THE RESOLUTION OF LOCAL SYNTACTIC AMBIGUITY
BY THE HUMAN SENTENCE PROCESSING MECHANISM.

Gerry Altmann
Department of Linguistics
University of Edinburgh
George Square
Edinburgh EH8 9LL.   (GB)

ABSTRACT

The resolution of local syntactic ambiguity by the Human Sentence Processing Mechanism is a topic which has provoked considerable interest in recent years. At issue is whether such ambiguities are resolved on the basis of syntactic information alone (cf. Minimal Attachment - Frazier, 1979), or whether they are resolved on some other basis. Crain & Steedman (1982) suggest that the resolution process is governed not by Minimal Attachment but instead by whether or not a referring expression provides sufficient information with which to identify a unique referent. Such an approach relies on the provision of adequate contextual information, something which has been lacking in experiments which have been claimed to support Minimal Attachment. In this paper I shall consider a number of such experiments, and the different patterns of results which emerge once contextual information is provided. Although the importance of contextual information will be stressed, I shall briefly consider reasons why parsing preferences arise in the absence of any explicit prior context. The conclusion is that computational models of syntactic ambiguity resolution which are based on evidence which has ignored contextual considerations are models of something other than natural language processing.

There has been much controversy recently surrounding the processes responsible for the "garden path" effect in the following kind of example:

The oil tycoon sold the off-shore
oil tracts for a lot of money
wanted to kill J.R.

The garden path effect arises here because the Human Sentence Processing Mechanism ("HSPM") encounters, during the processing of this sentence, a local syntactic ambiguity. The word "sold" is ambiguous: it can be interpreted either as a simple active, or it can be interpreted as a past participle, in a reduced passive. The only way to make the whole string into a sentence is to interpret it as a reduced form of the passive. However, what seems to happen in this (and similar) examples is that people tend to interpret the word "sold" as the main verb. This tendency leads them down a syntactic garden path.

So the HSPM exhibits a preference for one analysis over another when faced with a local ambiguity. But why? A number of suggestions have been made concerning this. One suggestion, originally proposed by Kimball (1973) and followed up more recently by Frazier (1979) and Rayner, Carlson, & Frazier (1983), is that the HSPM takes into account the syntactic structure of these sentences. There are two possible structures which could be assigned to the ambiguous sentence fragment

The oil tycoon sold the off-shore
oil tracts ...

The reduced passive interpretation requires an extra NP node as compared to the main verb interpretation. Kimball (1975) and Frazier suggest that when more than one interpretation is possible, one pursues that interpretation which creates the structure with fewest nodes. This is what Frazier calls the Principle of Minimal Attachment.

This structural hypothesis proposes, then, that an initial decision is made on grounds of syntactic structure alone. If it subsequently turns out to be the wrong decision (on grounds of "implausibility"), the alternative analysis (which is identified on the basis of "thematic selection" - Rayner et al., 1983) is then, and only then, attempted. In support of this claim, Rayner et al. collected reading times and eye movement data for sentences which, syntactically speaking, allow two attachment sites for a prepositional phrase: one attachment, to an NP, requires an extra NP node as compared to the other attachment, which is to a VP.

The burglar blew open the safe with the dynamite (Minimal attachment to VP)

The burglar blew open the safe with the diamonds (Non-minimal attachment to NP)

In the case of the non-minimally attached version, the correct attachment (to the NP) would be attempted only <u>after</u> the minimal attachment to the VP had first been tried. As they had predicted, reading times to the non-minimally attached versions were significantly <u>longer</u> than to the minimally attached versions.

An alternative to Minimal Attachment is proposed by Ford, Bresnan, & Kaplan (1982), who suggest that these preferences arise from the order in which lexical/syntactic rules in the grammar can be accessed (cf. Wanner's "implementation" of Minimal Attachment, 1980). Ford et al.'s theory of Lexical Preference is more powerful than Minimal Attachment because this ordering can, in part, be determined by the actual lexical items which are involved. But these two proposals are both, in effect, structurally based. They take no account of the referential function of the particular constructions involved.

Stephen Crain (Crain, 1980, discussed in Crain & Steedman, 1982) noted that what many of the garden-path sentences have in common is that of the two possible analyses, one is functionally equivalent to, or identical with, a restrictive relative clause. Nounphrases are used by the speaker to refer to objects. The function of a restrictive relative is to give additional information as to who or what is being talked about. This additional information is necessary because without it, there would be insufficient evidence with which to determine who or what was being referred to. So restrictive relatives are used because if we'd just heard the expression "the oil tycoon" or "the safe" we would not have known just <u>which</u> candidate "oil tycoon" or which candidate "safe" was intended. But where do these different "candidate" "oil tycoons" and "safes" come from? Normally, they must presumably be introduced into the discourse some time <u>before</u> these target sentences are encountered, and represented by speaker and hearer in some kind of model of the discourse. There is a sense in which all of these examples are unnatural because each sentence is presented in isolation. We refer to "<u>the</u> oil tycoon" and "<u>the</u> off-shore oil tracts", but we've never mentioned them before. To control for this, we should really present these target sentences embedded in a context.

Stephen Crain did just this: using an incremental grammaticality judgement task, and a class of ambiguity which is different in form but the same in principle (see below), he showed that garden path effects could be overcome or induced depending on the referential nature of the context (i.e. depending on whether just one "oil tycoon" or more than one "oil tycoon" had been introduced in the preceding text). The work of Crain suggests that a major methodological shortcoming in the work of Frazier (1979), Rayner et al. (1983) and Ford et al. (1982) is the absence of any context within which the particular syntactic constructions they were studying could be allowed to function. There is thus a very definite sense in which the processing of these constructions will have been abnormal. It follows that a suitable test of Rayner et al.'s claims is to replicate their experiment using the same reading time task, but with the provision of contexts which are felicitous with one or other of the two versions of their examples.

The following contexts were devised for an experiment (Altmann, forthcoming, and Altmann & Steedman, forthcoming):

To <u>induce</u> attachment to <u>NP</u>

A burglar carrying some dynamite broke into an heiress's house. Once inside he found two safes. One of them had some diamonds inside whilst the other had several priceless emeralds.

To <u>induce</u> attachment to <u>VP</u>

A burglar carrying some dynamite broke into an heiress's house. Once inside he found a safe and a jewelry box. One of them had some diamonds inside whilst the other had several priceless emeralds.

Minimal (<u>VP</u>) attachment

The burglar blew open the safe with the dynamite.

Non-<u>minimal</u> (<u>NP</u>) attachment

The burglar blew open the safe with the diamonds.

These examples are "minimally-different" to the extent that the only difference between them is the change from "two safes" to "a safe and a jewelry box". This is a change which, in theory, affects only the cardinality of the set of "safes".

It was found that there was a strong

effect of referential context on the non-minimally (NP) attached targets (230 msec). Furthermore, reading times to these were considerably faster than reading times to the minimally (VP) attached versions (there was a difference of 348 msec in the "NP-inducing" conditions, and 190 msec overall[1]). This is of course the reverse of what would be expected on a Minimal Attachment or Lexical Preference account. Neither of these could account for this effect without having to discount the experimental evidence which currently supports them.

However, it was also found that there was no effect of context on the minimally (VP) attached targets (the difference in reading time across the two context conditions was only 78 msec). This was surprising given that the "VP-inducing" context should have been felicitous with this target, and the "NP-inducing" context infelicitous. It became apparent, however, that neither of these contexts was in fact felicitous with VP-attachment.

The function of a PP when attached to an NP, in these examples, is to provide additional and necessary information with which to identify a particular object in the discourse model. As such it must be providing GIVEN information (cf. Clark & Haviland, 1977). The function of a PP when attached to a verb, is to provide NEW information about the action denoted by the verb: the burglar didn't simply blow open the safe, he blew it open with the dynamite. This in turn, presupposes that the action denoted by the verb ("blow open") is GIVEN. In the so-called "VP-inducing" context, this was not the case: the co-text (that is, the preceding text) was not felicitous with the VP-attachment. A second experiment was run in which the "blowing open" was known about by subjects in advance of the target sentence (i.e. was Given), and this time, strong effects of context were found on both kinds of target (113 msec for NP-attached targets across the two conditions of context, and 358 msec for the VP-attached targets). Once again, the non-minimally attached targets were significantly faster than the minimally attached targets (486 msec in the NP-inducing condition, and 245 msec overall).

What seems to be important, then, is not so much the structure of a construction, but rather the presuppositions which are implicated by its use. If these presuppositions have

---

[1] All reported differences were significant at least at $p < 0.05$.

been satisfied by the preceding discourse, then that construction will be favoured over a construction whose associated presuppositions have not been satisfied. This notion is important because its application to another class of ambiguity phenomena suggests that evidence previously thought to favour lexical or structural accounts of the resolution process does not bear on the issue of ambiguity resolution at all.

In the following ambiguous example, the complement clause analysis of the "that-clause" is preferred to the relative clause analysis (e.g. Wanner, Kaplan, & Shiner, 1974).

The boy told the girl that he liked the story

Furthermore, even when the relative clause analysis is initially chosen, these examples take longer to process (via a reading time measure) than when the complement analysis is chosen (e.g. Wanner, Kaplan, & Shiner, 1974; Altmann, forthcoming). In other words, the relative clause analysis is not just the least preferred, but is also the more "complex". The general explanation is that, in the example above, the noun phrase "the girl" is preferentially treated as a simple NP, and not as the first NP constituent in a complex NP. Wanner et al. (1974) and Wanner (1980) model these effects using an ATN, and show that they can be made to arise from peculiarities of the order in which arcs leave certain states. Frazier & Fodor (1978) cite this observation in support of Minimal Attachment, whilst Ford et al. (1982) would predict this effect on the basis of their theory of lexical and syntactic preferences, in which the simple NP expansion is ordered before the complex NP expansion.

With regard to the preference for complements over relatives, Crain's original demonstration of referential context effects used examples which exhibited this same class of local ambiguity. However, by the nature of the task he employed, Crain did not address the issue of complexity.

Restrictive relatives provide Given information, and the information contained within the relative must therefore be matched against information which already exists in the hearer's model of the universe. This matching process presumably requires a certain amount of inferencing, or "bridging" (cf. Haviland & Clark, 1974; Sanford & Garrod, 1981). Complement clauses require no such matching process, and are therefore less complex. The inferencing process can only

125

be controlled for if the materials under study are preceded by felicitous co-texts. To assess the contribution of inferencing to processing time, an experiment was run (Altmann, forthcoming) using the following examples (which are similar to those used by Crain, 1980).

"INFERENCING"
A policeman was questioning two women. He was suspicious of one of them but not of the other.

"MINIMAL INFERENCING"
A policeman was questioning two women. He had his doubts about one of them but not about the other.

RELATIVE CLAUSE TARGET
The policeman told the woman that he had his doubts about to tell the truth.

COMPLEMENT CLAUSE TARGET
The policeman told the woman that he had his doubts about her clever alibi.

(The underlining was not present in the experimental items.) The amount of inferencing required to process the relative target was manipulated by changing the (underlined) wording in the preceding co-text from "was suspicious of" ("inferencing") to "had his doubts about" ("minimal inferencing"). Apart from finding strong effects of context (thereby replicating Crain's experiment but with a different methodology - reading time), it was found that there was no absolute difference between complement targets and relative targets once context and inferencing were controlled for (only 31 msec in the "minimal inferencing" condition vs. 385 msec in the "inferencing" condition).

It would appear, then, that much of the evidence cited in the literature concerning the resolution of local syntactic ambiguity has been misinterpreted because these studies have failed to consider the referential function of the constructions in question.

An account of syntactic ambiguity resolution has been alluded to which is based on a notion of referential context and discourse models. More specifically, it is an account based on the act of establishing what is, and what is not, already known to the hearer. But although it can explain parsing preferences when a target sentence is embedded in a discourse, can it also account for the preferences exhibited in isolated sentences (the "null context" - cf. the original "oil tycoon" example)?

In the absence of any preceding discourse, there can exist no discourse model within which to integrate the information contained within the isolated sentence. In such cases, nothing can be succesfully interpreted as Given information. It follows that all incoming material must be treated as if it provides New information. If the incoming material is ambiguous, between a reading which promises New information (e.g. a complement clause) and one which promises Given information (e.g. a relative clause), then in the null context the former interpretation must be chosen. In general, if there is a choice between a complex NP analysis, which implicates additional Given information by which to identify the intended referent, and a simple NP analysis, then in the null context$_2$ the simple NP analysis must be chosen.

Structurally based theories of local syntactic ambiguity resolution can account for the null context data, but cannot account for the data concerning contextual effects on ambiguity resolution. The present account can be applied to both sets of data.[3]

Many studies have, in the past, viewed syntax as being concerned with the relations which hold within single sentences. Such a view is short-sighted, however. Within a dialogue, the speaker's intention is to evoke a specific set of

[2] Although this explains the preference, in the null context, for complement clauses over relative clauses, it does not explain the increased complexity of relative clauses. This is explained as follows: the relative clause intepretation violates more presuppositions (concerning the state of the hearer's discourse model) than does the complement clause interpretation (see Crain & Steedman, 1982, and Altmann & Steedman, forthcoming, for discussion). The reported experiments demonstrated that such violations lead to increased reading times. The relative clause interpretation induces longer reading times because of the increased number of these violations.

[3] It is argued in Altmann (forthcoming) and Altmann & Steedman (forthcoming), that an account based on the distinction between what is and what is not already known to the hearer/reader (here defined as the distinction between Given and New) will also generalise to the examples which have, on "structural" accounts, been explained by Right Association (Kimball, 1973) and Late Closure (Frazier, 1979).

processes in the hearer. This is achieved, in part, by way of the syntactic constructions which the speaker chooses to adopt. The role of these processes is to establish a relationship between the information conveyed by the utterance, and the information already known to the hearer. Such processes must therefore address information which is both internal and external to the utterance. Studies which purport either to investigate syntactic processing empirically, or to model it computationally, should not ignore the role or the requirements of these processes. To do so is to study something other than natural language processing.

ACKNOWLEDGEMENTS

REFERENCES

Altmann, G.T.M. Forthcoming. Reference and the Resolution of Local Syntactic Ambiguity: The Effect of Context during Human Sentence Processing. PhD Thesis. University of Edinburgh. To be Submitted.

Altmann, G.T.M., & Steedman, M.J. The Garden Path in Context: Reference and the Resolution of Local Syntactic Ambiguity. In preparation.

Bresnan, J. 1982. The Mental Representation of Grammatical Relations. MIT Press.

Clark, H.H. & Haviland, S.E. 1977. Comprehension and the Given-New Contract. In (ed.) R.O. Freedle Discourse Production and Comprehension.

Crain, S. 1980. Contextual Constraints on Sentence Comprehension. PhD Thesis. University of Connecticut.

Crain, S. & Steedman, M.J. 1982. On not being led up the garden path: the use of context by the psychological parser. In (eds.) Dowty, D., Kartunnen, L., & Zwicky, A. Natural Language Parsing: psychological, computational, and theoretical perspectives. In press.

Ford, M., Bresnan, J., & Kaplan, R.M. 1982. A Competence-Based Theory of Syntactic Closure. In J. Bresnan (1982).

Frazier, L. 1979. On Comprehending Sentences: Syntactic Parsing Strategies. Ph.D. Dissertation, University of Connecticut.

Frazier, L. & Fodor, J.D. 1978. The Sausage Machine: A new 2-stage Parsing Model. In Cognition, 6, 291-325.

Haviland, S. & Clark, H.H. 1974. What's New? Acquiring New Information as a Process in Comprehension. J.V.L.V.B., 13, 512-21.

Kimball, J. 1973. 7 Principles of Surface Structure Parsing. Cognition, 2, 15-47.

Kimball, J. 1975. Predictive Analysis and Over-The-Top Parsing. In J. Kimball (ed) Syntax and Semantics, Volume 4.

Rayner, K., Carlson, M., & Frazier, L. 1983. The Interaction of Syntax and Semantics during Sentence Processing. In Journal of Verbal Learning and Verbal Behaviour, 22, 358-374.

Sanford, A.J. & Garrod, S.C. 1981. Understanding Written Language. Wiley.

Wanner, E. 1980. The ATN and the Sausage Machine: which one is baloney? In Cognition, 8, 209-225.

Wanner, E., Kaplan, R., & Shiner, S. 1974. Garden Paths in Relative Clauses. Unpublished ms. Harvard University.

127