Ambiguity, Parsing Strategies, and Computational Models

Gerry Altmann
Centre for Speech Technology Research, University of Edinburgh, Edinburgh, U.K.

A variety of computational models have been developed in recent years to model the behaviour of the Human Sentence Processing Mechanism (HSPM) when it encounters local syntactic ambiguities. The majority of these incorporate the assumption that the HSPM makes its initial choice of analysis according to a small number of exclusively syntactic principles. The arguments in favour of this structural approach range from the computational efficiency of parsers incorporating these structurally based parsing strategies to the empirical evidence which has been claimed to refute the alternative interactive account of ambiguity resolution, in which contextual information can be used in order to determine the initial choice of analysis. The present paper reviews some of these models, and argues that the assumptions concerning their underlying computational efficiency are flawed. Experimental evidence is presented which is suggestive of the alternative interactive account.

INTRODUCTION

The observation that the Human Sentence Processing Mechanism (HSPM) exhibits definite preferences when faced with local attachment ambiguities has led to the formulation of a variety of principles which have been claimed to govern the workings of the parsing mechanism (e.g. Altmann & Steedman, 1988; Crain & Steedman, 1985; Ford, Bresnan, & Kaplan, 1982; Frazier, 1978; Kimball, 1973; Wanner, 1980). The two examples that follow illustrate the two main families of preference.

(1) The girl picked the gladiolus that I had been trying to grow for my mother.

Requests for reprints should be addressed to G. Altmann, Centre for Speech Technology Research, University of Edinburgh, 80 South Bridge, Edinburgh EH1 1HN, U.K.

Portions of the work reported here were carried out on an S.E.R.C. postgraduate research studentship in the Centre for Cognitive Science at the University of Edinburgh. My thanks to Ellen Gurman Bard, Janet Fodor, Mark Steedman, Mike Tanenhaus, and two anonymous reviewers for their comments on earlier versions of this paper.

© 1988 Lawrence Erlbaum Associates Ltd. and V.S.P. Publications
(2) She tickled the man with the gladiolus.

In (1) there is a tendency to associate the PP for my mother as a modifier to the growing, rather than the picking, although both logically and linguistically either reading is plausible. In (2) there is a clear tendency to interpret the PP with a gladiolus as the means of the tickling, rather than as a description of the man, although again it might be either. The tendency in (1) to attach as low down and far right as possible (to borrow from notions of constituent structure), and hence to interpret for my mother as the indirect object of trying to grow, is captured in the apparently equivalent principles of “Right Association” (Kimball, 1973), “Late Closure” (Frazier, 1978), and “Final Arguments” (Ford et al., 1982), whereas the tendency in (2) to interpret the post-verbal prepositional phrase as a verb-modifier rather than as an NP-modifier is most succinctly captured by Frazier’s (1978) “Minimal Attachment” strategy in which the processor attempts to build the structure containing fewest nodes.

The existence of such preferences poses a question concerning their origins. Why might the HSPM evolve to exhibit just these tendencies (and not their opposites, say)? On the commonly held but as yet unjustified assumption that the HSPM is maximally efficient, the arguments have concentrated on the increased computational efficiency of a language processor incorporating these strategies. As pointed out by Wanner (1987), following arguments from Kimball (1973, 1975) and Frazier and Fodor (1978), Minimal Attachment and Right Association result in trees with the fewest non-terminal nodes (Minimal Attachment) each with the widest possible coverage of the input string (Right Association). The ensuing non-terminal-to-terminal node ratio minimises the computational steps to be taken by the parser (Miller & Chomsky, 1963).

Such arguments, though computationally sound, still fail to explain how the parser comes by these tendencies in the first place. If these tendencies are realised as a set of (external) rules governing the operation of the parser, then the question of how the parser evolved that particular rule set remains unanswered. Although descriptively adequate, such a model lacks explanatory power (cf. Chomsky, 1965). If, on the other hand, they are not realised explicitly in the model, but merely arise as a consequence of the

---

1See Kimball (1975) for a less general version of this strategy.

2In the case of a traditionally right-branching language such as English, attachment low down and to the right (i.e. Right Association) corresponds to attachment to the current constituent being parsed—hence the equivalence, in English at least, between Right Association and Late Closure. Right Association guarantees the widest possible coverage of an input string because of this equivalence: Late Closure by definition means wide coverage.
workings of that parser, that is, as a consequence of the parser's architecture, then a certain, albeit limited, degree of explanatory adequacy is achieved. This distinction between explicit and implicit realisation will become clearer during the following section when we consider a number of attempts to demonstrate a dependence between the HSPM's architecture and the existence of strategies like Minimal Attachment and Right Association. Underlying these attempts will be the assumption that parsing is essentially a modular process with no interaction with semantic or discourse processing. We shall see in later sections that the notion of computational efficiency, and the criteria by which it is judged, change quite dramatically if this assumption is violated.

**SYNTACTIC PARSING STRATEGIES AND THE ARCHITECTURE OF THE HSPM**

The Sausage Machine

The Sausage Machine (henceforth "SM") (Fodor & Frazier, 1980; Frazier & Fodor, 1978) is a 2-stage parser consisting of a Preliminary Phrase Packager (PPP) which assigns lexical and phrasal nodes to the words it has access to in the sentence, and a Sentence Structure Supervisor (SSS) which takes the packages produced by the PPP and combines them to form a complete phrase marker. Right Association arises because the PPP is given only limited access to the sentence; it views the sentence through a buffer, or window, of about six words, and a lexical node cannot be attached to any item on its left which may have passed out of the PPP's window. Minimal Attachment arises as a consequence of a race (Fodor & Frazier, 1980): The first combination of rules to incorporate the current lexical item into the phrase marker wins; the fewer the nodes, the sooner the item is incorporated. Although the two strategies are accounted for in different ways, each is claimed to arise as a natural consequence of some aspect of the parser's architecture.

Wanner (1980) pointed out, however, that the SM cannot account for the preference for right association in sentence (3) below, all of which should fit within the limited view window of the PPP:

(3) He said she tickled him yesterday.

This observation led Fodor and Frazier (1980) to concede that the tendency for Right Association does not arise directly from the model's architecture. The two strategies are thus features of their model which are independent of the 2-stage architecture of the model.
Arc-ordering in an Augmented Transition Network

Frazier and Fodor (1978) had argued that if one were to capture the principles of Minimal Attachment and Right Association in an ATN\(^3\) then the ordering of the arcs leaving a state (i.e. the ordering of the rules in the grammar) would have to be independently specified for each state in each network. Wanner (1980) showed, however, that there is in fact a more principled account of arc-ordering, and hence parsing strategies, in an ATN.

In an ATN, there are a number of different arcs, each of which has a particular function. There are WORD arcs which specify particular words which must appear in the input in order to traverse that arc; there are CAT arcs which are like the WORD arcs except that they specify a category of word which must be found; there are SEEK arcs which specify a subnetwork which must be traversed; there are SEND arcs which terminate a network and in effect return the ATN processor from a subnetwork to the network which called this subnetwork; and there are JUMP arcs which enable the processor to jump from one state to the next without requiring anything in the input (hence allowing for the optionality of certain constituents).

To implement Minimal Attachment, Wanner simply schedules all WORD arcs and CAT arcs before all SEEK arcs. This forces the processor to explore all the possibilities available within the one network before attempting to explore those in any subnetwork—the equivalent of attaching incoming material to the phrase marker using the fewest possible intervening nodes. To implement Right Association, Wanner schedules all SEND and JUMP arcs (which either do, or can in principle, terminate the current constituent) after all others. This ensures that the parser incorporates a new item into the current constituent being parsed (i.e. low down and to the right).

Minimal Attachment and Right Association thus arise as a consequence of the same mechanism: arc-ordering. Although the serial ATN model has been criticised on a number of grounds (see Fodor & Frazier, 1980), the immediate concern here is the relationship between the two parsing strategies and the ATN’s architecture. Although Wanner claims that this relationship is principled, the ordering of arcs is really just an arbitrary specification on the workings of the processor.

Conflict Resolution in a Shift-Reduce Parser

Shift-reduce parsers (Aho & Johnson, 1974) can be thought of as consisting of two basic components: a push-down stack and a grammar. Accompany-

---

\(^3\)See Wanner and Maratsos (1978) for a description of parsing within an ATN framework.
ing these are two basic operations: SHIFT, in which the next incoming lexical item is placed on the top of the stack (displacing any other items downwards); and REDUCE, in which consecutive elements on the stack (obligatorily including the top-most) can be reduced according to the rules of the grammar to just a single element (e.g. from \(<v>[\text{det}]\text{[n]}\>\) to \(<v>[\text{NP}]\>\) to \(<[\text{VP}]\>\). The grammar is represented as a "look-up" table, which is used to decide when to SHIFT and when to REDUCE.

There are several sources of ambiguity, or conflict, which can occur within such processors. Two of them concern us here. First, at any stage in the processing the processor might be in a position to either SHIFT or REDUCE. Secondly, more than one reduction might be possible, corresponding to the local applicability of more than one rule in the grammar. Minimal Attachment and Right Association can be modelled according to the following algorithm for resolving these conflicts:

**For Right Association:** resolve all SHIFT-REDUCE conflicts in favour of SHIFT.

**For Minimal Attachment:** resolve all REDUCE-REDUCE conflicts in favour of that which reduces most cells in the stack.

Given that reduction effectively closes the constituent being parsed, it follows that if one were to resolve SHIFT-REDUCE conflicts in favour of REDUCE then the next item to be shifted onto the stack could not associate with items within that constituent. In other words, right association would not be possible. Only by keeping the current constituent open as long as possible, which is ensured by resolving SHIFT-REDUCE in favour of SHIFT, are right associations ensured.

Resolving REDUCE-REDUCE conflicts in favour of the reduction which consumes most cells in the stack ensures the greatest number of daughter constituents to the mother constituent formed by each such reduction. This results in the flattest possible structures and, consequently, the minimum number of intervening nodes between any terminal node and any dominating non-terminal node. Hence Minimal Attachment.

It is claimed (cf. Wanner, 1987) that it is the conjunction of Minimal Attachment and Right Association which gives rise to the overall efficiency

---

4A third source of ambiguity occurs when an incoming item has more than one category, e.g. "that" as \(<\text{det}>\) or \(<\text{comp}>\).

5With an unambiguous grammar, there are no such conflicts: In the general case, SHIFTs take place once a constituent has been formed (by reduction). Ambiguity in the grammar permits a single lexical item to form part of either one constituent or another, in which case either SHIFT or REDUCE is permitted by the look-up table.

6The algorithm in Shieber (1983) for handling Minimal Attachment and Right Association is due to work by Pereira which subsequently appeared in Pereira (1985).
of the processor, and it is perhaps surprising, therefore, that they are realised independently of one another in the model. Of course, any expectations (however valid) concerning their interdependence rest on the so far unchallenged assumption that a processor incorporating both Minimal Attachment and Right Association is indeed maximally efficient. Before challenging this view, and the modular assumptions concerning the relationship between syntax and higher-level discoursal processes, we shall consider one more computational account which, unlike its predecessors, succeeds both in its implicit realisation of the two strategies and in the interdependence of the two within the model.

**Limited Parallelism in an ATN**

In a more recent proposal concerning ATNs, Wanner (1987) demonstrates that the arbitrariness of the earlier model can be eliminated by assuming a processor which exhibits what he calls *limited parallelism*, in which all the arcs leaving a particular state are explored simultaneously. Processing proceeds in this parallel fashion until an analysis path successfully incorporates the current lexical item into the parse tree. At this point all other analysis paths are abandoned, and parallelism resumes again at the state which was reached in the successful analysis path.

Given this limited parallelism, the analysis path which involves making fewest transitions will always win the race to attach the current word into the phrase marker. Such a model will always induce minimal attachments simply because nonminimal attachments always involve an extra SEEK arc (as well as an extra SEND arc subsequently). Moreover, it will always induce right associations because anything other than a Right Association can occur only if the current constituent is first closed, which requires an extra SEND arc.

Unlike its predecessors, this model accounts for both Minimal Attachment and Right Association without postulating separate mechanisms for each—they arise as a natural consequence of the limited parallelism, and not because of any *ad hoc* mechanisms. One shortcoming, however, is that it predicts that Minimal Attachment will always take precedence over Right Association in the event of a conflict—encountering an extra SEEK arc necessarily occurs prior to encountering any extra SEND arcs. The model, therefore, fails to account for the preference in (4) below (cf. Fodor & Frazier, 1980), in which Right Association takes precedence over Minimal Attachment.

(4) The girl picked the gladiolus that I had planned to include in a bouquet for my mother.

The same is true of the earlier arc-ordering model (Fodor & Frazier,
1980). Such criticism notwithstanding, it would be tempting to interpret the elegance of the model as justification for Minimal Attachment and Right Association as predictors of human linguistic behaviour. However, a further shortcoming with the model is that it assumes that the analysis path with fewest transitions will always be traversed faster than any other analysis path. It is by no means obvious, however, why this should necessarily be the case. Associated with each arc in the network will be an action, or possibly several actions, each of which must be executed successfully in order for that arc to be traversed. There is no a priori reason to suppose that all actions will take the same time to execute, just as there is no reason to suppose that an arc with several associated actions will take the same time to traverse as one involving just one action (or even none). Without such an assumption, however, the model can no longer explain the preference for minimal attachments and right associations.

In the following section, we shall consider an alternative view of syntactic ambiguity resolution, in which higher-level pragmatic information has a role in guiding the resolution process.

SYNTACTIC STRUCTURE AND INTERPRETATION

Pereira (1985) pointed out that both his own and Wanner's earlier (1980) ATN model rely on the specific rules of grammar employed: If instead of the recursive NP-modification rule in (5b) below, the grammar were to contain the rule in (6b), then neither model would make the correct predictions:

(5) a. NP $\rightarrow$ det n  
b. NP $\rightarrow$ NP PP
(6) a. NP $\rightarrow$ det n  
b. NP $\rightarrow$ det n PP

Both models would predict a preference for NP-modification (that is, for interpreting "with a gladiolus" in "she tickled the man with a gladiolus" as an NP-modifier rather than as a verb-modifier): Wanner's arc-ordering model would SEEK-PP rather than SEND after encountering "the" and "man", and Pereira and Shieber's shift-reduce algorithm would SHIFT "with" onto the stack rather than REDUCE "the" and "man" to create an NP. Similarly, for Wanner's more recent model, (6b) would be the winning path because it does not require a SEND arc prior to SEEK-PP. In fact,

The shift-reduce parser does not suffer this drawback—rather than REDUCE "in" and "a bouquet", the parser SHIFTS "for" onto the stack and, subsequently, REDUCEs "a bouquet" and "for my mother".
any model which incorporated Minimal Attachment, whether explicitly or implicitly, would suffer in this same way: Both attachments, to the NP or to the VP, would involve the same number of nonterminal nodes and would not therefore be discriminated between on the basis of a Minimal Attachment principle.

This observation, concerning sensitivity to the grammar, threatens Minimal Attachment itself as a principle which correctly describes the operation of the HSPM. Under the Strong Competence Hypothesis of Bresnan and Kaplan (Bresnan, 1982), there is a one-to-one correspondence between the rules of grammar and the operations which the processor goes through as it builds any one interpretation of the utterance. It follows that the difference between the alternative structures postulated for complex NPs in (5) and (6) above is accompanied by a difference in the way in which the interpretation of NPs proceeds. In (5b) the head NP of the complex is itself a complete constituent, and hence interpretable. Assuming a direct correspondence between the rules of grammar and the steps that the processor goes through as it builds an interpretation, there is no such interpretable constituent in (6b): There is no combination of rules which allows the sequence <det n> to be parsed into a complete (and interpretable) constituent and yet also be followed by PP. Given its reliance on rules such as (5b), it follows that although Minimal Attachment may correctly describe the behaviour of the HSPM, the explanation for this behaviour may reside elsewhere: What appears to be governed by a purely structural process may instead be governed by some higher-level interpretive process. If this is in fact the case, then any purely structurally based notions of computational efficiency must be reconsidered in light of the computational demands which may be placed on the parser by the interpretive system.

Crain (1980; see also Crain & Steedman, 1985) pointed out that the structural differences between simple and complex NPs are accompanied by differences in presuppositional complexity. The least preferred readings for sentences traditionally associated with the Minimal Attachment strategy involve a complex NP construction containing an NP-modifier whose function is to restrict over a set of entities of the type denoted by the head noun. The complex NP thus presupposes that there exists a set of potential referents and that the proposition expressed by the modifier restricts this set of entities to a single individual. The simple NP analysis presupposes neither a set of entities nor the prior existence of propositional information associated with any member of such a set. In this sense, the preferred readings for the Minimal Attachment family of ambiguities are presuppositionally less complex than the alternative complex NP analyses. Crain and Steedman (1985) suggest that in the absence of explicit referential context (the "null" context), that is, when the presuppositions associ-
ated with both the simple and complex NP analyses are violated, then the "minimally attached" interpretation is selected according to a Principle of Parsimony: the least presuppositionally complex analysis is favoured over all others.

It follows that the preferences which Frazier attributes to Minimal Attachment should be overcome by preceding the crucial sentences with contexts with particular referential characteristics. Crain (1980, reported in Crain & Steedman, 1985) showed exactly this effect. He preceded relative/complement sentence pairs such as (7) with the contexts in (8) and showed that "garden path" effects could be overcome or induced depending on which context preceded each sentence:

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

(8) a. Complement-supporting context
A psychologist was counselling a married couple. One member of 
the pair was fighting with him but the other one was nice to him.
b. Relative-supporting context
A psychologist was counselling two married couples. One of the 
couples was fighting with him but the other one was nice to him.

Minimal Attachment would predict that (7b) should always induce a 
garden path because the NP, the wife, will be interpreted as a simple NP 
(i.e. as the direct object of "told") and not as the head of a complex NP 
(i.e. as the subject of the embedded relative clause). In order to test this, 
Crain used an incremental grammaticality judgement task (cf. Frazier, 
1978), in which subjects were instructed to respond as soon as they first 
thought that the target sentence (presented one item at a time to a screen) 
ceased to be grammatical. He found, however, that the predictions based 
only on Minimal Attachment did not hold when (7b) was preceded by (8b), 
which supports the presuppositions of the relative clause: the majority of 
responses indicated that subjects did not experience any garden path. 
Furthermore, when (7a) was preceded by (8b) a garden path was induced, 
even though Minimal Attachment would predict no such effect because of 
the preference for the simple NP analysis for the wife.

Crain's results call into question the claims of Frazier (1978) and Rayner, 
Carlson, and Frazier (1983) who proposed that contextual information 
could not override the preferences predicted by Minimal Attachment until 
the initial syntactic analysis was complete. They claimed that the initial 
analysis to be pursued is selected on the purely structural criterion of 
Minimal Attachment, and that the alternative analysis is attempted only if
the first is found to be inconsistent with the context. Ferreira and Clifton (1986) argued that Crain's results do not threaten this structural position because the incremental grammaticality judgement task he used is sensitive only to the final analysis selected by the processor, and does not reflect the initial actions of the HSPM. They go on to provide evidence which, they claim, demonstrates that referential contexts of the kind used by Crain do not determine the initial analysis pursued by the HSPM.

One problem with Ferreira and Clifton's own study, however, was that in order to assess whether or not there was any effect of referential context they preceded their target sentences with either biasing or neutral contexts. If the decision about whether to interpret a noun phrase as either simple or complex is indeed made on referential grounds (that is, on whether a referent can or cannot be found), then it follows that neutral contexts should not be constructable. Ferreira and Clifton attempted to get around this problem by first introducing into the context two entities to which a target definite NP could refer, but then foregrounding (cf. Chafe, 1972) one of them so that the definite NP would none the less be able to refer. Altmann and Steedman (1988) point out that as a consequence of this foregrounding, the contexts were not neutral at all, but supported the minimal attachment analysis. It is therefore not surprising that overall, no effects of context were found. Moreover, it is unclear whether they were completely justified in asserting that the differences in eye movements they obtained were entirely due to syntactic reanalysis effects. We return to this point in the next section.

Despite any criticism which attaches to Ferreira and Clifton's experiment, Crain's results are still open to a number of criticisms, including that concerning the insensitivity of the incremental grammaticality judgement task to on-line effects. In addition, too few stimuli were used to make the results statistically generalisable, and no attempt was made to factor out the relative contributions of target structure and context type to the observed effects. To be able to establish these relative contributions is crucial if one is to admit the possibility of some residual effect of structure (however caused) once context has been controlled for.

But aside from Crain's own results, there are other considerations which argue against a purely structural principle such as Minimal Attachment (see also Altmann & Steedman, 1988). The assumption underlying the interactive accounts proposed by Crain and Steedman (1985) and Altmann and Steedman (1988) is that the relative clause interpretation of the that-clause, or the NP-attachment of the prepositional phrase, can be chosen as soon as the referring expression which these constructions could modify is found to be referentially unsuccessful (the principle of Referential Failure: see Altmann, 1986; Altmann & Steedman, 1988). The implication here is that referential success or failure can be determined as
soon as the target NP is encountered. For this to happen, the NP must be immediately available to semantics for reference evaluation (like Minimal Attachment, this relies on rules of the form expressed by (5) and not (6) above). Minimal Attachment would be well suited to a situation in which such information was not immediately available, in which case the processor would have to make a decision on some other basis, such as structure. But there is sufficient evidence in support of the real-time nature of sentence processing (e.g. Marslen-Wilson, 1975; 1986; Marslen-Wilson & Tyler, 1980; Tanenhaus, Carlson, & Seidenberg, 1985; Tyler & Marslen-Wilson, 1977) as well as the immediacy of nounphrase evaluation (e.g. Garrod & Sanford, 1985; Shillcock, 1982) to suggest that referential information is available to the processor in close to real-time; it would be a strange processor indeed which favoured information which would not necessarily lead to the correct choice of analysis over information which would.

RELATIVE CLAUSES, REFERENTIAL CONTEXT, AND INFERENCING

Crain pointed out that the complement clause (7a) and relative clause (7b) sentences differ not only in terms of their structure but also in terms of the presuppositions (or felicity conditions) which are implicated by the use of the two kinds of clause. But they also differ with respect to the kinds of processing which these presuppositions demand of them. The propositional content of the complement clause represents new information; it is information which is to be added to the hearer’s model of the discourse. The propositional content of the restrictive relative, however, represents information which must already be contained within that model (hence its restrictive function); it represents given information (Clark & Haviland, 1977; Halliday, 1967; and others). An attempt must therefore be made to match the propositional content of the relative with information already contained within the discourse model. Relatives and complements thus demand very different kinds of process. Moreover, the retrieval of given information may make different demands on processing time depending on the ease with which the proposition contained within the relative can be inferred to correspond to some internal proposition within the discourse model (cf. Haviland & Clark, 1974; Sanford & Garrod, 1981). It follows that there may be a difference in processing time between relative clauses and their complement counterparts simply because of this process of inference, and even when each is preceded by felicitous contexts of the kinds proposed by Crain.

The preceding discussion suggests that increases in processing load which have hitherto been explained in terms of syntactically induced
garden paths may in fact be due to the intrinsic differences in processing load which are associated with different syntactic analyses. For instance, in the reduced-relative construction used by Ferreira and Clifton (1986):

(9) The editor played the tape agreed the story was big.

A certain amount of work must be done to infer that the antecedent in the biasing context to the italicised noun phrase is in fact the editor mentioned in

(10) He ran a tape for one of his editors.

Ferreira and Clifton dismiss this as an explanation for the increased fixations to the disambiguating regions of their materials, claiming that any increases in fixation time due to inferencing would be localised on the critical noun phrase (italicised in 9, above) and, consequently the increased fixations they found subsequent to this phrase could only be due to syntactic reanalysis. This claim is made despite Ehrlich and Rayner's (1983) observation that the complex processes involved in integration do not necessarily manifest themselves as increased fixation times on the actual fixation that initiated the process. There is thus some uncertainty concerning what it is that eye-movements actually reflect, and without further data on the exact localisation of the integrative effects associated with referential failure and the complex processing of given information, it appears that at present, at least, eye-movement data do not provide the resolution required to establish just where in the processing sequence parsing decisions are actually made.

In the experiment that follows, global reading times are used in order to assess the contribution of inferencing to processing times for materials similar to those used by Crain. The intention is to show, among other things, that the inferencing-load associated with the processing of given information does lead to a residual difference in processing time between relative and complement clause constructions in felicitous contexts.

According to the referential hypothesis, referentially felicitous contexts should induce shorter times than infelicitous contexts, which will induce garden paths. According to the structural hypothesis, there might be an effect of context on reading time, but this would only reflect the referential anomaly of the relative clause in the complement-supporting context and the complement clause in the relative-supporting context; it would not reflect a garden path in the complement clause sentence. Moreover, given that the relative clause analysis will only be attempted after an initial

---

8The experiment is from Altmann (1986) and was reported in reduced form in Altmann (1985; 1987).
attempt at the complement analysis (as determined by Minimal Attachment), it follows that in principle there should always be a residual difference between relative clause sentences and their complement counterparts; the former should induce longer reading times.

Materials were constructed so that different degrees of inferencing were required. It was predicted that the contextual condition which required least inferencing would decrease or even eliminate any residual differences between the processing times for relatives and complements. In addition, it was predicted that this decrease would affect only the relative clause targets. This is because in the case of the complement clause targets, syntactic information intervenes to prevent the construction of a complete relative clause (the materials are only locally ambiguous). The existence of such context-dependent inference-load effects would lessen the force of the structuralist argument which would account for any difference between relatives and complements purely in terms of a garden-path effect.

According to Frazier (1978), Rayner et al. (1983), and Ferreira and Clifton (1986), relative clause targets will always (even in a relative-supporting context) require reanalysis after the minimally attached complement clause analysis has first been attempted. In the absence of data to confirm such a residual difference between the two kinds of analysis, the Minimal Attachment principle can no longer be accepted as the only adequate description of human linguistic behaviour.

Method

A total of 32 examples were constructed as follows, and the full set appears in the Appendix.

(11) Complement target sentence
The psychologist told the woman that he was worried about her marital problems.

Relative target sentence
The psychologist told the woman that he was worried about to visit him again.

The italicisation is for expository purposes only and was not included in the experimental presentations. Each target sentence was presented on a

---

If the manipulation of inferencing were somehow placed on the head noun, then the interactive hypothesis predicts (and indeed any incremental model would predict) that the manipulation would affect both target types equally, because in both cases the interpretation of the head noun must be checked against the discourse model before proceeding with the analysis.
single line, and not on two lines as shown here. The targets were preceded either by the contexts in (12), the "minimal inferencing conditions", or by the contexts in (13), the "inferencing conditions".

(12) **Complement-supporting context: "minimal inferencing"**

A psychologist was counselling a man and his wife.

*He was worried about* one of the pair
but wasn't concerned about the other.

**Relative-supporting context: "minimal inferencing"**

A psychologist was counselling two wives.

*He was worried about* one of the pair
but wasn't concerned about the other.

In this condition, the content of the relative clause is expressed using the same wording as was used to express the equivalent proposition in the context (against which the proposition expressed by the relative will be matched). In the following conditions different wordings are used, in an attempt to increase the degree of inferencing required during the matching process.

(13) **Complement-supporting context: "inferencing"**

A psychologist was counselling a man and his wife.

*He was particularly concerned with* the problems of
one of the pair but wasn’t concerned about the other.

**Relative-supporting context: "inferencing"**

A psychologist was counselling two wives.

*He was particularly concerned with* the problems of
one of the pair but wasn’t concerned about the other.

Target sentence-pairs were matched as closely as possible for length. The design was a factorial repeated measures design incorporating a Latin Square. There were eight versions of each example (2 Inference × 2 Context × 2 Target). Eight sets of stimuli were constructed, with only one version of each of the 32 examples appearing in any one set. Each stimulus set was presented to a different group of subjects. No subject was exposed to more than one version of the same example, each subject was exposed to all conditions, and each example was represented equally in each one of its eight versions. The 32 test items were embedded in 48 fillers which bore varying degrees of resemblance to the test items.

Materials were presented by computer, one sentence at a time. Where the length of a sentence exceeded 80 characters (the maximum permitted on one line of the screen) the sentence was presented on two lines. None of the target sentences exceeded 80 characters.
Subjects were instructed to press a response key as soon as they had read and understood the sentence presented on the screen. This caused the next sentence in the sequence to appear. At the end of each sequence, subjects were presented with a simple yes/no question in order to test and encourage comprehension. The following is a typical example:

(14) Was the psychologist counselling two women?

A total of 56 members of the University of Edinburgh student community were paid to participate in the experiment. All were native speakers of English and na"ive as to the purpose of the experiment.

Results

Reaction times to target sentences which fell more than 1.96 standard deviations (S.D.) away from the mean were replaced with the critical cut-off value. Less than 4.7% of the data were affected by this procedure. The results are summarised in Table 1 and Fig. 1.

Three-way ANOVAs, by subjects and by materials, were performed on the data. For the purposes of these analyses, the context dimension was transformed into a dimension of “Felicity”: For the relative clause targets, the relative-supporting context (more than one referent) is felicitous and the complement-supporting (only one referent) infelicitous; for the complement clause targets, the complement-supporting context (only one referent) is felicitous and the relative-supporting context (more than one referent) infelicitous. A main effect of Felicity thus corresponds to what would otherwise have been an interaction between referential context (as defined by the number of candidate referents introduced in the context) and Target. Also, what would have been a main effect of context now becomes an interaction between Felicity and Target.

| TABLE 1 |
| Reading Time per Sentence (msec) |

<table>
<thead>
<tr>
<th>Context</th>
<th>Felicitous</th>
<th>Infelicitous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement target</td>
<td>3116</td>
<td>3481</td>
</tr>
<tr>
<td>Relative target</td>
<td>3501</td>
<td>4051</td>
</tr>
<tr>
<td>Minimal inferencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement target</td>
<td>3190</td>
<td>3657</td>
</tr>
<tr>
<td>Relative target</td>
<td>3221</td>
<td>3812</td>
</tr>
</tbody>
</table>
Overall, a strong main effect of Felicity was found—min $F'(1,45) = 12.16, P < 0.005$. A weak effect of Target was found—$F_1(1,48) = 14.20, P = 0.0004; F_2(1,31) = 2.19, P = 0.1486$—but there was no overall main effect of Inference—$F_1(1,48) = 1.65, P = 0.2055; F_2 < 1$. The weak effect of Target is due to the expected interaction between Inference and Target—min $F'(1,59) = 5.69, P < 0.025$. There was no interaction between Felicity and Target, nor were there any other interactions.

Additional information is given by a series of Multiple Scheffé tests:

1. There was no effect of Inference on the complement targets. The data points representing the minimal-inferencing conditions were not significantly different from those representing the inferencing conditions, whether in the felicitous or infelicitous contexts.
2. There was a strong effect of Inference on relative clause targets preceded by felicitous contexts. The minimal-inferencing condition was significantly lower than the inferencing condition ($P < 0.05$). The effect failed to reach significance when the relative clause targets were preceded by infelicitous contexts: Here, the minimal-inferencing condition was not significantly lower than the inferencing condition ($P > 0.1$). Any effect may have been masked by the increase in latencies due to the garden-path effect, whether induced contextually or structurally.

Taken together, these results indicate that in felicitous conditions there was an effect of Inference only on relative targets, and not on complement targets.\(^{10}\)

3. In the minimal-inferencing condition, there were no differences between the two targets in either the felicitous conditions or the infelicitous conditions. The two targets in the felicitous conditions were both significantly faster than the targets in the infelicitous conditions ($P > 0.005$).

Finally, all pairs of points which represent conditions differing only along the Context/Felicity dimension (i.e. all joined pairs of points in Fig. 1) were significantly different from one another ($P < 0.005$).

Discussion

According to the structural position, there should be some residual difference between relative and complement targets: The former will take longer to process simply because Minimal Attachment dictates that the complement analysis be attempted first. In the minimal-inferencing conditions, however, there was no residual difference between relative and complement targets. It must be assumed, under the referential hypothesis, that the matching of given information against the discourse model in this condition was sufficiently fast not to show up in the global reading time measure used here.

The referential hypothesis would account for the effects of context in

\(^{10}\)The same data were separated by Target and separate ANOVAs performed. For complement targets, no main effect of Inference was found—$F_1(1,48) = 3.59$, $P = 0.0643$; $F_2(1,31) = 3.43$, $P = 0.0736$—though again there was a strong effect of Felicity—min $F(1,51) = 6.48$, $P < 0.025$. For relative targets, a main effect of Inference was found—$F_1(1,48) = 12.56$, $P = 0.0009$; $F_2(1,31) = 4.31$, $P = 0.0463$; min $F(1,52) = 3.21$—just failed to reach significance. Again, there was a strong effect of Felicity—min $F(1,50) = 11.58$, $P < 0.005$. 

LCP 3/2-B
terms of garden-path effects induced in the infelicitous conditions. Under a principle of Referential Failure (Altmann, 1986; Altmann & Steedman, 1988), the simple NP interpretation of the woman would be rejected in the relative-supporting context as soon as it was discovered that no unique referent existed for this expression. This would lead to the correct interpretation of the relative target sentence, but to the globally inconsistent analysis (and hence a garden-path) for the complement target sentence. In the complement-supporting context, Crain and Steedman's (1985) Principle of Parsimony predicts that the complex NP interpretation of the woman would be rejected as soon as it was discovered that a unique referent could be found for this expression. This would lead to the correct interpretation of the complement target sentence, and to the globally inconsistent (garden-path inducing) interpretation of the relative target sentence. The residual difference between relatives and complements which was found only in the inferencing conditions would merely reflect the time taken up by the matching and inferencing processes. As originally predicted, the change in inferencing only affected the relative targets.

Had the minimal-inferencing condition not been included, it would have been possible to account for the results without compromising Minimal Attachment; the initial analysis pursued would be determined on the basis of Minimal Attachment alone (however realised), hence the main effect of target structure, and the effect of referential context would reflect the contextual anomaly of the final analyses found. Consequently, it would not have been possible to distinguish between the structural and referential accounts.11

The effect of inferencing, although not compromising to the referential hypothesis, does compromise the assumption underlying Wanner's (1987) ATN model that processing time is independent of the actions associated

---

11It could be objected that even the minimal-inferencing conditions do not distinguish between the referential and structural positions: The data could be accounted for by supposing that the complement clause analysis was initially chosen on the basis of Minimal Attachment, but that this was then switched to the relative clause analysis once the context was found to be relative-supporting. The processor would then garden-path if the sentence ending conflicted with this choice (as would happen with the complement clause target sentence in the relative-supporting context). It would then have to be argued that the brief existence of the complement clause analysis, and the contextually induced switching over to the relative clause analysis (on the basis, presumably, of referential failure), does not contribute to reading times sufficiently for this to be detected with the global reading time measure used here.

This is certainly a valid objection, although it does lead to the following question: If Minimal Attachment does govern the initial choice of analysis, and if context can (and indeed does) override this choice before that analysis gets very far, then why have a strategy like Minimal Attachment in the first place? And if the emphasis is on explanation, rather than on description, how (and why) would a processor acquire such an apparently redundant strategy?
with each analysis path. This in turn compromises the ATN model's (albeit limited) ability to account for the claimed preferences for minimal attachments. In light of the lack of any residual differences between relative and complement targets, the present results suggest, however, that Minimal Attachment need not be the guiding principle during the resolution of local syntactic ambiguities of the kind studied here.\(^{12}\)

In summary, the results suggest that much of the variance that has previously been attributed to syntactic preferences may in fact be due to discourse-level variables, either in terms of their guiding the parsing process itself, or in terms of the differential processing costs associated with the alternative syntactic analyses.

### CONCLUSIONS

The departure from a purely structural approach to syntactic ambiguity resolution towards a more interactive approach suggests that the emphasis in computational models must shift towards the actual processes associated with each analysis path and away from the purely structural properties associated with each path. Right Association is no less compromised by such a move than Minimal Attachment. Within the descriptive framework of the ATN, Minimal Attachment is associated with the decision of whether or not to pursue a SEEK arc. The interactive hypothesis explored here provides a more principled account than Minimal Attachment on the basis on which this decision is made—namely that the decision is made on the basis of whether or not the target nounphrase successfully refers. On the other hand, Right Association is associated with whether or not to pursue a SEND arc; that is, whether or not to close, and hence complete the interpretation of, a constituent. Thus, despite the earlier, purely structurally based motivation for Minimal Attachment and Right Association arising as a consequence of the same architectural feature, the two classes of ambiguity are in fact associated with very different kinds of decision process.\(^{13}\) As such, there is no reason why they should be

---

\(^{12}\) The difference between the present findings and those of Ferreira and Clifton (1986), who found no effects of referential context (but see above), may in part be due to differences in the syntactic structures that were employed in the target sentences. In the present experiment, the crucial simple/complex nounphrase occurs in object position, whereas in Ferreira and Clifton's experiment it occurred in subject position (see example 9, above). Perhaps these differ in the ease with which their referents can be found. However, any predictions made on this basis would presumably favour the nounphrase in subject position, since the subject is probably in focus, and consequently its referential success or failure should be more easily determined.

\(^{13}\) This is irrespective of whether the decision process is explicit, as in the shift-reduce parsers of Shieber (1983) and Pereira (1985), or implicit, as in Wanner's model.
inextricably linked, as they are in Wanner’s parallel ATN model, once they are considered at a more pragmatically oriented level of description. The explanation underlying Right Association, and its computational justification, remains an open question, however, and is not addressed by the present results.

The initial arguments concerning the increased computational efficiency of parsing minimally attached structures still stand. As described earlier, Crain and Steedman’s (1985) Principle of Parsimony favours the analysis which leaves unsatisfied (and hence violates) fewest presuppositions, and on p. 80 it was shown that the “minimally attached” structure (that is, the simple NP analysis) does in the null context violate fewest presuppositions. In the case where the presuppositions underlying both analyses are violated, it would be clearly inefficient, in terms of the modifications to the discourse model which would need to be made in order to accommodate the unsatisfied presuppositions, to select the analysis which violated most presuppositions. However, this is not to say that as a general strategy, simple NPs should be preferred to complex NPs irrespective of the context. Given that their respective presuppositions are satisfied, it may be the case that the simple NP is indeed the least complex to interpret. But given that only the presuppositions underlying the complex NP might be satisfied, it does not follow that a processor which always favoured the simple NP when resolving local ambiguities of this kind would be the most computationally efficient.

Computational efficiency is a notion which is domain-specific; a processor which is efficient given one domain may be quite inefficient given another. The computational arguments in favour of models of the HSPM which incorporate Minimal Attachment and Right Association are similarly domain-dependent. The arguments rest on the assumption that certain tree structures are computationally more tractable than others. It has not been the intention here to challenge this claim. The kinds of computation involved during natural language processing, taken as a whole, are very different to those whose efficiency, or otherwise, is addressed by such arguments. When considering the resolution of local syntactic ambiguity, computational efficiency should be assessed not on the basis of the computational tractability of the parser’s initial choice of analysis, but on the likelihood of the chosen analysis corresponding to what the speaker originally intended.

Manuscript received October 1987
Revised manuscript received May 1988
REFERENCES

APPENDIX

The following are the materials used in the reported experiment. The parenthesised expressions in the contexts give the alternative phrasing required to convert from relative-supporting to complement-supporting, and from minimal-inferencing to inferencing. The parenthesised expression in the target sentence gives the alternative phrasing (to appear immediately after the post-verbal preposition) required to convert from the complement clause target to the relative clause target.

A psychologist was counselling two wives (a man and his wife). He was worried about (was particularly concerned with the problems of) one of the pair but wasn't concerned about the other.

The psychologist told the woman that he was worried about her marital problems (to see him again).
An electrician was explaining why a fuse had burnt out to two girls (a boy and a girl) living in a flat together. The electrician was pleasantly surprised by (couldn’t believe) how much one of them knew about electrical systems but was also disappointed by how little the other understood.

The electrician told the girl that he was surprised by the amount she knew (to help her flatmate).

An antique dealer was trying to buy a painting belonging to two elderly sisters (an elderly brother and sister). He was arguing (couldn’t come to any agreement) over the price with one of them even though he’d already agreed a price with the other.

The dealer told the woman that he was arguing with them for no good reason (to be reasonable).

A teacher was talking to the fathers (parents) of two children in his class. He was shocked (found it hard to believe) that one of them lacked all interest in education but was cheered up by the other’s positive attitude.

The teacher told the man that he was shocked by his negative attitude (to change his attitude).

A musician was talking to two guitarists (a guitarist and a drummer) about a group he wanted to form. He was very impressed by (liked the music written by) one of them but didn’t think much of the other.

The musician told the guitarist that he was impressed by his skilled playing (to meet him again).

A detective suspected two women (a man and a woman) of having committed a murder. He had his doubts about (was suspicious of) the alibi given by one of them and had definite proof of the guilt of the other.

The detective told the woman that he had his doubts about her story (to confess).

A photographer had just finished working with two models (an athlete and a model). He was sure he’d taken bad pictures of (over-exposed the pictures of) one of them but was happy with those he’d taken of the other.

The photographer told the model that he’d taken bad pictures of her that day (to pose again).

A fireman was telling two women (a man and a woman) how serious the situation had been when their house had caught fire. He had had to risk his life (fight his way through the flames) to rescue one of them while the other had stayed outside.

The fireman told the woman that he had risked his life for many people before (to stop smoking).

A piano teacher had been giving lessons to two boys (a boy and a girl). She was impressed with the playing of one of them (she thought one of them very talented) but thought the other quite ordinary.

The piano teacher told the boy that she was impressed with his piano playing (to practise a lot).

An auctioneer was talking to two men (a man and a woman). He had just sold a picture (auctioned a painting) for one of them but had never dealt with the other one before.

The auctioneer told the man that he had sold a picture for two hundred pounds (to collect his money).

A lawyer was asked to defend two men (a man and a woman). He was going to go into court with (represent) only one of them but not the other.

The lawyer told the man that he was going into court with him for the trial (to plead innocent).
A social worker visited the two fathers (the mother and father) of some children in her care. She was horrified by (just couldn’t believe) how aggressive one of them was towards children even though the other one seemed very kind. The social worker told the father that she was horrified by his aggressiveness (to change his ways).

A policeman was talking to two women (a man and a woman) claiming to be CIA agents. He had his doubts about one of them (he suspected that one of them was lying) but did not suspect the other. The policeman told the woman that he had his doubts about her true identity (to give him proof).

A doctor was treating two sisters (a brother and sister). He could find no cure for one of them (he diagnosed one of them as having an incurable disease) even though he could easily cure the other. The doctor told the girl that he could find no cure for her at all (to be brave).

A girl was drinking alone at a bar when two blokes (a bloke and his girlfriend) came up and started talking to her. She’d had a big row (bitter argument) with one of them but had always got on well with the other. The girl told the man that she’d had a row with him once too often (to leave her alone).

A couple were visiting two young boys (a young girl and boy) in hospital. They had brought a present (bought a toy) for one of them but they hadn’t brought anything for the other, less ill, child. The couple told the boy that they had brought a present for them both (to share it).

A teacher was talking to two boys (a girl and a boy) about their exam results. She was worried about one of them for having done badly (she was troubled by how badly one of them had done) although she was pleased with the other one’s results. The teacher told the boy that she was worried about his poor performance (to work even harder).

A lecturer was giving a tutorial to two sisters (a brother and sister). He was impressed with one of them (he found one of them to be very intelligent) though he thought the other pretty stupid. The lecturer told the girl that he was impressed with her good work (to keep it up).

A company executive was about to interview two women (a man and a woman) for a job. He had received a letter of application (a postal application) from one of them but the other had simply phoned. The executive told the woman that he had received a letter from her that day (to sit down).

A doctor was talking to two men (a man and a woman). He had made out a prescription (he had prescribed some pills) for one of them but not for the other. The doctor told the man that he had made out a prescription for some pills (to go to bed).

A baker was talking to two women (a man and a woman). He had baked a cake (made a gateaux) for one of them and scones for the other. The baker told the woman that he had baked a cake for himself as well (to pay him later).

A politician was talking to two boys (a boy and a girl). He disagreed with the views of (was arguing with) one of them but thought the other to be very sensible. The politician told the boy that he disagreed with him over his ideals (to try and see sense).

A film director was at a party talking to two actors (an actor and an actress). He had heard of (he’d been told about) one of them before, but knew nothing about the other. The director told the actor that he had heard of him from someone else (to audition for his film).
A hospital patient was having a discussion with two nurses (a doctor and a nurse). He was complaining about (was angry at) the way he had been treated by one of them and hoped the other would do something about it. The patient told the nurse that he was complaining about her treatment of him (to treat him better).

A burglar was in a pub talking to two men (a man and a woman). He'd been in prison (served a sentence) with one of them but only vaguely knew the other. The burglar told the man that he'd been in prison with some nice people (to keep in touch).

A ski instructor was talking to two girls (a girl and a boy). He was very happy (pleased) with the ability of one of them but he was worried about the other. The instructor told the girl that he was happy with her overall skiing ability (to try something harder).

A duchess was at a garden party talking to two men (a man and a woman). She was terribly bored by one of them (she found one of them to be terribly dull) though she quite liked the other. The duchess told the man that she was bored by his dull conversation (to go somewhere else).

A dentist was at a sherry party talking to two women (a man and a woman). He'd been at school (he'd been classmates) with one of them but had never met the other one before. The dentist told the woman that he had been at school with the other for years (to ring him sometime).

A little boy was being bullied by two older girls (an older boy and his sister). The little boy was afraid of (was worried by the threats of) one of them but wasn't too worried about the other. The little boy told the girl that he was afraid of the two of them (to leave him alone).

A woman was talking to two brothers (a girl and her brother). She was annoyed with (had got cross with) one of them but not with the other. The woman told the boy that she was annoyed with his behaviour (to go upstairs).

An art teacher told two girls (a girl and a boy) about an exhibition at a local museum. He wanted to go with one of them (he asked one of them to go with him) but not the other. The art teacher told the girl that he wanted to go with her to the exhibition (to meet him later).

A girl was talking to two of her teachers (her mother and her teacher). She had been having trouble with one of them (she was complaining because one of them had been ill-treating her) for a long time whilst the other had always been good to her. The girl persuaded the teacher that she was having trouble with her home life (to see reason).