

Chapter in M. Carreiras, J. Garcia-Albea & N. Sabastian-Galles (Eds.) Language Processing in Spanish, published in 1996 by Erlbaum.

PARSING IN DIFFERENT LANGUAGES.

Fernando Cuetos (*) Don C. Mitchell (**) Martin M.B. Corley (**)

(*) Universidad de Oviedo

(**) University of Exeter

Cuetos, Mitchell & Corley

Parsing in Different Languages

—

Contents

1. Introduction
2. Types of parsing theory
 - 2.1. Universal Parsers
 - 2.1.1. Traditional garden-path theories
 - 2.1.2. Refined Garden-Path Theories
 - 2.1.3 Construal Theory
 - 2.2. Parametric theories of parsing
 - 2.2.1. Parametric competition with special-purpose strategies
 - 2.2.1.1. Modifier-straddling strategy
 - 2.2.1.2 The Recency/Predicate Proximity model
 - 2.3 Tuning theory
 - 2.4 Totally data-shaped parsers
 - 2.5 Summary of models of modifier attachment
3. Evaluation of the theories in the light of empirical evidence
 - 3.1. Basic observations
 - 3.2 Universal accounts
 - 3.2.1. Evaluation of Garden-Path theory
 - 3.2.2. Refined Garden-Path theory
 - 3.2.3 Construal theory
 - 3.3 Evaluation of parameterized accounts
 - 3.3.1 Strategy competition models
 - 3.3.2 Tuning hypothesis
 - 3.3.3. Data-shaped parsers
 - 3.3.4. Overview of evaluation
4. Conclusions

1.- INTRODUCTION

A considerable body of modern science is concerned with the origins of structures and phenomena. Cosmologists engage in intense debate on the origins of the universe, biochemists speculate about the origins of life, and paleontologists and physical anthropologists discuss the origins of hominids and homo sapiens. Similar debates have played a central part in discussions of numerous phenomena in the study of language (e.g., Bickerton, 1990). Researchers have investigated change, evolution and innovation in language and have asked whether particular properties or features are innate - perhaps determined by the structure of the human mind - or whether they are acquired as a consequence the language user's individual experiences. While it is probably uncontentious that certain aspects of language are acquired (e.g. vocabulary, accents, specific grammars or dialects), there is a strongly held view among linguists in particular that certain structural principles of language are universal and innate (Chomsky, 1986; Radford, 1990; Pinker, 1994; Atkinson, 1992; Cook, 1988). If this is true, it raises the possibility that some of the procedures that people use to process language may also be universal and innate. This question may be addressed at many

different levels of language processing, but because of the prominent role of syntax in modern linguistic debate one of the central battlegrounds in the nature/nurture war-zone has been in the field of syntactic analysis or parsing. Inevitably much of the debate revolves around whether a fixed set of parsing strategies is used to process different languages of the world. The present chapter addresses this question - drawing heavily on empirical studies of sentence processing in Spanish and other romance languages.

The function of the parser is to compute the syntactic structure of sentences, allowing the reader or listener to determine "who did what and to whom" and more generally to infer appropriate relationships between statements and entities expressed by the sentence. Among other things its raw material consists of the order of occurrence of the individual words in the sentence, together with detailed information about these words (e.g., major category class, subcategorization information, inflections conveying information about number, gender and case and so on). Other inputs include intonation, prosody and punctuation.

The parsing process itself is typically completed very rapidly and without conscious awareness. For the most part its workings are only laid bare when it constructs an erroneous analysis, holding up the operation of the rest of the sentence processing system. In normal circumstances this occurs most frequently with ambiguous sentences, with the consequence that most of the empirical work on parsing focuses on trying to specify how the mechanism handles materials of this kind. For example a typical test sentence might be one like (1).

- (1) Vi a Juan cuando iba a la playa
(I saw John going to the beach)

Here there is ambiguity about whether it the "Juan" or the speaker who is "going to the beach" and information about the preferred interpretations can provide a window onto the internal machinery of the parser. For comprehensive reviews of the basic empirical findings and their associated theoretical conclusions the reader is referred to general surveys of the parsing literature (e.g. Mitchell, 1987b (translated into Spanish in Valle, Cuetos, Igoa and del Viso, 1990), Mitchell 1994; Frazier 1987; Frazier and Clifton, (Chap 1, in prep).)

To address the question of whether these processes operate in essentially the same way from language to language it is essential to spell out the major theoretical proposals in some detail.

2. TYPES OF PARSING THEORY.

Parsing theories introduced over the last 20 years can be classified into three broad categories based on their treatment of cross-linguistic variation (see Mitchell, Cuetos and Zagar, 1990, for a review). First, there are theories which are essentially universal, in the sense that they assume that the computational machinery and strategies for parsing are effectively identical in all languages (e.g. Bever, 1970; Kimball, 1973; Frazier, 1987; Pritchett, 1992; Crocker, 1992; Inoue and Fodor, in press). Of course, all such accounts have to make provision for the fact that, at the very least, detailed grammatical rules and items of vocabulary change from language to language. This is typically handled by assuming that language-specific information is made available to the parser in a manner analogous to the use of simple datafiles by a computer program - in the way, for example, a spell-checking routine might use files consisting of correctly spelt words. This view is apparently the one adopted by Frazier (1987, p.565) when she wrote.."we should be able to remove the grammar of English from our theory of sentence processing, plug in the grammar of some other language, and obtain the correct theory of processing of that language." The second class of models assumes that the "customization" of parsers goes some way beyond file exchanges which simply replace one grammar with another. In these models the "datafile" provided by each language includes not only the grammar, but also information about the valence or priority assigned to certain operations in the parser's portfolio of strategies. The eventual choice of strategy can be seen as a consequence of parameter-setting during grammar-acquisition, but the assumption is that it is the parsing strategy itself which changes, and not just the rules of the grammar which provide its raw material. These models range from those in which certain syntactic operations take precedence in individual languages (e.g., Gibson, Pearlmutter, Canseco-Gonzales and Hickok, submitted), through intermediate proposals (e.g., Frazier and Rayner, 1988) to those in which the whole parsing strategy is changed radically (e.g., Mazuka and Lust, 1990). The third, and perhaps most radical class of theories assumes that the "datafile" is subject to much more extreme forms of customization. In fact, they assume that the device starts off with a tabula rasa and goes on to use general learning algorithms during exposure to language to construct analysis procedures which can then be used to handle input (e.g. Bates and MacWhinney, 1987; St. John and McClelland, 1990). With theories of this kind the parsing decisions are wholly determined by a specific history of training and may be completely specific to the device that has received this input.

In present chapter we will consider the potential of these three different types of theory in relation to one fairly specific aspect of syntactic analysis - namely resolving the ambiguities that typically occur when modifying constructions, such as relative clauses, have to be attached to a word or phrase within a complex noun phrase in the manner illustrated in Sentence (2).

(2) Someone shot the servant of the actress who was on the balcony.

Sentences of this kind will be referred to as NP-PP-RC sentences (since they included constituent sequences of the form: noun phrase - prepositional phrase - relative clause).

The basic problem is that when a modifier (like "who was on the balcony") is preceded by a complex noun phrase in this way, it can typically be attached to more than one of the nouns within the NP-PP complex. Thus, Sentence (2) could be interpreted as implying that either the "servant" or the "actress" was the person who was "on the balcony". More technically, the modifier can be attached to more than one "head" or "host site" within the preceding constituent. To disambiguate the sentence it is necessary to determine whether the modifier proposition applies to the first potential head or the second. Resolutions of the first kind will be referred to as NP1-attachment (or N1-attachment), early-attachment or high-attachment (on the grounds that the first site is located at a higher point in the phrase marker for such structures). Correspondingly, links of the second kind will be termed NP2-, N2-, late- or low-attachments.

Since one of the important functions of a parser is to resolve ambiguities of this kind, it is possible to examine fundamental properties of syntactic processing by determining whether such ambiguities are handled in a uniform way or whether the processing strategies vary from language to language. The reason for focusing on this particular structure is that the head-attachment ambiguity, and the associated parsing problems, occur in numerous different languages across a range of typological classes. It therefore provides an ideal test-bed for evaluating the theories outlined above and investigating the degree to which parsing machinery varies from language to language. In the next section we consider a number of specific theoretical proposals concerning the resolution of ambiguities of this kind.

2.1 Universal parsers

First, we consider proposals premised on the assumption that parsing procedures are essentially the same for all languages.

2.1.1. Traditional Garden-path theories.

The most widely discussed universal account of parsing is one referred to as the Garden-path theory (Frazier, 1978, 1987; Frazier and Rayner, 1982). This theory is concerned most fundamentally with the question of how the parser operates when it encounters structural ambiguity in the sentences it is required to analyse. The central claim is that, faced with alternative readings, the parsing device very rapidly selects a single interpretation and uses this as the basis for its continuing analysis of the sentence. Where this choice proves to be mistaken, a selection of reanalysis procedures are used to recompute the structural relations within the region of ambiguity (Frazier and Rayner, 1992; Ferreira and Henderson, 1991). A crucial feature of the model is that in the initial analysis, computational effort is devoted exclusively to the interpretation selected by the conflict-resolution device.

Within this framework, a major focus of attention has therefore concerned the way in which this initial analysis is selected. Frazier (1978, 1987) proposes that the decision is based on a small number of structural principles all of which serve to increase the speed and efficiency with which new material can be incorporated into the ongoing analysis. Of these principles, the most important are Minimal Attachment and Late Closure, which were defined as follows:

(3) Minimal Attachment: "Do not postulate any potentially unnecessary nodes" (Frazier, 1987:562).

(4) Late Closure: "If grammatically permissible attach new items into the clause or phrase currently being processed (i.e. the phrase or clause postulated most recently)." (Frazier, 1987:562).

The principle of Minimal Attachment ensures that the parser selects for its first analysis a structure that is simple and quick to build, avoiding the additional work involved in assembling more complex structures. The Late Closure principle ensures that new constituents are immediately integrated with prior material, minimizing the chances of exceeding the memory limits of the sentence processing mechanism. Since these features would presumably produce benefits for all types of readers and listeners, Frazier (1987) argued that they should apply universally across all the languages of the world. Comparable claims for universal parsing strategies have been made by several other

investigators (e.g. Bever, 1970; Kimball, 1972; see Mitchell, Cuetos and Zagar, 1990, and Mitchell and Cuetos, 1991a, for reviews).

A comprehensive account of parsing has to say something about how the analysis proceeds once the initial structural choice has been made. In particular there is pervasive evidence that the syntactic analysis of sentences can be influenced by non-syntactic factors such as semantic and discourse information and other contextual sources of influence. According to garden-path theory these influences are handled in part by a device which is completely distinct from the mechanism responsible for making the initial structural choice. This device - originally called the "thematic processor" (Rayner, Carlson and Frazier, 1983) - introduces information about the thematic relations that the words of the sentence are capable of entering and, in later extensions of the model, was also assigned wider functions in dealing with discourse and other context information (e.g. Clifton and Ferreira, 1989; Frazier, 1990b). Although there are no clear statements about the generality of these operations, in the absence of discussion of the way they may vary across languages, it seems safe to assume that, like the parsing strategies considered above, they are intended to be viewed as universal procedures that apply without exception to the different languages of the world.

In relation to our specific concern in this chapter with modifier attachment, the Garden-path model predicts that in sentences like (2) the relative clause should initially be attached to the "phrase currently being processed" - that is, it predicts low- or N2-attachment to "actress" rather than high-attachment to "servant". In Section 3 we discuss the evidence concerning this prediction.

2.1.2 Refined garden-path theory.

Following a series of studies to be discussed at length in Section 3, Frazier (1990a) and De Vincenzi and Job (1993) have recently elaborated considerably on the role of the second phase of processing within the framework of the garden-path theory. Following the analysis in earlier versions of the model, Frazier (1990a) argued that once the initial structurally-determined choice has been made, the second mechanism kicks in very quickly and sometimes triggers a complete revision of the structural analysis. The refinement consisted largely of a more systematic statement of the discourse-related workings of the second mechanism. In particular, Frazier (1990a) proposed that the device operates in accordance with what she referred to as the "Relativized Relevance" principle outlined in (5).

(5) Other things being equal (e.g., all interpretations are grammatical, informative, and appropriate to discourse), preferentially construe a phrase as being relevant to the main assertion of the current sentence (Frazier, 1990a:321).

In certain cases this can result in a reading which is quite different from the one proposed by the first (structurally-guided) mechanism and the revised analysis is the one that will be most in evidence if the reader's commitments are checked at any point after the second-phase discourse mechanism has come into play.

In further elaborations of this account De Vincenzi and Job (1993) suggested that there may be certain constraints on the circumstances in which the discourse mechanism can impose its revisions. In particular, they suggested that a phrase or clause can only be re-construed as being "relevant" to a previously expressed assertion if this earlier assertion is in the same processing region as the constituent to which the clause is initially attached. The processing region they identified as being critical was one which they termed the "theta domain" - roughly speaking, a set of words linked either to the last verb or to another word that is capable of assigning semantic roles to noun phrases (see De Vincenzi and Job, 1993, for details).

As with the original garden-path theory, the features and operations of this refined version of the theory are also considered to be universal properties of human language systems.

What predictions does this refined version of the garden-path model make in connection with the process of resolving ambiguities of modifier attachment? First, it implies that the relative clause in sentences like (2) should initially be attached low to "actress". At a later point in processing, however, the model predicts that this link should be broken and the modifier should be re-attached high to "servant". This revision should occur because the first phrase ("the servant") is more likely to be closely connected with the "main assertion" of the sentence. The discourse/interpretation mechanism should therefore follow the Relativized Relevance principle and shift the attachment from "actress" to "servant". The end result of this sequence of operations is that high-attachment should prevail after the discourse effects have exerted their influence (cf. Clifton, 1988; Frazier, 1990a; De Vincenzi and Job, 1993).

2.1.3 Construal theory

In a much more radical refinement of garden-path theory Frazier, Clifton and their colleagues (Carreiras and Clifton, 1993; Gilboy, Sopena, Clifton and Frazier, in press; Frazier and Clifton, in preparation) have proposed that the original parsing strategies (such as Late Closure and Minimal Attachment) only apply to a certain subset of the structural ambiguities that occur in any given language - namely those concerned with clauses which are potentially or obligatorily dependent on the main predicate of the sentence (linkages which are termed "primary" relations). For other relations the theory holds that, rather than being linked in at a specific point in the developing phrase marker, the ambiguous material may merely be "construed" as being associated with the constituents within an entire theta-domain. Thus, Construal Theory postulates that with non-primary relations - such as relative clause attachment - there is no initial commitment to one structural analysis rather than another. The standard parsing strategies play no role at all and all structural biases or preferences are introduced by thematic or discourse processing operations outlined in the Garden-Path precursors to the theory.

This account, then, is explicitly characterized by the proposal that the standard parsing strategies do not apply universally. Instead, the suggestion is that they come into play exclusively for the purpose of analyzing a particular subset of ambiguous structures (namely "primary" relations). However, this restriction to specific structures does not extend to its application to different languages. The primary strategies and construal mechanisms are assumed to apply universally across languages.

The prediction for processing sentences like (2) is that the relative clause should initially be "construed" or attached to a region incorporating both "servant" and "actress" - with no bias or commitment between the two. A little later Gricean and discourse principles would favour low-attachment (to "actress") in English and high-attachment (to "servant") in most other languages (cf. Frazier and Clifton, in preparation; Gilboy et al, in press). The reason for predicting attachment to "actress" in English is that if the speaker/writer had intended to modify "actress" she could have avoided using the postnominal Norman form, selecting instead to employ the unambiguous Saxon genitive "the actress's servant who was on the balcony". Assuming the speaker observes Gricean Maxims such as "Be clear", the use of the alternative form might be taken as being deliberate, serving to reduce the tendency to attach the relative clause to "servant" and shifting the emphasis in favour of low-attachment (to "actress"). In the absence of the alternative form, other languages would reflect a more widespread bias in favour of high attachment (following principles like Relativized Relevance).

2.2. Parameterized theories of parsing.

These are theories based, in part, on the fact that there are important qualitative differences between languages - as discussed by Chomsky's (1981, 1986) Principles and Parameters theory - and in much work on language acquisition (e.g. Radford, 1990; Atkinson, 1992). For example, languages can be classified as being either "head-first" or "head-last", "pro-drop" or "non-pro-drop" and so on. The parsing theories considered in this section take into account one or more parameter settings for the language in question, and incorporate this parameterized information into the ambiguity-resolving process (see Frazier and Rayner, 1988 and Mazuka and Lust, 1990, for related proposals designed to handle ambiguities other than those highlighted here).

2.2.1 Parameterized competition with Special Purpose strategies

2.2.1.1. Modifier-straddling strategy

A simple, parameter-based proposal was put forward by Cuetos and Mitchell (1988) (see also, Mitchell & Cuetos, 1991a; Mitchell, Cuetos and Zagar, 1990). The suggestion was that, in certain very specific circumstances, general parsing strategies might not operate alone in determining the initial structural analysis to be pursued. Instead, from time to time they may have to compete with local special-purpose strategies which impose some pressure to select a competing analysis. The effectiveness of these local strategies may vary from language to language, and in particular, Cuetos and Mitchell (1988) suggested that their role might be determined in part by parameter-based linguistic properties of the languages under analysis.

The particular illustration used by Cuetos and Mitchell (1988) concerned strategies for handling the attachment ambiguities that are the subject of this chapter. With materials of the kind already discussed they suggested that a special-purpose strategy may be operative in some languages. The strategy involved "jumping over the modifier" to attach relative clauses to the first NP in structures of the form "NP-modifier-RC". A rationale was given for suggesting that this "modifier straddling" strategy should become an influential competing force in post-modifying languages (where adjectives follow the noun) but not in pre-modifying languages. The suggestion, then, was that influence of the modifier-straddling strategy might be "switched on" in post-modifying languages, perhaps as a result of parameter-setting processes in the course of normal language acquisition.

In a system of this kind, the effects of a general parsing strategy could easily be swamped and reversed by the local effects of the special-purpose strategy. Given the premise that the local effects are parameter-linked, a system of this kind would not be completely universal in the sense that it would not operate in precisely the same way when implemented in different languages. However, such a routine could be regarded as universal modulo a limited amount of parametric variation.

Models of this kind predict that high attachment preference should prevail in languages where the special-purpose strategy is "switched on". Cuetos and Mitchell (1988) suggested that this might occur in languages where adjectives generally follow the nouns they modify (e.g. Spanish, Italian, French etc.). On this account the first noun in a complex noun phrase should be selected as the attachment site in languages in this class, while the pattern should be reversed in English, Dutch, German and other pre-modifying languages.

2.2.1.2. The Recency/Predicate Proximity model.

Recently Gibson, Pearlmutter, Canseco-Gonzales and Hickok (submitted) have proposed a somewhat more elaborate model based on parameterized competition between strategies. According to this account, the parser may weigh up several different competing factors when making its preliminary structural decisions. For example, in trying to decide between two or more potential attachment sites for a modifier, it might be influenced by a preference to minimize site-modifier separations within the sentence on the one hand (Recency preference), and a tendency to try to attach the modifier as close as possible to an S node within the phrase-marker description of the sentence (Predicate Proximity). In sentences where the different tendencies pull in different directions the decision about hosting the modifier depends on the relative strengths of the competing influences. Gibson et al showed that it was possible to obtain cross-linguistic variation in host preferences by postulating that the relative weight of one of these competing factors has at least two values and varies across languages. Where the effects of Recency are weaker than those of Predicate Proximity, the model predicts a bias in favour of attaching modifiers high (to N1), whereas when the reverse is true the model predicts low (N2) attachment.

Like the model outlined in the previous section this proposal can be regarded as describing a universal parser, but again modulo parametric variation.

2.3 Tuning theory.

In the previous section we considered parsers that are subject to limited variation based on the settings of perhaps one or two parameters. It was in the spirit of these proposals that parser variation should be explained by assigning parameter values to entire languages - or arguably even employing the same settings for entire classes of languages.

We now turn to a much more open-ended proposal - a suggestion which we have called the Tuning hypothesis (Mitchell and Cuetos, 1991a; Mitchell, Cuetos and Corley, 1992; Mitchell, 1994). The suggestion is that the initial choice of structural analyses in parsing is determined not by general principles (whether parameterized or not) but rather by the experience the individual reader or listener may have had on previous encounters with ambiguities of the same kind. Put simply, the proposal is that, faced with an ambiguity, the reader/listener will initially opt for the resolution which has turned out to be appropriate most frequently in the past. Thus, the model proposes that ambiguities are not resolved by any fixed or universal rule, but that their resolution varies according to the individual's exposure to comparable examples in his or her prior experience with the language. Each time a person encounters and resolves a specific form of ambiguity, the syntactic processing mechanism adjusts itself marginally to take account the new solution. Once the ambiguity has been resolved in a given direction, the mechanism is adjusted in such a way that there will be a slightly greater possibility of choosing that resolution in subsequent encounters with the ambiguous form in question. The more an interpretation has appeared and worked successfully in the past, the higher the chance it will have of being chosen again in the future.

A statistical or experience-based mechanism of this kind would presumably iterate towards a different solution for each ambiguity in its repertoire. If the patterns of resolution in the raw material (i.e. the choices intended by speakers and writers) were themselves governed by some overriding linguistic principles then an exposure-based parser of this kind would mimic these principles - but in doing so it would not be incorporating the principles themselves. Faced with a totally arbitrary ensemble of resolution patterns, it would home in on the more probable choices just as readily as it would when handling a highly regulated training corpus.

A model of this kind can be regarded as incorporating a fairly extreme form of parameter-setting. In any given language, the conflict resolving mechanism for selecting the initial parse would be informed by the values of perhaps

several dozen parameters varying on continuous scales. Clearly an account of this kind would be less parsimonious than any model based on a much smaller number of parameters. However, parsimony aside, this type of model could provide a ready explanation of any cross-linguistic variation there may be in the process of ambiguity resolution.

Of course, the Tuning hypothesis is based on universal principles of its own. The suggestion is that for all languages the initial reading of an ambiguous constituent will be governed by the statistical properties of that particular ambiguity in the language under consideration. It is just the specific decisions that will be subject to cross-linguistic variation.

Turning to the predictions for relative clause attachment, the preferred host would be the one that is used most often when comparable ambiguities are resolved in the language under examination. Using a small-scale corpus analysis, Mitchell, Cuetos and Corley (1992) found that in English most of the ambiguities of the form "NP1-of-NP2-RC" were resolved in favour of NP2 attachment. (Indeterminate attachments were excluded from the analysis). In contrast, a Spanish corpus showed a bias in the opposite direction. The pattern of preferences to be expected on the basis of tuning, then, is one of low-attachment in English and high-attachment in Spanish, with predictions for other languages depending on the details of the corresponding corpus analyses.

2.4 Totally data-shaped parsers

According to this type of account, shaping goes beyond determining the initial analysis to be pursued in ambiguous structures. The entire computational machinery for syntactic analysis is assembled on the basis of sustained exposure to language and the final product is moulded by semantic and discourse factors as well as strictly syntactic considerations. Tuning processes of this kind play a major role in several language acquisition experiments carried out within the Competition model (MacWhinney, 1987, Bates & MacWhinney 1987). Similarly, in neural network models (McClelland & Rumelhart, 1981, Seidenberg & McClelland, 1989) processing preferences are also determined by exposure to language.

Like the Tuning hypothesis discussed above, the performance of data-shaped parsers would be expected to reflect the biases represented within the training materials. As before, then, the prediction for relative clause attachment is that the parsing bias should reflect the bias in the corpus statistics. That is, low attachment should prevail in English and high-attachment in Spanish.

2.5 Summary of models of modifier attachment

The proposals summarized above are concerned with explaining the workings of the rapid-acting parsing system that allocates modifier information to one or other of the potential hosts within a complex noun phrase. The presentation highlights theoretical differences in the extent to which the models view the parsing mechanism as being specially adapted for different languages. As a corollary of this, the various accounts differ markedly in the status they give to Late Closure or Recency. At one extreme Late Closure is treated as a central principle for guiding parsing decisions, whereas at the other it is assigned no real role in the process (though no-one rules out the possibility that there is a correlation between Recency and statistical prevalence of attachment). Between these limits, the principle is seen as a bias which has to compete with others to influence parsing choices. At the heart of the difference between the models is the question of whether these tendencies are hard-wired or honed by experience.

3. EVALUATION OF THE THEORIES IN THE LIGHT OF EMPIRICAL EVIDENCE.

We will start by outlining a number of general findings concerning how attachment ambiguity is handled in different languages. In later sections we will go on to consider the implications these results have for the various parsing theories outlined above. Within these sections we will also introduce and discuss further, more detailed, investigations which have been conducted to evaluate a number of more specific issues within the area.

3.1 Basic observations

Most of the current discussion of the competing parsing theories will revolve around their predictions concerning preferred attachments when complex noun-phrases are followed by modifiers (particularly relative clauses).

The issue was first examined by Cuetos and Mitchell (1988). In a questionnaire study using NP-PP-RC sentences like (6a,b), we simply followed each item with a question designed to establish how the subjects had resolved the ambiguity.

(6a) Someone shot the servant of the actress who was on the balcony.

(6b) Alguien disparó contra el criado de la actriz que estaba en el balcón.

The questions were of the form: "Who was on the balcony?" (for 6a) and "¿Quién estaba en el balcón?" (for 6b), and subjects were required to give a one-word answer in each case. The results showed that English readers showed a reliable preference for attaching the relative clause to the second site (i.e. to "actress"; selected on 58% of occasions), while Spanish readers opted predominantly for the first site ("criado"; chosen on 62% of trials).

The Spanish bias in favour of high attachment was corroborated in a self-paced reading study using materials in which an extra phrase was added to force a low-attachment resolution at the end of the sentence. In the case of (6b) this was achieved by adding the phrase "con su marido" ("with her husband"). Our assumption was that if our subjects had opted to attach the relative clause to "criado" (a masculine head-noun) prior to reading this phrase, they would then be faced with the anomaly of a male person being "con su marido". Assuming that this would trigger reanalysis (disconnecting the relative clause from "criado" and reattaching it to "actriz"), we predicted that the reading time would be relatively long. In fact, the results showed that the time taken to read this display was reliably longer than it was in control conditions in which subjects had no need to shift the attachment from one site to another. High-attachment preferences have subsequently been confirmed in numerous questionnaire and reading-time studies using comparable Spanish materials (e.g. Carreiras, 1992; Carreiras and Clifton, 1993; Gilboy et al, in press; Mitchell and Cuetos, 1991a,b; Mitchell, Cuetos and Zagar, 1990, as well as several further unpublished studies carried out by us in Oviedo).

The results show a clear preference for N1 or high-attachment preference in NP-de-NP-RC structures in Spanish. However the pattern of findings for English has been less consistent. Mitchell and Cuetos (1991b) reported two further questionnaire studies confirming the earlier evidence for low-attachment preference. However, Clifton (personal communication) obtained the reverse result when the identical questionnaire was used in Massachusetts. Furthermore, using a different questionnaire Clifton (1988) found additional evidence in favour of high attachment in American English. To complicate issues further, Clifton and his colleagues have shown clear evidence of low-attachment in at least two different on-line studies using relative clauses with reflexive pronouns (Clifton, 1988; Carreiras and Clifton, 1993; Frazier, 1990a). Several other studies have shown no on-line bias in either direction (e.g. Carreiras and Clifton, 1993; plus several unpublished studies from our own lab in Exeter). Overall, the results appear to indicate that the attachment biases in English are reliably different from those in Spanish (see also, Gilboy et al, in press). Several studies indicate that the shift in preference may be marked enough to reverse the Spanish bias (favouring low- rather than high-attachment). However, for reasons that are not well understood, this particular finding cannot always be replicated reliably.

Following the early investigations, similar on- and off-line studies have been carried out in several other languages. In most cases the results point to a high-attachment bias similar to that obtained in Spanish. High-attachment preferences have been found in questionnaire studies in French (Mitchell, Cuetos and Zagar, 1990), Italian (De Vincenzi and Job, 1993), Dutch (Brysbaert and Mitchell, 1993) as well as in German and Russian (Vera Kempe and Ralph Radach, personal communication). On-line results have typically corroborated these conclusions (e.g. French: Zagar and Pynte, 1992; Dutch - Brysbaert and Mitchell, 1993). In contrast with this, though, De Vincenzi and Job (1993) have claimed to have provided on-line evidence for an immediate low-attachment in Italian. However, as Carreiras and Clifton (1993) have pointed out, this particular result may be subject to segmentation artefacts (see Section 3.2.2) for further discussion of this issue). The position on attachment preferences in Italian therefore remains unclear at the time of writing.

The results considered so far have been concerned with two-site sentences including prepositional phrases with the preposition "of" (or its translation in various languages). For the sake of completeness it should be mentioned that several studies have shown that the pattern of preferences is altered when different prepositions are used in the PP in place of "of" (e.g. Clifton, 1988; Frazier, 1990a; Gilboy et al, in press; De Vincenzi and Job, 1993). The experimental evidence on "preposition effects" will be discussed in more detail in Section 3.3.1, below.

So far we have restricted ourselves to materials in which there are just two potential attachment sites within the complex relative clause. However, Gibson et al (submitted) have recently gone beyond this to examine the patterns of preferences in three-site materials of the form NP-PP-PP-RC.

They conducted two questionnaire studies in which subjects were required to indicate whether the sentence fragment was grammatically acceptable or not. On a proportion of the trials the verb in the relative clause agreed (in number) with the first noun, as in (7a). On others, it agreed with the second, as in (7b) or the third (7c).

(7a) ... la l%mpara cerca de las pinturas de las casas que fue da%ada en la inundaci%n.

(7b) ... las l%mparas cerca de la pintura de las casas que fue da%ada en la inundaci%n.

(7c) ... las l%mparas cerca de las pinturas de la casa que fue da%ada en la inundaci%n.

The results for both Spanish and English showed that subjects were most likely to judge sentences like (7b) to be ungrammatical, while (7a) was seen as being more ungrammatical than (7c). Exactly the same pattern of results turned up in the latency data in two on-line grammaticality judgement studies.

In every case considered up to now, the attachment ambiguity has involved linking a relative clause to a complex made up of a noun phrase followed by one or more prepositional phrases. However, this is not the only kind of structure which can offer competing attachment sites for a subsequent modifying constituent. The same problem may arise with NP-RC-RC sentences like (8).

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

To examine attachment preferences in cases like this, Mitchell and Cuetos (1991a) presented subjects with sentence fragments like (8) and instructed them to complete the sentence in any way they thought plausible. In these circumstances a continuation with a plural verb would indicate high-attachment to the plural noun phrase "los libros", while a singular verb continuation would suggest low-attachment to "la chica". The results showed that with these structures subjects had very strong propensity to attach low (i.e. to the second site). Cuetos, Mitchell and Corley (1994) have since corroborated this finding using carefully controlled materials and a different questionnaire technique (but see Section 3.2.3, for apparent discrepancies using on-line measures).

These findings suggest that attachment preferences depend crucially on the precise internal structure of the complex noun phrase embracing the competing attachment sites. Like all of the other findings outlined above, this is a result which will have to be explained by any viable model of attachment resolution. We now turn to a detailed evaluation of the way in which different models account for these data.

3.2 Universal accounts

3.2.1 Evaluation of Garden-path theory

Taken at face value these findings argue against an important aspect of Garden-path theory - namely that the initial attachment of the relative clause in sentences like (2) should be determined by exactly the same principles in all languages. More specifically, given that principles like Minimal Attachment do not discriminate between the alternative interpretations of the sentence, the theory specifies that the initial attachment should always be governed by the Late Closure principle. In other words, the relative clause should systematically be attached to the second of the two potential attachment sites rather than the first. The fact that this does not occur raises problems for the Garden Path theory.

In an early defence of the traditional form of the theory, Clifton (1988) identified several different potential problems in the Cuetos and Mitchell (1988) study. He suggested that the increased reading time for the final display might have occurred not because subjects were forced to reassess a mistaken initial attachment (as we had assumed), but because extra time was required simply to handle a constituent which could potentially be attached to more than one site. On this kind of interpretation, the original findings could not necessarily be taken as evidence against Late Closure.

In order to evaluate this criticism, Mitchell and Cuetos (1991b) conducted an experiment including conditions that forced Early Closure as well as Late Closure. In this study, there was also a control condition in which neither analysis was rejected in favour of the other. Overall there were four types of sentences: materials in which Late Closure was forced (e.g. 9a); sentences in which Early Closure was forced (9b); a control condition which could be interpreted by following either Early or Late Closure (9c), and a control non-ambiguous condition (9d).

(9a) Alguien dispar% contra el criado de la actriz / que estaba en el balc%n / con su marido.

(9b) Alguien dispar% contra la criada del actor / que estaba en el balc%n / con su marido.

(9c) Alguien disparó contra la criada de la actriz / que estaba en el balcón / con su marido.

(9d) Alguien disparó contra la actriz / que estaba en el balcón / con su marido.

The results showed that the reading time for the final display - "con su marido" (with her husband) was enhanced only when it forced the reader to adopt the Late Closure reading (i.e., in Condition (a)). There were no differences in the three remaining conditions indicating that neither ambiguity per se in Condition (c) nor forcing Early Closure (Condition (b)) had the same effect as forcing Late Closure. Contrary to Clifton's (1988) suggestions, this indicates that there is a genuine bias against late closure in materials of this kind, and that with materials like (a), readers have to expend effort in overcoming this bias in settling on an appropriate Late Closure analysis.

Another methodological problem raised by Clifton (personal communication) was that the segmentation employed in the early studies might have biased readers against Late Closure interpretations. In the Cuetos and Mitchell (1998) study the sentences were presented in three displays, and the segmentation of the first display always coincided with the end of the NP-PP phrase "el criado de la actriz" (the servant of the actress). Since segmentation is known to influence parsing decisions (e.g. Mitchell, 1987a) it is possible that this altered subjects' biases. However, in later experiments, Mitchell and Cuetos (1991b) and Carreiras (1992) found that the results were essentially unchanged when the segmentation of the materials was altered in such a way that the two first displays appeared together (e.g., "Alguien disparó contra el criado de la actriz que estaba en el balcón" - "Somebody shot the servant of the actress who was on the balcony"). Thus there is no evidence to suggest that artificial segmentations of the materials introduced biases which were not already evident from other sources.

Taken together the findings provide increasing evidence that the Late Closure strategy is not a procedure that is used to the exclusion of others to settle ambiguities involving modifier attachment. Evidence against low-attachment bias has since been corroborated in Spanish (Carreiras and Clifton, 1993) and has recently been reproduced in several other languages including French (Zagar and Pynte, 1992), Dutch (Brysbaert and Mitchell, 1993), Italian (De Vincenzi and Job, 1993), as well as German and Russian (Vera Kempe and Ralph Radach, personal communication). At the very least, other factors must be involved in interpreting modifiers in each of these languages.

3.2.2 Refined Garden-Path theories

Given this clear evidence in favour of high-attachment biases, Clifton and Frazier (Clifton, 1988; Frazier, 1990a) then focused on another potential weakness of the original study - namely that the bias was first tested several words after the point near the beginning of the relative clause where the initial commitment should theoretically have been made. In effect, they acknowledged that there was a high-attachment bias at this relatively late point in the sentence but attributed this to the effects of a second, discourse-based, phase of processing which reversed the initial bias (still held to be determined by Late Closure). As spelt out in Section 2.1.2, this second, lagging operation has long been accorded an important function in the Garden-Path theory. However, prior to the modifier-attachment studies its postulated activities had been restricted to those associated with the assignment of thematic roles (e.g. Rayner et al, 1983) and to certain context effects (e.g. Ferreira and Clifton, 1986). In 1990, apparently as a direct response to the Spanish attachment data, Frazier (1990a) elaborated on these proposals, introducing the notion that the second-phase operation also includes specifically discourse-related strategies designed to ensure that modifiers are normally linked up with material which is at or near the "focus" of the text being processed (cf. the Relativized Relevance Principle, Frazier, 1990a, see Section 2.1.2 above). Since this involves novel extensions to the original theory, we will refer to it as the "refined" version of the model (see also, De Vincenzi and Job, 1993, for further adjustments along these lines).

The Refined theory provides a straightforward account of the high-attachment biases that have emerged in the majority of modifier attachment studies. Within a complex noun phrase the first noun is the head of the entire phrase and is arguably the "main assertion" or focus of the constituent. According to the elaborated theory, therefore, this is precisely the site to which the discourse mechanism would be expected to attach the relative clause. So Frazier (1990a) accounted for the high-attachment data by maintaining that the demonstrated biases reflect the state of play in the sentence processing mechanism after the discourse operation has exerted its influence.

In support of a discourse element of processing, Frazier (1990a) pointed to evidence that attachment bias is apparently influenced by the informativeness of the relative clause. In particular, she reported a questionnaire study comparing people's interpretations of clause carrying very little information (e.g. locatives such as "who were on the balcony") with modifiers conveying more informative attributive information (e.g., the relative clause in "Julie met the friend of the man who reads news on Saturday Night Live"). In line with the Refined model, the results showed that the high-attachment preference was more marked in locative sentences (70%) than in the attributive materials (59%).

Although there appears to be some support for the Refined version of the Garden-Path model, there are also a number of respects in which this account is difficult to reconcile with the data.

The first problem concerns the absence of any clear evidence that reanalysis plays a systematic role in on-line studies of modifier attachment. In the past, wherever there has been persuasive evidence for reanalysis, the effects have been reflected in fairly sizable latency variations (e.g. Frazier and Rayner, 1982; Frazier, Clifton & Randall, 1983; Frazier, 1987; Mitchell, 1987a,b). If comparable operations had played a role in modifier attachment there should have been similar latency effects in appropriate circumstances. In particular, there should have been clear signs of revision effects (shifts from low- to high-attachment) in all sentences which ultimately show a high-attachment preference. However, there was no sign of such effects in several different studies (see Mitchell & Cuetos, 1991b, for detailed discussions and analysis). In fact, up to now, no study in Spanish - or indeed French (Zagar and Pynte, 1992) - has yielded any evidence that there is a phase of processing in which low-attachment bias systematically prevails, and so it is not even clear that the preconditions for these reanalysis effects can be demonstrated.

In contrast with the evidence for Spanish and French, there have been claims that there is a low-attachment (Late Closure) phase of processing in Dutch and Italian. However, as argued below, these claims are questionable and the evidence in support of attachment followed by re-attachment remains uncertain.

The evidence for Late Closure in Dutch was offered by Flores d'Arcais (1990). It was based on ambiguous sentences like (10):

- (10) Jan zag Anneke lopen op het strand
(John saw Anneke running on the beach)

In sentence of this kind it could be either Jan or Anneke, or both, who were "running on the beach". When the sentence was disambiguated with a new sentence indicating that the first person was the one who was running (e.g. "hij was moe" (he was tired)) it took subjects longer to read this test sentence than it did when it disambiguated in favour of the second attachment site (using continuations such as "ze was moe" (she was tired)). While this provides superficial support for the notion that there is a phase of processing in which low attachment dominates, it falls a long way short of providing definitive endorsement for the view that comparable biases occur in processing "NP-PP-RC" sentences. For a start, the sentence structure in (10) is totally different and it is not at all obvious that it is possible to make direct comparisons with the earlier studies. Secondly, the Flores d'Arcais (1990) experiment involved testing attachment bias after the end of the clause and so, within the framework of the Refined theory, it is implausible to argue that the results can be taken as reflecting biases in the initial rather than the later (discourse-based) stage of processing. (Hence, the high-attachment bias is in fact completely contrary to expectation rather than favouring the theory). Thirdly, and perhaps most important, the biases in sentences like (10) can be reversed by exchanging one verb for another in the matrix clause (Brysbaert and Mitchell, in progress) and cannot therefore be taken as reflecting Late Closure effects within the Refined Garden-Path theory.

The ostensible support for a Recency-governed phase of processing in Italian comes from the De Vincenzi and Job (1993) study mentioned several times above.

Subjects were presented sentences like (11) in a self-paced reading task with successive displays marked by the oblique lines (/) in the example.

- (11) L'avvocato diffida / del padre / della ragazza / che si e'tradita-o / al processo.
(The lawyer suspects / the father / of the girl / who betrayed herself-himself / at the trial.)

Although various aspects of the data appeared to suggest that subjects eventually elected to attach the ambiguous relative clause high (i.e. to "padre", in this example), the reading latencies for Display 4 turned out to be significantly longer when the (gender-related) contents of the display forced high-attachment (e.g., "che si e'tradito") than when they forced low-attachment ("che si e'tradita"). De Vincenzi and Job (1993) interpreted this as evidence that the relative clause is initially attached low (to "ragazza"), but later decoupled and reconnected to the higher site ("padre"). Going beyond the data, they argued that the parser initially adopts the Late Closure strategy, and that this preliminary analysis is later overruled by interpretative processes. However, there are problems with this inference, as various investigators have pointed out (e.g. Carreiras and Clifton, 1993; Frazier and Clifton, in preparation). Specifically, in contrast with all previous studies of relative clause attachment, the two potential attachment sites here ("padre" and "ragazza") were presented in different displays - with the first being removed from view before the second appeared. Since the two sites were not treated in an equivalent fashion, this may have produced an artificial bias in favour of "ragazza". In support of this possibility, Brysbaert and Mitchell (1993) have recently shown that parsing biases (in Dutch) can be

influenced by altering the way in which the material is segmented. The implication of this is that it is not yet clear whether De Vincenzi and Job's (1993) early low-attachment biases reflect anything other than a segmentation artefact. An unsegmented version of this experiment needs to be carried out before we can settle the interpretation of the findings.

In short, contrary to the Refined theory, there is no clear support for the view that there is a brief phase of analysis during which the relative clause is temporarily attached low prior to being re-attached to the competing site once the discourse processes come into play. Nor is there any evidence that time-consuming reanalysis takes place at the points in sentences where these revisions are hypothesized to occur.

This lack of support for the Refined theory is also evident in the failure to find evidence that attachment preferences are strongly affected by discourse and semantic considerations. If the putative second phase of processing had been dominated by such effects in the way proposed, then attachment preferences would have been expected to be strongly influenced by semantic and discourse manipulations. However, both Carreiras (1992) and Zagar and Pynte (1992) have failed to obtain such effects. Carreiras (1992) used NP-PP-RC sentences which were semantically biased towards a low attachment interpretation (e.g. "Alguien disparó contra el criado de la actriz que estaba en el escenario" - Someone shot the servant of the actress who was on the stage.) Contrary to the hypothesis that attachments are strongly affected by semantic considerations, he found that subjects continued to show a preference for high-attachment in the face of clear semantic biases. Similarly, Zagar and Pynte (1992) failed to induce changes in attachment preferences when they manipulated discourse context to favour low- rather than high-attachment.

In summary, several of the central features of the Refined theory have failed to find empirical support in recent studies and on present evidence this modified version of the theory does not seem to be substantially more viable than the Garden-Path theory in its original form.

3.2.3. Construal theory

The crucial difference between Construal theory and the Revised Garden-Path theory is that relative clauses are treated as non-primary relations which are therefore not initially attached to any particular site within the complex noun phrase. Instead they are "construed" or associated in a general way with an extended sentence region organized around a single verb (a region referred to as a theta-domain). In the absence of differential links within the theta-domain, Construal differs from Revised Garden-Path theory in that it is not undermined by failures to demonstrate successive phases of analysis in which low attachment is later replaced by high attachment. Also, in Construal theory the eventual attachment is determined by a wide variety of different considerations including discourse and focus effects, Gricean principles and semantic information. Given this multiplicity of influences, the theory is not compromised in any way by the failure to show evidence of any particular kind of effect in this stage of processing. Unlike the Refined Garden-Path theory, therefore, it is not seriously weakened by the failure to demonstrate that semantic and discourse biases are capable of influencing the final attachment.

Construal theory offers fairly simple accounts of the main findings on parsing preferences in modifier attachment studies. In structures with complex noun phrases including the preposition "of" (or its cognates) the general preference for high attachment in languages other than English is attributed to the effects of discourse focus operations which take place in the second phase of processing. Principles such as Relativized Relevance (see Section 2.1.2) are used to account for the fact that relative clauses typically end up being linked to the first, and theoretically more focussed, attachment site. The fact that there is less propensity for high attachment to occur in English (e.g. Cuetos and Mitchell, 1988; Carreiras and Clifton, 1993) is attributed essentially to the fact that the second phase of processing proceeds on the basis that the writer or speaker always makes an effort to be clear (i.e. that the material to be analysed conforms to the Gricean Maxim of Clarity). Frazier (1990a) pointed out that while there is ambiguity concerning the way modifiers should be attached to post-nominal genitives of the form "the servant of the actress", the same problem does not occur with the pre-nominal form (i.e. "the actress's servant"). In this second case subsequent modifiers can only be attached to "servant". Following the Gricean principle that writers and speakers always try to be clear and avoid ambiguity, Frazier (1990a) has argued that they would elect to use the pre-nominal form if they wanted to signal that the modifier is intended to be associated with the head of the complex noun phrase ("servant"). Using the post-nominal form, therefore, would be a signal that the modifier should be interpreted as being attached to the alternative site (i.e. the non-head - "actress"). Armed with this information about the alternative interpretations of different expressions, a reader or listener would be biased in the direction of interpreting modifiers following post-posed genitives as being attached to the non-head (i.e. the second site in the complex noun phrase). In contrast with this, languages which do not have the two alternative genitive forms (like Spanish) would not be subject to the same kind of bias or pressure to choose the non-head attachment with structures of exactly the same kind (see also, De Vincenzi and Job, 1993; Gilboy et al, in press).

Construal theory also has a ready explanation for Mitchell and Cuetos's (1991b) observation that Spanish readers seem to show a strong low-attachment preference in identifying modifier heads within complex noun phrases of the form NP-RC-RC. As indicated above, the initial construal process attaches a modifier to an entire theta-domain, and then, if there are alternative hosts within this domain, the interpretative processes in the second phase are left to make this final decision. With NP-PP structures, both potential modifier hosts will be within the critical theta-domain, and discourse/focus processes will typically cause the first to be selected (as outlined above). However, with NP-RC structures, the theta-domain will not extend beyond the relative clause. Since this included just a single potential host in the sentences used by Mitchell and Cuetos (1991b), the theory predicts that this noun phrase will be chosen unambiguously as the attachment site for the final modifier.

This provides a neat explanation of the original observations with "pertenecia" sentences. However, the results of a recent study suggest that this proposal may not be entirely adequate as a description of immediate processing in sentences of this kind. In particular, it seems that on-line, as opposed to off-line questionnaire measures fail to show any statistical preference for the second as opposed to the first attachment site (Mitchell, Cuetos and Brysbaert, 1994..to be changed to Cuetos and Mitchell, when the written paper is done). According to Construal theory a shift to the earlier noun phrase involves detaching the modifier from its preferred theta-domain: a change which should trigger a full-scale reanalysis, causing an increment in processing time which should theoretically be just as large as that associated with any other instance of parsing reanalysis. The fact that this did not occur suggests that the account of ambiguity resolution currently offered by Construal theory may be incomplete.

The theory also has difficulty in explaining certain other findings, including some of the cross-linguistic differences in attachment preference reported by Gilboy et al (in press). As indicated above, the theory predicts that such differences should be restricted to structures in which just one of the languages allows an alternative expression to be used to convey one of the competing interpretations - as in the example of the pre- and post-nominal genitives discussed above. Thus, the interpretative process opts to attach a modifier to "actress" in the phrase "servant of the actress" essentially because the string "..actress's servant who was.." could have been used to express "servant-attachment" if that had been the writer's intention. One implication of this proposal is that cross-linguistic differences should not show up in cases where the alternative expression is not available. This was true of at least one of the classes of sentence examined by Gilboy et al: sentences dealing with quantities or measures (Example 11a).

(11a) John asked for the glass of water that was on the table.

(11b) *John asked for the water's glass that was on the table.

In cases like this, the preposed genitive form is ungrammatical (see 11b) and therefore unavailable as a device for disambiguating modifier attachment (Footnote 1). It follows, therefore, that the theory predicts that the cross-linguistic difference should disappear in sentences of this kind. In fact, the Gilboy et al results showed that the difference in this case was the largest of all the structures they examined. This suggests that the effect cannot depend upon the availability of a specific alternative genitive form, but depends instead on the accumulated or averaged influences over a wider class of items (such as all post-posed genitives). A statistical account of this kind was discussed in Section 2.3 and will be considered in further detail in Section 3.3.2.

One further finding that poses problems for Construal theory (and, indeed, for any other universal account of parsing) is that there are stable individual differences in preferences for high- or low-attachment. If it were really true that host-selection is determined by interpretative procedures which have universal application, then one would not expect to find systematic differences in biases for different readers. However, questionnaire studies show that people do, in fact, differ significantly in their tendency to opt for interpretations consistent with high- and low-attachment (Mitchell, 1988; Brysbaert and Mitchell, 1993). In the most detailed examination of this phenomenon to date, one of us (MMBC) tested 52 subjects on two parallel forms of a questionnaire administered three weeks apart. Informants were given incomplete sentences like (12a,b) - with noun-order and number marking counterbalanced over questionnaire-forms.

(12a) The court heard about the investigator(s) of the murder(s)...

(12b) The court heard about the murder(s) of the investigator(s)...

The task was to provide a completion for the sentence starting with either "who" or "which" followed immediately by either "was" or "were". In the overwhelming majority of cases, the choice of relative pronoun and verb-number was sufficient to indicate whether the subject intended the relative clause to be attached high or low. The results showed a very high positive correlation ($r=0.799$, $df=50$, $p < 0.001$) between the proportion of low-attachment choices

in the first test and the corresponding figure in the second test three weeks later. This confirms not only that there are substantial individual differences in attachment preference (figures in this study ranged from 0% to 100% low-attachment), but also that these effects are stable over a substantial period of time and are therefore not likely to be attributable to temporary priming effects and other similar short-lived phenomena.

To account for this findings, the Construal theory would at the very least have to acknowledge that there must be individual differences in the way different people deploy their parsing and interpretative machinery - a concession which goes against the spirit of any truly universal processing mechanism. If a universal system can be configured in different ways by individual "users" then, rather than being completely universal, it must at least be subject to parametric variation. It is to this possibility that we turn in the next section.

3.3 Evaluation of parameterized accounts

3.3.1 Strategy Competition models

The original, highly speculative, version of this account was that there is competition between late closure and a special-purpose strategy that deals exclusively with attaching post-positioned modifiers to noun phrases. Specifically Cuetos and Mitchell (1988) speculated about the role of what we have referred to as the "modifier-straddling" strategy - which is switched ON in languages in which adjectives normally precede the noun and switched OFF in languages with post-nominal adjectives. On this account late closure would be unchallenged in languages like English, whereas in post-modifying languages this strategy would be counterbalanced and potentially overwhelmed - culminating in a high-attachment bias and thus providing a satisfactory account of the early data. However, subsequent research has shown that this particular variant of the model is almost certainly unviable. For a start, De Vincenzi and Job (1993) observed that in Italian (a post-modifying language in which the straddling strategy should be active) it is easy to find sentences structured like the "servant of the actress who.." examples but which contradict the hypothesis by being interpreted in such a way that they show a low-attachment bias. They pointed out that this can occur even when semantic considerations strongly favour the alternative attachment (e.g. in "Il vino dello zio allungato con l'acqua" - "The wine of the uncle diluted with water.") Thus the straddling strategy does not seem to be active in certain sentences in a language where it should be influential. Furthermore, recent data from Dutch show that the putative special-purpose strategy appears to be active in a language in which the straddling strategy should be switched off. Specifically, Brysbaert and Mitchell (1993) found that there were clear high-attachment biases despite the fact that Dutch is a pre-modifying language. Recent evidence suggests the same may be true of German (Ralph Radach and Vera Kempe, personal communication). Clearly, if there is a parameter-controlled special-purpose strategy which competes with late closure, then the activation of this strategy is not intimately connected with adjective-noun order.

A potentially more plausible candidate for the special-purpose procedure is the one guided by Gibson et al's (submitted) Predicate Proximity principle - which introduces a bias in favour of attaching modifiers as close as possible to the root of a predicate (such as a verb phrase). With structures of the form NP-PP-RC, such a principle would bias the final clause attachment in favour of the first NP since this node is closer to the predicate incorporating the complex noun phrase as a whole. Gibson et al (1994) proposed that the cost associated with attaching a modifier to a noun-phrase at a fixed distance away from a predicate may be a value that varies parametrically from one language to another. Thus languages like Spanish and French might be characterized by high parameter values (or costs) - with the consequence that non-proximal attachments are strongly discouraged, effectively supporting high-attachment resolutions. Other languages (like English) might have lower values resulting in weaker biases in favour of high-attachment.

According to the model presented by Gibson et al (submitted) the eventual choice of attachment is governed by the balance of costs imposed by these and other kinds of competing influences. With modifier attachment sentences of the kind being examined here the main competitor would be Recency Preference - a bias which favours recent over more distant attachment sites and operates with equal force across all languages. To account for the data Gibson et al merely assume that Predicate Proximity is strong enough to overwhelm the effects of Recency Preference in some languages (such as Spanish and French), while in other languages (e.g. English) its effects are relatively weak - allowing Recency to prevail.

To investigate this hypothesis Gibson et al (submitted) conducted a series of studies using sentence segments including three potential attachment sites (as opposed to the two sites used in all other studies to date).

- (13) ... la lámpara cerca de las pinturas de las casas que fue dañada en la inundación.
(The lamp near the paintings of the houses which was damaged in the flood.)

According to the model the first site (N1) should be favoured over N2 and N3 by Predicate Proximity, while Recency preference should bias attachments in favour of N3 more than N2, and in turn on N2 more than N1. By making certain simplifying mathematical assumptions Gibson et al (submitted) predicted that at the outcome of these competing processes subjects overall preference would be to attach to N3 more than N1, with both preferred to N2.

As outlined in Section 2.2.1.2, Gibson et al (submitted) carried out on- and off-line grammaticality judgement studies in Spanish and English to assess the attachment preferences in three-site sentences. As indicated earlier, the result confirmed the predictions outlined above. Gibson et al therefore took the data as support for their proposal that attachment preferences are decided by the balance of two competing principles (Predicate Proximity and Recency). By hypothesis the same kind of competition would occur in two-site structures of the kind discussed at length in this chapter. In particular the suggestion is that the universal attachment bias (Recency Preference) is counterbalanced to different degrees by the varying effects of a language-related special-purpose strategy (Predicate Proximity Preference).

Unfortunately there are signs that this version of the strategy competition model may fare little better than the earlier "modifier-straddling" proposal.

The first problem concerns the model's attachment predictions for NP-RC1-RC2 sentences like (8) - repeated here as (14) for the reader's convenience.

(14) Pedro miraba los libros que pertencian a la chica que...

In this kind of sentence both potential attachment sites are close to a predicate, and so the Predicate Proximity principle should not distinguish between the two candidates. According to the model, it follows that the question of modifier attachment should be decided by Recency alone. Hence the prediction is that there should be unequivocal support for low attachment in this structure. In practice, although this has been confirmed in off-line studies (Mitchell and Cuetos, 1991b; Mitchell, Cuetos and Brysbaert, 1994), a recent self-paced reading study has shown that there is no reliable bias of this kind in on-line processing (Mitchell, Cuetos and Brysbaert, 1994 - or Cuetos and M, in progress). Given that numerous studies have shown stable effects when, according to the model, different strategies are in competition with one another, it is difficult to see why the biases should not be at least as emphatic when a single principle (Recency) is allowed to act without counter-effects. This finding suggests that attachment decisions may be influenced by factors other than those highlighted in the Gibson et al model.

A second problem is that there is increasing evidence that attachment preferences in NP-PP-RC structures may vary depending upon precisely which preposition is incorporated within the prepositional phrase (Clifton, 1988; Frazier, 1990a; De Vincenzi and Job, 1993; Gilboy et al, in press). For example, prepositions like "with" ("con") or "near" ("cerca") appear to encourage a higher proportion of low-attachments than "of" ("de") does. Such changes are particularly marked with sentences like "Alguien disparo el criado de aquella actriz que estaba en el balc%on" (the servant of that actress who was on the balcony), where there is virtually no doubt that it is "la actriz" (the actress) who was on the balcony. Since lexical changes of this kind would have a very marginal effect, if any, on Recency and should have no effect at all on Predicate Proximity, it seems unlikely that the Gibson et al model is capable of accounting for these findings as it stands. Of course, Gibson and his colleagues could question these findings on the grounds that these investigators used different experimental tasks and different materials from their own. We therefore carried out a study to examine the lexical effects in further detail.

We administered questionnaires to 112 Spanish speaking subjects - with 28 participants in each of four conditions). Two groups were required to make attachment judgements following instructions similar to those used by Cuetos and Mitchell (1988) (see examples below). The other two groups were required to make grammaticality judgements using instructions modelled on those used by Gibson et al (1994). In each pair, one of the two groups was presented with sentences similar to those used by Gibson et al. (i.e. "N1 cerca del N2 del N3" - "N1 near N2 of the N3" - see 15a and 16a). In the other group the word "cerca" ("near") was removed (i.e. "N1 del N2 del N3" - "N1 of the N2 of the N3" - 15b and 16b). In the attachment-judgement forms of the questionnaire (15a,b) the test material was followed by a question designed to reveal how the subject had resolved the attachment ambiguity. In the grammaticality judgement forms (16a,b), the subject simply had to mark "Si" or "Non" to indicate whether the constituent was grammatical or not.

(15a) N1 del N2 del N3

... el jard%n del garage del edificio que estaba tan deteriorado.
%Qu% estaba deteriorado? _____

(15b) N1 cerca de N2 de N3

... el jard%n cerca del garage del edificio que estaba tan deteriorado.

%Qu% estaba deteriorado? _____

(16a) N1 del N2 del N3

... el jard%n de los garages de los edificios que estaba tan deteriorado.

SI NO

(16b) N1 cerca de N2 de N3

... el jard%n cerca de los garages de los edificios que estaba tan deteriorado.

SI NO

The results of the study showed marked preposition effects with both kinds of judgement task. The detailed findings are shown in Tables 1 and 2.

Table 1: Percentage of judgments favouring N1, N2 and N3 sites in the attachment judgement task.

	N1	N2	N3
"Cerca" form of questionnaire	51.19%	13.69%	34.52%
"De" or "Del" form of questionnaire	34.82%	32.74%	32.44%

Table 2: Percentage of items judged non-grammatical when the relative clause is resolved as being attached to sites N1, N2 or N3 in the grammaticality judgement task.

	N1	N2	N3
"Cerca" form of questionnaire	28.57%	59.82%	23.21%
"De" or "Del" form of questionnaire	58.03%	70.53%	30.36%

The results show that performance was markedly affected by changing the preposition associated with the middle site (N2). In the formal statement of the Gibson et al model the parameters for neither Recency nor Predicate Proximity should be affected by this manipulation (see Gibson et al, 1994, for precise definitions of the parameter weights). The present findings therefore corroborate earlier evidence that preposition-identity influences attachment preference and confirms that certain alterations are needed before the model can account for the data.

A notable feature of these findings is that the two tasks (attachment judgment and grammaticality judgement) appeared to show markedly different patterns of attachment preference. To investigate these differences further we used the grammaticality judgement task to examine attachment biases in two-site materials more similar to those used by Cuetos and Mitchell (1988) and others. There were four versions of the questionnaire each completed by 20 subjects. In two conditions the verb within the relative clause agreed in number with the noun in the first site (high-attachment sentences) and in the other two conditions with the second noun site (low-attachment sentences). In one of the two versions the subordinate verb was singular and in other it was plural. Examples of the four conditions are given in (17a,b,c,d):

(17a) el criado de las actrices que estaba en el balc%n (High-attachment)

(17b) el criado de las actrices que estaban en el balc%n (Low attachment)

(17c) los criados de la actriz que estaba en el balc%n (Low attachment)

(17d) los criados de la actriz que estaban en el balc%n (High attachment)

The results are shown in Table 3

First noun	Second noun	Verb	Judged as ungrammatical
S	P	S (High)	30%
S	P	P (Low)	15%
P	S	S (Low)	25%
P	S	P (High)	17%

In contrast with other experimental procedures which show a uniform bias in favour of high-attachment, the grammaticality judgement task reveals no reliable differences. 23% and 20% of the high- and low-attached materials respectively were judged to be unacceptable. Interestingly, however, there was a substantial effect of verb number. For some reason people are more likely to say that the sentence is ungrammatical if the verb is singular (27%) than if it is plural (16%).

These results confirm that the grammaticality-judgment task generates patterns of data which are markedly different from those obtained in other investigations of modifier attachment. Moreover, anomalous effects such as the influence of verb-number just described, suggest that performance in this task may be subject to influences other than those associated with modifier attachment. The verb-plurality effect is reminiscent of similar non-grammatical effects that have been shown to affect performance in certain other linguistic tasks. For example, Bock and Miller (1991) have demonstrated that subjects show a (plurality-related) tendency to try to make a verb agree with the closest noun, even when this is grammatically incorrect. The current findings suggests that extraneous, non-grammatical biases may also play a role in the grammaticality-judgement task. If so, it suggests that it is unsafe to use such tasks to draw direct conclusions about attachment preferences in the way Gibson and his colleagues have sought to do.

We are now in a position to evaluate the Recency/Predicate Proximity version of the strategy-competition model. The evidence suggests that the model as it stands is not capable of providing a full account of the data. In particular, its predictions for NP-RC-RC structures are disconfirmed and it fails to explain systematic changes that occur when prepositions within the complex noun phrase are replaced by others. Moreover, the grammaticality judgement task used by Gibson et al (1994) seems to be subject to various extraneous biases and may not provide a secure basis for evaluating models of attachment preference.

However, none of the experimental findings to date specifically rule out any role for either Recency preference or Predicate Proximity, and it is quite possible that these principles will figure in subsequent, fuller accounts of attachment phenomena. A positive feature of the model is that it explicitly predicts a U-shaped function for attachment preferences in 3-site materials - a finding which has been corroborated by methods other than grammaticality-judgment (as outlined above). A second advantage is that, by postulating cross-linguistic variation in the relative strengths of the two competing principle, it provides a straightforward explanation of differences in attachment preference across languages and this extends the scope of the model in providing accounts of parsing effects in different languages. At present there is no clear indication how these parametric variations might come about. However, Gibson et al (1994) have outlined various alternatives and this is an issue that could easily be explored further in the future. Overall, then, the Recency/Predicate Proximity model seems likely to provide a useful framework for extending our understanding of attachment phenomena.

3.3.2 Tuning hypothesis

The tuning hypothesis allows one to make two different kinds of prediction. The basic proposal is that ambiguities are initially resolved in line with the statistical prevalence of the alternative readings in the language as a whole. So, on the one hand it is possible to estimate these statistical values (using corpus studies) and hence predict the patterns of preference to be expected with individual linguistic structures. However, in addition to these static, well-established biases the model can also be used to predict more local or temporary biases. A central feature of any full account of tuning would be a description of the mechanism used to keep tally of the occurrence of events in different categories. In order to respond to cumulative frequencies such a mechanism would have to be subject to incremental change, and if there was a shift in the input distribution this would have to be reflected in some way in the device's internal settings. If sufficient weight is given to recent samplings, the tuning hypothesis will predict that preferences will show short-term changes on the basis of exposure patterns over the preceding minutes, days or weeks. Thus, the model predicts that parsing preferences will change if, during some period prior to testing, the reader or listener has been exposed to an unusual preponderance of one ambiguity resolution rather than another. In addition to making predictions about long- and short-term biases the model is easy to reconcile with any evidence that there are individual differences in bias from one person to another. In principle, each "ambiguity-resolving device" must be tuned on the basis of its own input distributions (experience). In these circumstances different devices would be tuned to different sub-samples of the language as a whole, and the different histories of training could well lead to different parameter settings (particularly if distributions differ, say, in written and spoken language, in broadcast versus immediate social forms, and so on). Thus, tuning processes could well result in individual differences in preferences which remain stable over time.

Recent corpus studies of long-term effects have provided some encouraging support for the Tuning hypothesis. In connection with modifier attachment preferences, Mitchell, Cuetos and Corley (1992) reported a small-scale study indicating that in two-site ambiguities a 450,000-word Spanish corpus revealed that 60% of resolvable relative clause attachments were linked to the first site whereas in English the pattern was reversed (38% high-attachment). If these patterns are confirmed in larger-scale studies, this will indicate that the Tuning hypothesis makes the correct predictions for on- and off-line attachments in Spanish and English. To our knowledge no corpus studies of two-site attachment have been carried out in other languages.

Recently, Gibson and Pearlmutter (1994) and Gibson and Loomis (in press) have examined attachment distributions in materials with three potential attachment sites in two different English corpora. The results show that modifiers are attached most frequently to Site 3, then to Site 1, with Site 2 links occurring on only a small proportion of cases. The on- and off-line data reported by Gibson et al (1994) shows that the same pattern of preferences turns up in the results of their grammaticality judgement data. If the same pattern is reproduced in parsing, then this result, too, would be consistent with the Tuning hypothesis.

Evidence from other structures is also beginning to accumulate. For example, Tabossi, Spivey-Knowlton, McRae and Tanenhaus (in press) have compared the corpus statistics and latency data for sentences beginning with a noun phrase followed by a verb with the inflection "-ed". In certain circumstances such verbs are ambiguous, in the sense that they could either be the main verb of the sentence or a subordinate verb within a reduced relative clause. Tabossi et al (in press) presented evidence that it is the first (i.e. main verb) interpretation which prevails in corpus occurrences of these forms - exactly in line with the well-known on-line evidence that this interpretation is given precedence in first-pass parsing (see also, Trueswell, Tanenhaus and Garnsey, in press; MacDonald, in press, for further evidence in relation to this structure and Trueswell, Tanenhaus and Kello, 1993, for comparable evidence with a different form of ambiguity).

Another structure in which we have examined the relation between statistical-corpus patterns and parsing preferences involves contractions of the English auxiliaries "had" and "would" (Corley, Mitchell & Cuetos, 1993). In Example (18) the contracted form "she'd" can either represent "she had" or "she would" and in cases like this there is no prior syntactic or discourse information to indicate which reading is more appropriate.

(18) The gambler told everyone she'd bet the money to win a fortune. (La jugadora dijo a todos que apostar%a/apost % el dinero para ganar una fortuna)

A corpus analysis revealed that "had" is more frequent than "would", so presumably the past-tense ("she had") interpretation of "she'd" is more common than the conditional one ("she would"). Given this statistical evidence the Tuning hypothesis predicts that the former interpretation should initially be selected in on-line sentence processing. It follows, therefore, that readers should be garden-pathed more when they read sentences like (19a) (which is resolved in favour of the "would" interpretation) than they are with sentences like (19b) ("had" resolution).

(19a) He'd come to the house and paint the exterior.

(19b) He'd come to the house and painted the exterior.

(19c) He would come to the house and paint the exterior.

(19d) He had come to the house and painted the exterior.

Subjects were required to read sentences of this kind - together with control conditions (like 19c,d) - in a self-paced reading task. Reading times for the disambiguating display "and paint(ed)" confirmed that there were longer latencies in the (a) than in the (b) condition (982ms versus 893ms) and that this was reliably different from the (reverse) effect in the control conditions (c - 806ms; d - 929ms).

The next set of predictions concern tuning effects and parsing biases which are expected to change over time. According to the Tuning hypothesis, in the absence of statistical evidence there should be no a priori reason for choosing one resolution over any of its competitors. This means that prior to being exposed to the language in question, people should theoretically respond in an unbiased fashion, and then as contact with the language increases the preferences should steadily asymptote towards the values represented in the corpus. In the case of Spanish two-site attachments it follows that we can predict that preferences should shift from being unbiased at first (with 50% high and 50% low attachment), but gradually follow the corpus statistics and shift in the direction of high-attachment preference. Thus, we predicted that Spanish children should progressively favour high-attachment as they get older and are increasingly exposed to biased material. (English children in contrast, should start in the same position and move in the opposite direction). To test the Spanish part of this prediction we carried out questionnaire studies with three groups of 24 children in their 2nd, 3rd and 4th years of school attendance (7, 8 and 9 years-old) respectively. As expected the results showed that there was a significant positive correlation between the children's chronological age and the percentage of high-attachment choices they made in the questionnaire ($r = 0.3789$, $p < .001$). The overall group averages are given in Table 4.

Table 4 - Percentage of high-attachment judgements at different ages

Year 2	59.1
Year 3	72.7

Another prediction that can be made from the short-term effects of statistical tuning hypothesis is that if it is really the frequency with which the structures appear that determines their structural choices, then if people are exposed to an artificial preponderance of one particular attachment, it should be possible to demonstrate that this has an effect on their preferences. To test this prediction we carried out an intervention study in which we exposed Spanish children to a highly biased sample of attachment materials. One group was systematically exposed to high-attachment sentences and the other to low-attachment materials. The study was carried out in three main stages. First, a group of children was pretested by filling out an attachment questionnaire similar to those used by Cuetos and Mitchell (1988). On the basis of the results, children were allocated to two groups of 24 matched closely for attachment bias. Following this, over a period of two weeks (10 school days) both groups of children were required each day to read three short stories each including two NP-PP-RC sentences (i.e. each child read 60 two-site "exposure" sentences in all). In the High-attachment bias group the relative clauses were always resolved as being attached to the first noun, whereas in the Low-attachment bias group it always modified the second noun. After a further week's interval without any activity connected with the topic of research, the subjects were required to complete a second, matched form of the questionnaire. The results showed that the exposure regime had a reliable effect on the children's attachment preferences. Following the different experimental treatments the two groups differed significantly in the proportion of high-attachment choices they made the final (post-test) questionnaire ($F(1, 46) = 11.1, p < .005$) (see Table 5, below). The subjects exposed to a preponderance of high-attachment materials shifted in the direction of making a higher proportion of high-attachment choices. However, for some reason the reverse result did not occur for the subjects who had been exposed to low-attached sentences. The reason for asymmetry may lie in the exposure materials used with this group or in the fact that our subjects, although quite young, already had high-attachments firmly fixed at the beginning of the experiment, in which case the 60 new materials may not have been enough to alter their choices.

	Pretest	Post-test
	(matched)	
Group 1 (High-attachment bias)		
High-attachment preference	71.44%	92.17
Low-attachment preference	28.56	7.83
Group 2 (Low-attachment bias)		
High-attachment preference	70.80	72.19
Low-attachment preference	29.20	27.81

These intervention data are difficult to explain on the basis of theories which postulate that parsing choices are governed by fixed, linguistically determined strategies that operate in the same way for all subjects in all languages.

As a general proposal, the Tuning hypothesis is capable of providing a satisfactory account of all the observations concerning parsing preferences that occur in modifier attachment ambiguities. (A strong case can be made for extending the application of the proposal to account for initial preferences in all other forms of structural ambiguity. However, this is beyond the scope of the present chapter.) Despite the positive support from corpus analyses and from intervention studies, however, this account of immediate decision-making has been criticised.

One objection (expressed for example by Carreiras and Clifton, 1993; Gibson et al, 1994) is that in its most general form the hypothesis does not specify the precise level of detail at which statistical records are kept for later use in guiding parsing processes. Are two-site data stored (and consulted) separately from three-site data? Are the attachment-tallies for complex noun phrases of the form NP-PP distinct from those for NP-RC complexes? Where the noun phrase includes a prepositional phrase are detailed statistics filed away for attachment patterns in different subclasses of prepositions? Are the records further subdivided when the noun-phrases are modified by adjectives?

Clearly a fully predictive account would have to make a commitment on each of these distinctions. The Tuning hypothesis as it stands really represents a family of models with the individual members differing from each other with respect to the grain of analysis. Viewed from this perspective it is an empirical question whether particular variants of the model are viable or not. Experimental findings already discussed indicate that the coarser-grained versions of the model would not be capable of accounting for the data. A model which merged the statistical counts for NP-PP structures with those for attachments to NP-RC forms, would not have any explanation for the different patterns of

attachment that appear to characterise these two forms. Equally, existing data suggest that any viable version of the model would have to take account of preposition identity in subdividing its records.

The obvious danger in allowing the issue to be settled by empirical evidence in this way, is that it may eventually prove to be impossible to falsify the hypothesis. Each new discovery might be used to justify a new form of record-keeping, allowing the hypothesis itself to survive unscathed. However, it is unlikely to be a straightforward matter to keep the hypothesis on "life-support" in this way. If the proposal is to remain persuasive it will always be necessary to find a level of detail or grain in which corpus statistics and behavioural evidence pattern in the same way - and there is no guarantee that this will inevitably prove to be possible. Moreover, to provide a convincing account of the data, rather than a mere redescription, there would have to be a clear restriction on the freedom to vary grain in providing explanations for different findings. Given these provisos there does seem to be any case for viewing the hypothesis as less falsifiable than any of the existing competitors.

A second objection to the Tuning hypothesis is that by concentrating on statistical patterns in the language it directs attention away from the underlying explanations for those patterns - the linguistic phenomena which may be the ultimate determinants of parsing preferences (see Carreiras and Clifton, 1993, for a version of this argument)(Footnote 2). This observation presupposes that an account which "explains" the distribution of structures in the language is necessarily a "better" or more satisfactory account. However, it is not at all evident that this is the case. The relative merits of the different accounts depend on precisely which phenomena they are trying to explain. In proposing a mechanism for using distributional information in parsing, the Tuning hypothesis does not place any restrictions on the development of theories concerning the origins of the distributional patterns. The proposal can co-exist readily with theories of this kind. For example there may be some truth in the Gricean account of the preference for low-attachments to English structures of the form NP-of-NP (see Section 3.2.3 for details). For the reasons discussed earlier, one resolution may be associated with the post-nominal genitive, while the other is typically expressed by the pre-nominal form (e.g. "the colonel's daughter"). However, it is not necessary to propose that parsing biases occur because the Gricean computations are carried out "on the fly". Such factors may merely influence the statistical record of resolutions, and the real-time parser may then act very rapidly to convert this intermediate record into an appropriate parsing preference. The function of the statistically-driven parser is to explain the rapid mechanisms employed in real-time parsing and not to go beyond that an account for the prevalence of different structures in the language. Of course, a theory which is capable of operating over both domains would rightly be seen as a "better" theory. However, such a theory would not be capable of accounting for statistical effects which are truly arbitrary or conventional, or for effects that are introduced at an experimenter's whim - such as the intervention findings discussed above. No explanation which concentrates exclusively on computational and linguistic effects (ignoring "unmotivated" statistical effects) will be able to explain results of this kind.

In short, we simply do not agree that the Tuning hypothesis is any less adequate as a theoretical framework than most of its competitors. It is capable of generating testable predictions - as in the case of the developmental and intervention studies discussed earlier - and it is capable of providing a plausible account of most of the existing findings within its domain of application.

3.3.3. Data-shaped parsers

We shall limit ourselves to a very brief observation about the viability of models in this category and that is that they do not provide a natural account of the fact that attachment preferences do not appear to be strongly influenced by semantic and discourse biases. As indicated above, Carreiras (1992) has failed to show semantic effects and Zagar and Pynte (1992) found no evidence of discourse influences. In contrast with these findings, models which parsing decisions are made by non-modular neural nets should be subject to non-syntactic influences of this kind. The absence of such effects suggests that there is a degree of modularization - a feature which is better captured by all of the accounts considered in more detail above.

3.3.4. Overview of evaluation

Although the evidence is sketchy and still only covers a minute proportion of the world's 6000 plus languages, it is possible to use these investigations to draw some tentative conclusions about the nature of the parsing process. The results pose severe problems for theories which maintain that parsing is a completely universal process. The various forms of the Garden-Path theory and the Construal theory are undermined by evidence that ambiguity resolution procedures can be altered by experience and can vary from individual to individual. To accommodate such variation, these accounts need to be expanded or developed in some way. These models are also incomplete in their account of cross-linguistic differences and apparently show shortcomings in explaining attachment preferences in NP-RC-RC structures.

Parameterized accounts explicitly incorporate devices for simulating variation, and are therefore less challenged by evidence that attachment preferences vary between people and can be changed by exposure to language. However, these models encounter problems of their own. Modifier-straddling seems beyond salvage, having made the wrong predictions for at least two of the six languages for which we have preliminary evidence. The Recency/Predicate Proximity model has considerably greater potential, but faces severe problems in the accumulating evidence that some attachment preferences may be "preposition-specific".

The Tuning hypothesis provides a satisfactory account of the data - particularly the intervention and developmental findings. As indicated at several points in the discussion, statistical devices are perfectly capable of explaining most of the findings traditionally interpreted in terms of linguistic or computational simplicity. It only needs to be shown that ambiguity resolution distributions in corpora are skewed in the same way as the processing biases in parsing studies. At present, positive support is restricted to the structures considered in Section 3.2.3. However, we are currently carrying out corpus studies on a wider range of structures in an attempt to establish whether the hypothesis continues to provide a plausible explanation of the data. These findings should allow us to be more specific about issues such as the grain of record-keeping and we anticipate that this work will provide a basis for formulating the proposal as a fully predictive model of parsing choices.

4. CONCLUSIONS

We can now return to the questions raised at the beginning of the chapter. Is the parsing mechanism essentially "preconfigured" or fixed for all languages, or do its internal workings vary either as a function of the language user's contact with linguistic material or on some other unrelated basis? If the parser's configuration is related to language contact, is it appropriate to view the parser for a particular language as a computational process in which grammar parameters set during acquisition are used to customize the parsing mechanism? Or, is the putative customization process unrelated to the acquisition of grammar and linked instead to more basic principles of efficiency?

Some of these questions can be addressed by the investigations of the way in which people process sentences including modifiers that have to be attached to one of two or more sites within a complex noun phrase. Our survey of the work in a number of different languages raises numerous problems for models which assume that the parser is a fixed, universal, generic mechanism (e.g. the Garden-Path and Construal theories). In contrast it shows that the models which fare best are those which allow a degree of customization within the system. In principle, this tailoring of the parsing routine could be limited to setting the value of a single parameter - as in the Recency/Predicate Proximity model. However, this particular model does not appear to be flexible enough to account for existing data and it seems unlikely that a single-parameter account of this kind will ultimately prove to be viable. More plausible are multi-parameter accounts in which contact with language is used to adjust the weights associated with a number of different combinations of linguistic units (as in the Tuning hypothesis). Thus the evidence argues against both completely universal parsers (e.g. Frazier, 1987; Crocker, 1992; Inoue and Fodor, in press; Bever, 1970; Kimball, 1973 and Pritchett, 1992) and against parsers that are universal apart from variations that are parametrically linked to grammar settings (e.g. Frazier and Rayner, 1988; Mazuka and Lust, 1990).

In drawing these conclusions, we are making no claims about the universality of parsing processes other than those used to select the initial structural interpretation of linguistic material. Linguistic experience may or may not have an influence of the processes used to assemble structures, to check that they are internally consistent and compatible with what has gone before, and to reanalyze material following garden-pathing. The results reviewed here are compatible with the proposal that these decisions are determined entirely by language users' individual experiences - suggesting that parsing processes may be characterized by a considerable degree of adjustment or customization. At the same time it is not yet feasible to rule out the possibility that a significant proportion of parsing operations other than those highlighted here are either fixed and universal, or parameterized within a universal framework. The fundamental questions with which we introduced the chapter therefore remain open to further debate. However, it is encouraging to note that empirical findings are beginning to place constraints on the kinds of answers that can reasonably be given to these questions.

Acknowledgments

The research reported here was supported by grant No. DF-93-206-20 from the University of Oviedo and by grant No. R 0023 4062 from the Economic and Social Research Council (UK).

Footnote 1 - In this case it might be argued that the alternative expression would be "water glass" as opposed to "water's glass". However, adjectival noun phrases of this kind are not genuine alternatives since they cannot be used

interchangeably with the post-modified form. The expressions "water glass" "beer can" "jam jar" refer to the receptacles themselves, whereas the alternative embraces both receptacle and contents with some emphasis on the latter. Most people would prefer to be offered a "can of beer" rather than a "beer can"! Of the nine Sentence Types examined by Gilboy et al (in press) this particular one appears to have no alternative to the NP-PP form and so the authors should predict (incorrectly) that the attachment data for English should be comparable with those for Spanish.

Footnote 2 - This criticism is comparable to objecting to DNA-based accounts of medical and biological phenomena on the grounds that they offer no explanation of the provenance of different distributions of DNA patterns. However, it is uncontroversial that DNA is the vehicle of genetic transmission and therefore perfectly respectable to formulate hypotheses in terms of DNA-distinctions. Moreover, hypotheses formulated at this level can, and do, co-exist with speculations about ultimate causes of DNA distributions. Like DNA, though with vastly more modest scope, the Tuning hypothesis is concerned to provide a low-level molecular account of the immediate mechanism by which biases and tendencies are implemented.

References

- Atkinson, M. (1992) *Children's syntax: Introduction to principles and parameters theory*. Oxford: Basil Blackwell.
- Bickerton, D. (1990) *Language and species*. London: University of Chicago Press.
- BATES, E. & MacWHINNEY, B, (1987) Competition variation and language learning. In B. MacWhinney (Ed.): *Mechanisms of Language Acquisition*. Hillsdale, LEA.
- BEVER, T.G. (1970) The cognitive basis for linguistic structures. In J.R. Hayes (Ed): *Cognition and the development of language*. New York, John Wiley and Sons.
- BOCK, K. & MILLER, C.A. (1991) Broken agreement. *Cognitive Psychology*, 23, 45-93.
- Brysbaert, M. and Mitchell, D.C. (1993) Sentence parsing in Dutch: Another departure from Late Closure. Paper presented to the Dutch Psychonomic Society, Noordwijkerhout, December 18-19.
- CARREIRAS, M. (1992) Estrategias de análisis sintáctico en el procesamiento de frases: Cierre temprano versus cierre último. *Cognitiva*, 4 (1), 3-27.
- CARREIRAS, M. & CLIFTON, C. (1993) Relative clause interpretation preferences in Spanish and English. *Language and Speech*, 36, 353-372.
- Chomsky, N (1981) Principles and parameters in syntactic theory. In N. Horstein and D. Lightfoot (Eds.) *Explanations in Linguistics*. London: Longman.
- Chomsky, N. (1986) *Knowledge of language: Its nature, origin and use*. New York: Praeger.
- 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, Sydney, August 19-21.
- Clifton, C., Jr., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, 4, 77-103.
- Cook, V.J. (1988) *Chomsky's Universal Grammar: An introduction*. Oxford: Basil Blackwell.
- Corley, M.M.B., Mitchell, D.C., & Cuetos, F. (1993) Parsing biases for non-structural differences: Evidence for the use of statistics in parsing.

Presented at the Conference on Psychology of Language and Communication, Glasgow, 31 Aug to 3rd Sept 3.

Crocker, M.W. (1992) A logical model of competence and performance in the human sentence processor. PhD University of Edinburgh. Published as Research Paper HCRC/RP-34. Human Communication Research Centre, Glasgow, UK.

CUETOS, F. & MITCHELL, D. (1988) Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30, 73-105.

De Vincenzi, M, & Job, R. (1993) Some observations on the universality of the Late-Closure strategy. *Journal of Psycholinguistic Research*, 22, 189-206.

FERREIRA, F. & CLIFTON, C. (1986) The independence of syntactic processing. *Journal of Memory and Language*. 25, 348-368.

Ferreira, F., & Henderson, J.M. (1991). Recovery from misanalysis of garden-path sentences. *Journal of Memory and Language*, 30, 725-745.

FLORES D'ARCAIS, G.B. (1990) Parsing principles and language comprehension during reading. In D. Balota, G.B. Flores d'Arcais & K. Rayner (Eds): *Comprehension processes in reading*. Hillsdale, LEA.

Frazier, L. (1979). On comprehending sentence: Syntactic parsing strategies. PhD Dissertation, University of Connecticut. Bloomington: Indiana University Linguistics Club.

FRAZIER, L. (1987) Sentence processing. In M. Coltheart (Ed): *Attention and Performance XII*. Hillsdale, LEA

FRAZIER, L. (1990a) Parsing modifiers: Special purpose strategies in the human sentence processing mechanism. In D. Balota, G.B. Flores d'Arcais & K. Rayner (Eds): *Comprehension processes in reading*. Hillsdale, LEA.

Frazier, L. (1990b). Exploring the architecture of the language system. In G. Altmann (Ed.) *Cognitive models of speech processing: Psycholinguistics and computational perspectives* (pp. 409-433). Cambridge, MA: MIT Press.

FRAZIER, L. & CLIFTON, C Jr. (in preparation) Construal. Manuscript, University of Massachusetts, Amherst, M.A.

FRAZIER, L., CLIFTON, C. & RANDALL, J. (1983) Filling gaps: Decision principles and structure in sentence comprehension. *Cognition*, 13, 187-222.

FRAZIER, L. & RAYNER, K. (1982) Making and correcting errors during sentence comprehension. *Cognitive Psychology*, 14, 178-210.

Frazier, L. and Rayner, K (1988) Parameterizing the language processing system: Left- versus right-branching within and across languages. In J. Hawkins (Ed.) *Explaining Language Universals*. pp 247-279. Oxford: Basil Blackwell.

Gibson, E. & Loomis, J (in press) A corpus analysis of recency preference and predicate proximity. To appear in the Proceedings of the 1994 Cognitive Science Society. Hillsdale, N.J.: Erlbaum.

Gibson, E. & Pearlmutter, B. (in press) A corpus-based analysis of psycholinguistic constraints on prepositional phrase attachment. In C. Clifton, L. Frazier & K. Rayner (Eds.) *Perspectives in Sentence Processing*. Hillsdale, N.J.: Erlbaum.

GIBSON, E., PEARLMUTTER, N., CANSECO-GONZALES, E. & Hickok, G. (submitted) Cross-linguistic attachment preferences: Evidence from English and Spanish.

Inoue and Fodor, J.D. (in press) Information-paced parsing of Japanese. To appear in R. Mazuka and N. Nagai (Eds.) Japanese Sentence Processing. Hillsdale, N.J.: Erlbaum.

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

MacDonald, M.C. (1994) Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9, XXX-XXX.

MacWHINNEY, B. (1987) The competition model. In B. MacWhinney (Ed): *Mechanisms of Language Acquisition*. Hillsdale, LEA.

Mazuka, R. and Lust, B. (1990) On parameter setting and parsing: Predictions for cross-linguistic differences in adult and child processing. In Frazier, L. and de Villiers, J. (Eds.) *Language Processing and Language Acquisition*, pp, 163-205, Kluwer Academic Publishers: The Netherlands.

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

Mitchell, D.C. (1987a) Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 601-681). Hillsdale, N.J.: Erlbaum.

MITCHELL, D.C. (1987b) Reading and syntactic analysis. In J. Beech & A. Colley (Eds): *Cognitive approaches to reading*, Chichester, Wiley.

Mitchell, D.C. (1988) Natural language processing - A psychological perspective. Institute of Electrical Engineers Computing and Control Division. Digest No. 1988/118.

Mitchell, D.C. (1994) In M. Gernsbacher (Ed.) *Handbook of Psycholinguistics*. Academic Press

MITCHELL, D.C. & CUETOS, F. (1991a) The origins of parsing strategies. In C. Smith (Ed): *Current issues in natural language processing*. Texas, University of Austin.

Mitchell, D.C. & Cuetos, F. (1991b) Restrictions on Late Closure: The computational underpinnings of parsing strategies in Spanish and English. Unpublished paper, University of Exeter.

MITCHELL, D.C. & CUETOS, F. (1993) Restrictions on Late Closure. The computational underpinnings of parsing strategies in Spanish and English. Unpublished manuscript.

Mitchell, D.C., Cuetos, F. & Brysbaert, M. (1994) Modifier attachment in Dutch and Spanish: Evidence favouring statistical rather than universal "linguistic" accounts. To be presented at the 4th Workshop of Language Comprehension, Giens, France, May 13-14.

Mitchell, D.C., Cuetos, F., & Corley, M.M.B. (1992) Statistical versus linguistic determinants of parsing bias: Cross-linguistic evidence. Paper presented at the Fifth Annual CUNY Conference on Human Sentence Processing CUNY, NY.

- MITCHELL, D.C., CUETOS, F. & ZAGAR, D. (1990) Reading in different languages: Is there a universal mechanism for parsing sentences? In D. Balota, G.B. Flores d'Arcais & K. Rayner (Eds): Comprehension processes in reading. Hillsdale, LEA.
- Pinker, S. (1994) The language instinct: The new science of language and mind. London: Allen Lane.
- Pritchett, B.L. (1992) Grammatical competence and parsing performance. Chicago: University of Chicago Press.
- Radford, A. (1990) *Syntactic theory and the acquisition of English syntax*. Oxford: Basil Blackwell.
- 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.
- SEIDENBERG, M.S. & McCLELLAND, J.L. (1989) A distributed developmental model of visual word recognition and naming. *Psychological Review*, 96, 523-568.
- MacDonald, M.C. (in press) Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*.
- St. John, M.F., & McClelland, J.L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46, 217-257.
- TABOSSI, P., SPIVEY-KNOWLTON, M.J., McRAE, K., & TANENHAUS, M.K. (in press) Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In *Attention and Performance*, XV.
- 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.
- Valle, F., Cuertos, F. Igoa, J.M. & del Viso, S. (1990) *Lecturas de Psicolinguística: 1. Comprehension y producción del lenguaje*. Madrid: Alianza Editorial.
- ZAGAR, D. & PYNTE, J. (1992) The role of semantic information and of attention in processing syntactic ambiguity: Eye-movement study. Paper presented to the 5th conference of the European Society for Cognitive Psychology, Paris, September 12-16.