associated with more than one representation of a particular type have all representations listed as alternatives. In the model we propose, ambiguities arise when such words appear in an input. As a simple example, *calf* has two alternatives at the level of semantic representation ("young cow," "lower part of the leg"), and when the word is encountered as part of an input, the comprehender must determine which one is appropriate. A more complicated example is *cook*, which is ambiguous at several levels: It has two alternative grammatical categories (noun and verb), which are associated with different though related semantics (roughly, "an individual who prepares food by heating it" vs. "to prepare food by heating it"). In addition, in the verb case, it can be singular or plural (*I/We cook mussels every morning.*), finite or infinitive (*I cook mussels.* vs. *To cook mussels is a great pleasure.*), and first, second, or third person (*We/You/They cook mussels.*). The verb form of *cook* will also be ambiguous with respect to argument structure, as shown in (14) above, and the noun form will be ambiguous with respect to its thematic role. When *cook* is encountered in an input, a single representation for each of these levels must be selected.

Table 2.4: A Sample Lexical Entry (for *Cooked*)

Level of Representation	Representation
Phonology	(cooked)
Orthography	cooked
Semantics	"to have prepared food by heating it"
Morphology	past tense past participle
Voice	active passive
Finiteness	finite
Tense	past
Grammatical Category	verb adjective
Argument Structure (core)	<Agent, Theme> <Agent> <Theme>
Argument Structure (non-core)	<Instrument> <Manner> <Location> etc.
X-bar Structure	(VP structure for each core argument structure; also AP (adjective phrase) structure for core <Theme> grid)

Processing Assumptions

Having specified the types of knowledge structures in the lexicon, we turn to how they are used in processing. In keeping with the idea that syntactic ambiguities are lexically-based, our processing assumptions are taken directly from recent models of word recognition, particularly the interactive-activation models of McClelland and Rumelhart (1981) and Elman and McClelland (1984) and the lexical ambiguity resolution model of Kawamoto (1993). Our main goal is to identify basic characteristics of the system, rather than to describe how these characteristics can be implemented in specific computational models. All of the mechanisms we propose have been explored in previous modeling research, which provides a basis for assuming that they are not unrealistic with regard to their computational demands. There are no novel ideas about processing being introduced here; rather, there is the application of some familiar computational concepts in a new domain, syntactic ambiguity resolution. For an example of how parts of such a model might be implemented, see Pearlmutter, Daugherty, MacDonald and Seidenberg (1993), who describe a constraint satisfaction model of some aspects of thematic role assignment.

1. Activation. We assume a basic activation metaphor, in which units (or sets of units) corresponding to each type of information are activated in the course of processing. Although it appears innocuous and well-grounded in previous modeling research, this assumption carries with it some potentially controversial implications concerning the treatment of syntactic information. Syntactic structure, in our view, consists primarily of pieces of X-bar structure encoded with individual lexical items, and like meanings, tenses, argument structures, and every other part of a lexical representation, X-bar structures are assumed to be activated in the course of processing. Thus, our approach entails the idea that syntactic structures can be partially activated, a concept strictly at odds with the earlier view that syntactic processing involves constructing one or more representations incrementally. The parser developed by Marcus (1980), for example, explicitly built a phrase structure representation in the course of processing a sentence; Marcus' "determinism hypothesis" held that

once the parser built a piece of syntactic structure, it could not be revised. The debate over serial versus parallel parsers (e.g., Frazier & Rayner, 1982; Gorrell, 1987; Kurtzman, 1985) similarly concerned whether the parser (human or computer) built one tree structure or multiple structures on-line. In our approach, pieces of syntactic structure associated with lexical items can be activated to differing degrees; the only requirement is that the system settle on one of the alternative X-bar structures eventually for comprehension to occur. The notion of partially activating syntactic structures is quite foreign to the earlier approaches, although it follows naturally from the assumption that syntactic structures are stored as parts of lexical entries.

2. Frequency. Each component of a lexical entry carries information about its frequency of occurrence in the language. For example, the representation of the verb *raced* includes the information that it has two associated morphological tense forms, simple past tense and past participle; the frequency with which *raced* occurs with each form would be encoded as part of this representation. If frequencies are encoded for the syntactic aspects of lexical representation in the same way the frequencies of alternate meanings are encoded, lexical frequency information can affect syntactic ambiguity resolution in the same way that meaning frequencies affect lexical ambiguity resolution. This information plays a large role in resolving verb-based syntactic ambiguities. That humans encode information about the frequencies of occurrence for a broad range of entities is well-established (see, for example, Jacoby & Brooks, 1984); here we merely assume that they encode frequencies for all the types of information stored in the lexicon. As in many models of word recognition, we will assume that frequency information is encoded by activation level, although there are other ways this aspect of the system might be implemented (e.g., by encoding individual instances).

3. Constraint satisfaction. In resolving ambiguities, the processing system exploits three main characteristics of language: First, the different types of information associated with a word are not independent of one another. Thus, progress in resolv-

ing an ambiguity at one level provides information relevant to resolving ambiguities elsewhere in the system. Second, a word will not necessarily be equally ambiguous at all levels of representation. For example, a word can be ambiguous as to tense but unambiguous with regard to grammatical category. Third, even when the grammar admits multiple alternatives at a given level of representation, they often differ substantially in frequency and thus a priori probablity of occurrence. Ambiguity resolution is therefore a classic example of a constraint satisfaction problem (Rumelhart & McClelland, 1986): Multiple interdependent, partially redundant, probabilistic sources of information interact to allow the system to settle on an interpretation at each level. The fact that the processing system exploits these characteristics allows ambiguities to be resolved efficiently.

The computational properties of constraint satisfaction mechanisms have been studied in considerable detail. Rumelhart (1977) provided an early, informal discussion of this type of processing in language comprehension. St. John and McClelland (1990) developed a model that applied to some of the same phenomena we consider here, and our account draws on observations in McClelland (1987). MacWhinney and Bates (1989) have also explored constraint satisfaction mechanisms in language comprehension in their Competition Model. Constraint satisfaction concepts are beginning to have a significant impact on linguistic theory as well, particularly in morphology and phonology (e.g., Prince & Smolensky, in press; and references cited therein); Stabler (1993) describes an implementation of Government-Binding theory as a constraint satisfaction system. Connectionist models provide one way of implementing constraint satisfaction, but there are nonconnectionist schemes as well. In fact, the development of constraint satisfaction systems predates the emergence of contemporary connectionist theory (see Hummel & Zucker, 1983; Mackworth, 1977). Our focus is on how the constraint satisfaction process applies to the issue of syntactic ambiguity resolution; our claims are largely neutral with respect to whether this process is realized by a connectionist network, a set of symbolic rules or simultaneous equations, a production system, or some other means.

For descriptive purposes, and because we have hypothesized that syntactic am-

biguity resolution can be subsumed under principles identified in connection with lexical processing, we will assume the general framework of an interactive-activation model (McClelland & Rumelhart, 1981; Elman & McClelland, 1984). A fragment of a lexical representation is provided in Figure 2-5 for an ambiguous verb that can trigger the MV/RR ambiguity (e.g., *raced*, *pushed*). Only three levels of representation are shown, and each type of information is represented by a set of units representing the alternatives that are available for a given word. Thus the tense morphology level has the alternatives past tense and past participle, the voice level has the alternatives active and passive, and the argument structure level is shown with the transitive (assigning both an Agent and a Theme thematic role) and intransitive (assigning only an Agent) core thematic grids, although many verbs will have more than two argument structure options. It should be apparent that these units are placeholders for more complex representations whose details are important but not immediately relevant to the level at which we are describing the behavior of the system. As in the TRACE model (Elman & McClelland, 1984), there are inhibitory connections (shown as heavy, curved lines) between the alternatives within a given type. The connections between types of information reflect dependencies dictated by the grammar of the language: Where the grammar permits a relationship between two types of information (e.g., passive and past participle, active and transitive), the connections are assumed to be excitatory; where the grammar rules out a relationship between two alternatives (e.g., past tense and passive), the connection between the relevant units is inhibitory.[11]

Figure 2-5 captures some important interdependencies between these three levels of representation, particularly that the passive voice, as in the reduced relative construction or in simple passives such as *The horse was raced*, is permissible with only one alternative at each of the other two levels of representation — the transitive argument structure and past participle tense morphology, respectively. Passive sentences with past tense verbs are ungrammatical, as in **The horse was gave to the man*, as are passive uses with intransitive verbs, as in **The horse was slept*. Active

[11] This is a theory about how some grammatical constraints will be represented, not where they come from. We are not concerned here with issues of acquisition.

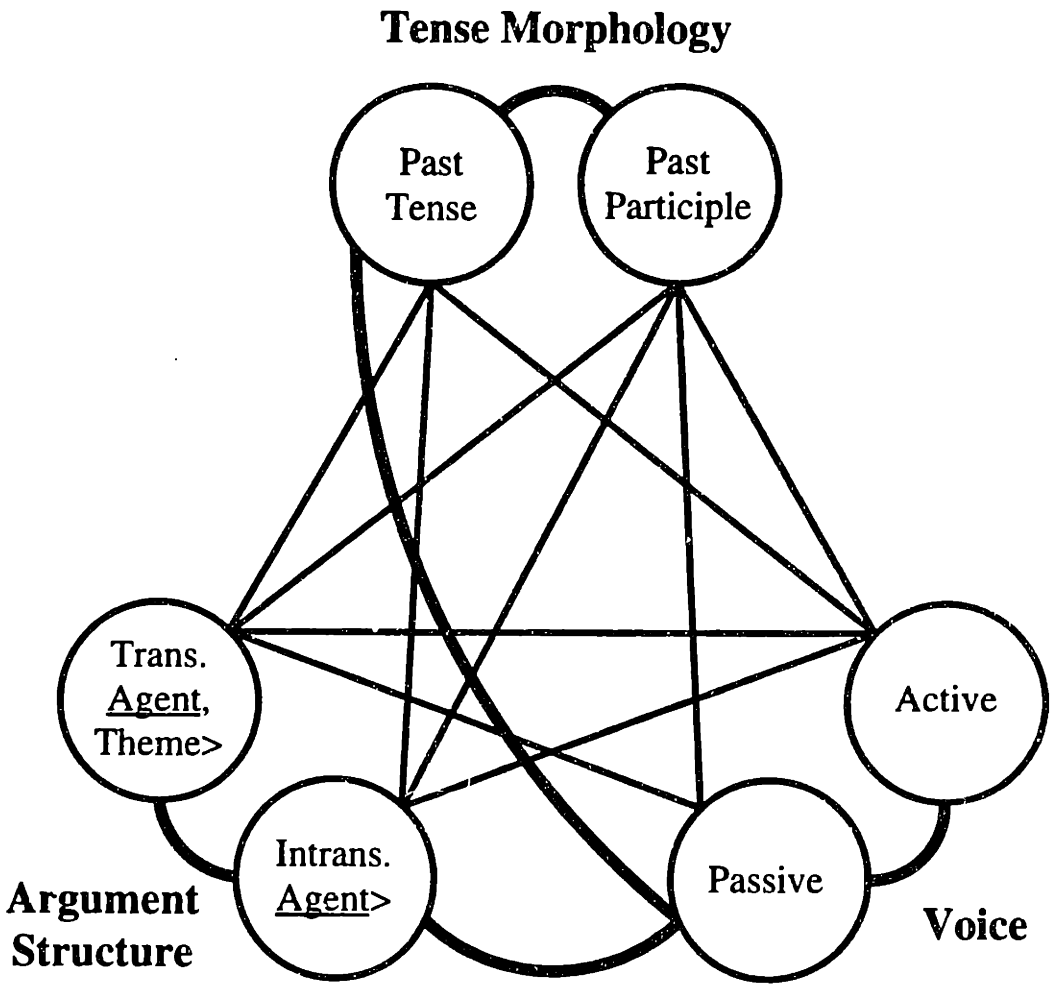


Figure 2-5: Fragment of a lexical entry for an ambiguous verb.

sentences, however, may appear with either transitive (*The horse pulled the wagon*) or intransitive argument structures (*The horse slept*). And they may occur with past tense verbs, as in these last two examples, or with a past participle when a form of *have* is present, as in *The horse had pulled the wagon* or *The horse has slept.*

Figure 2-5 also illustrates several of our claims about basic aspects of processing. First, processing involves the spread of activation between units in the system (initiated by input from orthographic or phonological representations not illustrated in the figure). The degree of activation is an index of the amount of evidence in favor of a particular hypothesis (e.g., that the morphological form is past tense or past participle), and the behavior of each unit in the system depends on the behaviors

of the other units to which it is directly or indirectly connected. The net activation of a unit is dynamically updated to reflect these effects. Second, disambiguation involves activating one alternative of a given type and inhibiting all others of the same type. We envision this as a winner-take-all process of the sort implemented in previous interactive-activation models (e.g., Elman & McClelland, 1984). Third, processing is *analytically exhaustive*, meaning that disambiguation must be achieved at all levels of lexical representation. The idea here is that the comprehension system does not have the option of ignoring individual types of lexical information. Thus, sentences are processed and ambiguities are resolved at all possible levels of linguistic representation. This strong assumption will probably have to be relaxed in order to accommodate cases in which the communicative goals of the reader or listener can be achieved with only a partial analysis of a sentence. We view these as deviations from the normative situation, however.

By applying these principles to Figure 2-5, it becomes clear that although the active voice has excitatory connections to both alternatives at the other two levels of representation, during the course of processing, the past tense and intransitive argument structure units will end up more strongly supportive of the active voice than the past participle and transitive argument structure. This is because the past tense and intransitive nodes have inhibitory connections to the passive voice, which in turn has an inhibitory connection to the active voice. By their inhibition of this inhibitor of the active voice, activation of the past tense and intransitive nodes allows the active voice interpretation to accrue more activation than it would otherwise be able to do. This is a standard sort of effect in interactive-activation models.

4. Context effects. So far we have described constraints that hold between various aspects of lexical representations. Words occur in sentence contexts, however, and it is therefore necessary to consider how the processing of a word is affected by the context in which it appears (see Rayner & Pollatsek, 1989, for review). The literature is divided on this issue. Some theories suggest that word processing is highly dependent on context (e.g., McClelland, 1987; Marslen-Wilson, 1975), whereas others

stress its relative autonomy (Forster, 1979; Fodor, 1983). When context effects have been observed, they have been interpreted as evidence for an "interactive" model of processing (McClelland, 1987). One of the weaknesses of this approach, however, has been the lack of specificity about the kinds of information that can potentially interact and the dynamics of interactive processing. Questions about what kinds of contextual information facilitate processing, what kinds of processes are facilitated, and what factors influence the degree of contextual effect all remain unanswered. Many studies have assessed the processing of words in "congruent," "incongruent," or "neutral" contexts (Stanovich & West, 1979). Congruency is defined operationally (e.g., on the basis of ratings or intuition) rather than in terms of the types of information that cause the context to be congruent or incongruent. Such studies provide important evidence concerning the existence of contextual effects, but not about the mechanisms underlying the effects or about their scope. A similar approach attempts to relate contextual effects to statistical measures such as Cloze probability, a measure of the predictability of a word in context (e.g., Fischler & Bloom, 1979). Rather less attention has been paid to the factors that determine the Cloze value in any given case. The MacWhinney and Bates (1989) Competition Model, which is similar in spirit to our proposal, raises similar issues. The model assumes that languages provides "cues" that interact ("compete") with one another in the course of processing. The model does not provide an independent characterization of what kinds of information could potentially serve as cues, however (Gibson, 1992).

One way to view these "cues" is in terms of the relevance of contextual information to the task of resolving the different types of ambiguity in the lexicon. By hypothesis, contexts are informative to the extent they facilitate deciding between alternatives at one or more levels of lexical representation. A "highly constraining" context is one that provides information that disambiguates at several levels. The limiting case is a context in which a target word is entirely predictable; that is, the context provides information that disambiguates the identity of the word itself. A less-constraining context provides information relevant to some of the ambiguities associated with a word. For example, it might constrain the grammatical category of a word but not

its identity.

These considerations suggest that there are two senses in which ambiguity resolution involves "interactive" processing. First there are *vertical* dependencies among the different types of information stored as part of the lexical entry for a word; some of these dependencies are represented in Figure 2-5. Thus, for example, morphological form is correlated with voice and argument structure. Seen in this light, the McClelland and Rumelhart (1981) interactive-activation model of word recognition was concerned with vertical interactions among orthographic representations (features, letters, and word-forms). We now extend this notion to other types of information associated with a word. Second, there are what can be termed *horizontal* dependences between a word and the context in which it occurs. These dependencies obtain wherever contexts provide information that is relevant to resolving an ambiguity at one of the levels of representation in the lexical entry.

5. Lexical dominance and contextual constraint. Given this theoretical account of how contexts affect lexical processing, the functional characteristics of the ambiguity resolution process depend on facts about the structure of the lexicon (e.g., the extent to which words exhibit ambiguities across different types of information) and about the kinds of information provided by sentences. They are also determined by how the lexical and contextual sources of constraint are integrated in the course of processing. The constraints that come from within the lexicon and from the broader context vary continuously in strength. For example, contexts vary from being very unconstraining (following *The next word in the sentence is...* almost any word could occur) to very constraining (*I drink my coffee with cream and...*). Thus, we would expect to observe a broad range of outcomes from the interaction between lexical and contextual sources of constraint with regard to the course of the ambiguity resolution process. In the well-studied area of meaning ambiguity, this broad range of outcomes has been obtained; thus, some studies show activation of multiple meanings of ambiguous words in biasing contexts (e.g., Swinney, 1979; Tanenhaus et al., 1979), whereas other studies using similar methodologies have yielded evidence for selective

activation of the contextually appropriate meaning (Simpson, 1981; Tabossi et al., 1987).

Given that the processing system allows this broad range of outcomes, the functional characteristics of lexical-contextual interactions will depend on the structure of both the lexicon and the natural language sentences that are actually produced. Some general tendencies regarding these interactions have already been noted in the literature. Recall Marslen-Wilson's (1987) observation about "bottom-up priority" in lexical processing. Working in the domain of auditory word recognition, Marslen-Wilson observed that the computation of lexical information seems to dominate over the effects of contextual constraints on lexical processing. The Kawamoto (1993) model can be seen as instantiating this view, insofar as it was structured so that the associations between different types of lexical information (alternative meanings, orthography, phonology) were stronger than the associations between contextual information and specific parts of the lexical entry (e.g., one of the meanings). We can now see that "bottom-up priority" derives from two general characteristics of the processing system. First, there are facts about the nature of contextual constraints. The kinds of information provided by natural language contexts tend to be very useful in deciding between alternatives at a given level of ambiguity. However, they tend to be less effective at pre-selecting one of the alternatives (Seidenberg et al., 1982). Consider, for example, how contexts might affect the resolution of the morphological ambiguity associated with *raced*. Given the structure of English, it is often the case that a context will admit only the past tense (or only the past participle) form. Following *The man who...*, for example, *raced* can only be a past tense form. The question then is how the processing system makes use of this information. One possibility is that knowledge that only a past tense would be grammatical in this context obviates the morphological ambiguity entirely. Thus when *raced* is recognized, only the past tense is activated. This is the way that context was assumed to work in "selective access" models of lexical ambiguity resolution.

Several considerations suggest that this is not how contextual constraints function in general, however. First, the context *The man who...* does not dictate that the

next word must be *raced* or even that it must be a verb: Many other continuations are possible (*The man who people think...*, *The man who slowly...*, etc.). Whereas the likelihood that the next word in *The man who...* will be a past participle is effectively zero, the likelihood that it is a past tense is not much higher, and it is unlikely that the processing system is set up to anticipate relatively low frequency events. Second, in order to make use of this information about the relative likelihood of the past tense, the lexicon would have to be organized in such a way as to prime the activation of all past tenses, not merely the past tense form of *raced*. There are no substantive proposals about how this might be accomplished computationally. Finally, given the rapid time course of processing events, it is not clear that pre-activating all of the past tenses would be computationally efficient or provide a net benefit. Within a short amount of time, the actual word will be decoded, providing additional information relevant to resolving the ambiguity.

One way to summarize these observations is to say that although the architecture of the system affords the possibility of continuous interactions between contextual information and the lexicon, the *effects* of contexts tend to be more retroactive than proactive. Whereas the information provided by a context is often only weakly constraining, it is typically quite sufficient to allow the comprehender to distinguish between a small number of alternatives (see Seidenberg et al., 1982, and McClelland, 1987, for discussion). Although the context *The man who...* may not specify that the next word must be a past tense, once *raced* is identified as the next word, the only alternative morphological tense forms will be past tense and past participle, and the context will provide enough information to determine which is appropriate. In short, less information is needed in order to discriminate between two alternatives than to preselect one of them out of a whole range of possibilities. Thus, the nature of contextual constraints is such that they will typically provide useful information once lexical processing has yielded partial activation of multiple alternatives, but they typically do not provide sufficient information in advance to restrict activation to the contextually-appropriate alternative. Of course, these are general tendencies rather than absolute restrictions. Clearly, there are some contexts that are so constrain-

ing as to make a particular word highly predictable (*I drink my coffee with cream and...*), and the processing system can apparently exploit this degree of contextual constraint when it is available (Fischler & Bloom, 1979). However, this high degree of predictability is apparently achieved relatively infrequently, at least in English texts (Gough & Cosky, 1977).

In summary, lexical items differ with respect to degree of ambiguity and the frequencies of the alternatives at each level of ambiguity; contexts differ in terms of the degree to which they are informative about these ambiguities. The combination of these two general properties of language yields a processing system that is *contextually constrained* but *lexically dominated.* This account largely emerged from studies of meaning ambiguity. Given our view that syntactic ambiguities derive from other types of lexical ambiguities, it follows that the processing of these ambiguities should exhibit similar characteristics vis a vis the interactions of lexical structure and contextual constraint. We discuss evidence bearing on this prediction below.

6. Syntactic processing. The processing mechanisms described above were assimilated from previous work on lexical processing but with slight elaboration, they will provide a framework for the computation of syntactic structure. We saw above that it is possible, by treating X′-structure as yet another level of representation in the lexicon, to "compute" syntactic structure in the lexicon, by activating the X′-structures associated with individual words in the input and connecting them in the appropriate manner.

With the processing assumptions outlined above, we can conceptualize this as shown in Figure 2-6 (cf. Stevenson, 1990, 1993a, 1993b, for a related proposal): All the structural positions where the two X′-structures shown can attach to each other (the circled positions) are connected, so attaching one X′-structure to the other is a matter of activating the appropriate connection between them.[12] Because X′-structures are linked to each other by attaching one X′-structure's XP node into another X′-

[12]Each of the connectable positions in these X′-structures will of course also have connections not shown, to all the other X′-structures in the lexicon (modulo certain syntactic constraints; see below).

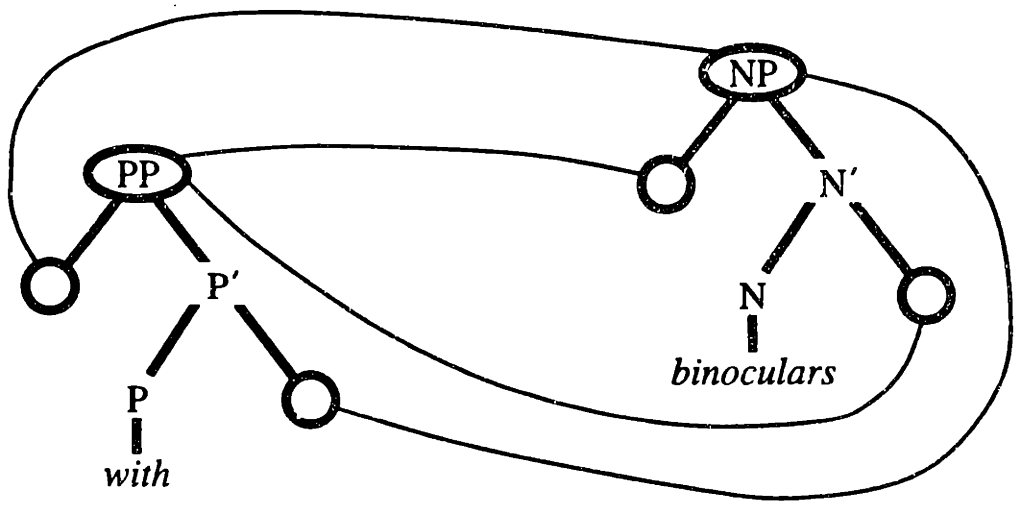


Figure 2-6: Connections between two X′-structures in the lexicon.

structure's specifier or complement position, a particular connection between two X′-structures in the lexicon will always link a specifier or complement position with an XP node. Thus activating a particular connection corresponds to attaching the XP node at one end of the link into the specifier or complement position at the other end of the link.

For example, in the top panel of Figure 2-7, activating the highlighted connection (as opposed to the other possibilities shown) indicates that the PP headed by *with* attaches into the complement position of the NP headed by *binoculars*, as in *binoculars with a scratch*. The corresponding fragment of a phrase structure tree is shown in the bottom panel of Figure 2-7; the highlighted connection in the top panel corresponds to the connection between the N′ and PP nodes in the bottom panel. In the top panel of Figure 2-8, a different connection between the same two X′-structures is activated, and this corresponds to a different (fragment of) structure: The NP headed by *binoculars* attaches into the PP headed by *with*, as in a fragment like *the spy with binoculars*. The corresponding phrase structure tree is shown in the bottom panel of Figure 2-8; the highlighted connection in the top panel is the connection between the PP and NP nodes in the bottom panel.

Thus, by combining the notion of activation with the idea that X′-structures are

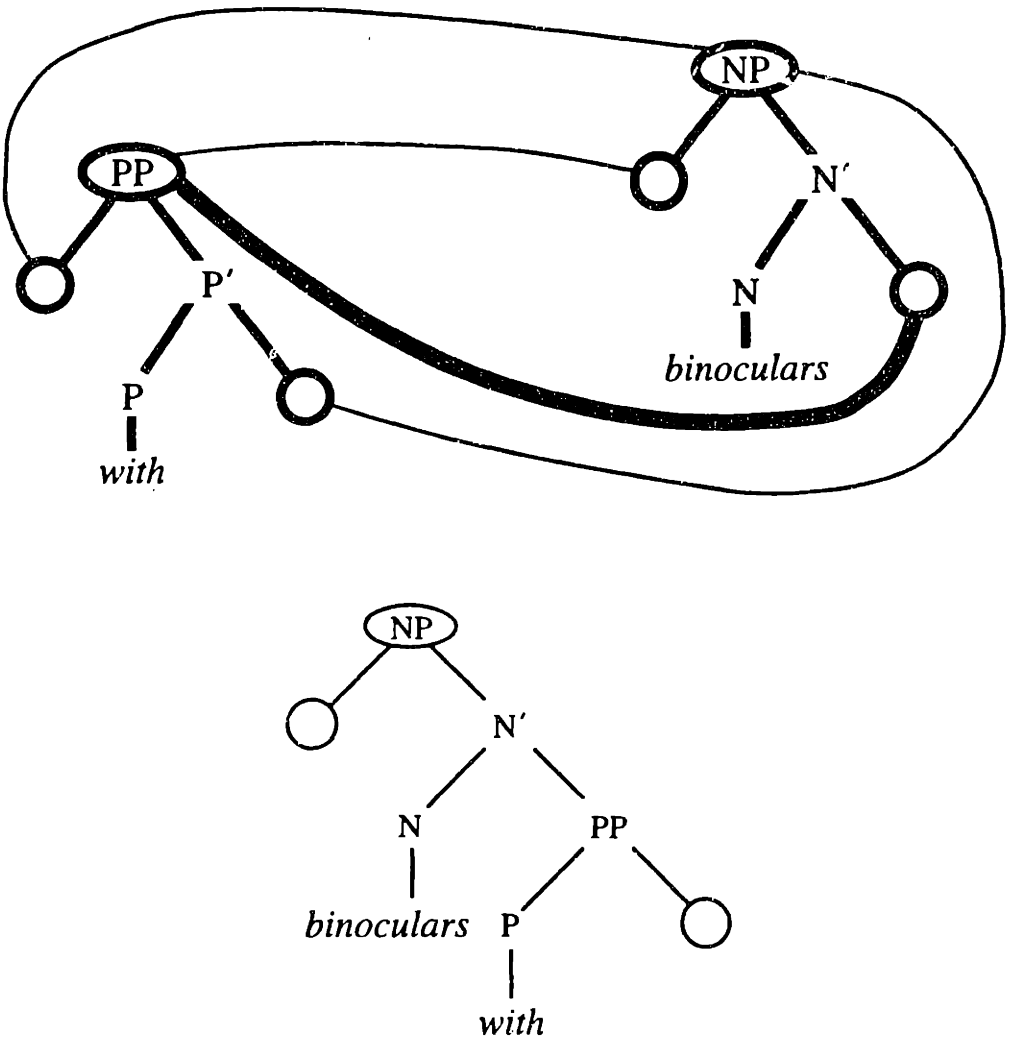


Figure 2-7: Lexical syntax representation (top) and corresponding phrase structure fragment (bottom) for ... *binoculars with*...

a part of lexical entries and can be connected together to form syntactic structures within the lexicon, we get a system in which X′-structures and larger pieces of syntactic structure can be activated to varying degrees. Furthermore, as with other representations within the lexicon, mutually exclusive X′-structures and mutually exclusive connections between X′-structures will inhibit each other, so that there is competition between alternative syntactic analyses, and the system will be forced to choose one.

In this conceptualization of syntactic tree-building, syntactic rules or principles

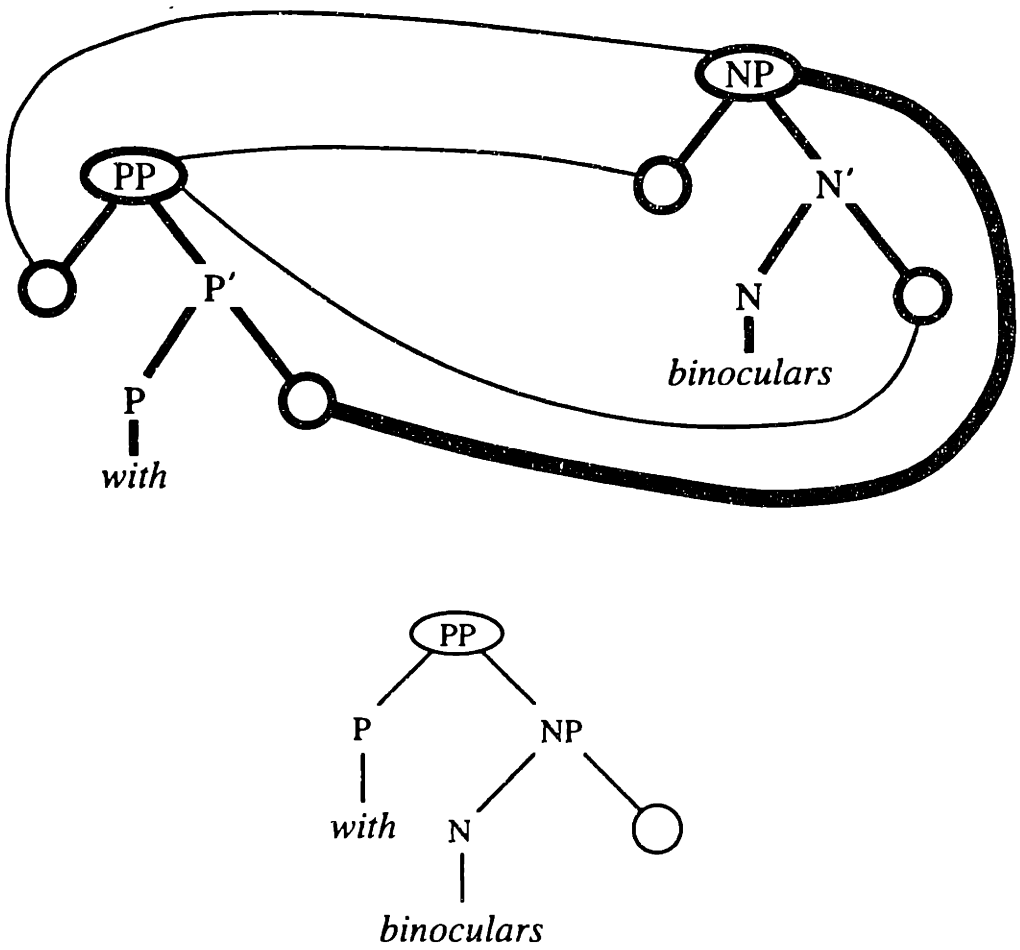


Figure 2-8: Lexical syntax representation (top) and corresponding phrase structure fragment (bottom) for ... *with binoculars*...

operate as strong constraints on which attachments between X′-structures are permitted, either by preventing certain connections from becoming highly activated during processing, or by preventing even the existence of certain connections. In a theory like Government-Binding, in which syntactic principles rule out certain structural configurations, any connection between X′-structures would be allowed in general, but the syntactic principles will rule out those connections corresponding to ill-formed structures. In a theory like Generalized Phrase Structure Grammar, which provides phrase structure rules that directly generate syntactic trees, the only allowed connections between X′-structures would be those listed in the phrase structure rules of the grammar. In either case, the grammar will provide strong constraints and will

rule out most possible attachments in many situations; and these constraints will be enforced by the pattern of connectivity among the X′-structures in the lexicon.

A Simple Processing Example

Having now described both the representations and the processes at work in our model, we can illustrate how they interact a bit more dynamically by describing the processing of a simple example. As we saw above, the verb *cook* has, as part of its lexical entry, three different core thematic grids, repeated below, along with example sentences.

(14a) cook:	<Agent, Theme>	John cooked the stew.
(14b)	<Agent>	John cooked.
(14c)	<Theme>	The stew cooked.

Furthermore, each of these core grids is associated with an X′-structure, and each of the roles in each grid is linked to some position in the associated X′-structure.[13] Thus during the course of processing, the system must settle on a choice of a single argument structure and X′-structure for *cook*. The grammar will also force a thematic role to be assigned to each noun phrase, and the thematic role assigned will have to be the role linked to the position where the noun phrase (or a trace of the noun phrase, as described above; see below as well) attaches in the syntactic structure.

For example, when *John* is encountered in the input (as in (14a) or (14b)), its associated X′-structure (an NP) will be activated, and because of its semantics (animate, human, etc.), as well as, perhaps, some frequency bias, the representation of the Agent role in its lexical entry will also receive some activation, although other roles will also be activated to at least some degree (Experiencer, Theme, Goal, etc.): No particular role is likely to win out at this point. In addition, of course, various other

[13]We will assume that the X′-structures for the lexical entries of verbs are linked directly to IP X′-structures, so that when a VP X′-structure is activated, the associated IP structure is also activated. This is particularly relevant to the claim that thematic roles are linked directly to positions in X′-structures, because we assume that external arguments (and in some cases (see below) internal arguments) are linked to IP specifier positions. This stipulation could be eliminated by adopting certain syntactic conventions (i.e., the VP-internal subject hypothesis in Government-Binding theory; Koopman & Sportiche, 1988), but this is well beyond the scope of the current work.

elements of its lexical representation will also be activated, including third person, singular number, male gender, and its null core thematic grid.

Next *cooked* is encountered, and its permitted X′-structures and argument structures become activated (to varying degrees depending on frequency),[14] along with all the other components of its lexical entry. At this point, the relationship between *John* and *cooked* will be computed (or at least, relevant computation will begin): The system must determine how their respective X′-structures are connected as well as which thematic role *John* receives (and from which grid it is assigned), and in practice, because argument structures and X′-structures are tightly linked, both to each other and with the grammar, determining how *John* is assigned a role will determine how the X′-structures are linked, and vice-versa.

Large parts of these operations will depend on which thematic role *John* is assigned, and a number of constraints will enter into this decision: First, the available thematic role options provided by *cooked* will constrain the choice of role for *John*: The only possibilities are Agent, assigned either from the grid in (14a) or (14b), or Theme, assigned from the grid in (14a) or (14c). This constraint might be even stronger if there is a strong frequency bias in favor of one of the thematic grids over the other. If, for example, the intransitive grid in (14b) were much more frequent than the others, this would create a preference for the assignment of the Agent role (from the (14b) grid) over the alternatives. Second, as noted above, semantic constraints and/or frequency biases in the lexical entry of *John* will create some preference for certain roles over others (in particular, Agent is likely to be preferred over Theme). Third, various contextual constraints might have an effect on the preference for one role over another: For example, suppose the input *John cooked...* is preceded by a discourse describing John's preparations for the dinner party that he is going to

[14]We remain vague on the question of exactly when the appropriate CP X′-structure becomes activated (as it must in all clauses). It is possible that any sentence-initial word triggers the activation of a CP, or, alternatively, that its activation is triggered by the activation of a VP X′-structure when no overt COMP head (*that*, for example) is present. See also footnote 13. Note that we assume the existence of null forms of INFL and COMP, with lexical entries like those of other words (including, importantly, the relevant IP and CP structures), with the exception that they have no phonology or orthography. When overt forms are not present in the input, these null forms will be activated.

throw later in the evening. Then simply seeing *John* at the beginning of a sentence is not necessarily likely to have any effect on preferences for thematic role assignment. However, when *cooked* follows *John*, assigning the Agent role becomes extremely plausible: The discourse has already made it clear that John is quite likely to be cooking something (for dinner). Thus a number of factors will enter into the determination of the appropriate thematic role for *John*, and they will operate as parallel constraints and will directly influence the activation of the various thematic role alternatives in the lexical entry for *John*. Similarly, the process of choosing an argument structure for *cooked* will involve constraints from frequency biases and possibly from contextual influences; in fact, of course, choosing a thematic role and choosing an argument structure will proceed in tandem, because they strongly constrain each other.

Excluding the possibility of some overpowering contextual bias, *John* will be assigned the Agent role under the above constraints, narrowing the choice of argument structure for *cooked* to (14a) or (14b). If *cooked* happens to have a frequency bias in favor of one or the other, then the frequency-favored thematic grid will be chosen. Otherwise, they may continue to compete with each other, each remaining partially activated, until later input forces a decision one way or the other,[15] as in the presence of *the* in (14a), which would indicate that the upcoming phrase is an NP and will likely require the assignment of another thematic role, supporting the transitive grid in (14a), which has such a role still available (and which is associated with an X′-structure to which an NP could attach to receive it).

In addition to choosing an argument structure, of course, the system must settle on an appropriate syntactic structure for the input, which means that the X′-structures associated with *John* and *cooked* must be appropriately connected. In most cases, the syntax itself will handle much of this work: In particular, just as in Figure 2-2 for the sentence *The horse raced past the barn*, the NP (*John*) must attach as the

[15] Alternatively, one might allow individual thematic grids to include optional arguments, in which case we might collapse (14a) and (14b) to a single core grid with an external Agent argument and an optional internal Theme argument. This would primarily be a matter of adjusting the argument structure formalism, but it might have measurable consequences in terms of argument structure frequency effects, for example. See Carlson and Tanenhaus (1988) for some potentially related empirical considerations.

specifier of the IP,[16] which will be activated by the time the VP is activated (see footnotes 13 and 14). This is the position of the subject and the only position to the left of *cooked* to which a thematic role is assigned (the external argument role when there is one, otherwise the direct internal argument role, as in (14c)).[17] Thus, once the choice of argument structure has determined the X′-structure choice for *cooked*, syntactic constraints will force the appropriate syntactic structure. Even prior to the selection of a single argument structure, however, only a few candidate syntactic structures (hypotheses) will be permitted by the syntax, each one corresponding to a different choice of X′-structure for *cooked*; they will each be partially activated to some degree, depending on the strength of their support from the syntax (which will not vary much in this particular case) and from other parts of the system (their associated argument structures, for example); and they will compete with each other just as other representations at the same level do within a single lexical entry. Thus throughout the processing of an input, current syntactic hypotheses (partially activated sets of connections between X′-structures) will continue to interact with other lexically-based frequency and contextual biases (over argument structure, tense, etc.).

2.3 Syntactic Ambiguity Resolution

The Main Verb/Reduced Relative Ambiguity

Given this conception of the structure of the lexicon and these processing assumptions, and having worked through a simple example, we can now examine the extensive and often contradictory literature concerning ambiguity resolution for the main verb/reduced relative (MV/RR) ambiguity. This structure is ambiguous over at least

[16] Another attachment is possible here, corresponding to a reduced relative clause (see Figure 2-3). However, because *John* is a proper noun, the only interpretation of the clause is as an *appositive*, or non-restrictive, relative clause, which functions as a parenthetical phrase, as in *John, cooked by the cannibals, was later eaten as a light snack.* Appositives are typically set off from the clause in which they are embedded by a pause in speech or by commas in text, so we ignore that possibility here.

[17] We are, obviously, glossing over some significant syntactic detail here, as elsewhere. The detailed explication of the operation of the syntactic component of the system will depend hugely on a choice of syntactic theory, and as our focus is on lexical mechanisms, we remain neutral on syntactic detail.

five different levels of lexical representation, three of which were shown previously in Figure 2-5:

(1) syntactic structure — the Main Verb interpretation has a different syntactic structure (i.e. different linkages of X-bar structures) than the Reduced Relative interpretation;

(2) tense morphology — the *-ed* ending on the ambiguous verb is interpreted as a past tense marker in the MV interpretation and as a past participle marker in the RR interpretation;

(3) voice — active for MV, passive for RR;

(4) argument structure — the RR interpretation requires the transitive (<Agent, Theme>) argument structure, whereas the MV interpretation may take a variety of structures — transitive, intransitive, etc.; and

(5) assignments of thematic roles to nouns in the sentence — the noun (or NP) preceding the verb is the external argument in the MV sentence (receiving the Agent role when the grid is <Agent, Theme>), whereas it is the direct internal argument (Theme in the above grid) in the RR sentence.

These alternative interpretations over the different representations are not independent of one another: Some pairs of alternatives cannot co-occur, some only rarely co-occur, some frequently co-occur. For example, consider the input *the witness examined*, which could be part of either an MV or RR structure. Because the verb *examined* is obligatorily transitive, there is no argument structure ambiguity, as had been shown in Figure 2-5, but the four other ambiguities listed above still remain. The interrelationships between these four remaining ambiguities in *the witness examined* are illustrated in Figure 2-9.

This figure looks very similar to the previous Figure 2-5, but Figure 2-5 illustrated a portion of a lexical representation for an individual word, whereas Figure 2-9 does not. Instead, Figure 2-9 depicts (in shorthand form) some of the constraints provided by the grammar of English that are brought to bear during the parsing of *the*

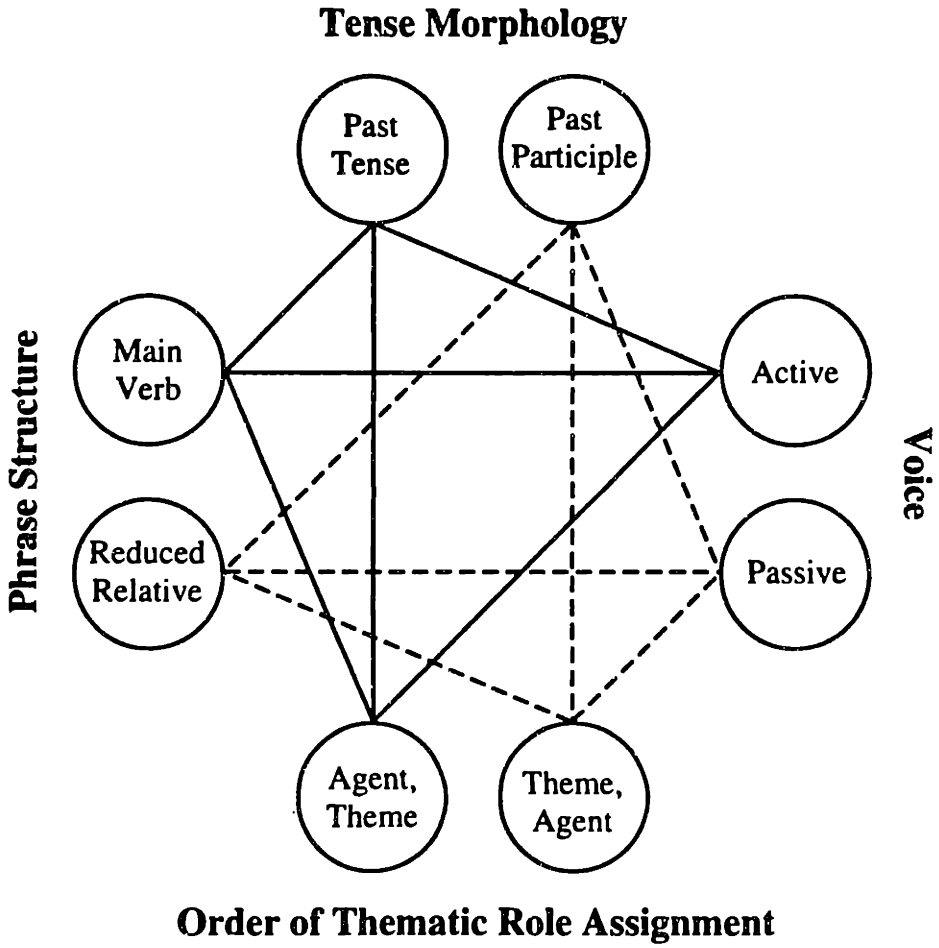


Figure 2-9: Input properties allowed to co-occur by grammatical constraints in the MV/RR ambiguity. Lines indicate permitted co-occurrences. Solid lines correspond to the main verb interpretation; dashed lines correspond to the reduced relative interpretation.

witness examined... The lines in Figure 2-9 connect properties of an input that are allowed to co-occur by the grammar of the language; the actual representation of these properties in the system is more complex: For two of the pairs of properties, tense morphology and voice, Figure 2-9's representation is just that of Figure 2-5: These are levels of representation within the lexicon (as in Figure 2-5), but they also happen to be relevant features for the MV/RR ambiguity. The other two pairs of properties, however, are not part of any lexical entry. The phrase structure units, Main Verb and Reduced Relative, represent the corresponding phrase structures, so each unit is shorthand for a particular pattern of X′-structures and connections between X′-structures which can become activated during processing. The other units (for order of thematic role assignment) represent another property of an input, namely which role from the <Agent, Theme> grid (the only core grid available for *examined*) is assigned to the noun preceding the verb, and which will be assigned to some later constituent (the input at this point is only *the witness examined*). Figure 2-9 is therefore a representation of which sets of properties are grammatically permitted to co-occur, given the input. The connections thus do not encode frequency; they simply show the constraints provided by the grammar.

The connections in Figure 2-9 admit two different parses. The first of these is the main verb interpretation, indicated by the solid line connections, and the second is the reduced relative interpretation, indicated with dashed lines. Because Figure 2-9 represents the grammatical constraints that apply and the alternative interpretations that are available by the point that *the witness examined* has been encountered during parsing, some possibilities that are admitted by the verb *examined* in isolation have been ruled out by the syntactic context. For example, an active past participle interpretation of *examined*, as in *the witness had examined*..., is eliminated because no form of *have* is in the input. Similarly, the passive MV interpretation (as in *the witness was examined*) is eliminated by the absence of a form of *be* in the input, and the unreduced relative (*the witness who was examined*...) is eliminated by the absence of *who* and a form of *be*.

With these additional possibilities already ruled out, each node (property) in Figure 2-9 participates in only one parse — the past tense node is needed for the MV parse, the past participle for the RR, etc. Thus any information concerning which of the two alternative possibilities at any level is correct will have effects not only for that level of representation but for the other levels as well. This situation is illustrated in Figure 2-10, in which two different kinds of information affect the choice between past tense and past participle morphology. The first of these is straightforward morphological information, which is available for those verbs with different past tense and past participle forms. Thus for the string *the witness ran*, the verb *ran* is unambiguously a past tense (the past participle is *run*), with the consequence that all the other attributes of the main verb interpretation are forced by the grammar, as shown with the heavier lines. The second example involves a weaker influence: the frequency of the alternative morphological tense forms for a morphologically ambiguous verb (*kissed*). In the example *the witness kissed*, the verb *kissed* is much more frequent as a past tense than as a past participle form (Francis & Kučera, 1982). Although the frequency information is not as definitive as morphological information, the frequency asymmetry in *kissed* does constrain the three other levels, so that the active voice, Agent-Theme role assignment ordering, and MV phrase structure are preferred over their respective alternatives. Such frequency asymmetries and the relationships shown in Figure 2-10 yield a prediction about the process of MV/RR ambiguity resolution: MV/RR ambiguities containing verbs like *kissed*, which are more frequent in past tense than past participle uses, should be more difficult to resolve in favor of the reduced relative structure, compared to verbs such as *examined*, which are more equibiased between the two alternative morphological forms.

Frequency Effects

Although no experiments have explicitly manipulated the relative frequency of past tense and past participle uses for ambiguous verbs, data reported by Spivey-Knowlton et al. (1993) can be used to address this prediction. Spivey-Knowlton et al. collected sentence completion norms in which subjects read a sentence fragment (sometimes in

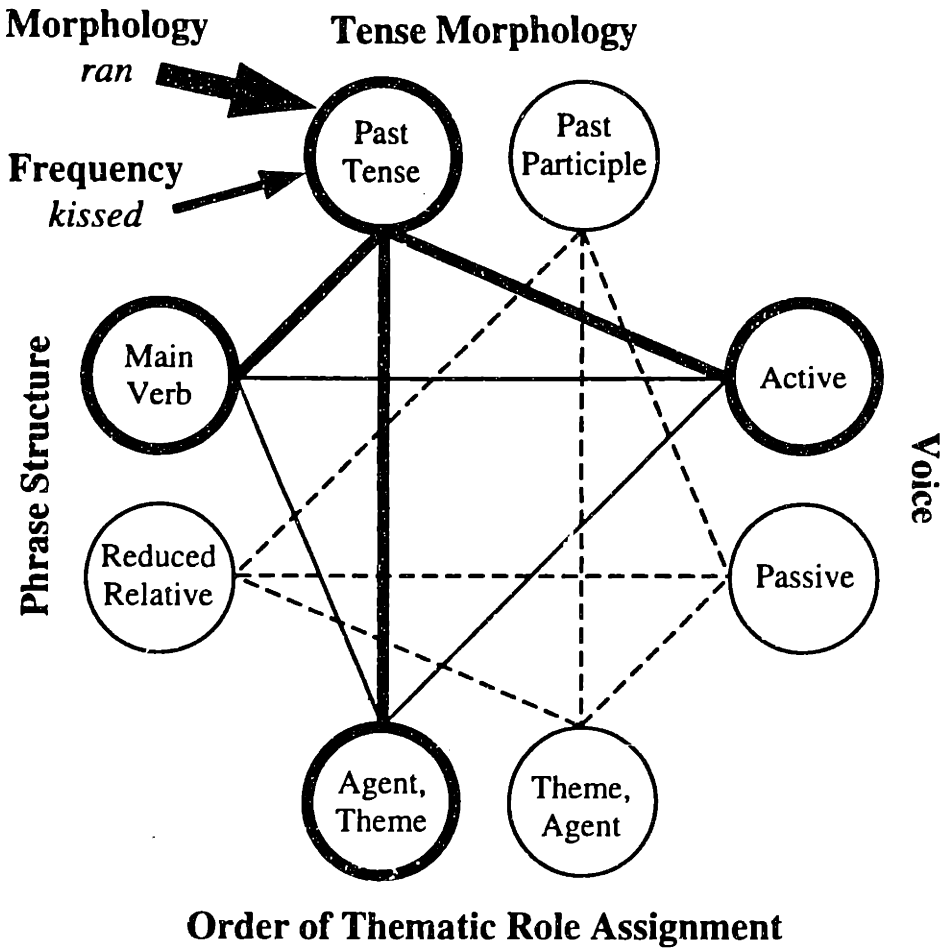


Figure 2-10: Input properties allowed to co-occur by the grammar in the MV/RR ambiguity, with morphology (*ran*) or frequency (*kissed*) supporting past tense morphology.

context) up to the MV/RR ambiguous verb (e.g. *The witness examined*) and wrote a completion. We compared the percentage of reduced relative completions that subjects produced (averaged across the three different norming procedures Spivey-Knowlton et al. used) with relative frequency of past tense and past participle forms for their verbs, as found in Francis and Kučera (1982). This analysis supported the prediction: Frequency of reduced relative completions was reliably correlated with verb past participle frequency — the higher the past participle frequency of the ambiguous verb, the more often the structure was interpreted as a reduced relative ($r(14) = .50$, $p \stackrel{\cdot}{<} .05$). The relative frequency of the alternative morphological

tenses therefore appears to be an important component of the viability of the RR interpretation.

A similar effect can be found with the relative frequency of alternative argument structures. Figures 2-9 and 2-10 show the order of thematic role assignments for obligatorily transitive verbs like *examined*, but as noted above with Figure 2-5, the MV interpretation, which uses the active voice, can accommodate several argument structures, including both transitive and intransitive. By contrast, the RR interpretation requires the passive voice, which requires the transitive argument structure. Argument structure frequency should therefore have a substantial effect on MV/RR ambiguity resolution, because verbs that do not permit a transitive argument structure (e.g. *sleep*) cannot appear in the RR construction, and verbs which only rarely appear with a transitive argument structure only rarely appear in the RR construction. Thus the extent to which an ambiguity is resolved in favor of the RR construction is predicted to be (in part) a function of the frequency of a verb's alternative argument structures: The lower the frequency of the transitive argument structure for a verb, the less likely the ambiguity will be resolved with a reduced relative interpretation. MacDonald (in press, Experiment 3) tested this prediction by investigating comprehension difficulty for ambiguous sentences resolved with the RR interpretation, in which the verbs varied in argument structure frequency. In one condition, the verbs (e.g. *pushed*, *carried*), had higher frequency transitive than intransitive argument structures (as measured in Connine, Ferreira, Jones, Clifton & Frazier, 1984). In another condition, the verbs (e.g. *moved*, *raced*) were more frequent in their intransitive than transitive structures. MacDonald found that comprehenders had difficulty resolving the ambiguous RR sentences only in the second condition, where the required transitive argument structure was lower in frequency than the intransitive structure. When the ambiguous verbs were more frequent in the transitive than in the intransitive structure, the ambiguous sentences were no more difficult than unambiguous controls. This example conforms to the pattern of effects in lexical ambiguity resolution: Item-specific information, in this case the frequency of a verb's alternative argument structures, can determine the extent to which one

interpretation is preferred over the other.

The literature provides no data for how these two frequency effects (tense morphology and argument structure) might be related to one another, as no study has explicitly manipulated both of them. However, a simple example reveals that the conjunction of the two should be quite powerful in guiding the ambiguity resolution process, because the reduced relative interpretation requires the conjunction of both past participle morphology and the transitive argument structure. Consider again the example of *The horse raced past the barn fell*, and compare it to the situation in which *carried* replaces *raced*. The two verbs differ on past tense versus past participle frequency (Francis & Kučera, 1982): *Raced* is used as a past participle only 8% of the time, whereas *carried* is a past participle 52% of the time. The verbs also differ on argument structure frequencies: *Raced* is transitive about 11% of the time, whereas *carried* is transitive about 94% of the time (Connine et al., 1984). Table 2.5 shows the probabilities of different combinations of these tense morphology and argument structure alternatives. The table shows that *raced* is overwhelmingly unlikely to be introducing a reduced relative, as the probability of *raced* having both the transitive argument structure and past participle morphology is less than 1%. By contrast, the conjoint probability of *carried* being both a past participle and transitive is 49%. This result does not mean that *carried* appears in reduced relatives 49% of the time — there are still other uses of transitive past participles that do not involve a reduced relative, including uses like *had carried*, *was carried*, etc. The point is that some enabling conditions exist that make *carried* a possible candidate to be introducing a reduced relative clause, whereas these conditions do not exist for *raced*.[18] The conjunction of several different kinds of frequency information is therefore likely to be extremely important for resolving this ambiguity.

[18]We have made the simplifying assumption here that tense morphology and argument structure probabilities are independent, but this need not be true for some or all verbs.

Table 2.5: Tense Morphology and Argument Structure Conjoint Probabilities

	Verb	
Tense Morphology & Argument Structure	*raced*	*carried*
Past tense, intransitive	.82	.03
Past tense, transitive	.10	.45
Past participle, intransitive	.07	.03
Past participle, transitive[a]	.01	.49

[a]Tense morphology / argument structure combination necessary for the reduced relative construction.

Context Effects

In addition to lexical frequency information, the MV/RR ambiguity is also affected by constraints from the surrounding context. Obvious examples of contextual effects include the elimination of alternative interpretations such as *the witness had examined* and *the witness was examined* in Figure 2-9, by virtue of the fact that the crucial words *had* or *was* are not in the input. More subtle effects are also possible with contexts that do not definitively eliminate alternatives but merely make one interpretation more plausible than the other(s). Consider again the animacy manipulation investigated by Trueswell et al. (in press) in sentences like *The witness/evidence examined by the lawyer was useless.* These sentences provide a context that influences the plausibility of the order of thematic role assignments: While *witness* is a good Agent for *examined*, *evidence*, being inanimate, is a highly implausible Agent. Thus the Agent, Theme order is favored for *witness examined*, whereas the Theme, Agent order is favored for *evidence examined.* This latter situation is illustrated in Figure 2-11, in which plausibility information affects the order of thematic role assignments, which in turn constrains the representations at other levels, resulting in all the subcomponents of the RR interpretation being favored.

Trueswell et al.'s (in press) results support this account; they found that ambigu-

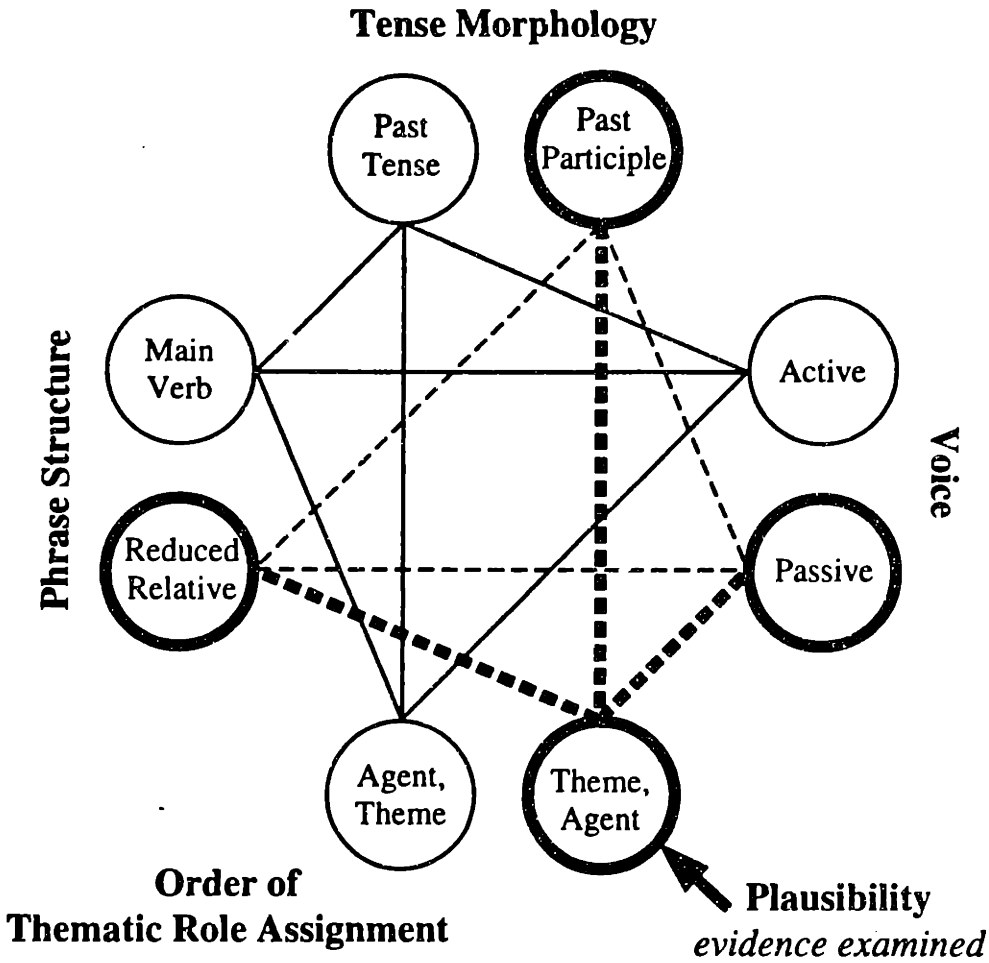


Figure 2-11: Input properties allowed to co-occur by the grammar in the MV/RR ambiguity, with plausibility information (*evidence examined*) supporting the Theme-Agent role ordering.

ous sentences resolved with the RR interpretation were more difficult than unambiguous controls when animacy information supported the incorrect main verb interpretation (*defendant examined*), but not when the subject noun was inanimate (*evidence*), supporting the RR interpretation. These animacy effects have been replicated and extended to plausibility effects that do not depend on animacy (MacDonald, in press; Pearlmutter & MacDonald, 1992; Tabossi, Spivey-Knowlton, McRae, & Tanenhaus, 1993), and similar kinds of context effects have been demonstrated (Spivey-Knowlton et al., 1993; Trueswell & Tanenhaus, 1991; 1992) for discourse contexts which make the RR interpretation the more felicitous one in a discourse, as initially suggested by

Crain and Steedman (1985).[19]

These results argue persuasively for the role of contextual information during MV/RR ambiguity resolution, but other studies in the literature have failed to find such effects (Britt et al., 1992; Ferreira & Clifton, 1986; Rayner et al., 1983; 1992). The resolution of these conflicts in the literature appears to rest with lexical frequency effects of the sort shown in Figure 2-10. As discussed in the introduction, Duffy et al.'s (1988) empirical investigation and Kawamoto's (1993) model of context and frequency effects in lexical ambiguity resolution revealed that context was able to cause a single interpretation to be favored only in cases where strong frequency asymmetries were not present (i.e., for equibiased words). In words with a strong frequency bias, contextual support for the lower frequency meaning was generally not strong enough to eliminate the frequency advantage of the higher frequency meaning. The same effects should hold for other aspects of lexical representation, including those relevant to the MV/RR ambiguity. That is, contextual information should be able to affect the interpretation of the MV/RR ambiguity most clearly when the ambiguous verb is roughly equibiased in its alternative interpretations for tense morphology, argument structure, voice, etc. Figure 2-11 above represents this situation: No frequency biases are present at any level of representation, and the contextual information influences the order of thematic role assignment, promoting the reduced relative interpretation. When one alternative interpretation at some level of representation is much more frequent than another, however, contextual factors supporting the subordinate interpretation will be much less effective. This situation is presented in Figure 2-12, which depicts a verb, such as *raced*, that is more frequent as a past tense than as a past participle. The frequency information strongly promotes the MV interpretation, despite the fact that the contextual information favors the thematic role assignment

[19]As noted in the introduction, Crain (1980; see also Crain & Steedman, 1985; Altmann & Steedman, 1988) argued that the RR interpretation requires noun modification, which is superfluous unless the noun has more than one possible referent in the discourse model. Thus the isolated sentences which are the typical stimuli of sentence processing experiments make ambiguity resolution in favor of the reduced relative interpretation artificially difficult, compared to natural speech or written contexts, where the ambiguous RR construction is used felicitously. We will return to this issue in conjunction with a consideration of the PP attachment ambiguity below.

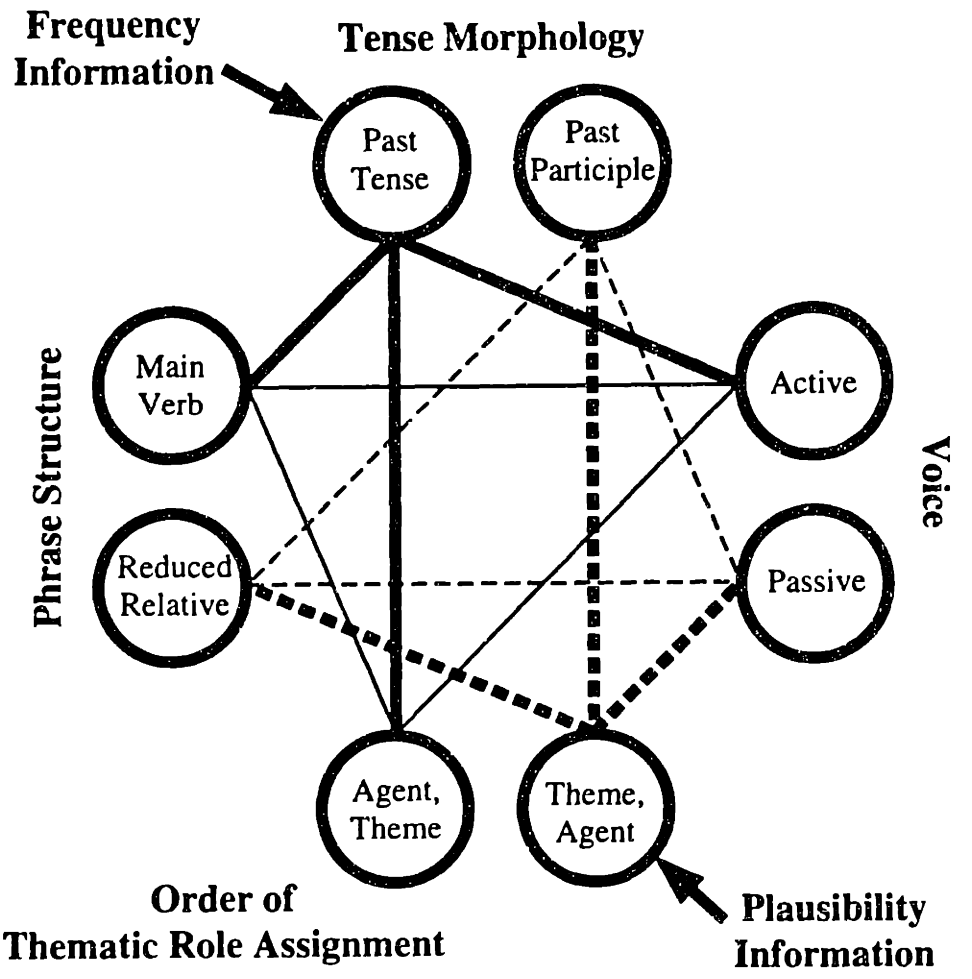


Figure 2-12: Input properties allowed to co-occur by the grammar in the MV/RR ambiguity, with frequency supporting past tense morphology and plausibility supporting the Theme-Agent role ordering.

corresponding to the reduced relative. This situation is comparable to a sequence like *The evidence raced by the lawyer...* Here *raced* is biased to be past tense, but *evidence* does not make a plausible Agent of *raced.* The result (as in Figure 2-12) is that the contextual information does not successfully promote the components of the reduced relative interpretation. This result corresponds to our intuitions, that *the evidence raced by the lawyer...* does not seem to promote a reduced relative interpretation but rather yields an odd-sounding active voice, past tense, main verb interpretation in which the evidence is (perhaps) metaphorically going by too quickly for the lawyer to follow it.

If contextual information does have a greater effect on MV/RR ambiguity resolution for equibiased ambiguous verbs than for strongly biased verbs, then it may be possible to reconcile conflicts in the literature by examining the argument structure and morphological tense frequency biases of the ambiguous verbs that have been used in different studies. The prediction here is that those studies which did find context effects tended to use more equibiased verbs, whereas those studies which found that helpful context did not affect MV/RR ambiguity resolution used many items that were strongly biased (by frequency) to the MV interpretation. Ideally both argument structure and tense frequency biases should be examined for verbs that have been used in previous experiments, as was done for the verbs *raced* and *carried* in Table 2.5. Unfortunately, there is currently not enough information available about the argument structure frequencies of English verbs to assess even 20% of those that have been used in previous context studies. An analysis using only morphological tense information (Francis & Kučera, 1982) was therefore conducted.

Table 2.6 reviews the published studies of context effects in the MV/RR ambiguity. The second column of the table reports the percentage of past participle uses in English for the verbs in each study (Francis & Kučera, 1982); the studies are ordered from lowest to highest mean past participle frequency. Additional columns indicate the type of context manipulation and the reading time measure. The last column of the table indicates the results obtained in conditions when context favored the reduced relative interpretation: Four studies found that subjects still had difficulty with the reduced relative interpretation in the presence of helpful context information (the "minimal attachment" results, in which helpful context is ignored), whereas six other studies found that comprehenders had little or no difficulty resolving the ambiguity in the presence of helpful context. As in the lexical semantic ambiguity case, the effectiveness of the context manipulation varied with (past participle) frequency: The mean past participle frequency for the four studies that found no context effects was 49.7%, whereas the six studies that did find a context effect had a past participle frequency of 64.5%, a reliable difference ($t(8) = 2.59$, $p < .05$). As predicted, the studies which found that helpful context made it easy to interpret ambiguous sen-

tences in favor of the RR interpretation are the studies that used verbs that more closely approximated equibiased items.[20] Studies which used verbs more strongly biased to the MV interpretation, however, did not find that context helped resolve the ambiguity in favor of the subordinate interpretation. This relationship holds despite wide variations across the studies in type of context, strength of context, and reading time measure.[21]

In summary, the interpretation of the MV/RR ambiguity depends first on two pieces of lexical information: the frequencies of the past participle form and the transitive argument structure. Given a verb with frequency biases that make the reduced relative interpretation a viable option, contextual information can guide the comprehender to one or the other interpretation. This is the result obtained in six studies in Table 2.6, where helpful context eliminated any difficulty with the reduced relative interpretation. The same contextual information will likely have little effect, however, if the lexical biases of the ambiguous verb overwhelmingly favor the main verb interpretation. Thus lexical and contextual information can jointly determine the course of ambiguity resolution.

Of course the MV/RR ambiguity is only one of a number of syntactic ambiguities that arise in English. We have focused on this ambiguity because it is probably the best known and the most intensely studied, but if syntactic ambiguity resolution really does reduce to lexical ambiguity resolution in the way we have discussed, then other syntactic ambiguities should also be viewed in this way. We next consider two other syntactic ambiguities, the Noun Phrase/Sentential Complement (NP/S) ambiguity and the Prepositional Phrase Attachment ambiguity (often called PP-attachment). Interpretation preferences for these ambiguities have been taken as evidence for the

[20]Recall that it is the conjunction of several factors that is important for the reduced relative interpretation, including argument structure frequencies, for which little data are available. Thus while a 49.7% past participle frequency is quite close to equibiased for the tense morphology level of representation, a 64.5% past participle frequency is more likely to yield an equibiased *conjunction* of factors.

[21]It is worth noting, in light of the extensive criticism of self-paced reading, that Table 2.6 does not reveal a pattern in which eyetracking data support minimal attachment and self-paced data (by-word or by-phrase) support an interactive model. Instead, each of the three methods has produced data in support of minimal attachment, and the same three methods have also yielded early context effects.

Table 2.6: Effects of Prior Context Modulated by Morphological Tense Frequency

Study	%Past Part[a]	Prior Context Manipulation[b]	Measure[c]	Early Effects in Helpful Context Condition[d]
Rayner, Garrod & Perfetti, 1992	44.4	Pragmatic	Eyetracking	Minimal Attachment
Britt, Perfetti, Garrod & Rayner, 1992	47.0	Pragmatic	SP-word	Minimal Attachment
Ferreira & Clifton, 1986	47.4	Thematic-anim	Eyetracking, SP-phrase	Minimal Attachment
Trueswell & Tanenhaus, 1991, 1992	47.6	Pragmatic	SP-phrase, 1991; Eyetracking, 1992	Context helps (weak)
Rayner, Carlson & Frazier, 1983	60.0	Thematic-plaus	Eyetracking	Minimal Attachment
Trueswell, Tanenhaus & Garnsey, in press	61.0	Thematic-anim	Eyetracking, SP-word	Context helps
MacDonald, in press, Experiment 3	64.1	Thematic-anim	SP-word	Context helps
Tabossi, Spivey-Knowlton, McRae & Tanenhaus, 1993	66.9	Thematic-plaus & -anim	SP-phrase	Context helps (weak)
Spivey-Knowlton, Trueswell & Tanenhaus, 1993	71.3	Pragmatic	SP-phrase, SP-word	Context helps, SP-phrase; Context helps (weak), SP-word
Pearlmutter & MacDonald, 1992	75.8	Thematic-plaus	SP-word	Context helps (weak)

(table continues)

[a]Mean percentage of past participle (vs. past tense) use of ambiguous verbs in stimulus materials, as measured by Francis & Kučera (1982).

[b]Thematic-anim: thematic role manipulation using animacy; Thematic-plaus: thematic role manipulation using plausibility (with animacy controlled); Pragmatic: pragmatic manipulation of felicity of reduced relative interpretation.

[c]SP-word: self-paced reading, one word at a time; SP-phrase: self-paced reading, one phrase (usually 2–3 words) at a time.

[d]Minimal Attachment: presence of helpful context did not prevent garden pathing; Context helps: helpful context prevented garden pathing; Context helps (weak): helpful context condition was easier than unhelpful context condition but did not remove all difficulty associated with the ambiguity.

operation of Minimal Attachment (Frazier, 1987), but we argue that ambiguity resolution for these ambiguities can also be analyzed as the combined effects of lexical frequency and contextual information.

The Noun Phrase/Sentential Complement Ambiguity

This ambiguity, like MV/RR, is triggered by an ambiguous verb, in this case a verb with both a transitive argument structure and an argument structure assigning the Proposition role to a direct internal argument. (15) shows one set of examples.

(15)

a. knew: <Agent, Theme>
John knew the answer.

b. <Agent, Proposition>
John knew the answer was correct.

c. John knew that the answer was correct.

In (15a), the NP interpretation, the NP *the answer* is assigned the Theme role from *knew*'s transitive argument structure, while in (15b), the S (sentential complement) interpretation, *knew* assigns the Proposition role to the entire embedded clause (CP) *the answer was correct.*[22] Thus as in the MV/RR ambiguity, the alternative phrase structures (NP vs. S) are tied to alternative representations of an ambiguous verb. Much like the MV/RR ambiguity, the NP/S ambiguity can be eliminated for one interpretation with the addition of an optional word, *that*, as illustrated in (15c).[23]

The Minimal Attachment algorithm, when applied to the NP/S ambiguity, pre-

[22]The traditional description of this ambiguity distinguishes the NP and S phrase structures, whereas we are stressing the importance of the <Agent, Theme> vs. <Agent, Proposition> argument structure. We will continue to use the traditional NP/S terminology, but in doing so, we refer to a collection of representations, including both argument structure, phrase structure, and the appropriate corresponding discourse constructions, rather than just the relevant phrase structure, unless otherwise noted.

[23]As shown in Figure 2-2, the phrase structure of a clause includes a CP maximal projection headed by C, which must be phonologically and/or orthographically empty for matrix clauses in English. For embedded clauses, as in Figure 2-3 (the reduced relative), it can sometimes be empty, and that is the case for the S version of the NP/S ambiguity (15b). The unambiguous version (15c) differs only in that C is not empty: The CP is headed by *that*.

dicts that the ambiguity will always be initially analyzed with the NP interpretation (the phrase structure required for a clause is more complex than that required for a noun phrase), but a lexically-based analysis makes a different prediction: If the relative frequency of the alternative argument structures is coded in the lexicon for each verb, then the extent to which the NP or S interpretation will be preferred for an ambiguous sentence should vary with the argument structure frequencies of the verb. In particular, an ambiguous verb that is used more frequently with the <Agent, Proposition> grid than with the <Agent, Theme> grid should be more easily comprehended in the S interpretation, independent of whether the disambiguating *that* is present in the sentence.

Frequency Effects

This clear contrast between the lexical-frequency-based and minimal attachment predictions has been intensively investigated in previous studies, with stubbornly mixed results. A number of experiments have found evidence for the frequency-modulated view, that resolution of the ambiguity in favor of the S interpretation is difficult when the lexical bias of the ambiguous verb favors the incorrect NP interpretation, but with verbs that are used more frequently with S-complements (assigning the Proposition role) than NP direct objects, there is little or no difference between ambiguous and unambiguous sentences requiring the S interpretation, as in (15b) versus (15c) (Garnsey, Lotocky & McConkie, 1992; Holmes, Stowe & Cupples, 1989; Trueswell et al., 1993). Other studies have found support for Minimal Attachment — that even when verbs are biased towards the Proposition argument structure, ambiguous S-interpretation sentences are more difficult than their unambiguous counterparts (Frazier & Rayner, 1982; Ferreira & Henderson, 1990). As with the MV/RR ambiguity, these different results do not appear to be due to the choice of reading time measure, as several researchers have found that both eyetracking and self-paced reading produced similar patterns on the same set of materials (Ferreira & Henderson, 1990; Trueswell et al.,

1993).[24] Instead, the answer seems to be that lexical frequency information affects ambiguity resolution in a slightly more complicated way than was originally thought.

Research by Trueswell et al. (1993) has uncovered other important frequency effects for this ambiguity. First, Trueswell et al. noted that argument structure frequency information was not simply a matter of the relative frequencies of the two alternative argument structures in (15). They showed that some verbs such as *pray* and *agree*, which some researchers had included in the "S-bias" class (verbs that are more frequent with the S than with the NP interpretation), are actually quite rare in both the S and NP interpretations. These verbs instead appear more frequently in some third construction, such as an intransitive structure (e.g., *John prayed every night*) or an infinitival complement structure (e.g., *John agreed to wait in line*). When a comprehender encounters an NP after these verbs (e.g., *John prayed/agreed the book...*), the presence of the NP *the book* indicates that the verb's preferred intransitive or infinitival argument structure is not correct and that some non-preferred argument structure is needed for this input. Trueswell et al. argue that slowed reading times in such cases, which have previously been attributed to the operation of Minimal Attachment (that the NP interpretation is initially adopted and found to be implausible, Ferreira & Henderson, 1990), actually reflect a lexical frequency effect: Contextual information (the NP) conflicts with a strong frequency bias (in favor of the intransitive or infinitival complement argument structure), resulting in a slowdown; note that this is just what Duffy et al. (1988) found when a semantic context conflicted with a strong frequency bias in favor of one meaning of a word.

The second lexical frequency effect that modulates comprehension of the NP/S ambiguity concerns co-occurrence information between verbs and the complementizer *that*. Trueswell et al. collected sentence-completion norms to determine argument structure preferences for a number of verbs, in which subjects read a sentence fragment up to an NP/S ambiguous verb (e.g., *John hinted...*) and wrote a completion for the

[24]See Holmes, Stowe, and Cupples (1989) for some possible differences across self-paced measures, and see Trueswell, Tanenhaus, and Kello (1993) for detailed analyses of how the addition of an extra word (*that*) in the unambiguous condition can change reading patterns independent of ambiguity, so that special care is needed when interpreting eyetracking data.

sentence. They observed that on those occasions when S interpretation completions were written, the extent to which the optional complementizer *that* was included ((15b) vs. (15c)) depended on the verb. For example, 100% of the S interpretations of *hinted* included *that*, whereas only 17% of the S interpretations for *wished* included *that*. Trueswell et al. found that a verb's *that-preference* accounted for a sizable portion of the variance in reading times at the post-verbal NP in ambiguous strings such as *hinted/wished the book...* When the NP (e.g. *the book*) followed a verb with a low *that*-preference (one that was only rarely followed by *that*), reading times on the NP were short, but when the NP followed a verb with a high *that*-preference, such as *hinted*, reading times on the NP increased. Again, previous findings of long reading times for NPs following S-bias verbs had been interpreted as evidence for Minimal Attachment, but Trueswell et al. showed that this effect varies across individual verbs, specifically with the verb's preference to include an overt *that* when used with the S interpretation.

Juliano and Tanenhaus (1993) investigated the source of *that*-preference effects and found that the size of a verb's *that*-preference was inversely related to the frequency of the verb in the language: Higher frequency verbs omit *that* more frequently from S interpretation sentence completions. This result ties these phenomena even more closely to the frequency-coded representations of individual verbs in the lexicon. Thus the extent to which an NP/S ambiguity is interpreted with the S interpretation is a function of (a) the relative frequencies of the alternative argument structures — not just the argument structures corresponding to the NP and S interpretations, but also the argument structures for other interpretations such as intransitive and infinitival complement sentences; and (b) a verb's *that*-preference, which is itself dependent on the verb's frequency in the language. All of these effects follow naturally from the theory proposed here.

Context Effects

Although our theory suggests that NP/S ambiguities should exhibit effects of lexical structure and context very similar to those obtained for the MV/RR ambiguity, the

literature provides very little evidence bearing on this prediction. As in the case of the MV/RR ambiguity, one source of contextual constraint is the plausibility of thematic role assignment. In the NP/S case, however, it is not thematic role assignment to the NP preceding the ambiguous verb, but rather to the NP after the verb. Consider the pair *John admitted the error/book...* In one case, the NP *the error* makes a plausible Theme for the verb *admitted*, whereas in the other case, *the book* is a highly implausible Theme. Because the Theme role can only be assigned by *admitted* from its transitive grid, requiring the NP interpretation, this interpretation is favored in the *error* example, and it is discouraged in the *book* example. Only one investigation of such contextual effects exists in the literature, Holmes et al. (1989), which found some small effects of context, although the ambiguous items resolved with the S interpretation still yielded longer reading times than unambiguous controls, even in the presence of helpful context. The lessons of the Trueswell et al. (1993) and Juliano and Tanenhaus (1993) investigations may be applicable here, in that the Holmes et al. verbs were not controlled with respect to the frequency information that we now know to be important for this ambiguity. Indeed, if ambiguity resolution for the NP/S ambiguity works in the same way as resolution of the MV/RR ambiguity and other lexical ambiguities, then a specific pattern of frequency and context effects is predicted: Context effects of the sort studied by Holmes et al. (1989) should be observed only when the ambiguous verb is roughly equibiased in its alternative argument structures. If a verb is overwhelmingly biased to one interpretation, then the relative plausibility of thematic role assignments should have minimal effects. This and similar hypotheses suggested by the lexical view of syntactic ambiguity resolution have yet to be tested.

The Prepositional Phrase Attachment Ambiguity

The third syntactic ambiguity we consider is the Prepositional Phrase Attachment ambiguity, which arises when a PP appears following a verb and its direct object noun. The ambiguity is whether the PP modifies the verb (the *verb attachment* interpretation) or the direct object noun (*noun attachment*). In contrast to the other

syntactic ambiguities discussed above, PP attachment ambiguities usually do not contain a syntactic disambiguation later in the sentence; only the relative plausibility of the alternatives suggests the preferred interpretation. Some examples can be seen in (16), in which (16a) is most plausible with noun attachment, as shown by the bracketing; (16b) is most plausible with verb attachment; and (16c)–(16d) show the alternative interpretations for a more ambiguous sentence.

(16)

a. Fred [ordered [a pizza with pepperoni]].	N-attachment	Attribute
b. Susan [read [the book] in a hurry]].	V-attachment	Manner
c. The spy [saw [the cop with the binoculars]].	N-attachment	Attribute
d. The spy [saw [the cop] with the binoculars].	V-attachment	Instrument

Because prepositions assign thematic roles, the attachment ambiguity is also an ambiguity of argument structure, and the thematic role assigned by each preposition to its object is also shown in (16). Thus in (16a), *with pepperoni*, when attached to the NP *a pizza*, is an Attribute of the pizza. Similarly in (16c), the PP *with the binoculars* modifies *the cop*, and *with* assigns an Attribute thematic role to *the binoculars*. In the verb attachment example (16b), the PP *in a hurry* modifies *read*, and *a hurry* is assigned the Manner thematic role. In the other verb attachment example (16d), *the binoculars*, rather than being an Attribute of *the cop* as in (16c), is assigned the Instrument role — it is the instrument of the seeing event.

Although the bracketings in (16) do not suggest that either interpretation is more syntactically complex than the other, at least one syntactic analysis suggests that noun attachment ((16a), (16c)) involves the more complex syntactic structure (Frazier, 1990; cf. Smith, 1988). Given this particular analysis, application of the Minimal Attachment algorithm yields a preference for verb attachment over noun attachment.[25] As with the other syntactic ambiguities, however, the literature does

[25]In fact, the question of exactly what the Garden Path theory predicts for the case of PP Attachment ambiguities is somewhat difficult to answer. However, as our interest here is not with disproving the Garden Path theory, but rather with detailing a viable alternative, we will not attempt to further consider the possibilities. See Frazier (1990) and the references cited therein for

not provide a consistent picture of whether the verb attachment analysis is initially preferred: Some studies have found a verb attachment preference (Clifton & Ferreira, 1989; Clifton, Speer & Abney, 1991; Rayner et al., 1983), but others have found that a verb attachment preference obtains only under certain lexical or discourse conditions, and under other conditions, a noun attachment preference is observed (Altmann & Steedman, 1988; Britt et al., 1992; Rayner et al., 1992; Sedivy & Spivey-Knowlton, 1993; Taraban & McClelland, 1988). Most of these studies have made the assumption that the verb attachment interpretation is the dominant one, but several investigations of text corpora have found that the noun attachment interpretation is actually the more frequent interpretation in naturally-occurring text, with the verb attachment interpretation comprising only about 35% of the tokens (Hindle & Rooth, 1993; Whittemore, Ferrara, & Brunner, 1990). Again, these conflicts in the literature may be resolved by a close scrutiny of the lexical and context effects that enter into ambiguity resolution in this case.

Frequency Effects

Viewed from the perspective of lexical frequency effects, there are three potentially highly constraining sources of biasing information for interpreting this ambiguity prior to encountering the NP within the PP: the verb, the direct object noun, and the preposition. As noted above, the ambiguity relies on the choice of argument structure, and most verbs and nouns allow a number of possible non-core thematic grids with varying frequencies. For example, consider the verbs *cut* and *walk*. *Cut* occurs quite naturally with an Instrument, as in *cut the bread with the knife*, but is rather odd with a temporal PP, as in *cut the bread for three minutes*. *Kiss*, on the other hand, is somewhat odd with an Instrument (*kissed the guy with the ???*) but is perfectly natural with a temporal PP, as in *kissed the guy for half an hour*. Although the range of possibilities is often more limited, nouns behave quite similarly. For example, nouns derived from verbs (e.g., *destruction*) have biases closely related to those of their corresponding verbs; nouns related to communication (*mail*, *message*,

the most recent discussion of which we are aware.

etc.) occur often with Theme (*mail about the parking situation*) and/or Goal PPs (*message to Jane*), and, while most nouns are quite natural with Location PPs (*in the room, next to the nightstand*), very few occur frequently with a temporal PP. Thus both the verb and the noun in a given PP attachment ambiguity often have preferences about which thematic role is likely to be assigned by the preposition.[26]

Prepositions also often provide highly constraining information for resolving this ambiguity, particularly in combination with the constraints provided by nouns and verbs. Like nouns and verbs, prepositions are associated with particular argument structures, reflecting the different thematic roles which the preposition is allowed to assign. *Of*, for example, nearly always attaches to a preceding noun (Hobbs & Bear, 1990) and assigns an Attribute (*book of poems*) or Theme role (*destruction of the city*), while prepositions such as *into* (*onto, to*) nearly always assign a Goal role (*took the dog into the house*). *With*, on the other hand, assigns a range of possible roles, including Manner (*pointed with fascination*), Instrument (*pointed with the mouse*), Attribute (*pointed at the mouse with a long tail*), and Location (*pointed at the mouse with the hamsters*), among others. Thus prepositions, like nouns and verbs, have a range of preferences for different argument structures. According to our theory, disambiguating between noun and verb attachment is a matter of combining preposition-, verb-, and noun-based frequency biases to determine (1) which role the preposition is assigning and (2) whether the role assigned by the preposition is associated with a grid in the argument structure of the verb, or, alternatively, with a grid in the lexical entry of the noun.

These predictions have not been directly addressed in the psycholinguistic literature, but Taraban and McClelland (1988) investigated PP attachment ambiguities and found that the largest determinant of difficulty in PP attachments was whether or not the role assigned by the preposition conformed with expectations derived from

[26]It is quite possible that the frequency biases that we have described derive from deeper relationships between the semantics of the verbs and nouns involved and the semantics of the thematic roles being assigned, as in, for example, Levin and Rappaport Hovav (1992) and Pinker (1989). This would be perfectly compatible with the current theory; the crucial points here are simply that potentially useful frequency asymmetries exist and that the system is able to keep track of them.

the prior sentence context (including the verb, direct object NP, and the preposition itself, as well as the subject NP). An unexpected NP (as object of the preposition) that did not violate the role-assignment expectation also caused some difficulty, but an unexpected attachment site (noun vs. verb) did not cause difficulty beyond that created by an unexpected role assignment. Although Taraban and McClelland did not differentiate among effects due to the verb, the direct object noun, and the preposition, the results are clearly consistent with a lexically-based theory. Taraban and McClelland also showed that the Rayner et al. (1983) stimuli strongly promoted verb attachments, concluding that this verb attachment preference, which Rayner et al. had attributed to the Minimal Attachment algorithm, should instead be attributed to biases created by the individual items in the experiment, because the preference could be completely reversed with a different set of items.

Some evidence for the importance of lexical frequency information is also available in the computational linguistics literature. Hindle and Rooth (1993) selected a corpus of 880 sentences, each containing a PP attachment ambiguity, from the 1989 Associated Press newswire. They extracted the verb – direct object noun – preposition triple (e.g., *saw – cop – with* in (16c)–(16d)) from each of these sentences and attempted to guess the appropriate attachment of the PP in the original sentence on the basis of the three words in the triple. The Minimal Attachment principle (which dictates the verb attachment) yielded the correct interpretation on only 33% of the sentences, whereas the information in the word-triples was sufficient to predict the correct attachment with over 85% accuracy. Whittemore et al. (1990), using a different corpus of 724 sentences,[27] also note that structurally-based attachment strategies are not particularly effective and propose instead an algorithm that first makes use of verb and noun lexical preferences for particular prepositions, and then makes use of a preposition's lexical preferences to occur with particular classes of nouns and verbs

[27]Whittemore, Ferrara and Brunner's (1990) corpus was a set of travel-related dialogues and was therefore rather more restricted than the corpus of Hindle and Rooth (1993). In addition, their PP attachment ambiguities included constructions in addition to the verb – direct object – PP sequence that we have been considering (including examples involving multiple verb and/or noun attachment sites). For these reasons their conclusions should probably be considered preliminary with respect to the current questions. See Gibson and Pearlmutter (1993) for some discussion of these concerns.

(temporal PPs prefer to attach to events, etc.). This portion of the algorithm, in combination with a preference for attachment to the rightmost site when more than one possible site has a lexical preference, accounts for correct attachment in approximately 95% of Whittemore et al.'s tokens. Thus in both the Hindle and Rooth (1993) and Whittemore et al. (1990) studies, as in the Taraban & McClelland (1988) study, although the relative contributions of the verb, noun, and preposition are difficult to determine, it is clear that a range of local lexical information can provide powerful constraints on the ambiguity resolution process.

Context Effects

The fact that information in the word triples still yielded 15% attachment errors in Hindle and Rooth's (1993) study does suggest that this local information is not capturing all factors relevant to the ambiguity resolution process. As with the other ambiguities, contexts often provide additional relevant information. As with the MV/RR ambiguity, one of the interpretations of the PP attachment ambiguity (noun attachment) involves noun modification. As described above, Crain and his colleagues (Crain, 1980; Crain & Steedman, 1985; Altmann & Steedman, 1988) have argued that such modification is infelicitous in isolated sentences. For example, it makes no sense to distinguish the cop in (16d) as the one who had binoculars when this cop is the only one in the discourse. In the context of several different cops, however, modification becomes much more felicitous, and a number of studies have demonstrated that attachment preferences are modulated by the extent to which it is felicitous to modify the critical noun: When noun modification is felicitous, the noun attachment interpretation is easily comprehended, but when noun modification is infelicitous, the verb attachment interpretation is preferred (Altmann & Steedman, 1988; Britt et al., 1992; Rayner, et al., 1992; Sedivy & Spivey-Knowlton, 1993).

These studies provide ample evidence for the effects of context on PP attachment ambiguity resolution, and several researchers have noted that context effects seem to be more powerful in this ambiguity than in the MV/RR ambiguity (Britt et al., 1992; Rayner et al., 1992). These researchers have suggested that this is a result of

differences between the two syntactic constructions (e.g., perhaps the PP Attachment ambiguity is not really guided by Minimal Attachment, while the MV/RR ambiguity is). The lexical model presented here offers an alternative interpretation: Variations in the effectiveness of context are a result of the nature of the frequency asymmetries in the lexical items that trigger these two kinds of ambiguities. We have seen that context effects are most apparent when the alternative interpretations of an ambiguous word are roughly equibiased; when an item is strongly biased, context effects are unable to override the bias. For the MV/RR and PP Attachment ambiguities, both strongly biased and equibiased lexical items exist, as shown in Table 2.5 for the MV/RR ambiguity, and as discussed with the examples like *cut*, *of*, *into*, etc. for the PP attachment ambiguity. Thus both types of ambiguities are triggered by lexical items that span the range between equibiased and strongly biased; however, we suspect that the distribution of items is quite different in the two cases. The combination of verb, noun, and preposition biases in the PP attachment case tends to result in relatively equibiased ambiguities, so that context effects tend to be more evident for this ambiguity than for the MV/RR ambiguity. Although data concerning the relative frequency biases are scant, several pieces of information support this interpretation.

The more equibiased nature of the lexical items that trigger PP attachment ambiguities compared to those that trigger an MV/RR ambiguity can be seen first in the two corpus studies of PP attachment discussed above (Hindle & Rooth, 1993; Whittemore et al., 1990), both of which found that noun attachments made up about 65–70% of the sample, compared to about 30–35% verb attachments. Although the alternatives are obviously not exactly equal, the lower frequency verb attachment interpretation is still quite common in the language. Although we know of no formal studies of the frequency of the reduced relative interpretation, our own investigation of the MV/RR ambiguity reveals that it is present in less than 1% of the Wall Street Journal sentence tokens (which probably overestimate the frequency of relatively complex constructions) that we have examined for over 100 verbs with a variety of argument structures and a range of past participle frequencies. These re-

sults suggest that while some verbs that trigger the MV/RR ambiguity may be more equibiased than others in terms of the lexical representations that enter into the MV and RR interpretations, as Table 2.5 attests, the majority of verbs are strongly biased to the MV interpretation. This is exactly the situation in which we would not expect context effects to be evident, and this situation simply appears more often for the lexical items that trigger the MV/RR ambiguity than for those that trigger the PP Attachment ambiguity.

2.4 General Discussion

We have presented a generalized theory of ambiguity resolution that subsumes both "syntactic" and "lexical" ambiguity. The major step in developing this theory was the introduction of a greatly enriched conception of the mental lexicon. This involved two major adjustments to previous views of the lexicon in psycholinguistics: We assumed that the lexicon is the repository for all types of knowledge associated with words, including their syntactic functions, and that rather than merely listing these different types of information, the lexicon encodes the grammatical and probabilistic relationships that hold among them.

Our account of ambiguity resolution follows from pursuing implications of this conception of the lexicon. First, this approach permits a radical rethinking of the "syntactic" ambiguities that have been at the center of research on language comprehension. Rather than deriving from ambiguities in the construction of phrase structure trees, they can be seen as deriving from ambiguities inherent in lexical items. Second, knowledge of the grammar — e.g., the fact that certain features of representations can or cannot co-occur; the fact that only some phrase structures are well-formed — can be encoded by the structure of the lexicon itself (e.g., whether the connections between units representing different types of information are excitatory or inhibitory). Third, we can begin to derive a theoretical account of the role of "context" in language processing. Contextual information is relevant to the extent that it provides information that facilitates resolving the different types of ambiguities in

the lexicon. Finally, insofar as syntactic structures have a lexical basis, we predicted that syntactic processing should be governed by independently-established principles concerning lexical processing, in particular "lexical" (i.e., meaning) ambiguity resolution. The factors implicated in meaning ambiguity resolution include properties of the ambiguous word (e.g., the relative frequencies of meanings) and the extent to which the context provides information relevant to distinguishing between the alternative meanings. The meaning ambiguity literature also shows that the effects of contextual constraints are modulated by lexical factors, and this was predicted to carry over to the resolution of syntactic ambiguities. These predictions were supported by a series of analyses of results already in the literature.

Together these assumptions yield a picture in which ambiguity resolution (and language processing in general) is a constraint satisfaction problem in which multiple, overlapping constraints are used to resolve ambiguities at different levels of representation. Processing involves activation of different alternatives and settling into a stable pattern in which only one alternative is active at each level of representation. Thus, the comprehender can be said to converge on the correct interpretation or "relax" into it.

We showed that this theory provides a framework for understanding a wide range of findings concerning the processing of syntactic ambiguities. The main achievement is that the theory provides a basis for reconciling and integrating a set of findings that are contradictory from the perspective of the Garden Path Theory, which predicts that syntactic ambiguities are resolved by pursuing the simplest structural alternative sanctioned by the grammar, with reanalysis occurring in cases where the simplest structure turns out to be incorrect. Under this theory, behavioral studies focused on whether the predicted "garden path" effect occurred in contexts that provided information favoring the structurally more complex ("non-minimal attachment") alternative. Deviations from the predicted outcome had to be attributed to methodological problems of one sort or another.

The present theory suggests that the mix of results supporting and contradicting the Garden Path theory is a consequence of the use of stimulus materials that varied

with respect to factors that govern ambiguity resolution. The factor implicated by the Garden Path theory, structural simplicity, is confounded with frequency: Structures that are simpler also tend to be used more often. The two factors can be deconfounded by looking at how the relative frequencies of alternatives vary across items. Thus, most verbs that participate in NP/S ambiguities occur more often in the simpler NP construction; however, verbs such as *dream* violate this general pattern. Some verbs (e.g., those used by Rayner et al., 1983) support attachment of certain prepositional phrases (the minimal attachment pattern), but others favor nonminimal attachment (Taraban & McClelland, 1988). Findings supporting the minimal attachment pattern are thus not surprising, simply because frequency biases also favor this pattern, particularly in the NP/S and MV/RR ambiguities. However, the entire range of outcomes, including both minimal and non-minimal attachment results, can be understood in terms of how frequency information associated with individual lexical items affects processing.

This theory also sheds light on the question of the scope of contextual effects on processing. Previous research assessed whether information provided by the context could be used to "override" the minimal attachment pattern, with the studies yielding a mixed pattern of results. From our perspective, framing the question in this way ignores two critical factors: (a) the nature of the information provided by the context, and (b) facts about the lexical items creating the ambiguity that constrain context effects. Contexts obviously differ in the extent to which they provide disambiguating information; those that are only weakly constraining would not be expected to have a large impact on anyone's theory. In the absence of an account of the factors that determine the strength of a contextual manipulation, the mere failure to "override" minimal attachment with a particular set of stimulus materials is not very informative. We have attempted to provide a characterization of degree of contextual constraint in terms of the extent to which it differentiates among alternatives at different levels of representation. We have also suggested that most contexts probably are weakly constraining, in the sense that they provide an effective basis for deciding between a small number of alternatives but are less effective in isolating a single alternative in

advance. Moreover, the effects of contextual information are limited by lexical factors, specifically the frequencies of the alternatives at different levels. Putting these factors together yields a system that is contextually constrained but lexically driven. The theory admits a broad range of possible outcomes, including "override" of minimal attachment with appropriate stimulus materials. Thus, Trueswell et al. (1993), who developed a careful analysis of factors that influence preferences in the NP/S ambiguity, were able to construct materials that eliminated the garden path effect entirely. Similarly, Taraban and McClelland (1988) used their account of some of the factors that determine prepositional phrase attachment preferences to generate materials that either did or did not favor minimal attachment. In summary, with the emergence of a theory of relevant aspects of lexical structure and contextual information, we are able to replace the question of whether the parser does or does not follow the minimal attachment principle with questions concerning the ways in which a variety of grammatical and probabilistic constraints interact to yield the range of outcomes that are observed.

Chapter 3

Cross-Linguistic Attachment Preferences: Evidence from English and Spanish[1]

3.1 Introduction

Much of the research on the human sentence processing mechanism (HSPM) over the past twenty-five years has involved the search for computationally-based universal principles underlying the HSPM's operation (e.g. Bever, 1970; Kimball, 1973; Frazier & Fodor, 1978; Frazier, 1979; among numerous others). One such proposed principle is Late Closure (Frazier, 1979; Frazier & Rayner, 1982; cf. the principle of Right Association (Kimball, 1973)):

(17) Late Closure: When possible, attach incoming lexical items into the clause or phrase currently being processed (i.e. the lowest possible nonterminal node dominating the last item analyzed).

In other words, attachments to structures associated with the most recent words in the input stream are to be preferred over attachments associated with words further

[1]This work was conducted in collaboration with Edward Gibson, Enriqueta Canseco-Gonzalez, and Greg Hickok.

back in the input stream (cf. Gibson, 1991). This principle correctly predicts the preferred readings of the adverbial attachments in the examples in (18):

(18)

a. John thought Bill died (# will die) yesterday.

b. Tom will say that the Red Sox will win (# won) tomorrow.

The preferred reading of each of these sentences links an adverbial (*yesterday* in (18a), *tomorrow* in (18b)) to the second (lower) clause: the clause currently being processed. Furthermore, this preference is so strong that semantic factors have little apparent effect on it. For example, when the lower clause is in the future tense in (18a) (*will die*), so that it is semantically incompatible with the past tense modifier *yesterday*, people still initially prefer to associate the adverbial with the incompatibly tensed verb, resulting in processing difficulty.[2] A similar effect occurs in (18b), where the future adverbial *tomorrow* cannot be semantically associated with the past tense verb *won*. Further evidence for the Late Closure principle is given in the examples in (19) and (20):

(19) Joe called (# looked) the friend who had smashed his new car up. (Church, 1980)

(20)

a. # Since Jay always jogs a mile seems light work. (Frazier & Rayner, 1982)

b. # Because she washes her sheets never look dirty.

In (19), the particle *up* can attach to either the VP headed by *smashed* or the VP headed by *called*. Attachment to the more recent VP *smashed* is preferred by Late Closure. Furthermore, even if the attachment to the higher VP is necessary for a grammatical reading of the sentence to result (e.g. when the first VP is headed by *looked*), this attachment is still initially blocked, yielding a garden-path effect.

The garden-path effect in (20a) appears as a result of a strong preference to

[2]Following standard practice, we will prefix sentences that cause noticeable processing difficulty with the symbol "#".

attach the NP *a mile* as direct object of the verb *jogs*, rather than as subject of the matrix clause to come. This preference is explained by Late Closure, because the direct object attachment is to the clause currently being processed, unlike the matrix subject attachment. A similar effect in (20b) is explained in the same way. See Frazier & Rayner (1982) and Mitchell (1987) for on-line evidence of this strong local preference.

Effects such as those given above have, until recently, provided the bulk of the evidence for Late Closure as a universal parsing principle. While these effects are strong, it is important to note that this evidence is incomplete in at least two important respects: 1) the evidence comes solely from English; and 2) the evidence comes entirely from VP (and S) attachment preferences. Cuetos & Mitchell (1988) attempted to address both of these concerns by examining the attachment preferences of a relative clause (RC) to potential NP attachment sites in both English and Spanish. In particular, Cuetos & Mitchell (hereafter C&M) examined NPs of the following form, where the preposition was always *of* in English (*de* in Spanish):

(21) NP_1 Prep NP_2 RC

In order to give subjects complete sentences to read, these noun phrases were placed in object positions of matrix verbs, yielding examples as in (22):

(22)

a. The journalist interviewed the daughter of the colonel who had had the accident.

b. El periodista entrevisto a la hija del coronel que tuvo el accidente.

The English relative clause *who had had the accident* (*que tuvo el accidente* in the Spanish version) can modify the NP headed by *colonel* (*coronel*), as depicted in Figure 3-1, or it can modify the NP headed by *daughter* (*hija*), as depicted in Figure 3-2.

C&M argued that if the principle of Late Closure is indeed a computationally-based universal principle, then it should apply in these constructions in both languages. This hypothesis therefore predicts a preference for the attachment of the RC to the NP headed by *colonel* (*coronel*), because this NP occurs most recently in the

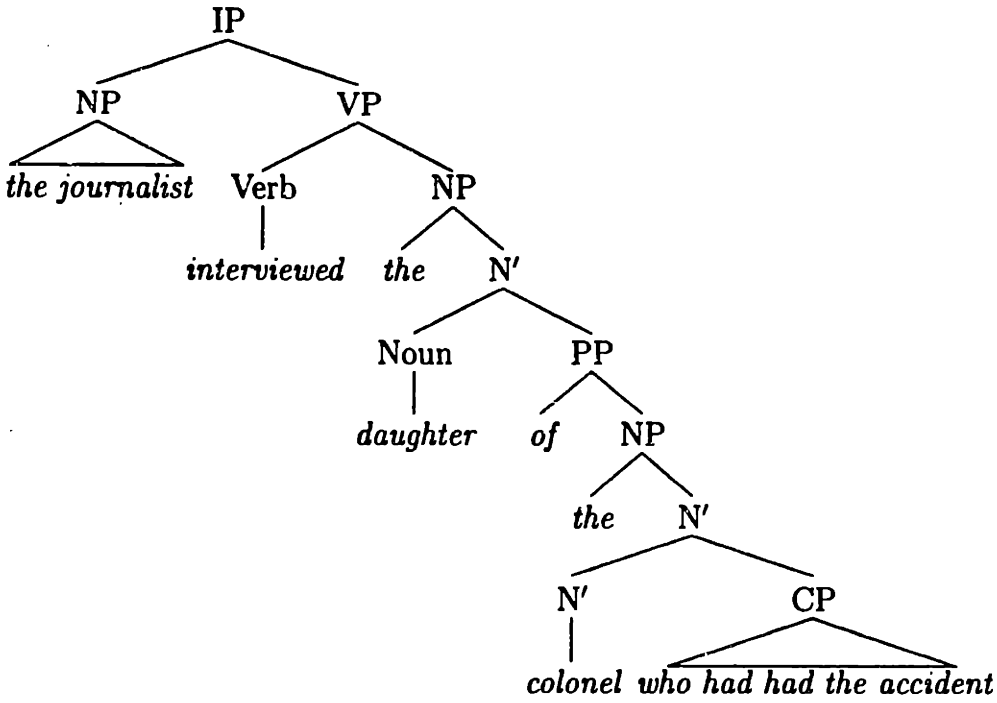


Figure 3-1: Phrase structure tree corresponding to the Late Closure interpretation of *The journalist interviewed the daughter of the colonel who had had the accident.*

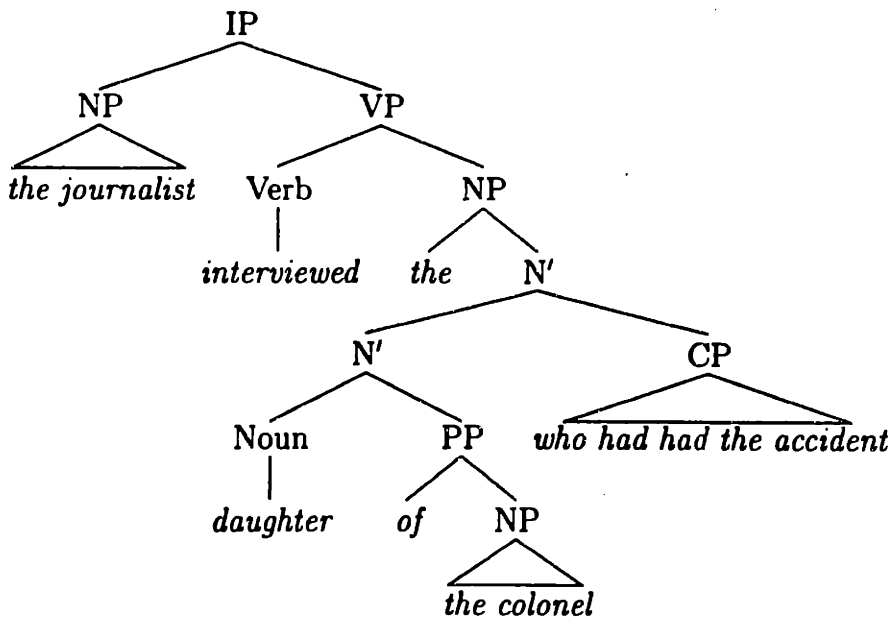


Figure 3-2: Phrase structure tree corresponding to the Early Closure interpretation of *The journalist interviewed the daughter of the colonel who had had the accident.*

input string, parsing from left to right.

An alternative possibility put forward by C&M is that Late Closure is a language-particular strategy, one that English-speakers have adopted for essentially accidental reasons, having nothing to do with computational grounding. In particular, C&M propose that comprehenders may apply Late Closure because it has been successful in the past (the *linguistic tuning* hypothesis) (Mitchell & Cuetos, 1991a; cf. Bates & MacWhinney, 1987; MacWhinney, 1987). According to this view, other attachment strategies are equally likely. For example, people might instead prefer to attach to the least recent (first) available attachment site, so that the attachment to *daughter* (*hija*) in the above examples would be preferred. Following C&M, we will refer to the strategy that prefers the least recent attachment sites over more recent sites (primacy preference) as Early Closure (cf. Kimball, 1973; Church, 1980), the opposite of Late Closure.[3] Thus under the linguistic tuning view, English speakers might follow one strategy (Late Closure) in the processing of the above NPs which results in a recency preference, while Spanish speakers follow a different strategy (Early Closure) which results in a primacy preference.

In order to test these hypotheses, C&M performed a number of experiments, both off-line and on-line, in Spanish and in English. The English experiment was a questionnaire study in which subjects were presented with examples like (22a) and were then asked who was the subject of the relative clause modifier. For example, a subject who was presented with (22a) would be asked who had had the accident. The answer "the colonel" would indicate a preferred low attachment (following Late Closure, recency preference), while the answer "the daughter" would indicate a preferred high attachment (following Early Closure, primacy preference). C&M found that English subjects preferred the attachment to the second (later, lower) NP 58% of the time

[3]Note that Kimball's Early Closure principle (originally called simply "Closure"; first referred to as Early Closure by Frazier (1979)) does not actually predict primacy preference in examples like (22). Kimball's Early Closure principle states that nodes are closed for attachment purposes as soon as possible, but it says nothing about the order in which attachment sites are considered for attachment. In fact, Kimball's principle of Right Association works in tandem with the Early Closure principle to give the same effects as *Late* Closure on NP attachment ambiguities like those in (22).

(as in Figure 3-1) as opposed to 37% for the attachment to the first NP (as in Figure 3-2).[4] C&M concluded from these results that English speakers tend to follow a Late Closure strategy when dealing with ambiguous attachments to NP sites.

Although C&M did not conduct any studies to verify that Late Closure is used on-line in these structures in English, some on-line evidence for the use of Late Closure in NPs similar to those in (22a) is provided by Clifton (1988) (also described in Frazier (1990)). Clifton and Frazier report self-paced reading studies in which subjects read sentences like those in (23):

(23)

a. The doctor called in the son of the pretty nurse who hurt herself.

b. The doctor called in the son of the pretty nurse who hurt himself.

If the final word in the relative clause modifier is the pronominal anaphor *herself* as in (23a), then the relative clause *who hit herself* can only attach to the NP headed by *nurse*, since this is the only attachment site which is possibly feminine in gender (ignoring the matrix subject NP (*the doctor*)). Similarly, if the relative clause contains the pronoun *himself* as in (23b), then the relative clause is most likely to modify the NP *the son*, since *son* is the only attachment site with overtly masculine gender. The other likely attachment site is headed by *nurse*, which is usually assumed to be associated with feminine gender (given no context).

If the human parser follows Late Closure, then (23a) should be processed without any reanalysis, since the relative clause *who hurt herself* initially attaches to the rightmost available attachment site as required. However, the relative clause *who hurt himself* in (23b) must attach to a less recently available site, the NP headed by *son*. Hence, Late Closure predicts that this attachment will involve reanalysis and therefore predicts that processing a sentence like (23b) will be slower than processing a sentence like (23a). This is exactly what Clifton and Frazier found.

Thus there appears to be some evidence that Late Closure makes the appropriate

[4]The percentages do not add to 100% because a small fraction of the subjects reported that the relative clause modified the matrix subject NP.

prediction for attachments to NP sites in English.[5] In contrast, C&M found that it does not work in Spanish. In the Spanish versions of the sentences in (22), C&M found that the attachment of the relative clause to the first (earlier, higher) NP was preferred 62% of the time versus 36% for the second NP, showing a primacy preference pattern following Early Closure. C&M also found evidence for Early Closure in on-line studies. In particular, C&M performed self-paced reading studies on Spanish subjects using materials such as the following:

(24)

a. Alguien disparó contra el criado de la actriz que estaba en el balcón con su marido.
"Someone shot the servant(MASC) of the actress who was on the balcony with her husband."

b. Alguien disparó contra la criada de la actriz que estaba en el balcón con su marido.
"Someone shot the servant(FEM) of the actress who was on the balcony with her husband."

c. Alguien disparó contra la actriz que estaba en el balcón con su marido.
"Someone shot the actress who was on the balcony with her husband."

Each of (24a) and (24b) has two possible NP attachment sites for the relative clause beginning with the words *que estaba en el balcón* ("who was on the balcony"). In (24b), the two attachment sites are both feminine in gender: *la criada* ("the servant") and *la actriz* ("the actress"). In (24a), the two possible sites differ in gender: one is masculine (*el criado*), and the other is feminine (*la actriz*). Because the full relative clause modifier in both (24a) and (24b) concludes with the PP *con su marido* ("with her husband"), the relative clause can only pragmatically attach to a feminine attachment site, ruling out *el criado* in (24a). C&M found that the PP *con son marido* in (24a) was processed significantly more slowly than in either of the controls (24b) or (24c). These results suggest that people are processing these constructions

[5]Several more recent studies, however, have failed to find strong (or in some cases, any) effects of Late Closure in English in constructions like those used by C&M. We discuss some of these studies below (Carreiras & Clifton, in press; Gilboy, Sopena, Clifton & Frazier, 1993). For the moment, we can at least say that, to the extent that English subjects display any preference, it is in favor of more recent attachments.

using an Early Closure strategy, necessitating a reanalysis in the processing of (24a), but not (24b) and (24c), making (24a) significantly slower to process.

Mitchell & Cuetos (1991a) replicated this result with an additional control condition, which ruled out the possibility (suggested by Clifton (1988) and Frazier (1990)) that the difficulty in (24a) compared to (24b) and (24c) was simply the result of the need to make a decision between two possible sites. Such a decision need not be made in (24b) because it is globally ambiguous, and in (24c), only one attachment site exists (*la actriz*).

A number of other researchers have since found similar differences between Spanish and English relative clause attachment preferences. For example, Carreiras & Clifton (in press) performed on-line studies in Spanish and English and observed a primacy preference in Spanish, but no preference one way or the other in American English. (C&M's studies involved British English.) In a questionnaire study investigating constructions like those used by C&M, Gilboy, Sopena, Clifton, and Frazier (1993) found similar differences between English and Spanish attachment preferences, such that English-speaking subjects displayed a stronger preference to attach to the more recent of two sites than did Spanish-speaking subjects, although Gilboy et al. also observed a great deal of variation among preferences within each language.

Based on their own results, C&M concluded that the Spanish parser operates by means of an Early Closure strategy rather than Late Closure on constructions like those in (24) and (22b). Because of the difference between the English and Spanish results, C&M conclude that there may be no universal computational base for processing principles like Late Closure: "It may be that different languages make use of parsing strategies in an essentially arbitrary way." (Cuetos & Mitchell, 1988, p. 72)

However, before accepting the linguistic tuning hypothesis, it is worth checking to see 1) how pervasive the preference for attachments to non-recent sites is in Spanish; and 2) whether there are alternative explanations of the effects observed by C&M in Spanish. Let us first examine some of the original English motivations for Late Closure, translated into Spanish. For example, consider (18b) and (20b) along with

their Spanish translations:[6]

(18b) John thought Bill died (# will die) yesterday.

(25) Juan dijo que Bill se murió (# morirá) ayer.

(20b) # Because she washes her sheets never look dirty.

(26) Dado que ella lava sus sábanas nunca se ven suciás.

The effect in (25) is very similar to that in its English translation. In particular, there is a strong preference to associate the adverbial *ayer* ("yesterday") with the most recently occurring VP (clause). Just as in the English example, this preference is manifested as a garden-path effect if this attachment is blocked because of a semantic clash between the adverbial *ayer* and the future tense verb *morirá*.

Similarly, there is a strongly preferred reading associated with the ambiguous (26), a preferred reading that is predicted by Late Closure. In particular, the NP *sus sábanas* ("her sheets") prefers to be attached as object of the verb *lava* ("washes"), resulting in an interpretation in which the VP *nunca se ven suciás* ("never look dirty") has a non-lexical pronoun as its subject.[7] The alternative reading, in which the verb *lava* has no object and the NP *sus sábanas* is the matrix subject, is not noticed by most Spanish speakers.

Thus, based on intuitive judgments, it appears that Late Closure applies in many constructions in Spanish, just as it does in English, at least those involved in VP attachment site ambiguities.[8] One possible account of the data then is that Late Closure applies generally in Spanish, just as in English, but there are additional factors (perhaps in both languages) involved in the NP constructions that C&M tested, which give rise to the differences that they found. If such factors exist, then resolving an attachment ambiguity between two NP sites would differ significantly from resolving

[6]Verb-particle constructions such as in (19) do not occur in Spanish, so particle attachment is not a possible test in Spanish.

[7]This interpretation is available in Spanish because Spanish is a so-called "pro-drop" language, allowing non-lexical pronouns in subject positions.

[8]In fact, Cuetos & Mitchell (1988, pp. 92-93) agree that this is the case. See also Carreiras & Gutierrez (1992) for evidence that Spanish speakers follow Late Closure in constructions other than relative clause attachment.

an attachment ambiguity between two VP sites. Intuitive evidence for this position comes from a comparison of the effect associated with violating Late Closure in each case. As observed above, the preference for attachments to recent VP sites is very strong, leading to significant difficulty when semantic/pragmatic information contradicts the recency preference (e.g., (18b) and (20b) above). In contrast, semantic and pragmatic factors can override a recency preference in ambiguous attachments to NP sites in English. For example, C&M found that the preference to attach the complementizer *that* to a non-human site overrides the general pattern to prefer recent attachments in English:

(27) the book of the girl that was in the living room

In (27), the relative clause *that was in the living room* prefers to attach to the non-human NP site *the book*, despite the fact that there is a more recent attachment site, *the girl*. Thus selectional restrictions can serve to override a recency preference in ambiguous attachments to NPs. Furthermore, many other kinds of semantic/pragmatic information can resolve these ambiguities in favor of less recent attachment sites without causing noticeable difficulty, as demonstrated in (28), which is easy to understand despite the fact that the prepositional phrase *with 300 pages* does not attach to the most recent available site:

(28) the book on the table with 300 pages

Indeed, C&M's off-line results partially confirm these intuitions for both English and Spanish. That is, while there is a preference across speakers for the most recent attachment in English and for the less recent attachment in Spanish, this preference is not very strong in either language. In particular, in 37% of the trials, subjects chose the less recent attachment site in the English experiment, and 36% chose the most recent site in the Spanish experiment.

Thus it may be that the factor in English that makes the non-recent attachment easier in the NP cases than in the VP cases is also at work in the Spanish NP attachment examples (perhaps to a stronger degree), making the non-recent attachment sites more available than they otherwise would be. That is, in determining the best

attachment site for a modifier that may grammatically attach in more than one location, the parser may consider a number of independent factors and weigh them together in order to make a decision.[9]

There are a number of candidates for additional factors involved in preference determination in the range of data we have discussed so far. One possible candidate is a preference for items to attach to the elements bearing the core semantic content of a clause — i.e., the components of the head verbal θ-grid of a clause, consisting of a verbal role-assigner and its role-receivers. If resources are short, it is natural that the elements of this grid should be ranked highly as attachment sites, because the hearer/reader needs to keep the components of this grid in memory in order to correctly interpret the clause. A general statement of this preference principle is given below (cf. Milsark, 1983; Abney, 1989; the Relativized Relevance Principle of Frazier (1990, p. 321); the Verbal Argument Attachment Preference property of Gibson (1991, p. 137); and the Construal Principle of Gilboy et al. (1993) and Frazier & Clifton (in preparation)):[10]

(29) Verbal θ-Grid Preference: Prefer attachments to elements that can be members of a verbal θ-grid.

The Verbal θ-Grid Preference principle predicts that the complementizer *that* in the NP *the books on the table that...* prefers to attach to *books* rather than *table*, because the NP headed by *books* can be a member of a verbal θ-grid (as the subject of a following verb) whereas the NP headed by *table* cannot. Similarly, if the NP headed by *books* is in the direct object position of a verb as in *I read the books on the*

[9]Note that the hypothesis that structures are evaluated with respect to multiple factors says nothing about whether the parser proceeds serially, or whether the parser can retain two or more structures in parallel. That is, the parser might initially follow up on only the most highly ranked structure first, and then try others in a reanalysis phase. Under such a serial hypothesis, presumably structures that were relatively highly ranked initially would be easier to re-access. But it could also be that the parser retains two or more structures, as long as they are close together in terms of cost (Gibson, 1991). Both the serial and the parallel hypotheses are compatible with the framework as specified thus far.

[10]The Verbal θ-Grid Preference factor might be extended to apply to non-verbal θ-grids that make up the core semantic meaning of clauses, such as in sentences which are syntactically headed by a copula verb (e.g. *Mary is a doctor*, or *Mary is tall*). We leave to future work the issue of whether or not this extension is empirically correct.

table that..., then attachment to *books* is also preferred by Verbal θ-Grid Preference.[11]

Of course, Late Closure (recency preference) by itself predicts exactly the opposite effects of Verbal θ-Grid Preference in these examples. So assuming that both preference factors are used to evaluate possible attachment sites, they will tend to negate each other, giving the result that neither attachment site is strongly preferred, which is the observed outcome. In order to formalize this idea, we might think of each of these preferences contributing positive weight (facilitation) to the attachment sites that each favors. Alternatively, these preferences might be manifested as costs counting against those attachment sites that violate preferences. The second approach is the one that we will follow here, where cost is given in terms of an abstract unit, the Processing Load Unit (PLU) (Gibson, 1991). The first of the two cost-bearing properties that we hypothesize is an instantiation of Late Closure, the Property of Recency Preference (Gibson, 1991):

(30) The Property of Recency Preference (PRP):
The load associated with the structure resulting from an attachment of structure X = a decay function of (the number of more recent words that would also allow an attachment of structure X) $*$ x_{RP} PLUs.

The definition of the property of Recency Preference given here extends the one given in Gibson (1991) in assuming that cost increases with distance according to an exponential decay function, rather than according to a linear function.[12] We assume an exponential decay function following standard claims in the short term and working memory literature, wherein activation of some element in memory decays according to an exponential function of time, approaching an asymptote in the limit (see e.g. Anderson, 1983, 1985; McClelland & Rumelhart, 1981; and the references in each).

[11] As mentioned above, there are a number of other possible factors that could help to explain the range of data that we have looked at so far. For the moment, we will discuss the effects only in terms of the Verbal θ-Grid Preference factor. Other candidates for additional preference factors, including Frazier's Relativized Relevance Principle and Mitchell et al.'s (1990) proposal are discussed in the General Discussion section.

[12] The definition of the property of Recency Preference given in Gibson (1991) also distinguished thematic from non-thematic attachments. Since we will not discuss data points which are relevant to this distinction, we will not complicate our definition of Recency Preference by including such a distinction, although it will probably be necessary eventually.

Table 3.1: Costs (PLUs) Associated with Attachment Sites According to Recency

Attachment Site	Cost
Most recent	$0x_{RP}$
Second most recent	$1x_{RP}$
Third most recent	$1.5x_{RP}$
Fourth most recent	$1.75x_{RP}$
Fifth most recent	$1.875x_{RP}$

In our cost terms, we assume that the cost of attaching to a particular site increases with distance from the current word, but by smaller and smaller increments according to an exponential decay function. According to such a function, the attachment of a modifier to the third most recently available site will be more costly than the attachments to either the most recent or the second most recent sites. However, the difference in cost between attachment to the second- and third-ranked sites will not be as large as the difference between the first- and second-ranked sites.

For simplicity, we will assume an exponential decay function whose cost increment divides in half as distance increases. Under this assumption, if the initial cost increment is x_{RP} PLUs, then the cost associated with the nth most recent attachment is $(2 - (1/2)^{n-2}) * x_{RP}$ PLUs. The costs associated with various attachment sites are given in Table 3.1.

The property of Verbal θ-Grid Preference, whose cost is also measured in terms of Processing Load Units, is given as follows:

(31) The Property of Verbal θ-Grid Preference:
Associate a cost of $x_{\theta+V}$ PLUs with the attachment of a structure to a site which cannot be a member of a θ-grid heading a clause.

Given these definitions of cost-assigning properties, note that the strong VP attachment preferences observed as the original motivating evidence for recency preference are unaffected by the presence of a Verbal θ-Grid Preference, because in all of these ambiguities, both attachment sites are verbal θ-grid sites, so that Verbal θ-Grid

Preference does not create any relative preference.[13]

Now let us see how the properties of Recency Preference and Verbal θ-Grid Preference can be used to explain the pattern of data observed by C&M in the NP constructions in (22), repeated below:

(22a) The journalist interviewed the daughter of the colonel who had had the accident.

(22b) El periodista entrevisto a la hija del coronel que tuvo el accidente.

In (22a), The total cost associated with the attachment of the relative pronoun *who* to the NP headed by *daughter* is x_{RP}, because the attachment is to the second most recent site, while the cost of attachment to the NP headed by *colonel* is $x_{\theta+V}$, because this attachment (while most recent) is not to an element that can be a member of a verbal θ-grid. In order to account for the fact that the more recent attachment is preferred in English constructions like these, we assume that the cost associated with violating Recency is greater than that for not attaching to a verbal θ-grid. Furthermore, we assume that the costs associated with the two properties are close together, in order to account for the fact that the preference for one attachment over the other is not particularly strong in either language: e.g. only 58% versus 37% preference off-line in English (Cuetos & Mitchell, 1988).

Finally, to account for the fact that the preferences are reversed (as measured by C&M) in Spanish constructions like these, we also assume that the cost associated with at least one of the cost-assigning properties can vary cross-linguistically, so that in Spanish the cost associated with violating Recency (x_{RP}) is *less* than that for attaching to a non-verbal θ-grid element ($x_{\theta+V}$). Because facts about VP attachment preferences are very similar in the two languages, as evidenced by the similar judgments in (18b) and (25) as well as (20b) and (26), we hypothesize that it is the cost associated with the Verbal θ-Grid Preference Property that varies from language to

[13]Note that the VP attachment ambiguities also make a property of Primacy Preference unworkable in general. To see this, suppose that there existed a Primacy Preference property, much like the Recency Preference property, so that attachments to the first possible site would be facilitated to some degree by this property as recent attachment sites are. Such a facilitation might then be used to explain the relative ease with which the relative clause in examples like (28) attaches to the first available NP attachment site. However, such a principle would also have no way of explaining the contrast between relatively easy examples like (28) and strong garden-path effects as on (18)–(20). Thus a Primacy Preference property seems empirically unworkable.

language. In particular, we propose that the cost associated with the Verbal θ-Grid Preference Property ($x_{\theta+V}$) is greater in Spanish than in English, enough so that it is greater than that associated with Recency Preference (x_{RP}).

The partial theory of attachment preferences just described is so far empirically indistinguishable from a particular version of a linguistic tuning hypothesis, in which attachment preferences are learned on a construction-by-construction basis from patterns observed in the input, such that English speakers happen to apply Late Closure in both NP and VP attachment constructions, while Spanish speakers apply Late Closure in VP attachment constructions and apply Early Closure in NP attachment constructions.[14] However, the two-factor theory (Recency and Verbal θ-Grid Preference) makes a very different set of predictions from the linguistic tuning hypothesis in more complex constructions involving a third NP attachment site, as in (32) and the examples in (33) for English and in (34) for the corresponding Spanish:

(32) NP_1 Prep NP_2 Prep NP_3 RC

(33)

a. the lamps near the paintings of the house that was damaged in the flood

b. the lamps near the painting of the houses that was damaged in the flood

c. the lamp near the paintings of the houses that was damaged in the flood

(34)

a. las lámparas cerca de las pinturas de la casa que fue dañada en la inundación

b. las lámparas cerca de la pintura de las casas que fue dañada en la inundación

c. la lámpara cerca de las pinturas de las casas que fue dañada en la inundación

[14] A tuning hypothesis in which an attachment pattern is set for a language as a whole cannot account for the fact that in Spanish NP attachments the HSPM seems to generally obey Early Closure, while in VP attachments, the HSPM follows Late Closure. Thus to be empirically satisfactory, a tuning hypothesis must distinguish at least NP and VP attachment sites (Cuetos & Mitchell, 1988; Mitchell et al., 1990). In general, a tuning theory might distinguish a number of possible factors in determining attachment preferences, including the category of the attachment sites. Tuning theories that make additional distinctions are discussed in the General Discussion.

Each of these constructions consists of an NP (*the lamp(s), la(s) lámpara(s)*) followed by a PP (*near the painting(s), de la(s) pintura(s)*), a second PP (*of the house(s), de la(s) casa(s)*), and a relative clause (*that was damaged in the flood, que fue dañada en la inundación*). Although there are initially three possible NP attachment sites for the RC, only one is permitted when the singular verb *was* (*fue*) is processed, because only one of the NP sites is singular in number.

Because there are three attachment sites in these constructions, a complete model of the relevant components of the HSPM must predict a ranking (partial order) among the sites corresponding to how easily people accept attachments to each. In a parallel model, this ranking might come about as a result of the application of a general complexity metric, as suggested in Gibson (1991). On the other hand, if the proposed model is serial — or if it allows parallel structures in general, but only one is built for the constructions under consideration — then a reanalysis metric is required that predicts an ordering of the two attachment sites ranked below the initial choice. Thus in a serial model it is necessary to model not only which attachment site(s) is (are) accessed first (as a principle like Late Closure predicts), but also which sites are accessed next if the first attempt(s) fail.

Of the various alternative theories, we can first consider Late Closure, which predicts that the NPs in (33a) and (34a) should be the easiest to process, because the RC attaches to the rightmost attachment site in these NPs. Late Closure is a serial processing principle involved only in the selection of an initial attachment site, however, so by itself it predicts nothing about possible differences between the processing complexities of the other two NPs in each language. That is, Late Closure dictates only that the RC in (32) will be attached to NP_3; it says nothing about what factors will interact in the reanalysis to allow attachment to NP_1 or NP_2, so Late Closure predicts no difference between the processing of (33b) and (33c), nor does it predict a difference between (34b) and (34c). However, if it is assumed that Late Closure is also involved in a metric for ranking structures for reanalysis, the principle predicts that the NPs in (33b) and (34b) are the next easiest to process, because the RC in each is attached to the second most recent NP site, which is the

second most available attachment site via Late Closure. Similarly, Late Closure as a reanalysis principle predicts that (33c) and (34c) are the most difficult to process of the two sets of examples, because in these NPs the RC attaches to the least available site, the initial (highest) NP. This ranking of the NPs — NP_3, NP_2, NP_1 — also applies to a parallel interpretation of Late Closure, in which structures that result from attachments involving more recent words are ranked more highly, as in a system that uses the property of Recency Preference alone to evaluate such structures.

Early Closure makes exactly the opposite predictions of Late Closure with respect to the constructions in (33) and (34). In particular, the NPs (33c) and (34c) are predicted to be easiest to process, since these involve attachment to the NPs which are closed earliest. Furthermore, if Early Closure is involved in determining both the first and subsequent analyses, attachment to (33b) and (34b) is predicted to be next easiest, with attachment to (33a) and (34a) hardest of the three possibilities in each language.

Of course, neither Early Closure nor Late Closure alone can account for even the data that we have so far seen from English and Spanish. However, the combination of the two principles under linguistic tuning (where VP attachment cases are tuned separately from NP attachment cases) can be made to fit the data for cases in which only two potential attachment sites are available. Furthermore, we can extend the predictions of the linguistic tuning hypothesis to the three attachment site cases so that linguistic tuning applies generally across all ambiguous NP attachments in each language. Under this version of the linguistic tuning hypothesis in English, Late Closure would apply, so that attachments to NP_3 in (32) should be preferred over attachments to NP_2, which should in turn be preferred over attachments to NP_1. In Spanish, Early Closure would apply, and thus the pattern of preferences should reverse: NP_1 would be preferred to NP_2, which would in turn be preferred over NP_3.

In contrast, the combination of Recency Preference and Verbal θ-Grid Preference predicts a very different set of results for three attachment site cases. As discussed above, the processing difficulty of attaching to a particular site is predicted by the cost associated with that site according to the two preference factors. In particular,

Table 3.2: Costs (PLUs) Associated with Attachments to NP_1, NP_2 and NP_3

Attachment Site	Cost
NP_1	$1.5x_{RP}$
NP_2	$x_{RP} + x_{\theta+V}$
NP_3	$x_{\theta+V}$

th ost associated with attachment to NP_1 is $1.5x_{RP}$ PLUs, since this site is the least recent of three available attachment sites. No cost is associated with attachment to this site via the property of Verbal θ-Grid Preference, because this attachment site can be a member of a verbal θ-grid. By a similar calculation, the cost associated with attachment to NP_2 is $x_{RP} + x_{\theta+V}$ PLUs, because 1) this site is the second most recent possible attachment site; and 2) this site cannot be a member of a verbal θ-grid. Finally, the cost associated with attachment to NP_3 is just $x_{\theta+V}$ PLUs, because this site cannot be a member of a verbal θ-grid. No cost is associated with attachment to this site via Recency Preference, since this is the most recent possible attachment site. The costs associated with the three attachment sites are summarized in Table 3.2.

Recall that the evidence from two-site attachments indicates that in English the cost associated with the property of Recency Preference is greater than that associated with the property of Verbal θ-Grid Preference. Furthermore, we assume that these two costs are close together, because the preference is not very strong in two-site NP attachment cases. Given these initial constraints, the predictions for English are as follows:

- Attachment to NP_3 should be much the easiest, since $x_{\theta+V}$ PLUs (the cost associated with attachment to NP_3) is much less than either $x_{RP} + x_{\theta+V}$ PLUs (the cost associated with attachment to NP_2) or $1.5x_{RP}$ PLUs (the cost associated with attachment to NP_1).

- From our assumption that the costs associated with the two properties are close together, it is reasonable to assume that $x_{\theta+V} > 0.5x_{RP}$. From this assumption

it follows that the cost associated with the attachment to NP_2 is more than $x_{RP} + 0.5x_{RP} = 1.5x_{RP}$ PLUs, so that attachment to NP_1 ($1.5x_{RP}$ PLUs) should be the next easiest, with attachment to NP_2 hardest. Thus the predicted increasing difficulty order in English is NP_3, NP_1, NP_2.

In order to account for the two-site Spanish facts, it was hypothesized that the cost in Spanish associated with the property of Verbal θ-Grid Preference ($x_{\theta+V}$) is higher than the cost associated with violating Recency (x_{RP}). This constraint results in three possible orderings in Spanish:

- We know that $x_{RP} < x_{\theta+V}$ in Spanish. If $1.5x_{RP} > x_{\theta+V}$ (most natural under the assumption that the costs associated with the two properties are close together), then attachment to NP_3 would be easiest (cost = $x_{\theta+V} < 1.5x_{RP}$), NP_1 would be the next easiest (cost = $1.5x_{RP}$) and NP_2 would be the hardest (cost = $x_{\theta+V} + x_{RP} > 2x_{RP}$), resulting in a difficulty ordering of (NP_3, NP_1, NP_2), the same ordering as in English.

- In contrast, if $x_{\theta+V} > 1.5x_{RP}$ then NP_1 would be the easiest attachment to make (cost = $1.5x_{RP}$), NP_3 would be the next easiest (cost = $x_{\theta+V} > 1.5x_{RP}$), and NP_2 would be the hardest (cost = $x_{\theta+V} + x_{RP} > 2.5x_{RP}$), resulting in a difficulty ordering of (NP_1, NP_3, NP_2).

- Finally, if $x_{\theta+V} = 1.5x_{RP}$ then NP_1 and NP_3 would be equally difficult attachments to make, with NP_2 still the hardest.

Hence under the two-factor theory (Recency and Verbal θ-Grid preferences) there is a single predicted difficulty order among the three attachment sites in English, and there are three such possible orderings in Spanish, one of which is the predicted English order (NP_3, NP_1, NP_2). The predictions of the various theories are listed in Table 3.3.

For examples like those in (33) and (34), native speakers' intuitions suggest that either NP_1 or NP_3 is easiest to attach to, but deciding between the two is rather difficult. However, both English and Spanish speakers have a strong intuition that

Table 3.3: Attachment Difficulty Ordering Predictions of Various Theories

Theory	Difficulty ordering[a]
Late Closure (Recency Preference)	NP_3, NP_2, NP_1
Early Closure (Primacy Preference)	NP_1, NP_2, NP_3
Linguistic Tuning (English)[b]	NP_3, NP_2, NP_1
Linguistic Tuning (Spanish)[b]	NP_1, NP_2, NP_3
Recency and Verbal θ-Grid Preferences (English)	NP_3, NP_1, NP_2
Recency and Verbal θ-Grid Preferences (Spanish)	Any order of $\{NP_1, NP_3\}$; NP_2 hardest

[a]Ordering is least to most complex. [b]This is the tuning theory that applies across all ambiguous NP attachments.

attaching to the middle site (NP_2) is quite difficult, a finding not accounted for by either Late or Early Closure alone, because both principles predict that the middle site should be intermediate in difficulty. In order to quantify these intuitions and to test these theories of initial parsing preferences and reanalysis in English and Spanish, we performed three experiments using NPs like those in (33) and (34): 1) an English grammaticality-judgment experiment; 2) an English self-paced reading experiment; and 3) a Spanish self-paced reading experiment.

3.2 Experiment 1: English Off-Line Grammaticality Judgment Task

This experiment was intended as an initial, off-line, test of our intuitions with respect to processing difficulty. Subjects were asked to perform an acceptability judgment for items like those shown in (33) above.

Method

Subjects

Fifteen subjects participated in the experiment as volunteers. All were native speakers of English.

Materials

Twenty-one complex NP items of the form shown in (32) and below in (35) were constructed. The head nouns of all three potential attachment sites in all items were non-human and had regular plural forms. The complementizer in the relative clause was always *that*. The items are listed in Appendix B.1. Each item contained an initial NP followed by two PPs and a relative clause which was a plausible modifier for any of the three preceding NPs. However, the relative clause was always singular in number (marked on the second word of the relative clause in all items), and only one of the preceding NPs was also singular, so the relative clause could grammatically modify only one of the three preceding phrases. Thus each item had a low attachment version, a middle attachment version, and a high attachment version, depending on which NP was singular, as in (35) and (33).

(35)

a. Low: $NP_{1(plural)}$ Prep $NP_{2(plural)}$ Prep $NP_{3(singular)}$ $RC_{singular}$

b. Middle: $NP_{1(plural)}$ Prep $NP_{2(singular)}$ Prep $NP_{3(plural)}$ $RC_{singular}$

c. High: $NP_{1(singular)}$ Prep $NP_{2(plural)}$ Prep $NP_{3(plural)}$ $RC_{singular}$

The stimuli were constructed so that the second (low) PP always preferentially attached to NP_2 rather than NP_1. This insured that attaching the relative clause to NP_2 would be allowed by the grammar (i.e. no crossed branches). The strong preference for the second PP to attach to NP_2 was created syntactically, semantically, and/or pragmatically. Examples are shown in (36):

(36)

a. the computer(s) near the model(s) of the building(s) that was destroyed in the fire

b. the sofa(s) beside the rug(s) under the table(s) that was dropped by the movers

In (36a), the attachment to NP_2 is preferred for syntactic reasons: a PP headed by the preposition *of* is a commonly occurring argument of the noun *model*, whereas the noun *computer* is not often modified by a PP headed by *of*. In (36b), the PP *under the table* is much more likely to modify the noun *rug* than the noun *sofa* for pragmatic reasons: sofas are not normally found under tables.

The 21 experimental items created as just described were combined with 60 fillers to form three lists. Half of the fillers were grammatical and acceptable complex NP fragments of English, and half were ungrammatical or unacceptable (e.g., incorrect agreement, crossed branches, missing words, multiple center embeddings or subjacency violations). The fillers were of approximately the same length and complexity (number of words and constituents) as the experimental items. The experimental items were counterbalanced across the three lists so that each list contained exactly one version (high, low, or middle attachment) of every item. One pseudo-random ordering of all fragments was generated for each list, with the stipulation that no two experimental items were adjacent.

Procedure

Subjects were instructed to read each fragment and circle either "Good" or "Bad" (printed below the fragment). They were told to make a decision on the basis of their first impression about whether the fragment was a possible English description of a thing or set of things; they were instructed to ignore questions of plausibility. Subjects were also given five examples, two of which were presented as "Bad," one because it contained a number agreement violation, and the other because it was missing a word. Most subjects completed the study in approximately 15 minutes.

Table 3.4: Experiment 1 Ungrammaticality Judgments

Attachment Site	Ungrammaticality Judgments (%)
High (NP_1)	34
Middle (NP_2)	71
Low (NP_3)	31

Results

Three of the original 21 experimental items (marked with an asterisk in Appendix B.1) were excluded from the analysis because post-experiment conversations with several subjects revealed that the meaning of the second PP shifted subtly depending on whether it was singular or plural. Including these items had no effect on the pattern of results.

The mean percentage of sentences judged ungrammatical in each attachment condition (high, middle, and low) is presented in Table 3.4. A one-way ANOVA on these data revealed a main effect of attachment (by subjects, $F1(2, 28) = 8.923$, $p < .005$; by items, $F2(2, 34) = 13.798$, $p < .001$). Individual mean comparisons indicated that while the low and high attachment conditions did not differ (both $ps > .65$), the middle attachment condition was judged ungrammatical more often than either of the other two attachment conditions (vs. low: $F1(1, 14) = 21.000$, $p < .001$; $F2(1, 17) = 29.708$, $p < .001$; vs. high: $F1(1, 14) = 13.304$, $p < .005$; $F2(1, 17) = 14.853$, $p < .005$).

Discussion

The results of Experiment 1 clearly support the intuition that middle attachment is difficult in English, relative to low and high attachment: the English HSPM does not allow easy access to attachment sites like NP_2. As noted earlier, a reanalysis procedure based on a metric of ranking structures solely by either Late or Early Closure is not consistent with this result. In both cases, the reanalysis metrics predict that NP_2 should be the second site accessed, and thus attachments to this site should be rated

grammatical more often than attachments to one of the other sites. Because this is clearly not the case, we must abandon the hypothesis that the HSPM's reanalysis procedure is based solely on a closure principle. Similarly, the version of the linguistic tuning hypothesis that we are initially considering here makes the same predictions for English as Late Closure alone does, and so it too is unable to account for these results.

In contrast, the two-factor theory predicts that attachments to NP_2 are the most difficult, as is observed. This theory also predicts that attachments to NP_3 are easier than attachments to NP_1, a result not observed in this experiment. However, it could be that this off-line experiment is simply not sensitive enough to measure the difference in difficulty between attachments to NP_1 and NP_3. That is, reanalysis might be fast or easy enough to cover up a difference between these two sites. A more sensitive measure of difficulty than off-line grammaticality judgments is therefore necessary to reveal a difference, if one exists. Experiment 2 provides an on-line measure of difficulty, which is much more likely to identify smaller differences between the various attachment conditions.

3.3 Experiment 2: English On-Line Grammaticality Judgment Task

While Experiment 1 provided an initial confirmation of English-speakers' intuitions, Experiment 2 should provide a more powerful test of the predictions of Late and Early Closure, linguistic tuning, and the combination of Recency Preference and Verbal θ-Grid Preference. This experiment uses a self-paced word-by-word reading task combined with a continuous grammaticality judgment task and thus should allow a comparison of both percentage of ungrammaticality judgments and of reading times, in each region of the items from Experiment 1.

Method

Subjects

Thirty Massachusetts Institute of Technology undergraduates were paid $4.00 each for their participation in the study. All were native speakers of English, and none had participated in Experiment 1.

Materials

The same items from Experiment 1 were used in the current experiment, with a number of exceptions: The three items whose meaning altered noticeably with number were not used (items marked with an asterisk in Appendix B.1). In addition, half of the items were altered so that the disambiguating verb in the relative clause was plural and only one of the three potential attachment sites was plural. All items were modified to have exactly the same number of words and to fit on a single 80-character display line. Items which were changed for Experiment 2 are listed in Appendix B.2.

A written survey was also conducted, using the 18 experimental items and 22 (grammatical) fillers, to insure that in all the experimental items, the second PP (PP_2) preferentially attached to NP_2 and not NP_1. Each item in the survey consisted of an NP followed by two PPs. Immediately following each item, the subject was asked to circle the noun which was modified by the second PP. The two choices (the noun from the first NP and the noun from the first PP) were listed to the right of the question, and for half of the items, the noun from the NP was listed first. For the other half, the noun from the first PP was listed first. One random ordering of the 40 experimental items and fillers was created so that 10 items were on each page. The order of the pages was randomized for each subject. Fifteen subjects, none of whom participated in the reading time portion of the experiment, completed the survey.

For 14 of the 18 experimental items, every subject indicated that PP_2 attached to NP_2 (as desired). On 3 of the remaining 4 items, 14 of the 15 subjects (93%) indicated that PP_2 attached to NP_2, and for the remaining item, 13 of the 15 subjects (86%) indicated that PP_2 attached as desired. Across all 18 experimental items,

the preference for PP_2 to attach to NP_2 is thus 98%, indicating that, as intended, the possibility of attachment to the intermediate site in the three attachment site constructions is not ruled out by the grammar. The percentage of subjects preferring attachment of PP_2 to NP_2 over NP_1 is shown with each item in Appendix B.1 (for items which were in the survey) and Appendix B.2.

The 18 experimental items were combined with the 60 fillers from Experiment 1 (modified to fit on an 80-character display line) to form three lists. Experimental items were counterbalanced as in Experiment 1.. Ten practice items were also constructed to be similar to the fillers.

Procedure

An IBM PS/2 personal computer was used to present the stimuli. A subject saw two screens of instructions, followed by the 10 practice items and then the 78 experimental and filler items. The experimental and filler items were presented in a different random order for each subject.

On every practice item, and on one third of the other items, the subject was asked to paraphrase the item for the experimenter to insure that subjects were attending to the stimuli. On the practice items, the experimenter corrected the subject's errors.

At the beginning of a trial, the item was displayed on the screen with all non-space characters replaced by dashes. When the subject pressed the space bar, the first word of the item was displayed, replacing the corresponding dashes. When the subject pressed the space bar a second time, the first word reverted to dashes, and the second word was displayed in place of the appropriate dashes. Each subsequent press of the space bar revealed the next word and removed the previous word. Pressing the space bar on the last word of the item caused the item to be replaced by a request to paraphrase the item just read or a request to press the space bar again to continue with the experiment. The computer recorded the time between each button-press.

In addition, the subject was instructed to press a key marked "NO" just above the space bar if, at any point during the display of an item, the item became "unacceptable" (ungrammatical) in English. Both the instructions and the practice items

1	2	3	4
NP_1	PP_1	PP_2	complementizer
the computer(s)	*near the model(s)*	*of the building(s)*	*that*

5	6	7
disambiguation	remainder	final word
was destroyed	*in the*	*fire*

Figure 3-3: Regions used for analysis. (Items were always presented to subjects one word at a time, with no line breaks.)

included some clearly ungrammatical fragments as examples. Display of an item was terminated whenever the subject indicated that it had become ungrammatical. Subjects could still be asked to paraphrase an item if they had decided it was ungrammatical.

Most subjects completed the experiment in approximately 20 minutes.

Results

For the purposes of analysis only, items were separated into seven regions, as shown in Figure 3-3. Trials on which the subject indicated that the item was ungrammatical prior to Region 5 (the second (number-marked) word of the relative clause) were excluded from all of the following analyses. This affected less than 1.0% of the trials.

Cumulative Grammaticality Judgments

The cumulative percentage of ungrammaticality judgments in each region (including those trials judged ungrammatical prior to Region 5) is shown in Figure 3-4. An item was scored as ungrammatical in a given region if the subject indicated that the item was ungrammatical on any word of the region or in any preceding region. An omnibus ANOVA on these data (excluding those judged ungrammatical prior to Region 5) crossing the factors attachment (low, middle, high) and region (1 through 7) yielded a highly significant interaction ($F1(12, 348) = 6.768$, $p < .001$; $F2(12, 204) = 4.199$, $p < .001$) and significant main effects of both attachment ($F1(2, 58) = 7.710$, $p < .005$; $F2(2, 34) = 4.602$, $p < .05$) and region ($F1(6, 174) = 89.661$, $p < .001$;

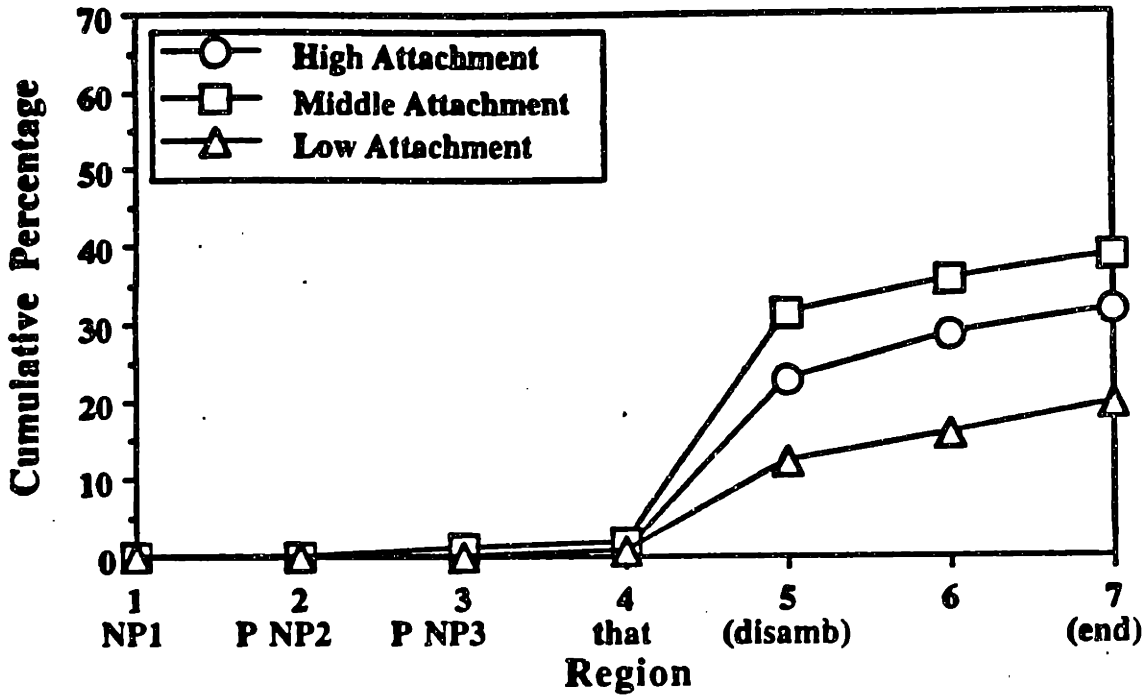


Figure 3-4: Cumulative percentage of items judged ungrammatical by attachment condition in each region in Experiment 2 (English).

$F2(6,102) = 153.541$, $p < .001$).

No effects are present in Regions 1 through 4, of course, because data of relevance in these positions was excluded. Even including such data, no effects are significant (all ps $> .13$).

In Region 5, the first two words of disambiguation in the relative clause, a significant main effect of attachment was present ($F1(2,58) = 7.856$, $p < .001$; $F2(2,34) = 4.990$, $p < .05$), and individual mean comparisons indicated that the low attachment condition was judged ungrammatical less often than either the high attachment ($F1(1,29) = 6.250$, $p < .05$; $F2(1,17) = 4.509$, $p < .05$) or the middle attachment condition ($F1(1,29) = 14.107$, $p < .001$; $F2(1,17) = 8.649$, $p < .01$). However, although the high attachment condition was judged ungrammatical less often numerically than the middle attachment condition, this difference was not significant ($F1(1,29) = 2.546$, $p > .12$; $F2(1,17) = 1.585$, $p > .20$).

The pattern of data in Region 6 was very similar to that in Region 5: A main effect of attachment was present ($F1(2,58) = 7.353$, $p < .005$; $F2(2,34) = 4.758$,

$p < .05$), and the low attachment condition was judged ungrammatical less often than either of the other two attachment conditions (vs. high: $F1(1,29) = 6.918$, $p < .05$; $F2(1,17) = 4.399$, $p = .051$; vs. middle: $F1(1,29) = 12.521$, $p < .005$; $F2(1,17) = 8.670$, $p < .01$). Also as in Region 5, the high attachment condition was better numerically than the middle attachment condition, but again the difference was not significant ($F1(1,29) = 1.662$, $p > .20$; $F2(1,17) = 1.053$, $p > .30$).

The same pattern persisted into Region 7 (the last word of the item): A main effect of attachment was present ($F1(2,58) = 6.411$, $p < .005$; $F2(2,34) = 3.629$, $p < .05$), and the low attachment condition differed from both the high attachment ($F1(1,29) = 4.999$, $p < .05$; $F2(1,17) = 2.887$, $p = .108$) and middle attachment conditions ($F1(1,29) = 10.376$, $p < .005$; $F2(1,17) = 7.075$, $p < .05$). While the high attachment condition was numerically better than the middle attachment condition, the difference was not significant ($F1(1,29) = 2.009$, $p > .16$; $F2 < 1$).

Reading Times

A regression equation predicting reading time from word length was constructed for each subject, using all items (fillers and experimental items) except for those experimental items judged ungrammatical prior to Region 5 (the first two words of disambiguation). For any item judged ungrammatical, the word on which the judgment was made (the last word presented) was also excluded. At each word, the reading time predicted by the subject's regression equation was subtracted from the actual measured reading time, and all analyses were performed on these differences (residual reading times). This transformation has the effect of removing extraneous variance by subtracting out a baseline for each subject, and by controlling for noise due to length effects. See Ferreira & Clifton (1986) and Trueswell & Tanenhaus (1991) for some discussion of this procedure. The raw reading times by condition in each region are reported in Appendix B.4; although the significance levels are consistently weaker in the raw times, the pattern of data is the same as that in the residual reading times.

Items which were judged ungrammatical at any point were excluded from the remainder of the analyses, and one subject, who had judged all items in the high

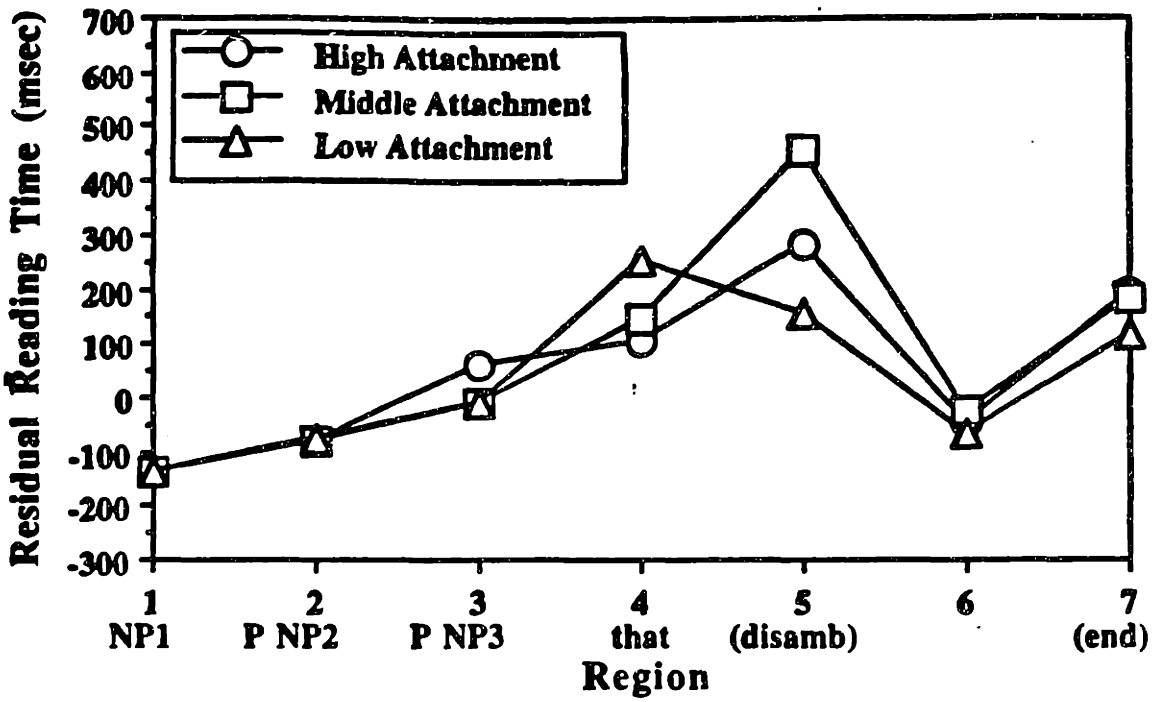


Figure 3-5: Residual reading time per word by attachment condition in each region in Experiment 2 (English).

attachment condition ungrammatical, was also excluded. The remaining data were trimmed by excluding residual reading times beyond 4 standard deviations from the appropriate attachment condition × position cell mean. This trimming affected 1.1% of the data.

The data, plotted in Figure 3-5, were analyzed in an omnibus ANOVA, crossing attachment (high, middle, and low) and region (1 through 7). This ANOVA revealed a strong main effect of region ($F1(6,168) = 22.204$, $p < .001$; $F2(6,102) = 35.053$, $p < .001$) and a strong interaction ($F1(12,336) = 4.663$, $p < .001$; $F2(12,204) = 4.127$, $p < .001$). The main effect of attachment was non-significant by subjects ($F1(2,56) = 1.817$, $p > .17$) but significant by items ($F2(2,34) = 4.218$, $p < .05$).

One-way ANOVAs (high, middle, low attachment) in each individual region separately showed no effects of attachment in Regions 1 and 2 ($ps > .70$). In Region 3 (the second PP), however, the high attachment condition tended to be slow, and this was reflected in a main effect of attachment ($F1(2,56) = 5.163$, $p < .01$; $F2(2,34) = 3.390$, $p < .05$) and in individual mean comparisons: While low and middle attachments did

not differ (Fs < 1), high attachments were significantly slower than middle attachments ($F1(1,28) = 5.702$, $p < .05$; $F2(1,17) = 4.893$, $p < .05$) and also slower than low attachments ($F1(1,28) = 6.552$, $p < .05$; $F2(1,17) = 3.261$, $p < .09$). Although we are not particularly concerned with effects in this region, it is worth noting that the various conditions do differ at this point: In the high attachment condition, the noun phrases in Regions 2 and 3 agree in number, while in the other two conditions, they disagree. Given that subjects are expected to attend to number information in order to perform the grammaticality judgment task, it is not surprising that it affects reading times as attachments become complex.

In Region 4 (the complementizer introducing the relative clause (and the ambiguity)), there was a slight tendency for the low attachment condition to be slower than the high attachment condition: The one-way ANOVA including all three conditions revealed a marginal effect by subjects ($F1(2,56) = 2.615$, $p < .09$), but the effect was not significant by items ($F2(2,34) = 1.846$, $p > .17$). Individual comparisons showed that the low attachment condition tended to be slower than the high attachment condition ($F1(1,28) = 4.517$, $p < .05$; $F2(1,17) = 3.804$, $p < .07$) but did not differ from the middle attachment condition ($F1(1,28) = 2.003$, $p > .15$; $F2 < 1$). The high and middle attachment conditions also did not differ (Fs < 1). As with Region 3, we are not particularly concerned with effects in this region, but it is not surprising that some (weak) effects are present.

In the theoretically interesting Region 5, where attachment of the relative clause is disambiguated, the main effect of attachment was highly significant ($F1(2,56) = 10.197$, $p < .001$; $F2(2,34) = 7.766$, $p < .005$), and all three conditions differed from each other, with low attachment fastest, high attachment intermediate, and middle attachment slowest. Low attachment was significantly faster than both high attachment ($F1(1,28) = 6.217$, $p < .05$; $F2(1,17) = 7.646$, $p < .05$) and middle attachment ($F1(1,28) = 12.872$, $p < .005$; $F2(1,17) = 12.470$, $p < .005$), and high attachment was faster than middle attachment, though this difference was only marginal by items: $F1(1,28) = 7.950$, $p < .01$; $F2(1,17) = 3.655$, $p < .08$. Thus as in the on-line grammaticality judgment data, subjects display an immediate preference

for low attachments. However, the reading time data also differentiate between the high and middle attachment conditions: Although neither attachment is initially preferred, subjects have less difficulty attaching to the highest site than to the middle site. Finally, in Regions 6 and 7 (the last three words of the relative clause), there were no main effects of attachment (all $ps > .20$) and no expected effects, so individual mean comparisons were not performed.

Discussion

Both the grammaticality judgment data and the reading time data in this experiment show that English-speaking subjects have an initial preference for low attachments. This manifested as relative difficulty with both high and middle attachments as soon as these attachments were forced by syntactic agreement constraints. In the case of the grammaticality judgment data, the low attachment condition was judged ungrammatical less often than either the high or middle attachment condition, and in the reading time data, the low attachment disambiguation was read more quickly than either the high or middle attachment disambiguation, all of which contained exactly the same words. These data thus clearly indicate a preference among English speakers for Late Closure over Early Closure as a statement of initial attachment preference.

However, as suggested by intuition and the results of Experiment 1, the Late Closure strategy alone cannot fully account for the data in Experiment 2: Late Closure itself predicts nothing about preferences for attachments other than to the lowest (most recent) site, and although subjects preferred low attachments overall, they also differentiated between high and middle attachments. Furthermore, the simplest extension of Late Closure, which would indicate that the order of preference for attachment sites should run from the lowest to the highest, cannot account for the data, because middle attachments, not high attachments, are the most difficult of the three in the reading time data. In the grammaticality judgment data as well, although the middle versus high difference did not reach significance, the middle attachment condition was numerically the most difficult. Similarly, these results are problematic for the linguistic tuning hypothesis, because it predicts that English speakers should

be relying (only) on Late Closure in the construction tested.

As in Experiment 1, the difficulty of middle attachments compared to high and low attachments is predicted by the two-factor (Recency and Verbal θ-Grid Preference) theory. Furthermore, the expected advantage of low attachments compared to high attachments is found in Experiment 2, whereas the two conditions did not differ in Experiment 1. This provides support for the two-factor theory and suggests (not surprisingly) that the on-line task of the current experiment provides a more powerful methodology for determining the relative difficulty of various different attachments. We now turn to an examination of these same three-site constructions in Spanish, where again the two-factor theory and linguistic tuning make differing predictions.

3.4 Experiment 3: Spanish On-Line Grammaticality Judgment Task

Experiments 1 and 2 provided clear evidence that low attachments are less difficult than higher ones in English, and this is consistent with work on various other English constructions in which attachment ambiguities are claimed to be resolved on the basis of Late Closure. The linguistic tuning hypothesis, as described above, predicts that while this is true for English, the opposite is true for Spanish: High attachments are preferred and should therefore be easier than low ones. This was shown by C&M in constructions with two potential NP attachment sites, and Experiment 3 was conducted in order to determine whether Spanish speakers display this same preference in the more complex three-attachment-site construction.

Experiments 1 and 2 also demonstrated that Late Closure (a simple recency preference) will have to be supplemented in some non-trivial manner to account for the full range of data, particularly the relative difficulty of attachments to an intermediate site compared to attachments to a high site. Given the more complex English pattern, the Spanish subjects in Experiment 3 could vary from the English subjects in Experiment 2 along two different dimensions: First, Spanish subjects might show a general preference for high attachments overall (replicating C&M), or they might

prefer low attachments overall, as in English in Experiment 2. Second, regardless of whether they follow Late or Early Closure, Spanish subjects may (as in Experiment 2) or may not show the most difficulty with middle attachments in comparison to either high or low ones.

Method

Subjects

Twenty-six native Spanish speakers from the Boston, Massachusetts academic community (primarily undergraduates and graduate students) participated in the study, either as volunteers or for $4.00. Two of these subjects were excluded because of poor performance or difficulty in the paraphrase task. Most of the remaining 24 subjects were from Mexico (8 subjects), Puerto Rico (6), Spain (3), Argentina, Chile, El Salvador, Guatemala, Peru, or Venezuela; the remaining subject was a native of the United States, with a Mexican father, and was bilingual.

Materials

The experimental, filler, and practice items from Experiment 2 were translated into Spanish. However, a number of changes to the experimental items were necessary: Because some prepositions in Spanish have differing forms depending on the grammatical number of their object, the total number of words in each condition of each item was not precisely matched; these differences were confined to the first three regions of each item. As in English, the disambiguating word immediately following the complementizer in the relative clause always had a different form depending on the number of the noun phrase with which it agreed. Finally, the Spanish versions of the materials tended to be longer than their English counterparts, and because the items were kept to a total length of 80 characters or less, the relative clause was sometimes changed to be only three or four words long. The disambiguating verb in the relative clause was therefore sometimes not *was/were* (*fue/fueron*) and was instead a form with more semantic content. The Spanish experimental items are shown in

Appendix B.3 along with their English translations.

As in Experiment 2, a separate survey was conducted to insure that in the experimental items, PP_2 preferentially attached to NP_2. The surveys were constructed exactly as in Experiment 2. Twenty subjects participated as volunteers from the Spanish-speaking community of Pittsburgh, Pennsylvania. For 14 of the 18 items, the percentage of subjects preferring attachment of PP_2 to NP_2 was 80% or higher. Of the remaining items, two displayed the desired attachment across 75% of the subjects, one across 65% of the subjects, and the remaining item displayed the desired attachment across 20% of the subjects. The percentage of subjects preferring attachment of PP_2 to NP_2 for each item is shown in Appendix B.3.

Whereas in the English stimuli, PP_2 attached to NP_2 98% of the time, in the Spanish, the corresponding value is only 83%. A number of subjects completing the Spanish survey noted that the phrasing of the question about attachment in the survey was unnatural for the item with only a 20% preference. Excluding this item, the overall preference for attachment of PP_2 to NP_2 for the Spanish stimuli rises to 86%, which is nevertheless somewhat low. Given the undesirable possibility that some of the stimuli may occasionally have been treated as ungrammatical, we conducted additional analyses excluding the four items for which the attachment of PP_2 to NP_2 was preferred below 80% (the items marked with an asterisk in Appendix B.3). These analyses are discussed in the Results section below, where relevant.

Procedure

The procedure was exactly as in Experiment 2. All instructions, materials, and feedback were presented in Spanish. Most subjects completed the experiment in less than 30 minutes.

Results

The items were broken into seven regions as in Experiment 2 (see Figure 3-3). Trials on which an item was judged ungrammatical prior to Region 5 (the disambiguation) were excluded from all analyses, affecting approximately 3.0% of the data. As noted

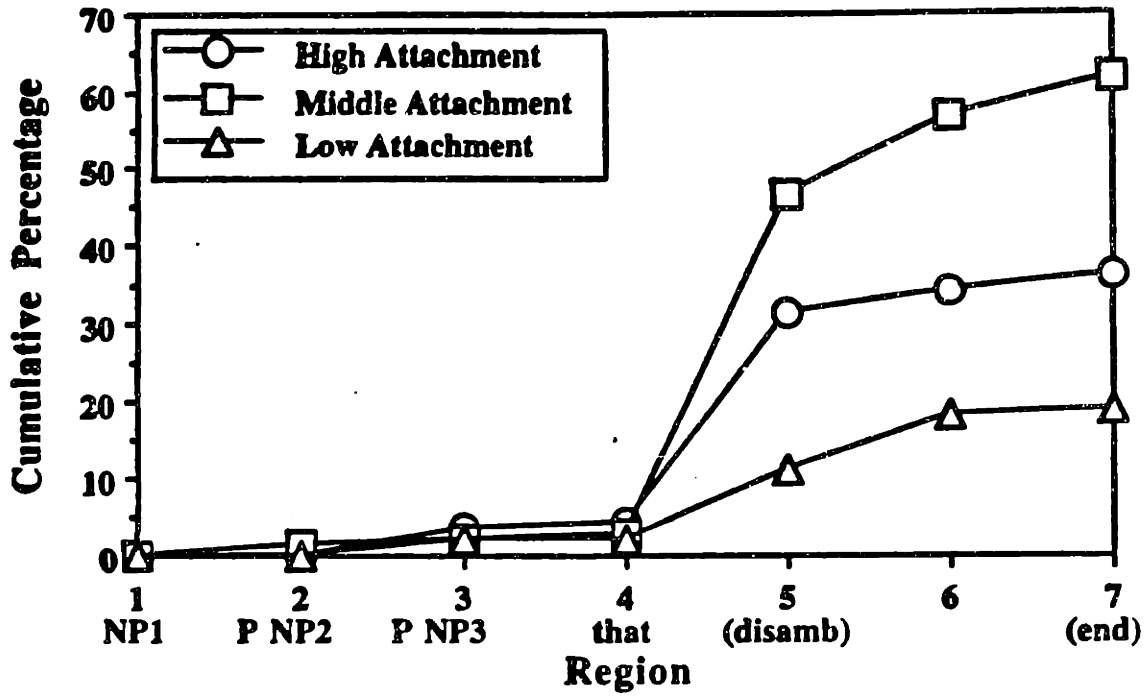


Figure 3-6: Cumulative percentage of items judged ungrammatical by attachment condition in each region in Experiment 3 (Spanish).

above, we also conducted analyses excluding 4 of the 18 items, and these results are discussed for Region 5 (the disambiguation) for both cumulative grammaticality judgments and for reading time analyses.

Cumulative Grammaticality Judgments

The cumulative percentage of ungrammaticality judgments by region (including those trials judged ungrammatical prior to Region 5) is shown in Figure 3-6. An omnibus ANOVA on these data (excluding those judged ungrammatical prior to Region 5) crossing the factors attachment (low, middle, high) and region (1 through 7) yielded a highly significant interaction ($F1(12,276) = 22.316, p < .001$; $F2(12,204) = 23.630$, $p < .001$), as well as significant main effects of both attachment ($F1(2,46) = 25.028$, $p < .001$; $F2(2,34) = 29.295$, $p < .001$) and region ($F1(6,138) = 92.920$, $p < .001$; $F2(6,102) = 116.551$, $p < .001$).

As in Experiment 2, no effects are present in the first 4 regions, because relevant data in those regions is excluded. Including such data, all effects are still

nonsignificant ($ps > .13$). The pattern of effects in each of Regions 5 through 7 is exactly the same, and the results are very similar to those from Experiment 2: A main effect of attachment is present in each region (Region 5: $F1(2,46) = 18.635$, $p < .001$; $F2(2,34) = 23.132$, $p < .001$; all ps in Regions 6 and 7 $< .001$, we do not report detailed statistics for the sake of brevity), and the low attachment condition is judged significantly better than either of the other two conditions: For the low versus middle attachment comparison, in Region 5, $F1(1,23) = 47.222$, $p < .001$; $F2(1,17) = 44.286$, $p < .001$; in Regions 6 and 7, all $ps < .001$. For the low versus high attachment comparison, in Region 5, $F1(1,23) = 10.653$, $p < .005$; $F2(1,17) = 16.834$, $p < .005$; in Regions 6 and 7, all $ps < .05$. In addition, whereas in English (Experiment 2), the high attachment condition was judged numerically but not significantly better than the middle attachment condition, the two conditions did differ significantly (high better than middle) in this experiment: In Region 5, $F1(1,23) = 6.483$, $p < .05$; $F2(1,17) = 8.256$, $p < .05$. In Regions 6 and 7, all ps $< .005$.

To insure that the relative difficulty of the middle attachment condition was not simply a result of its being ruled out by the grammar, we also performed the above analyses excluding the four items (marked with an asterisk in Appendix B.3) most likely to be ruled out by the grammar on the basis of the survey described above in the Materials section. Most importantly, in Region 5, where such an effect ought to be most apparent, the pattern of results is unchanged: The low attachment condition is judged better than either of the high or middle attachment conditions, and the high attachment condition is judged better than the middle attachment condition (all $ps < .05$). In Regions 6 and 7, the pattern of results is also the same, with the single exception that the difference between the low and high attachment conditions is no longer significant. However, note that the concerns which motivated these additional analyses predict only that the middle attachment condition should be affected; neither the high attachment condition nor the low attachment condition can become ungrammatical in these stimuli. Furthermore, the difference between the analyses is primarily the result of a relatively large change in the high attachment

condition mean (toward the low attachment condition mean), so it is quite clear that this is unrelated to concerns about the grammaticality of the middle attachment condition. In any case, the fact that the results in Region 5 are not at all affected indicates that the relative difficulty of the middle attachment condition is not simply a result of its being ruled out by the grammar.

Reading Times

Residual reading times were calculated for each subject's data exactly as in Experiment 2. Items judged ungrammatical were excluded from further analyses, and three subjects were excluded: Two had judged all the items in the high attachment condition ungrammatical, and one had judged all the items in the middle attachment condition ungrammatical. Trimming was performed as in Experiment 2, affecting 1.2% of the data. As in Experiment 2, raw reading times are reported in Appendix B.4, and the data patterns discussed below for the residual reading times are also present in the raw reading times.

Figure 3-7 plots the residual reading time data by condition and region. An omnibus ANOVA crossed the factors attachment (high, middle, and low) and region (1 through 5). Regions 6 and 7 were not included because not all items had words in these regions, we had no particular predictions about reading times in these regions, and no effects had been found in these regions in Experiment 2. The ANOVA revealed main effects of attachment ($F1(2,40) = 11.182, p < .001$; $F2(2,34) = 9.038, p < .005$) and region ($F1(4,80) = 25.719$, $p < .001$; $F2(4,68) = 20.100$, $p < .001$), as well as a highly robust interaction ($F1(8,160) = 7.811$, $p < .001$; $F2(8,136) = 5.878$, $p < .001$).

Separate ANOVAs in each region showed no effects of attachment in either Region 1 or 2 (all $ps > .12$). In Region 3, analyses by subjects revealed no effects ($F1(2,40) = 1.51$, $p > .23$), but the items analysis showed a marginal effect of attachment ($F2(2,34) = 3.100, p < .06$), and individual mean comparisons showed that the middle attachment condition was marginally slower than both the high attachment ($F2(1,17) = 3.315, p < .09$) and low attachment conditions ($F2(1,17) = 3.504$,

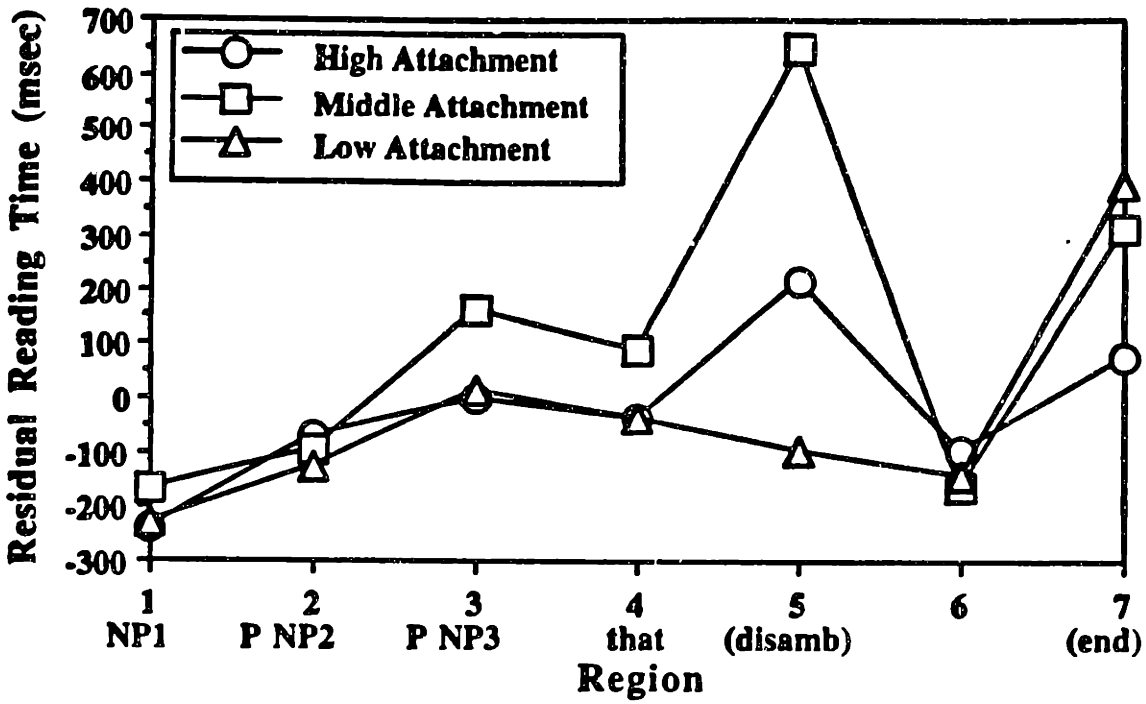


Figure 3-7: Residual reading time per word by attachment condition in each region in Experiment 3 (Spanish).

$p < .08$). The high and low attachment conditions did not differ ($F2 < 1$). As in Experiment 2, we thus find some limited effects in Region 3, although this pattern is different from that in Experiment 2, in which the high attachment condition is slower than both the middle and low attachment conditions. As in Regions 1 and 2, no effects were present in Region 4 (the complementizer) (all $ps > .17$).

In Region 5, the disambiguation point, a main effect of attachment was present ($F1(2,40) = 15.264$, $p < .001$; $F2(2,34) = 13.074$, $p < .001$). Furthermore, all three conditions differed significantly: Low attachments were faster than both high attachments ($F1(1,20) = 15.066$, $p < .005$; $F2(1,17) = 13.995$, $p < .005$) and middle attachments ($F1(1,20) = 21.958$, $p < .001$; $F2(1,17) = 24.486$, $p < .001$), and high attachments were faster than middle attachments ($F1(1,20) = 8.788$, $p < .01$; $F2(1,17) = 5.311$, $p < .05$). Thus, exactly as in Experiment 2 at the disambiguation, low attachments are fastest, middle attachments are slowest, and high attachments are intermediate.

As was done for the cumulative grammaticality judgment analyses, we re-analyzed

the Region 5 reaction time data excluding the four items most likely to be judged ungrammatical in the middle attachment condition. The analysis by subjects duplicated the effects described above: low attachments were fastest, middle attachments were slowest, and high attachments were intermediate (all $ps < .05$). The items analysis also duplicated the effects of the analysis including all 18 items, with the exception that the high attachment versus middle attachment comparison did not reach significance ($p < .12$). However, this lack of significance is not the result of a change in the middle attachment condition, as would be expected if the difficulty of the middle attachment condition was caused by the ungrammaticality of some of the items. In fact, the middle attachment condition mean changes by only 10 msec, whereas the high attachment condition mean changes by 28 msec. The lack of effect is thus the result of an increase in the standard deviation of both conditions, along with a corresponding decrease in degrees of freedom (and power). Thus, the middle attachment condition is clearly more difficult than either of the low or high attachment conditions, and this difference is not the result of the middle attachment being ruled out by the grammar.

Discussion

The results of Experiment 3 quite clearly replicate those of Experiment 2: Spanish-speaking subjects have less difficulty with low attachments than with either middle or high attachments, and high attachments are in turn easier than middle attachments. All of these effects are present in both cumulative grammaticality judgments and in reading times. Thus there appears to be no tendency for subjects to follow Early Closure over Late Closure, as was the case in C&M's data involving constructions with only two potential attachment sites. As in Experiment 2, this pattern of preferences is not predicted by either a single principle (such as Late or Early Closure), or a linguistic tuning hypothesis along the lines of that proposed by C&M. However, a combination of Recency Preference and Verbal θ-Grid Preference, as described above, does predict this pattern, for both English, as in Experiment 2, and for Spanish, as in the current experiment.

Before going on, it is also reasonable to consider one potential objection to these results: We have shown that the Spanish pattern of data replicates the English pattern on these constructions quite closely. However, although our subjects were native Spanish speakers and were not native English speakers, they all lived in or were visiting the United States and spoke at least some English. Fernando Cuetos (personal communication) reports that attachment preferences among Spanish students learning English as a second language vary from the characteristic Spanish preference (for high attachment) to a preference intermediate between Spanish and English, depending on the number of years spent learning English. Thus one might object that the results of Experiment 3 were influenced by the subjects' exposure to English.

However, there are several reasons to doubt this possibility. First, the pattern of results in this experiment is not a blend of the English pattern from Experiment 2 with a Spanish pattern in which high attachments are preferred over low ones (as in C&M), as would be expected if the subjects' native Spanish preferences had been muddied by exposure to English. In fact, if there is any difference, the relative difficulty of the high attachment condition compared to the low attachment condition is more pronounced in Spanish than in English, which is exactly the opposite of what would be expected if exposure to English was responsible for the Experiment 3 results. Furthermore, Cuetos, Mitchell, and Corley (1993), using an off-line grammaticality judgment task as in Experiment 1, have found that native Spanish speakers with no exposure to other languages display the same preferences as are reported here: Attachments to a middle site are rated ungrammatical most often, and, to the extent that there is any preference, high attachments are rated ungrammatical more often than low attachments. These results with monolingual Spanish speakers are nearly identical to the results (for English) of Experiment 1 and are entirely congruent with the results of Experiment 3, indicating that exposure to English cannot explain the Experiment 3 subjects' performance.

3.5 General Discussion

The results from the three experiments show that English and Spanish pattern nearly identically with respect to attachment preferences in constructions involving three potential attachment sites. First, in all three experiments, the middle site was consistently the most difficult to attach to: In Experiment 1, conducted in English, middle attachments were rated unacceptable more often than either high or low attachments; in Experiment 2, with an on-line reading and grammaticality-judgment task, also in English, middle attachments took longer than either high or low attachments, and they were judged unacceptable more often than low attachments. They were also judged unacceptable numerically, though not significantly, more often than high attachments. Finally, in Spanish (Experiment 3, with the same on-line task as Experiment 2), attachments to the middle site both took longer and were rated as unacceptable more often than either attachments to high or to low sites, replicating the results from English (with the addition of the significant middle vs. high attachment difference for grammaticality judgments).

In addition to the similarity between English and Spanish with respect to the relative difficulty of middle attachments, the two on-line experiments showed that low attachments are most preferred in both languages. In both Experiment 2 and Experiment 3, low attachments were processed more quickly and were judged unacceptable less often than either high or middle attachments. Although this difference did not appear in Experiment 1, in which low attachments were judged unacceptable less often than middle attachments but no less often than high attachments, this is not surprising given the relative insensitivity of the off-line task used in Experiment 1.

This pattern of data, in which for both English and Spanish speakers, low attachments are easiest, middle attachments are hardest, and high attachments are intermediate, is unexpected on the hypothesis that either Late or Early Closure alone is the universal choice of attachment heuristic for the HSPM: Each of these heuristics (when extended to predict an ordering of the three attachment sites) predicts incorrectly that middle attachments should be intermediate in difficulty between high and

low attachments. Furthermore, the simplest extension of linguistic tuning for three-site constructions does no better: It incorrectly predicts that Late Closure should apply in English and that Early Closure should apply in Spanish.

In contrast, the theory which includes an interaction of two independent factors -- the properties of Recency and Verbal θ-Grid Preference — predicts the appropriate pattern of data for English, and is also compatible with the Spanish results. In particular, a complexity ordering of NP_3, NP_1, NP_2 is predicted in English, which is the pattern that is observed. The observed complexity order in Spanish — NP_3, NP_1, NP_2 — is also predicted by this theory (although other orders of NP_3 and NP_1 would also be possible). The current evidence therefore favors the two-factor theory over a one-factor theory (e.g. Recency or Verbal θ-Grid Preference alone).

Alternative Tuning Hypotheses

Before accepting the two-factor theory, we can compare it to other instantiations of the general linguistic tuning framework put forward by C&M. In particular, setting preferences based on finer-grained categories than those discussed thus far would lead to the prediction of alternative complexity orders for three-attachment-site cases. For example, a tuning theory might be sensitive to at least the following possible factors: 1) the grammatical category of the attachment sites (NP vs. VP); 2) the category of the attaching item (relative clause vs. prepositional phrase, for example); and 3) how many attachment sites (of each category) are currently available. If the tuning process is sensitive to the number of attachment sites that are available, then any of the six possible complexity orderings of the three NP attachment sites in question would be theoretically possible in a given language.

In fact, however, a tuning theory can make a more specific prediction about the complexity ordering of attachment sites. In the tuning framework, comprehenders' attachment preferences are a product of the input they have received: Comprehenders will prefer to attach to the site which has most often been attached to in the past. Thus determining the predicted complexity ordering for a language is a matter of examining language samples of the appropriate sort and calculating the frequency of

attachment to each possible site.

Gibson and Pearlmutter (1993) calculated these frequencies for English, by extracting and hand-parsing all of the instances of the pattern in (37) from the Brown corpus of about one million words of English text (Kučera & Francis, 1967).

(37) NP_1 [$_{PP_1}$ $Prep_1$ NP_2 [$_{PP_2}$ $Prep_2$ NP_3 [$_{PP_3}$ $Prep_3$ NP_4

Gibson and Pearlmutter then analyzed those 329 instances of (37) in which PP_1 modified NP_1, PP_2 modified NP_2, and PP_3 modified one of NP_1, NP_2, or NP_3. They found that attachments of PP_3 to NP_3 were most common, attachments to NP_1 were next most common, and attachments to NP_2 were least common. They also found that this preference ordering (NP_3, NP_1, NP_2) was even more pronounced when cases involving thematic or lexical preferences were omitted. If we assume that the preferences associated with the attachment of a prepositional phrase like PP_3 are the same as those associated with the attachment of relative clauses, then the version of the tuning hypothesis that allows separate tuning for constructions with different numbers of potential attachment sites predicts the English complexity ordering observed in Experiments 1 and 2.

However, many details of the tuning hypothesis still must be more clearly specified. In particular, although we suggested above that the categories relevant for tuning were essentially syntactic (e.g., NPs, VPs, PPs, relative clauses), this is not a requirement of the hypothesis. The categories might in fact be semantic, pragmatic, morphological, phonological, etc. Furthermore, the grain-size of the categories must be made explicit: In the case of syntactic categories, we have assumed that NP-site cases are tuned separately from VP-site cases, but the system might also tune active VP-site cases separately from passive VP-site cases, singular NP-site cases separately from plural cases, PPs headed by *with* separately from those headed by *of*, etc.

Thus, a tuning theory could consider a potentially huge range of information in determining categories within which to tune preferences: The identity of attachment sites, the identity of attaching constructions, and the number of attachment sites are all potentially important. Furthermore, the system cannot keep track of every

possible feature co-occurrence fact. Some constraints must be present, because as the number of categories tuned separately increases, the amount of input data in each category decreases and the amount of information that has to be remembered about each input sentence increases, making the tuning of any particular category more difficult and less reliable.

However, even supposing that a version of the linguistic tuning hypothesis could be specified which answered the above concerns and accounted for the data in the two- and three-site attachment constructions in both English and Spanish (on the basis of corpus studies in each language), there are still reasons to prefer the two-factor theory over linguistic tuning: First, the two-factor theory *predicts* the complexity ordering observed in the English corpus analysis. In contrast, linguistic tuning makes no predictions as to what patterns should be most frequent in corpus analyses and must have the frequency facts first before it can make any predictions at all.

Second, the two-factor theory predicts that in no language will the middle site (NP_2) in the three-NP-site construction ever be ranked above the most recent site (NP_3), and it also predicts that in constructions involving only VP attachment sites, the most recent site will always be most preferred (because only Recency applies). Linguistic tuning predicts in both of these cases that any ordering of sites is possible a priori. Because only a few languages have been investigated so far, it is obviously premature to draw any conclusions about the validity of the two-factor theory predictions, but they are clearly stronger than the corresponding predictions of the linguistic tuning hypothesis.

Finally, the linguistic tuning hypothesis states only that language learners imitate the preference patterns that they hear, and so it cannot explain the origin of such patterns. It is possible that attachment preferences arise as a result of historical accident, but this hypothesis would predict roughly equal numbers of Early and Late Closure attachment preferences across constructions and languages, and although again, only a very limited amount of data is available, this so far seems not to be the case: Late Closure preferences tend to be much more common than Early Closure preferences (Mitchell & Cuetos, 1991b). The two-factor theory, although it does not

specifically predict any origin for cross-linguistic differences, at least allows for the possibility that variation in processing preferences is directly related to parameterized variation across grammars. We will discuss the question of the possible origins of cross-linguistic differences below, but it should be clear that in this case, the two-factor theory allows for the possibility of a stronger prediction.

Thus the linguistic tuning hypothesis, though currently in need of more careful specification, could probably be modified to account for the range of processing data currently available. The two-factor theory also accounts for this data, however, and it also predicts data from corpus studies; furthermore, it makes strong predictions that certain preference patterns should not appear in any language. To the extent that there is any data at all, these predictions are supported, and this is unexpected under the linguistic tuning hypothesis.

Alternatives and Extensions to Verbal θ-Grid Preference

One of the main ramifications of this work is that, despite initial appearances, Recency Preference is important in the processing of Spanish NP attachment constructions, just as in other Spanish constructions and in English constructions in general. While it is possible that the strength of Recency Preference varies across these two languages (although our proposal does not suggest this), Recency plays an important role in the processing of both languages. Thus contrary to the suggestions made by C&M, Recency Preference is probably a universal computational component of the HSPM.

However, the evidence for Verbal θ-Grid Preference is not yet as compelling. As mentioned earlier, the range of data presented may be compatible with alternatives to the version of Verbal θ-Grid Preference we have described. In order to capture the pattern of similarities and differences between English and Spanish identified by C&M and related work, as well as the results described here, the companion factor to Recency Preference must associate some (uniform) cost with attachments to sites other than the least recent site (NP_1 in both (32), repeated below, and in the corresponding two-site construction).

(32) NP_1 Prep NP_2 Prep NP_3 RC

Furthermore, this factor must not apply to VP attachment sites, or it must apply uniformly to all VP sites, because by hypothesis, the factor does not play a role in VP attachment constructions: Recency alone seems to handle them.[15] Finally, whatever factor is involved must apply to both subjects and direct and indirect objects: The complex NPs in C&M's materials were embedded as either direct or indirect objects of matrix verbs, and we assume that the current stimuli are treated as subjects, at least with respect to θ-role assignment, because they are presented without any preceding verb.

This last requirement is relevant to ruling out an alternative to Verbal θ-Grid Preference which assumes that NPs which have not yet been assigned θ-roles are necessarily held in working memory, with the side effect that they are more easily accessed for attachment purposes. Under such an assumption, a cost would be associated with structures requiring attachment into NPs that have already been assigned θ-roles. For the three-site constructions (in (32)) of the current paper, this hypothesis makes the same predictions as Verbal θ-Grid Preference: High attachments (to NP_1) will not be assigned any cost, because NP_1 has not yet received a θ-role and will therefore be available in working memory for attachments; but both NP_2 and NP_3 are assigned θ-roles by their prepositions, and thus some (fixed) cost factor will

[15]Because no cost is associated with VP attachments by the Verbal θ-Grid Preference Property, the theory presented here predicts an attachment preference ranking of most recent to least recent in attachment examples with two or more VP attachment sites. (Thanks to Lyn Frazier for making this observation.) Specifically, the attachment preference ordering of the adverbial *yesterday* is predicted to be VP_3, VP_2, VP_1 in an example like (i), from Frazier & Fodor (1978):
(i) Joe [$_{VP_1}$ said that Martha [$_{VP_2}$ claimed that 1984 will [$_{VP_3}$ be blissful yesterday.

It is clear that *yesterday* prefers to attach to the most recent VP, VP_3, headed by *be*. However, this attachment is blocked for semantic reasons: a tense clash. As observed by Frazier & Fodor (1978, p. 301), the preference ranking of the other two attachment sites is not clear. However, it is clear that *both* attachments are very difficult to make.

The lack of an intuitive ranking between these two sites can be explained in two possible ways in the theory discussed here. First, it could be that the costs associated with the two sites are close enough together that the difference might not be detectable on the basis of conscious judgments. Recall that the difference in preference between attachment to NP_1 and NP_3 in three-site NP attachment cases like (32) is small enough that no difference is measured off-line, yet a difference can be found on-line. The same might be true here.

A second possibility is that the HSPM might not be able to distinguish between representations associated with large costs relative to a better inexpensive structure (Gibson, 1991). That is, the HSPM may not be able to construct any ranking at all of two representations, if they are sufficiently worse than the best available representation. Further experimental work is necessary in order to determine which one (if any) of these two explanations is correct.

be assigned to middle or low attachments. However, because in the C&M stimuli, both potential attachment sites have already been assigned θ-roles at the point when attachment of the relative clause is being considered, the same cost will be assigned by this factor to either possible attachment, and so the difference between the costs associated with the two sites will be solely a matter of Recency Preference, predicting that attachments into object NPs should pattern in the same way as attachments involving only VP sites in Spanish, which seems to be false.

This same requirement also seems to rule out perhaps the most obvious discourse-based alternative factor, based on the notion of discourse focus or salience. If we suppose that the focus of a discourse is simply "what the discourse is about" (determined operationally, by asking subjects for judgments), we can propose an alternative to Verbal θ-Grid Preference, a discourse salience property that associates a cost of x_{focus} PLUs with the attachment of a structure to a site which is not the focus of the current discourse. This property is closely related to an on-line version of Frazier's (1990) Relativized Relevance Principle, which states that, other factors being equal, a phrase should be construed "as being relevant to the main assertion of the current sentence." (Frazier, 1990, p. 321)[16] In the case of the stimuli in the current experiments, the assumption that the first NP encountered is the current focus would be necessary, and then this factor would behave exactly like Verbal θ-Grid Preference: NP_1 is assigned no additional cost, and NP_2 and NP_3 are assigned an additional cost of x_{focus} PLUs; this is of course the correct set of predictions. In C&M's stimuli, however, assuming that the first NP encountered is the focus will result in the subject NP being focused (which is probably correct), and this predicts that a cost of x_{focus} will be assigned to both potential attachment sites for the relative clause, again with the resulting incorrect prediction that C&M's materials will pattern like attachments involving VP sites.

These discourse-based factors also incorrectly predict that the preference pattern in constructions of the form in (38) should be the same as for NP_1 Prep NP_2 RC,

[16]Frazier's Relativized Relevance Principle as originally proposed is not an on-line principle: It applies *after* Late Closure. However, Cuetos & Mitchell (1991a, 1991b) argue that this principle cannot account for their Spanish high attachment observations if it applies after Late Closure.

as long as the Wh-relative construction, like a prepositional phrase, does not affect discourse focus. An example of such a construction in Spanish, attached as a direct object of *miraba* (*was looking at*), is shown in (39).

(38) NP_1 Wh-rel Verb NP_2 RC

(39) Pedro miraba los libros que pertencian a la chica que...
(Peter was looking at the books which belonged to the girl who/which...)

Mitchell & Cuetos (1991b) report that Spanish speakers overwhelmingly prefer to attach the final relative clause to the more recent site (NP_2: *la chica*) in examples like (39). While this result is not expected under any other current proposals, it follows directly from the two-factor theory including Verbal θ-Grid Preference. The cost associated with both NP_1 (*los libros*) and NP_2 (*la chica*) is the same under Verbal θ-Grid Preference, because both are in positions that can receive a role from a verbal θ-grid. Recency Preference, therefore, is most prominent in making the attachment decision.

The definition of the Verbal θ-Grid Preference Property as originally stated can also be extended to account for additional empirical effects. In particular, Verbal θ-Grid Preference may be defined so that attachments further away from a verbal θ-grid are associated with increasing cost, modulated by a decay function (as for Recency Preference). This extended definition allows a simple explanation for why there is no strong syntactic preference for either NP or VP attachment of a following PP in a Verb NP PP sequence, so that plausibility and lexical factors are the most important in making such attachments (Taraban & McClelland, 1988; Hindle & Rooth, 1993).[17] Consider (40):

(40) I saw the man on the hill.

By the extended Verbal θ-Grid Preference Property, the PP *on the hill* prefers to

[17]Of course, if the PP is an argument of exactly one of the preceding Verb or NP, then there is a syntactic preference for the PP to be attached in the argument position (see Pritchett (1988) and Gibson (1991) and the references in each; cf. Clifton et al. (1991)).

attach as close as possible to the verbal θ-grid headed by *saw*, so that VP attachment is preferred. Recency Preference results in a preference to attach to the more recent phrase, the NP *the man*, so that the two preferences cancel each other out to a large degree. This lack of a strong syntactic pattern is exactly what is observed in on-line experiments (Rayner et al., 1983; Taraban & McClelland, 1988) and in examinations of large corpora of these kinds of phrases (Hindle & Rooth, 1993).

Similarly, the extended Verbal θ-Grid Preference Property explains the contrast between (41a) and (41b), a contrast noted in Ford et al. (1982) (see also Milsark (1983) for further relevant examples):

(41)

a. Martha notified us that Joe died by express mail.

b. Martha notified us of Joe's death by express mail.

The PP *by express mail* prefers to attach to the VP headed by *died* in (41a) by both Recency and Verbal θ-Grid Preferences, so that the semantically odd reading is much the highest ranked, in spite of its semantic peculiarity. In (41b), the PP *by express mail* prefers to attach to the NP *Joe's death* according to Recency Preference, but Verbal θ-Grid Preference also creates a preference for this PP to attach to the VP headed by *notified*. These two preferences therefore partially cancel each other out, allowing semantic factors to help favor attachment to the VP *notified*, which results in a much more acceptable sentence. Thus the Verbal θ-Grid Preference Property has some independent empirical support.

On the Explanation of Cross-Linguistic Differences

Finally, although the present experiments provide evidence favoring a human parsing mechanism based on certain universal principles, the question remains about why certain cross-linguistic differences exist. That is, even the two-factor theory proposed here includes a "parameter" allowed to vary across languages, namely the cost associated with the Verbal θ-Grid Preference factor ($x_{\theta+V}$), and it is necessary to consider why such a parameter might exist, and how children learning a language might acquire

the appropriate value for it.

Three basic possibilities seem to exist: First, the cost associated with a preference factor might simply be a matter of historical accident in any given language. In this case, no relationship is expected between parameter settings in the grammar of a language and cross-linguistic variations in attachment preferences, and the two sets of settings are acquired independently. Children acquiring a language must determine preference factor costs based on adult utterances that they can 1) identify as ambiguous and 2) disambiguate on the basis of other information (plausibility, etc.). Therefore the linguistic parameters of the grammar are likely to be set before the learner can start to store information about ambiguities. Acquisition of the cost associated with Verbal θ-Grid Preference could be accomplished by noting how often attachments in the two NP-site ambiguity (for example) are to the high site versus the low one.

The second possibility is that preference costs are directly determined by one or more parameters of Universal Grammar. So, for example, linguists in the Government Binding framework have proposed that a parameter in Universal Grammar determines whether or not a language allows a null subject in certain situations, and it happens to be the case that English and Spanish differ on this parameter setting (Chomsky, 1981; Jaeggli & Safir, 1989; cf. Hyams, 1986). Suppose that the setting of this parameter is perfectly correlated with whether or not the cost associated with Verbal θ-Grid Preference is more or less than that associated with Recency, perhaps because assigning a θ-role to a null NP is more complex than assigning a role to a phonologically realized NP (one has to do more work to determine what it is, etc.), so that verbal θ-grids in null-subject languages (like Spanish) will necessarily be maintained more strongly in memory, and attaching to an NP not associated with a verbal θ-grid will be relatively more costly. In this case, the setting of the parameter in Universal Grammar also fixes the cost associated with the preference factor, so the language learner need only set the Universal Grammar parameter; the system need not be sensitive to the relative frequencies of different attachments in the input. Without additional stipulations, this type of relationship between parameters

of the grammar and costs associated with preferences therefore predicts that children acquire appropriate attachment preferences at the same time that they set grammar parameters.

The third alternative is an indirect relationship between preference factor costs and one or more parameters of the grammar. In this case, although costs and parameter settings are correlated (perhaps not as strongly as in a direct relationship, but more so than in the historical accident case), setting parameters of the grammar does not set costs. Instead, parameter settings in the grammar have only the usual effect: Various constructions are allowed or disallowed. However, because different parameter settings license different sets of constructions, different languages will have different relative frequencies of various ambiguities and different relative frequencies of possible resolutions (attachments). Thus acquisition must proceed as in the historical accident case: The child must identify ambiguities and be able to disambiguate them in order to form attachment preferences.

In the only other discussion of any of these possibilities of which we are aware, C&M (and Mitchell et al., 1990) consider a hypothesis of the latter, indirect sort: They suggest that the English-Spanish attachment preference difference could be attributed to the grammatically-specified difference in the permissibility of prenominal modifiers (adjectives in particular) in the two languages. They predict that because the Spanish grammar forces adjectives to intervene between a noun and any of its prepositional or relative clause modifiers, NP attachment sites will on the average be farther away from the phrases which attach to them than in English, and thus relatively higher attachments will be more expected and less difficult in Spanish.

This predicts that all languages with both pre-nominal adjectival modifiers and post-nominal relative clause modifiers should favor low attachment of relative clauses in ambiguous cases, as in English. However, Brysbaert (1993) reports that high attachment is preferred in just these circumstances in Dutch, which is like English in that adjectives precede the noun they modify. C&M's formulation also has difficulty explaining the preference for low attachment in constructions like (39), and it is not obvious under this account why low attachment preferences in English VP

constructions are much stronger than the corresponding preferences in English NP constructions.

Despite the failure of this particular account, it is nevertheless possible that some parameter or set of parameters in the grammar could be directly or indirectly responsible for determining the relative strengths of the proposed preference factors. A particular hypothesis will have to await the results of further investigation of relevant constructions, and of parametric variation in general, in a wider range of languages.

References

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research, 18*, 129–144.

Allport, D. A., & Funnell, E. (1981). Components of the mental lexicon. *Philosophical transactions of the Royal Society of London, B295*, 397–410.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191–238.

Anderson, J. R. (1980). *Cognitive psychology and its implications.* New York: W. H. Freeman.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Balota, D. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Erlbaum.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development.* Cambridge: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing.* Cambridge: Cambridge University Press.

Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognitive development of language.* New York: Wiley.

Boland, J. E., & Tanenhaus, M. K. (1991). The role of lexical representations in sentence processing. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 331–366). Amsterdam: North-Holland.

Bresnan, J. (Ed.). (1982). *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.

Brysbaert, M. (1993). *Report: Revised Dutch version of Cuetos & Mitchell's questionnaire.* Unpublished manuscript, University of Leuvan, Belgium.

Burgess, C. (1990). *Interaction of semantic, syntactic and visual factors in syntactic ambiguity resolution.* Unpublished doctoral dissertation, University of Rochester.

Burgess, C., Tanenhaus, M. K., & Seidenberg, M. S. (1989). Context and lexical access: Implications of nonword interference for lexical ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 620–632.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Syntax and semantics: Vol. 21. Thematic relations* (pp. 263–288). San Diego, CA: Academic Press.

Carreiras, M., & Clifton, C., Jr. (in press). Relative clause interpretation preferences in Spanish and English. *Language and Speech.*

Carreiras, M., & Gutierrez, M. (1992). *On-line parsing of ambiguous sentences: The role of verb information.* Paper presented at the 5th European Conference on Developmental Psychology, Seville, Spain.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar.* Waltham, MA: Ginn.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht, The Netherlands: Foris.

Chomsky, N. (1986a). *Barriers.* Cambridge, MA: MIT Press.

Chomsky, N. (1986b). *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Chomsky, N., & Lasnik, H. (1991). Principles and parameters theory. To appear in J. Jacobs, A. von Stechow, W. Sternefeld, & T. Vennemann (Eds.), *Syntax: An international handbook of contemporary research.* Berlin: Walter de Gruyter.

Church, K. W. (1980). *On memory limitations in natural language processing.* Unpublished master's thesis, MIT.

Clifton, C., Jr. (1988). *Restrictions on late closure: Appearance and reality.* Paper presented at the 6th Australian Language and Speech Conference, University of New South Wales.

Clifton, C., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes, 4*, SI77–104.

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. (1987). Functional architecture of the language-processing system. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language.* London: Erlbaum.

Connine, C. M., Ferreira, F., Jones, C., Clifton, C., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research, 13*, 307–319.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition, 2*, 130–138.

Crain, S. (1980). *Contextual constraints on sentence comprehension.* Unpublished doctoral dissertation, University of California, Irvine.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language processing: 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.

Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1993). *El efecto frecuemcia en el procesamiento de oraciones ambiguas.* Paper presented at the Symposium de Psicolinguistica, Tenerife, Spain.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*, 547–619.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Elman, J. L., & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.), *Speech and language* (Vol. 10). New York: Academic Press.

Ferreira, F., & Clifton, C., Jr. (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 case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–90). New York: Holt, Rinehart & Winston.

Fischler, I., & Bloom, P. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 18*, 1–20.

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

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. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms.* Amsterdam: North-Holland.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.

Foss, D. (1982). A discourse on semantic priming. *Cognitive Psychology, 14*, 590–607.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

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

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.

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. (1990). 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: Erlbaum.

Frazier, L., & Clifton, C., Jr. (in preparation). *Construal.* Manuscript, University of Massachusetts, Amherst.

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.

Garnsey, S. M., Lotocky, M., & McConkie, G. W. (1992). *Verb-usage knowledge in sentence comprehension.* Poster presented at the 33rd annual meeting of the Psychonomics Society, St. Louis, MO.

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry, 12*, 155–184.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown.* Unpublished doctoral dissertation, Carnegie Mellon University.

Gibson, E. (1992). On the adequacy of the competition model [Review of *The crosslinguistic study of sentence processing*]. *Language, 68*, 812–830.

Gibson, E., & Pearlmutter, N. J. (1993). *A corpus-based analysis of psycholinguistic constraints on PP attachment.* Manuscript submitted for publication.

Gilboy, E., Sopena, J. M., Clifton, C., Jr., & Frazier, L. (1993). *Argument structure and association preferences in Spanish and English complex NPs.* Unpublished manuscript, University of Barcelona, Spain.

Glucksberg., S., Kreuz, R. J., & Rho, S. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12*, 323–335.

Gorrell, P. (1987). *Studies in human syntactic processing: Ranked-parallel versus serial models.* Unpublished doctoral dissertation, University of Connecticut.

Gough, P., & Cosky, M. J. (1977). One second or reading again. In N. J. Castellan, Jr., D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 271–288). Hillsdale, NJ: Erlbaum.

Grosz, B., Pollack, M., & Sidner, C. (1989). Discourse. In M. Posner (Ed.), *Foundations of cognitive science.* Cambridge, MA: MIT Press.

Gruber, J. S. (1976). *Lexical structures in syntax and semantics.* Amsterdam: North-Holland.

Hale, K. L., & Keyser, S. J. (1986). *Some transitivity alternations in English* (Lexicon Project Working Paper No. 7). Cambridge, MA: MIT, Center for Cognitive Science.

Hale, K. L., & Keyser, S. J. (1987). *A view from the middle* (Lexicon Project Working Paper No. 10). Cambridge, MA: MIT, Center for Cognitive Science.

Hale, K. L., & Keyser, S. J. (1992). The syntactic character of thematic structure. In I. M. Roca (Ed.), *Thematic structure: Its role in grammar* (pp. 107–143). Berlin: Foris.

Hindle, D., & Rooth, M. (1993). Structural ambiguity and lexical relations. *Computational Linguistics, 19,* 103–120.

Hinton, G. E. (1986). Learning distributed representations of concepts. *Proceedings of the 8th Annual Conference of the Cognitive Society,* 1–12.

Hinton, G. E., & Shallice, T. (1991). Lesioning a connectionist network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Hobbs, J. R., & Bear, J. (1990). Two principles of parse preference. *Proceedings of the 13th International Conference on Computational Linguistics, 3,* 162–167.

Hockett, C. (1963). The problem of universals in language. In J. Greenberg (Ed.), *Universals of language.* Cambridge, MA: MIT Press.

Hogaboam, T. W., & Perfetti, C. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 14,* 265–274.

Holmes, V. M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language, 28,* 668–689.

Hummel, R., & Zucker, S. (1983). On the foundations of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (Vol. PAMI-5).

Hyams, N. (1986). *Language acquisition and the theory of parameters.* Dordrecht, The Netherlands: Reidel.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar.* Cambridge, MA: MIT Press.

Jackendoff, R. (1977). *X-bar syntax: A study of phrase structure.* Cambridge, MA: MIT Press.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures.* Cambridge, MA: MIT Press.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanlaytic cognition: Memory, perception and concept learning. In G. H. Bower, (Ed.), *Psychology of learning and motivation* (Vol. 18). New York: Academic Press.

Jaeggli, O., & Safir, K. J. (Eds.). (1989). *The null subject parameter.* Dordrecht, The Netherlands: Kluwer.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 3,* 228–238.

Kawamoto, A. (1988). Distributed representations of ambiguous words and their represenation in a connectionist network. In S. Small, G. W. Cotrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution* (pp. 195–228). San Mateo, CA: Morgan Kaufman.

Kawamoto, A. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32,* 474–516.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2,* 15–47.

Knight, G. P. (1984). A survey of some important techniques and issues in multiple regression. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research.* Hillsdale, NJ: Erlbaum.

Koopman, H., & Sportiche, D. (1988). *Subjects.* Manuscript, University of California at Los Angeles.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Kurtzman, H. K. (1985). *Studies in syntactic ambiguity resolution.* Unpublished doctoral dissertation, MIT (Distributed by Indiana University Linguistics Club).

Levin, B. (1989). *English verbal diathesis* (Lexicon Project Working Paper No. 32). Cambridge, MA: MIT, Center for Cognitive Science.

Levin, B., & Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry, 17,* 623–661.

Levin, B., & Rappaport Hovav, M. (1992). The lexical semantics of verbs of motion: The view from unaccusativity. In I. M. Roca (Ed.), *Thematic structure: Its role in grammar* (pp. 247–269). Berlin: Foris.

MacDonald, M. C. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes, 4*, 35–56.

MacDonald, M. C. (in press). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes.*

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24*, 56–98.

Mackworth, A. K. (1977). Consistency in networks of relations. *Artificial Intelligence, 8*, 99–118.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

MacWhinney, B., & Bates, E. (Eds.). (1989). *The cross-linguistic study of sentence processing.* Cambridge: Cambridge University Press.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science, 189*, 226–228.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375–407.

McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4*, SI287–335.

McRae, K., de Sa, V., & Seidenberg, M. S. (1993). *The role of correlated properties in accessing conceptual memory.* Manuscript submitted for publication.

Merrill, E. C., Sperber, R. D., & MacCauley, C. (1981). Differences in semantic encoding as a function of reading comprehension skill. *Memory & Cognition, 9*, 618–624.

Meyer, D. E., & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

Milsark, G. L. (1983). On length and structure in sentence parsing. *Cognition, 13*, 129–134.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum.

Mitchell, D. C., Corley, 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. (1991a). The origins of parsing strategies. *Conference proceedings: Current issues in natural language processing.* University of Texas at Austin.

Mitchell, D. C. & Cuetos, F. (1991b). *Restrictions on late closure: The computational underpinnings of parsing strategies in Spanish and English.* Unpublished manuscript, University of Exeter, UK.

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: Erlbaum.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2). Hillsdale, NJ: Erlbaum.

Morris, R.K. (in press). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Morton J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

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

Pearlmutter, N. J., Daugherty, K., MacDonald, M. C., & Seidenberg, M. S. (1993). *Constraint satisfaction in main verb / reduced relative ambiguities.* Poster presented at the 6th annual CUNY Sentence Processing Conference, Amherst, MA.

Pearlmutter, N. J., & MacDonald, M. C. (1992). Plausibility effects in syntactic ambiguity resolution. *Proceedings of the 14th Annual Meeting of the Cognitive Science Society,* 498–503.

Perfetti, C. (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: Erlbaum.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: MIT Press.

Plaut, D., & Shallice, T. (in press). Deep dyslexia: A case study in connectionist neuropsychology. *Cognitive Neuropsychology.*

Posner, M., & Snyder, C. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Hillsdale, NJ: Erlbaum.

Prince, A., & Smolensky, P. (in press). *Optimality theory.* Cambridge, MA: MIT Press.

Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. *Language, 64*, 539–576.

Rappaport, M., Laughren, M., & Levin, B. (1987). *Levels of lexical representation* (Lexicon Project Working Paper No. 20). Cambridge, MA: MIT, Center for Cognitive Science.

Rappaport, M., & Levin, B. (1988). What to do with θ-roles. In W. Wilkins (Ed.), *Syntax and semantics: Vol. 21. Thematic relations* (pp. 7–36). San Diego, CA: Academic Press.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*, 618–660.

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

Rayner, K., Garrod, S., & Perfetti, C. A. (1992). Discourse influences during parsing are delayed. *Cognition, 45*, 109–139.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-Hall.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornič (Ed.), *Attention and performance VI* (pp. 573–603). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Sedivy, J. C., & Spivey-Knowlton, M. J. (1993). *Presuppositional complexity and local bias in attachment preferences.* Paper presented at the 6th annual CUNY Sentence Processing Conference, Amherst, MA.

Seidenberg, M. S. (1990). Lexical access: Another theoretical soupstone? In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 33–71). Hillsdale, NJ: Erlbaum.

Seidenberg, M. S. (1993). Connectionist models and cognitive theory. *Psychological Science, 4*, 228–235.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.

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

Simpson, G. B. (1981). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316–340.

Simpson, G. B., & Krueger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language, 30*, 627–643.

Smith, C. (1988). *Structural and other factors in parsing.* Unpublished manuscript, University of Texas at Austin.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology, 47*, 276–309.

Stabler, E., Jr. (1993). *The logical approach to syntax: Foundations, specifications, and implementations of theories of government and binding.* Cambridge, MA: MIT Press.

Stanovich, K., & West, R. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition, 7*, 77–85.

Steedman, M. J., & Altmann, G. T. M. (1989). Ambiguity in context: A reply. *Language and Cognitive Processes, 4*, SI105–122.

Stevenson, S. (1990). A parallel constraint satisfaction and spreading activation model for resolving syntactic ambiguity. *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, 396–403.

Stevenson, S. (1993a). A competition-based explanation of syntactic attachment preferences and garden path phenomena. To appear in *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics.*

Stevenson, S. (1993b). Establishing long-distance dependencies in a hybrid network model of human parsing. *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, 982–987.

St. John, M., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, *46*, 217–257.

Stowe, L. (1989). Thematic structures and sentence comprehension. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 319–357). Dordrecht, The Netherlands: Kluwer.

Stowe, L. (1990). *Resolving ambiguity: Evidence for delay.* Paper presented at the 3rd annual CUNY Conference on Human Sentence Processing, City University of New York.

Stowell, T. (1981). *Origins of phrase structure.* Unpublished doctoral dissertation, MIT.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–669.

Swinney, D., & Hakes, D. T. (1976). Effects of prior context upon lexical access during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *15*, 681–689.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, *27*, 324–340.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research*, *49*, 161– 167.

Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1993). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. To appear in *Attention and performance XV.* Hillsdale, NJ: Erlbaum.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language*, *32*, 359–372.

Tanenhaus, M. K., & Carlson, G. (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., & 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. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

Tanenhaus, M. K., Leiman, J., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior, 18*, 427–440.

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.

Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes, 6*, 303–338.

Trueswell, J. C., & Tanenhaus, M. K. (1992). Consulting temporal context during sentence comprehension: Evidence from the monitoring of eye movements in reading. *Proceedings of the 14th Annual Conference of the Cognitive Science Society*, 492–497.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (in press). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language.*

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 528–553.

Whittemore, G., Ferrara, K., & Brunner, H. (1990). Empirical study of predictive powers of simple attachment schemes for post-modifier prepositional phrases. *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, 23–30.

Williams, E. (1981). Argument structure and morphology. *Linguistic Review, 1*, 81–114.

Williams, J. N. (1988). Constraints upon semantic activation during sentence comprehension. *Language and Cognitive Processes, 3*, 165– 206.

Zernik, U. (1991). *Lexical acquisition: Exploiting on-line resources to build a lexicon.* Hillsdale, NJ: Erlbaum.

Appendix A

Chapter 1 Stimuli

A.1 Experiment 1 Items

Items are given in the order (a) Material-first conditions, and (b) Location-first conditions, with ambiguous and unambiguous verbs separated by a slash. The three ratings in parentheses following each item are, respectively, the subject, object, and the unambiguous rating. The ratings are on a seven-point scale, with 1=good, 7=bad. Sentences used as word control items in Experiment 2 are marked with an asterisk.

*1. Jim's father pointed out with a frown that

(a) the dirty slush splashed/frozen on the winding driveway made sure that going anywhere would be a real chore. (4.20, 2.20, 2.75)

(b) the winding driveway splashed/frozen with the dirty slush made sure that going anywhere would be a real chore. (6.30, 3.40, 3.30)

2. Mr. Smith claimed that no one cared that

(a) the newest posters hung/drawn in the upstairs room were too small and not too colorful. (5.20, 2.10, 3.65)

(b) the upstairs room hung/drawn with the newest posters was too small and not too colorful. (6.00, 4.80, 4.45) (This item was excluded from analyses based on its comprehension question scores.)

*3. Although it ended up being of no importance,

(a) the red paint squirted/thrown onto the large canvas caught Kelly's eye immediately. (2.70, 2.30, 1.90)

(b) the warning sign squirted/written with the red paint caught Kelly's eye immediately. (6.10, 3.40, 2.35)

*4. The children could pick out the spots where

(a) the green peas scattered/sown in the tiny garden were doing very well. (2.10, 2.60, 2.75)

(b) the tiny garden scattered/sown with the green peas was doing very well. (6.00, 3.20, 2.10)

*5. Craig told Jeff about the production, noting that

(a) the white sawdust showered/blown onto the dim stage made the play look like it occurred in the snow. (4.70, 3.50, 2.20)

(b) the dim stage showered/seen with the white sawdust made the play look like it occurred in the snow. (6.10, 2.70, 3.45)

*6. Julie said she just wanted to mention that

(a) the bread crumbs spread/shaken onto the pork chops would make the buffet dinner even more delicious. (3.70, 2.20, 2.30)

(b) the pork chops spread/shaken with the bread crumbs would make the buffet dinner even more delicious. (6.00, 2.90, 3.00)

*7. The inexperienced salesman seemed very annoyed just because

(a) the dead insects splattered/strewn across the grimy windshield made the car appear older than it was. (4.50, 2.00, 1.80)

(b) the grimy windshield splattered/strewn with the dead insects made the car appear older than it was. (5.40, 2.40, 3.00)

8. After Susan worked all morning on the breakfast,

(a) the cream cheese smeared/eaten on the garlic bagels tasted really wonderful to her. (3.00, 2.30, 1.75)

(b) the garlic bagels smeared/eaten with the cream cheese tasted really wonderful to her. (5.70, 2.70, 1.90)

9. The nurse was upset after each visit because

(a) the cotton bandage draped/thrown over the plaster cast got smellier every day. (3.80, 2.50, 3.15)

(b) the plaster cast draped/worn with the cotton bandage got smellier every day. (5.10, 3.00, 3.45)

10. The farmer became excited when he realized that

(a) the various seeds scattered/sown in the new fields were beginning to show some growth. (2.50, 2.30, 1.95)

(b) the new fields scattered/sown with the various seeds were beginning to show some growth. (6.30, 3.30, 1.60)

11. The famous architect noted with apparent interest that

(a) the ancient ivy wrapped/grown around the stone balcony provided the house with an air of grandeur. (4.50, 2.50, 2.75)

(b) the stone balcony wrapped/interwoven with the ancient ivy provided the house with an air of grandeur. (6.80, 3.40, 3.90)

12. Before the caterers could close up the kitchen,

(a) the overcooked macaroni splattered/strewn across the tile floor had to be cleaned by somebody. (3.40, 2.80, 2.20)

(b) the tile floor splattered/strewn with the overcooked macaroni had to be cleaned by somebody. (6.00, 1.80, 2.00)

13. The mounted policeman commented to his partner that

(a) the bright confetti showered/blown into the crowded streets caused the city to seem far more cheery than usual. (3.70, 3.30, 2.30)

(b) the crowded streets showered/seen with the bright confetti caused the city to seem far more cheery than usual. (6.50, 2.80, 3.35)

*14. It was quite clear from the sketches that

(a) the tattered hood draped/thrown over the iron helmet assured the knight a fair degree of anonymity. (4.50, 2.30, 2.55)

(b) the iron helmet draped/worn with the tattered hood assured the knight a fair degree of anonymity. (6.20, 3.40, 2.35)

15. Even when the weather was not too nasty,

(a) the dark mud splashed/frozen on the cement sidewalk was a nuisance to everyone. (2.60, 2.20, 2.40)

(b) the cement sidewalk splashed/frozen with the dark mud was a nuisance to everyone. (6.40, 3.10, 3.95)

16. It was clear to the executive secretary that

(a) the black ink squirted/thrown onto the new contract created some real problems for the clients. (1.50, 1.90, 2.70)

(b) the new contract squirted/written with the black ink created some real problems for the clients. (6.50, 5.40, 1.95)

17. Just as he was finishing, Eric realized that

    (a) the gooey mixture spread/shaken around the glass container was placed in the oven in the last step of the recipe. (2.70, 2.20, 4.50)

    (b) the glass container spread/shaken with the gooey mixture was placed in the oven in the last step of the recipe. (5.90, 5.30, 4.35)

*18. The real estate agent forgot to mention that

    (a) the old curtains hung/drawn across the narrow windows needed to be replaced as soon as possible. (2.70, 2.00, 1.60)

    (b) the narrow windows hung/shown with the old curtains needed to be replaced as soon as possible. (5.20, 4.30, 4.20)

*19. Marge was commenting to Pete and Jill that

    (a) the melted butter smeared/eaten on the hot roll looked really delicious but wasn't healthy. (3.60, 2.00, 4.50)

    (b) the hot roll smeared/eaten with the melted butter looked really delicious but wasn't healthy. (5.00, 2.50, 2.50)

*20. The repairman was rather surprised to discover that

    (a) the tough vines wrapped/grown around the abandoned gate could not be moved by anyone. (5.40, 2.00, 2.70)

    (b) the abandoned gate wrapped/interwoven with the tough vines could not be moved by anyone. (5.80, 3.00, 4.65)

A.2 Experiment 2 Items

Nouns separated by slashes are given in the order Agent/Theme in the sentences. The words in parentheses were inserted to form the unambiguous versions. The three ratings in parentheses following each item are, respectively, the subject rating, the object rating, and the comparative rating for the Agent NP, indicated by "A:", and the same for the Theme NP, indicated by "T:". The ratings were made on a seven-point scale. For the subject and object ratings, lower values correspond to better ratings. For the comparative rating, lower values correspond to a preference for the NP as subject, while higher values correspond to a preference for the NP as object.

1. The evening news on all the network stations reported that the assassin / politician (that was) shot by the mugger died in the ambulance outside the emergency room. (A: 1.50, 3.18, 2.59; T: 3.18, 1.31, 5.27)

2. Though the food in the hotel was not particularly good, the waitress / customer (that was) served by the bellhop was very pleased with the prompt room service. (A: 1.07, 3.71, 2.11; T: 5.00, 1.64, 5.86)

3. Although many people had already heard most of the story, the reporter / performer (that was) interviewed by the author wished to keep her part in events secret. (A: 1.05, 3.84, 2.25; T: 4.00, 1.73, 5.18)

4. The producer said that the live broadcast went smoothly, and the audience / musician (that was) applauded by the host enjoyed the show immensely from start to finish. (A: 1.07, 5.26, 1.67; T: 3.64, 2.09, 5.18)

5. The collision was very traumatic and violent, and not surprisingly the doctor / patient (that was) treated by the intern remembered nothing about the circumstances causing her injuries. (A: 1.27, 4.45, 2.09; T: 3.83, 1.39, 6.13)

6. Peggy noted that just before the first vandalism incidents began, the manager / secretary (that was) fired by the president stormed angrily out of the building during lunch. (A: 1.11, 2.07, 3.15; T: 4.86, 1.50, 5.91)

7. A small story in the Wall Street Journal reported that the accountant / taxpayer (that was) audited by the IRS owed almost two thousand dollars in back taxes. (A: 1.36, 2.44, 3.20; T: 5.91, 1.36, 6.27)

8. In the adventure that the librarian read to the children, the pirates / peasants (that were) terrorized by the dragon hired a wizard to come and kill it. (A: 2.13, 4.65, 2.13; T: 3.36, 1.32, 5.11)

9. Carl described how, in a legend on one Mesopotamian tablet, the army / victim (that was) enslaved by the emperor was forced to help rebuild an enchanted city. (A: 1.30, 3.52, 3.00; T: 6.09, 1.45, 6.64)

10. The secret agent reported right away to his superiors that the terrorists / hostages (that were) kidnapped by the mercenaries escaped just after the hidden compound was bombed. (A: 1.41, 5.26, 2.07; T: 5.27, 2.27, 5.45)

11. The attorneys had suggested the punishment would be mild, and the judge / murderer (that was) sentenced by the tribunal was obviously upset about the twenty-year prison term. (A: 1.04, 4.46, 1.82; T: 5.41, 1.19, 6.33)

12. The documentary television program made very clear the fact that the professor / jury (that was) instructed by the bailiff was already quite familiar with the trial procedures. (A: 1.78, 4.26, 2.65; T: 3.82, 1.33, 5.69)

13. The court was packed with photographers and curious observers, and the lawyer / criminal (that was) questioned by the press refused to make any comments about the hearing. (A: 1.11, 2.61, 2.86; T: 4.64, 1.27, 6.09)

14. Jacob told Catherine that he was surprised to learn that the artist / model (that was) painted by the students maintained the same pose for over two hours. (A: 1.36, 3.91, 2.09; T: 4.29, 1.71, 5.87)

15. The crowd that gathered outside the building could see that the orderly / invalid (that was) carried by the fireman was coughing violently from the smoke he inhaled. (A: 1.18, 5.27, 1.55; T: 5.56, 1.41, 6.48)

16. Marion told Rebecca how everyone had laughed out loud when the psychoanalyst / volunteer (that was) hypnotized by the magician began to crow like a rooster at dawn. (A: 1.45, 4.64, 1.95; T: 4.39, 1.07, 5.79)

17. Major Brown quietly told the shocked debriefing officers about how the hunter / prisoner (that was) captured by the rebels begged for freedom for weeks with no success. (A: 1.82, 3.41, 2.45; T: 4.02, 1.38, 5.57)

18. The National Geographic Society Special showed one clip in which the lion / deer (that was) stalked by the poachers sprinted away through some tall grass and escaped. (A: 1.04, 3.07, 2.93; T: 5.27, 1.27, 6.55)

19. The videotape made during one twenty-four hour period showed how the detective / suspect (that was) investigated by the FBI proved to be entirely innocent of the crime. (A: 1.07, 3.93, 2.56; T: 4.82, 1.36, 6.27)

20. The store owner who witnessed the accident could not believe that the driver / pedestrian (that was) struck by the ambulance needed only twenty stitches and was otherwise unharmed. (A: 1.45, 1.95, 3.84; T: 3.73, 1.64, 5.41)

21. Jake said that nobody was really shocked to hear that the patrolman / suspect (that was) arrested by the detectives spoke with an attorney before answering any questions. (A: 1.00, 4.09, 1.55; T: 5.25, 1.57, 6.21)

22. Everyone in the room thought it was extremely obvious that the warden / prisoner (that was) interrogated by the sergeant was very nervous and had something to hide. (A: 1.07, 3.67, 2.30; T: 5.20, 1.40, 6.40)

23. During the initial investigation into the theft of the documents, the prosecutor / criminal (that was) accused by the deputy denied having any personal knowledge of illegal activities. (A: 1.78, 3.78, 2.13; T: 2.96, 1.18, 5.11)

24. The letter received by the collection office clearly stated that the corporation / customer (that was) billed by the caterer refused to pay because the food was terrible. (A: 1.77, 2.05, 3.64; T: 5.15, 1.74, 6.33)

Appendix B

Chapter 3 Stimuli and Raw Reading Times

B.1 Experimental Items for Experiment 1

The items used in Experiment 1 are listed below. The three conditions were created by making one of the first three nouns singular and the other two plural. The three items marked with an asterisk were excluded from all Experiment 1 analyses and from Experiment 2.

These items were also used in Experiment 2, unless there is a correspondingly-numbered item in Appendix B.2, in which case the item shown in Appendix B.2 replaced the one shown here. For Experiment 2, the number of the verb in the relative clause (*was*) was changed from singular to plural for the even-numbered items, and thus the three conditions for the even-numbered items were created by making one of the first three nouns plural and the other two singular. The number in parentheses following some of the items is the percentage of subjects from the Experiment 2 written survey who indicated that the second PP attached to the second NP for that item.

1. the tree beside the cat with the kitten that was in the nearby park
2. the building near the dog with the carcass that was seen by many people

3. the brick by the rooster with the hen that was found in the yard
4. the sign above the memo to the committee that was ignored by the workers (93)
5. the photo beside the letter to the company that was consulted in the lawsuit (100)
6. the plaque with the gift to the museum that was very large and impressive (93)
7. the review of the book about the trial that was very long and complex (100)
8. the newspaper with the article about the movie that was criticized on the TV (100)
9. the videotape of the play about the murders that was discussed for many weeks (100)
10. the lamp near the painting of the house that was damaged in the flood (100)
11. the chair below the drawing of the bed that was purchased at the auction (100)
12. the boat near the model of the building that was finished just last week
13. the design on the kite above the house that was small but very beautiful (100)
14. the sofa beside the rug under the table that was dropped by the movers (100)
15. the road-sign near the puddle below the car that was by the street corner (100)
16. the chair by the bowl for the dog that was next to the door (100)
17. the dress beside the toy for the puppy that was left in the bedroom
18. the pond near the trail for the horse that was visible in the photograph

*19. the bike near the car with the new airbag that was damaged by the vandals

*20. the TV by the computer with the new hard-drive that was taken from the room

*21. the runway near the plane with the new engine that was ruined in the storm

B.2 Experimental Items for Experiment 2

The items in Experiment 2 which were changed from Experiment 1 are shown below. Even-numbered items appeared with a plural verb in the relative clause. The conditions were created as described in Appendix B.1. The number in parentheses after

each item is the percentage of subjects from the Experiment 2 written survey who indicated that the second PP attached to the second NP for that item.

1. the cushion beside the cat with the kitten that was in the nearby park (93)
2. the frisbee near the dog with the carcass that was found in the park (100)
3. the brick by the hen with the rooster that was kicked by the farmer (100)
12. the computer near the model of the building that was destroyed in the fire (100)
17. the hat beside the toy for the puppy that was received as a gift (100)
18. the pond near the barn for the horse that was visible in the photograph (86)

B.3 Experimental Items for Experiment 3

The items in Experiment 3 are listed below with their English translations. Because singular versus plural forms differed by more than just number-marking on the nouns, all three conditions (in the order high, middle, low attachment) are shown for the first two items. For the remainder, only the low attachment version is shown, with its English translation. Items marked with an asterisk were excluded from the supplemental analyses conducted in Experiment 3. (See the Materials and Results sections of the Experiment 3 description.) The number in parentheses following each item is the percentage of subjects from the Experiment 3 written survey who indicated that the second PP attached to the second NP for that item.

1. (a) los árboles junto al gato con el gatito que fueron admirados por el niño
   (b) el árbol junto a los gatos con el gatito que fueron admirados por el niño
   (c) el árbol junto al gato con los gatitos que fueron admirados por el niño
   (d) the tree next to the cat with the kittens that were admired by the child (95)
2. (a) el edificio cerca de los perros con los fósiles que fue claramente elogiado
   (b) los edificios cerca del perro con los fósiles que fue claramente elogiado
   (c) los edificios cerca de los perros con el fósil que fue claramente elogiado

(d) the buildings near the dogs with the fossil that was clearly praised (85)

3. (a) el tabique cerca del pollo con los gallos que fueron pateados por el granjero

   (b) the brick near the chicken with the roosters that were kicked by the farmer (100)

4. (a) los avisos arriba de los memos para el jefe que fue ignorado por los empleados

   (b) the notices above the memos for the boss who was ignored by the employees (90)

5. (a) el anuncio del legado a los museos que impresionaron mucho al público

   (b) the announcement of the legacy to the museums that impressed the public very much (90)

6. (a) las revistas junto a las cartas para la acusada que fue consultada en el juicio

   (b) the magazines near the letters to the defendant that was consulted in the trial (100)

*7. (a) la nota del libro acerca de las teorías que resultaron dificiles de comprender

   (b) the note about the book about the theories that resulted difficult to understand (65)

*8. (a) las revistas con las fotos de la película que fue criticada en T.V.

   (b) the magazines with the pictures of the movie that was criticized on TV (75)

9. (a) el comentario del libro acerca de los asesinatos que causaron furor

   (b) the commentary of the book about the murders that caused a furor (85)

10. (a) las lámparas cerca de las pinturas de la casa que fue dañada en la inundación

   (b) the lamps near the paintings of the house that was damaged in the flood (80)

11. (a) el escudo cerca del esquema de los edificios que fueron terminados el lunes

   (b) the coat of arms near the plan of the buildings that were finished on Monday (85)

*12. (a) los muebles debajo de los modelos del aparato que fue comprado en la subasta

   (b) the furniture under the models of the appliance that was bought at the auction (20)

13. (a) el listón en el papalote arriba de los árboles que ardieron por accidente

    (b) the ribbon in the kite above the trees that were burned by accident (95)

*14. (a) los libros detrás de los tapetes debajo de la mesa que desapareció en el viaje

    (b) the books behind the rugs under the table that disappeared on the trip (75)

15. (a) la señal cerca del charco debajo de los carros que causaron problemas

    (b) the signal near the puddle under the cars that caused problems (85)

16. (a) los gabinetes detrás de los platones para el perro que fue comprado en México

    (b) the cabinets behind the plates for the dog that was bought in Mexico (100)

17. (a) el paquete detrás del juguete para los gatitos que fueron olvidados el martes

    (b) the package behind the toy for the kittens that were forgotten on Tuesday (85)

18. (a) los lagos cerca de los senderos para el caballo que fue filmado en el documental

    (b) the lakes near the paths for the horse that was filmed in the documentary (80)

B.4 Raw Reading Times

Table B.1: Experiment 2 Raw Reading Times (msec/word)

	Region						
Attachment	1	2	3	4	5	6	7
High	399	454	610	618	870	451	783
Middle	498	446	534	695	1037	515	775
Low	406	466	543	844	785	453	757

Table B.2: Experiment 3 Raw Reading Times (msec/word)

	Region						
Attachment	1	2	3	4	5	6	7
High	713	897	960	891	1403	635	1282
Middle	742	847	1216	872	1935	570	1916
Low	793	829	984	818	1098	609	1704