The Effect of NP Definiteness on Parsing Attachment Ambiguities

Julie C. Sedivy
University of Rochester

Michael J. Spivey-Knowlton
University of Rochester

The Effect of NP Definiteness on Parsing Attachment Ambiguities

Julie C. Sedivy and Michael J. Spivey-Knowlton[1]

University of Rochester

1.0   Introduction

One of the biggest questions facing researchers attempting to understand how humans process language in real time is the problem of rampant temporary ambiguity in natural language. In other words, it is often the case that at some given point, an input string may be mapped onto more than one semantic or syntactic representation. Many of these indeterminacies appear to be quickly and painlessly resolved; in fact, as listeners, we rarely become aware of having misanalyzed a sentence and having to recompute its meaning. However, some sentences do produce a discernible effect of initial misanalysis. Much of the work in sentence processing has focused on identifying when and why such misanalyses occur.

The study reported in this paper deals with a classic structural ambiguity that has received a good deal of attention in the sentence processing literature: that of sentences containing ambiguously attached prepositional phrases. The following sentence is an example this construction:

(1)   The gang leader hit the lawyer with a whip before getting off the
       subway.

There are two possible interpretations of this sentence, corresponding to two possible structural assignments: if the prepositional phrase *with a whip* is attached to the verbphrase, the sentence is interpreted as meaning that the gang leader used a

447

448   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

whip in hitting the lawyer. If, on the other hand, the prepositional phrase is attached to the nounphrase *the lawyer*, the sentence is understood as meaning that the gang leader in some unspecified manner hit the lawyer who had a whip. Not all interpretations are created equal, however. For most people, the first and strongest interpretation is the one in which the gang leader used a whip to inflict violence upon the lawyer (i.e. the VP-attached structure). Furthermore, something strange seems to happen when people read sentences of the following sort:

(2)   The gang leader hit the lawyer with a wart before getting off the subway.

Many people experience a brief moment of surrealistic imagery before realizing that a much more sensible interpretation of this sentence is available. In fact, a number of reading time studies have shown that people systematically take longer to read the prepositional phrase when its semantics demands a NP-attachment than when it is compatible with a VP-attachment. (Altmann, 1986; Ferreira & Clifton, 1986; Rayner, Carlson & Frazier, 1983; but see also Taraban & McClelland, 1988). This is generally taken as evidence that people initially assign such ambiguous strings a structure in which the prepositional phrase is attached to the VP, and are forced to revise their analysis when this structure turns out to be meaningless or pragmatically deviant.

In an attempt to account for this phenomenon, a number of proposals have been made that differ widely in their flavour. One of the most influential proposals dealing with these and a number of other structural ambiguities is the Minimal Attachment Hypothesis (Frazier, 1978, 1987). This hypothesis is grounded within a general processing theory which assumes that parsing is carried out in a serial, highly modular fashion, with initial syntactic commitments being made without reference to semantic or pragmatic information. Specifically, Minimal Attachment states that there is a general parsing principle which initially computes only the least syntactically complex structure for any string, where syntactic complexity is defined in terms of the number of non-terminal nodes required in building a structure. Semantic or discourse-related information is not permitted to influence the initial parse. Rather, such information may only serve to evaluate the initial parse, and in the event that it turns out to be semantically or pragmatically unviable, to guide a subsequent reanalysis.

In a different vein, Taraban and McClelland (1988) suggest that the misanalysis effect for sentences such as (2) above results not from any differences in syntactic complexity between the possible structures, but from a violation of expectations of a thematic nature. They point out that a change in the attachment of a prepositional phrase implies a change in the thematic role that is assigned to the object of a preposition, and suggest that attachment preferences are guided by preferences for particular thematic roles associated with specific verbs and prepositions. They show evidence from a reading time study which indicates that certain combinations of verbs and prepositions produce expectations for NP-attached rather than VP-attached prepositional phrases. Furthermore, the occurrence of an unexpected thematic role results in a discernible misanlysis effect regardless of whether the change in the thematic role coincides with a change in the attachment of a prepositional phrase. It is possible that such thematic preferences

are driven by the frequency of occurrence of specific roles with individual lexical items.

Alternatively, proponents of the Referential Theory (Altmann & Steedman, 1988; Altmann, Garnham & Denis, 1992; Crain & Steedman, 1985; Ni & Crain, 1990) claim that the misanalysis effect for sentences such as (2) results from the way in which the parser attempts to establish referential links with a mental discourse model. They have stressed that the processing of sentences in the absence of linguistic context neverthless takes place with reference to the discourse model, and have suggested that evidence for the garden path effect for sentences containing NP-attached prepositional phrases has its roots in the discourse properties of definite nounphrases. This theory hinges on the notion that a "simple" (i.e. unmodified) definite NP presupposes the existence of a single referent denoted by the noun. On the other hand, a "complex" (i.e. modified) definite NP presupposes the existence of multiple referents denoted by the head noun, one of which must be distinguished by virtue of the property denoted by the prepositional phrase. The referential account assumes a theory of parsing in which all possible syntactic parses are made in parallel, and where the parser selects the option that has the fewest unsatisfied presuppositions. In the absence of preceding context, a complex NP interpretation requires adding more unsatisfied presuppositions to the discourse model than a simple NP interpretation, resulting in the initial selection of the wrong structure for sentences such as (2).

The most compelling evidence for the Referential Theory comes from the effect of introducing contexts containing two NP referents prior to the target sentence. This has the effect of producing a preference for a complex NP analysis, that is, a nounphrase modified by an attached prepositional phrase. Here are some examples of successful context manipulations from Altmann and Steedman (1988):

(3)   *Context supporting NP-Attachment:*

A fireman was running to the scene of a fire carrying a heavy axe.
He had to smash down a door. When he got to the scene of the fire ,
he found a door which had a rusty lock and a door which was nailed
shut.

Context supporting VP-attachment:

A fireman was running to the scene of a a fire carrying a heavy axe.
He had to smash down a door. When he got to the scene of the fire,
he found a door which had a rusty lock and a window which was nailed
shut.

Target sentence:

The fireman smashed down the door with the heavy axe but smoke
overcame him.
The fireman smashed down the door with the rusty lock but smoke
overcame him.

450   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

The effect of the referentially biasing contexts was not only to eliminate the misanalysis effect for NP-attached prepositional phrases in contexts supporting NP-attachment, but in fact to reverse it such that VP-attached prepositional phrases in complex NP-supporting contexts actually took longer to read than their NP-attached counterparts.

This is an interesting result that, of the three proposals outlined earlier, is predicted only by the Referential Theory. However, it does not necessarily follow that the parsing preferences exhibited by ambiguously attached prepositional phrases are in all cases reducible to referential factors. It is possible that referential factors form one of a number of constraints at work in parsing such ambiguities. In other words, there may exist a parsing preference for VP-attachment that is independent of referential factors, but may be overridden by them.

2.0   Experiment 1

The aim of our study was to determine whether there is evidence for a parsing preference when referential factors of the sort described above do not play a role. While we acknowledge that "null" contexts cannot be said to be neutral, in that even in the absence of context, interpretation must take place with relation to the building of discourse structure, it *is* possible to create sentences that are presuppositionally neutral in the relevant sense. To this end, we exploited the discourse properties of indefinite nounphrases, and used sentences in which indefinite nounphrases, rather than definite nounphrases appeared before the ambiguously attached prepositional phrase. Since the function of an indefinite NP is to introduce a new discourse entity, and carries with it no presuppositions of existence or uniqueness (Heim, 1982), a modified indefinite NP is presuppositionally equivalent to an unmodified one. Therefore, if referential factors really are responsible for the misanalysis effect found for sentences in the absence of context, we would expect to find a standard garden path effect for sentences containing definite nounphrases, and no such effect for sentences containing indefinite nounphrases.

There was an additional motivation to examine the effects of referential factors sentence-internally: that is, we were not convinced that a referential explanation best accounts for the context effect found in the Altmann and Steedman study. One possibility is, as they claim, that the presence of two NP referents satisfies the presuppositions required by definite nounphrases. However, there is a possible confound in the context manipulations that were used. The NP-supporting contexts typically introduced a minimal pair of entities into a discourse in which it was established that one of the entities would play a special role. For instance, with respect to the example in (3), the reader has been told in advance that the fireman will smash down a door. When we encounter the two NPs referring to doors of some type in the context supporting NP-attachment, we are unsure which is the relevant one in this event. This indeterminacy may create a powerful expectation that subsequent discourse will provide information to distinguish between these two doors. This is subtly different from the claim being made by Referential Theory. In other words, the context effect may not have so much to do with accommodating presuppositions associated with definite constructions, as with developing expectations as to the informational content that is likely to be relevant in subsequent discourse. The latter interpretation is supported in a paper by

THE EFFECT OF DEFINITENESS ON PARSING AMBIGUITIES   451

Spivey-Knowlton (1992). This study used contexts which created an expectation for information relating to the verbphrase rather than the nounphrase by introducing two similar possible events into the discourse preceding the target sentence. It was found that, following such contexts, the preference for VP-attachment was higher than for sentences which occurred following a context that did not contain two similar events. Since verbphrases do not carry the same existential presuppositions as nounphrases do, it is difficult to interpret these results in terms of the Referential Theory. For these reasons, we felt it would be advantageous to evaluate the role of referential factors by comparing constructions that require the satisfaction of referential presuppositions with ones that do not, rather than by manipulating the discourse context.

2.1   Method

The experiment was a self-paced reading task using materials based on the target sentences in Altmann and Steedman (1988) Unlike the Altmann and Steedman study, however, the sentences were presented in isolation. In order to manipulate the presuppositional requirements of the sentences, we varied the definitenessss of the nounphrase preceding the ambiguously attached prepositional phrase, as well as the attachment site of the prepositional phrase. Below is an example of the four versions of a single stimulus item:

(4) a. The fireman/ smashed down/ the door/ with the rusty lock/ but smoke/ overcame him.
b. The fireman/ smashed down/ the door/ with the heavy axe/ but smoke/ overcame him.
c. The fireman/ smashed down/ a door/ with a rusty lock/ but smoke/ overcame him.
d. The fireman/ smashed down/ a door/ with a heavy axe/ but smoke/ overcame him.

Examples (4a) and (4c) are examples of NP-attached phrases, whereas sentences (4b) and (4d) are examples of VP-attached phrases.

Thirty-two University of Rochester undergraduates served as subjects for the study. The sentences were presented on a computer screen one phrase at a time (slashes in (4) indicate phrasal units of presentation). At the beginning of each trial, every character in the sentence was covered by a dash. The subject would begin to read through the sentence at his/her pace by pressing a button to uncover the first phrase. After reading the first phrase, he or she would press the button again, at which point the next phrase would be uncovered, and the earlier phrase would be covered over with dashes once again, and so on, until the entire sentence had been read. The 32 target sentences were randomly embedded within 48 filler trials. All of the experimental sentences and half of the filler sentences were followed by questions which either probed the syntactic attachment made by the subject or checked for comprehension of some other aspect of the sentence. Subjects were instructed to read the sentences at a comfortable pace that approximated their normal reading speed.

452   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

2.2   Results

Table 1 displays the reading times for the four conditions at each phrase position. The results show that there is a preference for VP attachment over NP attachment for both sentences containing definite NPs and sentences containing indefinite NPs. This garden path effect begins at the disambiguating prepositional phrase, and persists into the next region. However, at the position of the prepositional phrase, the magnitude of this preference is substantially greater for sentences with definite NPs.

In order to statistically evaluate the results, an analysis of variance was first carried out on reading times collapsed across all phrase positions. The analysis revealed that sentences with indefinite NPs were read more quickly than sentences with definite NPs ($F1(1,28)=6.53$, $p<.02$; $F2(1, 28)=5.26$, $p<.05$). It also showed that VP-attached sentences were read more quickly than NP-attached sentences ($F1(1,28)=15.96$, $p<.001$; $F2(1,28)=12.32$, $p<.005$). An interaction of Definiteness by Attachment, however, did not approach significance in this analysis. It would appear, then, that the garden path effect for the NP-attached sentences is no stronger with a definite NP than with an indefinite NP. In fact, simple effects showed significant garden paths for NP-attached sentences with definite NPs ($F1(1,28)=11.66$, $p<.005$; $F2(1,28)=5.97$, $p<.05$) *and* the indefinite NPs ($F1(1,28)=6.83$, $p<.02$; $F2(1,28)=5.03$, $p<.05$).

Table 1.

	Exp. 1: Reading Time (ms) by Sentence Region				
	NP region	Verb region	NP region	PP region	Next region
Definite NP					
NP-Attached	559	549	533	835	704
VP-Attached	<u>532</u>	<u>541</u>	<u>532</u>	<u>728</u>	<u>661</u>
VP-Attachment					
preference	27	8	1	107	43
Indefinite NP					
NP-Attached	545	548	518	724	698
VP-Attached	<u>555</u>	<u>558</u>	<u>516</u>	<u>677</u>	<u>638</u>
VP-Attachment					
preference	-10	-10	2	47	60

In addition, an analysis of variance of the reading times at the prepositional phrase revealed that PPs following indefinite NPs were read faster than PPs

following definite NPs ($F1(1,28)=15.32$, $p<.002$; $F2(1,28)=10.74$, $p<.005$) Also, VP-attached PPs were read faster than NP-attached PPs ($F1(1,28)=11.5$, $p<.005$; $F2(1,28)=13.31$, $p<.002$). There was a marginally significant interaction of Definiteness and Attachment, with the garden path effect showing up as smaller for indefinite NPs than definite NPs ($F1(1,28)=3.97$, $p=.056$; $F2(1,28)=1.63$, $p=.21$) A post-hoc test revealed that, within the definite NP condition, the VP-attached PP was read faster than the NP-attached PP ($t(28)=5.07$, $p<.001$). Likewise, the same statistic applied to the indefinite NP condition revealed that the VP-attached PP was read faster than the NP-attached PP, but this difference was much less robust ($t(28)=2.25$, $p<.05$).

An analysis of variance of reading times at the phrase following the PP again revealed that when the PP was attached to the VP, reading times were shorter than when the PP was attached to the NP ($F1(1,28)=9.38$, $p<.01$; $F2(1,28)=6.74$, $p<.02$). However the main effect of Definiteness and the interaction between Definiteness and Attachment were completely absent at this phrase position. Statistical analyses of other phrase positions revealed no reliable differences.

2.3   Discussion

The result showing that reading times are longer for definite NPs than for indefinite NPs suggests that subjects are indeed attempting to establish referential linking of a NP to a discourse entity even in the absence of context, as argued by proponents of the Referential Theory. The main effect of Attachment indicates that people are showing evidence for misanalyzing prepositional phrases requiring NP-attachment, but it doesn't tell us what the garden path effect looks like for sentences containing indefinite versus definite NPs.

We were particularly interested in the effect of definiteness upon the attachment preferences in order to discriminate among competing hypotheses. The results suggest that Referential Theory alone is unable to account for parsing preferences. If parsing consists of computing all possible structures in parallel, and choosing those that require the fewest number of unsatisfied presuppositions, there is no reason to expect that modified indefinite nounphrases should be less preferred than unmodified ones, as neither case involves setting up referential presuppositions. However, we found evidence of a VP-attachment preference for sentences with *both* definite and indefinite NPs, suggesting that there is some parsing mechanism independent of referential linking that is in part responsible for the difficulty subjects had with PPs attached to the NP. On the other hand, Minimal Attachment claims that the garden path effect is due to parsing principles based on the initial computation of the least complex syntactic structure, and therefore predicts that the preference for VP-attachment should be equally strong for both Definiteness conditions. However, the results indicate that the magnitude of the garden path effect is not equal; sentences containing definite NPs showed a larger garden path effect than sentences with indefinite NPs. However, it is possible to maintain the Minimal Attachment hypothesis by arguing that the more robust garden path for sentences in the Definite condition is due not to a stronger preference for VP-attachment, but to the fact that recovery is more difficult for sentences in which the correct structure contains a modified *definite* NP because of the additional presuppositions that must be accommodated. Although unable to rule it out categorically, the pattern of results obtained in this study does not support this

454   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

interpretation. If the influence of definiteness is to be attributed to the recovery phase in parsing, rather than to initial structural assignment, we would expect the definiteness effect to show up at the latest point in the garden path. In fact, the definiteness effect shows up only on the disambiguating prepositional phrase itself, the first phrase at which evidence of a garden path is possible, whereas the effect of Attachment persists at the following phrase. This suggests that the parser is still correcting itself beyond the point at which definiteness has an effect. Furthermore, the Minimal Attachment hypothesis falls short when confronted with evidence that, given referentially appropriate contexts, the preference for a VP-attachment of the prepositional phrase can be eliminated altogether, or even reversed (Altmann and Steedman, 1988; Britt, Perfetti, Garrod and Rayner, 1992; but see also Ferreira and Clifton, 1986).

3.0   Experiment 2

In Experiment 1, we intended to look at the effect of varying the definiteness of the nounphrase preceding the prepositional phrase. Because it is infelicitous to modify an indefinite NP with a prepositional phrase containing a definite NP, the experiment simultaneously varied the definiteness of the noun within the PP as well. In order to be sure that the results obtained in the previous experiment were not due to the effect of changing the definiteness of the PP-internal nounphrase, which we considered theoretically irrelevant, we conducted a control experiment that varied only the definiteness of the NP inside the prepositional phrase.

3.1   Method.

This experiment used the same 32 experimental sentences and 48 filler sentences as in Experiment 2. However, the NP preceding the PP (see (5) below) was always definite, with attachment site and the definiteness of the NP *within the PP* as the only manipulations:

(5) a. The fireman/ smashed down/ the door/ with the rusty lock/ but smoke/ overcame him.
     b. The fireman/ smashed down/ the door/ with the heavy axe/ but smoke/ overcame him.
     c. The fireman/ smashed down/ the door/ with a rusty lock/ but smoke/ overcame him.
     d. The fireman/ smashed down/ the door/ with a heavy axe/ but smoke/ overcame him.

Twenty-eight University of Rochester undergraduates who had not participated in the first experiment served as subjects. Subjects were run in the same fashion, and on the same apparatus as in Experiment 1.

3.2   Results and Discussion

As in the previous experiment, the results show evidence of a preference for VP-attachment over NP-attachment, although in this case, only at the phrase following the disambiguating prepositional phrase rather than at the prepositional phrase itself. However, varying the definiteness of the NP within the potentially

ambiguous prepositional phrase did not result in a difference in the magnitude of the garden path. The can be seen in Table 2 below.

An analysis of variance was first carried out on reading times collapsed across all regions. In this analysis, no main effect of Definiteness was found, nor was there an interaction between Definiteness and Attachment. There was, however, a robust main effect of Attachment, with sentences containing NP-attached prepositional phrases being read significantly slower than sentences with VP-attached prepositional phrases $(F1(1,24)=7.96, p<.01; F2(1,28)=10.01, p<.005)$. The results of this analysis suggest that the manipulation of definiteness of the PP-internal nounphrase did not alter the size or the nature of the garden path induced by NP-attachment. Therefore, it was not responsible for the interaction observed in Experiment 1, where the definiteness of the NP preceding the PP was manipulated.

Next, we conducted an analysis of variance of the reading times at the disambiguating prepositional phrase. PPs containing indefinite NPs turned out to be read faster than PPs containing definite NPs, again suggesting that referential linking is happening on-line even in the absence of context.$(F1(1,24)=18.84, p<.001; F2(1,28)=18.78, p<.01$. But no effect of Attachment was observed at this position, and neither was there any hint of an interaction between Definiteness and Attachment. At the position following the PP, only a main effect of Attachment was observed, such that NP-Attached versions were read more slowly than VP-attached versions $(F1(1,24)=20.46, p<.001; F2(1,28)=28.91, p<.001)$.

Table 2.

	NP region	Verb region	NP region	PP region	Next region
Exp. 2: Reading Time (ms) by Sentence Region					
Definite NP within PP					
NP-Attached	523	544	526	783	809
VP-Attached	504	531	534	770	687
VP-Attachment preference	19	13	-8	13	122
Indefinite NP within PP					
NP-Attached	516	558	523	719	833
VP-Attached	522	556	549	712	680
VP-Attachment preference	-6	2	-26	7	153

456   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

The control experiment clearly shows evidence for a VP-attachment preference, together with the absence of an effect of definiteness of the PP-internal NP on the strength of the VP-attachment preference. This allows us to safely attribute the previously reported interaction between Definiteness and Attachment to the definiteness of the nounphrase preceding the prepositional phrase rather than within it.

4.   Corpus Analysis of Frequency

The results of the reading time experiments show that while referential factors play an important role in the parsing of ambiguous prepositional phrase attachment, they do not account for the attachment preference entirely. It appears that along with a discourse-driven parsing mechanism, there exists a more local, discourse-independent mechanism. This local component may be related to some notion of structural complexity, as suggested by the Minimal Attachment hypothesis. Alternatively, it may be of the nature proposed by the Thematic Expectations model in Taraban and McClelland (1988), where specific verb-preposition pairs create particular attachment expectations. It has been proposed that such expectations are driven by the relative frequencies with which lexical items occur in either VP- or NP-attached structures (Hindle & Rooth, 1991). Moreover, Mitchell (1992) has shown that there is a relationship between frequency counts and parsing preferences in the processing of ambiguous structures. We felt it worthwhile, therefore, to explore the possibility that frequency-based information might play a role in on-line parsing. While frequency counts are likely to be *explained* by grammatical or discourse-related factors, it is possible that parsing involves a mechanism which abstracts away from these underlying factors such that it remains sensitive only to the probabilities of the occurrence of certain structures. In order to examine this possibility, we carried out a corpus analysis based on half of the Brown corpus (Kucera & Francis, 1967). Since all of the stimulus items in the reading time experiments contained the preposition *with* as the head of the ambiguously attached phrase, we limited the analysis to instances of this preposition.

We found that there was an overall bias towards VP-attachment. This bias was not consistent across conditions, however. Sentences containing a definite NP preceding the ambiguously attached PP showed a greater incidence of VP-attachment than NP-attachment., whereas sentences in which the crucial NP was indefinite more frequently attached the prepositional phrase to the nounphrase. Frequency counts are given below in Table 3.

We initially analyzed a restricted set of sentences which closely mirrored the structure of the sentences used in the reading times experiment (i.e. a potentially ambiguous prepositional phrase following a verb with a NP complement). A classification of these sentences by attachment site revealed that overall, VP-attached prepositional phrases were significantly more frequent than NP-attached PPs, according to a chi-square goodness-of-fit test ($X^2(1)=11.51$, p<001). In addition, the frequency patterns were affected by the definiteness of the verb's complement NP; definite NPs followed by *with* were highly biased towards VP-attachment, wheareas indefinite NPs followed by *with* were biased towards NP-attachment. This interaction was statistically significant ($X^2(1)=43.38$, p<.0001.

Analysis of a less restricted set of sentences containing prepositional phrases headed by *with* showed a similar pattern. This set of sentences contained not only the sentences analyzed above, but also sentences where the *with* phrase followed an earlier post-verbal PP-adjunct, or where the potentially ambiguous PP followed a passive verb and an agent-by-phrase. The only criterion for inclusion in this set was that the *with*-phrase be structurally ambiguous with respect to its attachment site. Again, VP-attachments significantly outnumbered NP-attachments ($X^2(1)$=11.13, p<.001). A similar interaction between Attachment and Definiteness was observed ($X^2(1)$=56.78, p<.0001, with sentences containing definite NPs preceding the *with*-phrase biased towards VP-attachment, and indefinite NPs biased towards NP-attachment.

Table 3.

Frequency Counts of Sentences from the Brown Corpus

Restricted Set	NP-Attached	VP-Attached	Totals
Definite NP	12	82	94
Indefinite NP	53	29	82
Totals	65	111	176

Unrestricted Set	NP-Attached	VP-Attached	Totals
Definite NP	16	96	112
Indefinite NP	63	32	95
Totals	79	128	207

The results of the corpus study, when compared with the patterns of reading times in the experiments reported here, raise some interesting questions with respect to the way in which frequency counts might be encoded and used by the language processor. Frequency effects cannot stand alone: they must be stated with respect to something. In order to be able to make claims about the relationship between frequency counts and on-line parsing, it is necessary to know more about what sort of information frequency effects are conditionalized over. Clearly, there is no neat mapping between the frequency counts and the reading time data such that the behaviour of the garden path effect across all the conditions in the on-line experiments could be accounted for by appealing to the frequency counts alone. The most notable mismatch is that although the corpus analysis revealed a preference for VP-attachment only for sentences containing definite nounphrases, the reading time study showed a VP-attachment preference for both sentences containing definite *and* indefinite nounphrases. If frequency counts are influencing parsing, they are doing so in a more restricted manner, and in combination with

458   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

other parsing mechanisms. The nature of the information which frequency effects are contingent upon is an open question, but a very important one. Different answers to this question imply different accounts of the role of context-dependent and context-independent parsing mechanisms.

For instance, a possible scenario is one in which the frequency of NP-versus VP-attachment of prepositional phrases is contingent upon the definiteness of preceding NPs. This, in combination with a separate mechanism favouring VP-attachment (e.g. Minimal Attachment) would be sufficient to account for the pattern of results exhibited in the reading time studies: the Minimal Attachment strategy would require an initial VP-attachment structure assignment, with frequency effects serving to strengthen or weaken the initial structural expectation at a later point.[2] In this scenario, there is no need to posit a context-dependent, referentially-based mechanism. Alternatively, attachment biases based on frequency may be dependent upon specific lexical entries. For instance, the parser may be making use of information relating to the relative frequency of attachment sites for the preposition *with*, or, if the the granularity of frequency effects is finer, the frequency of attachment sites for the head of the VP preceding the ambiguous phrase. If frequency effects are conditionalized over lexical information of this type rather than over the definiteness of NPs preceding the ambiguous phrase, a very different explanation from the one just outlined is warranted; in this scenario, we would have to account for the effect of definiteness on the strength of the VP-attachment preference in the reading time data by appealing to a discourse-based explanation. The frequency-based information, then, would account for a general, context-independent VP-attachment preference. A structurally-based parsing mechanism such as Minimal Attachment would be theoretically redundant, although in principle, there is nothing to prevent the existence of such a mechanism in combination with a frequency-based mechanism. At present, we are not able to choose definitively among these possibilities. We simply point out that these are questions that need to be answered by future research before the effects of NP definiteness and specific attachment sites on parsing can be confidently attributed to any of the factors proposed by the parsing theories discussed in this paper.

5.   General Discussion

This study set out to evaluate the effect of referential properties of nounphrases on the on-line parsing of structurally ambiguous sentences. We found evidence both for influence of referential factors in parsing, and for limitations on this influence. Specifically, the reading time studies show that the magnitude of the misanalysis effect in potentially ambiguous prepositional phrases (i.e. sentences containing definite NPs) is greater when the sentence requires the accommodation of unsatisfied presuppositions than when it does not (i.e. in sentences containing indefinite NPs). However, the data indicate that there remains a discernible garden path effect for constructions which do not require the accommodation of unsatisfied referential presuppositions (i.e. in sentences containing indefinite nounphrases). This shows that the parsing preference documented for these constructions cannot be reduced to referential factors alone.

This particular combination of context-dependent and -independent factors cannot be accounted for by any one of the hypotheses outlined at the beginning of

THE EFFECT OF DEFINITENESS ON PARSING AMBIGUITIES   459

this paper. For instance, the Referential Theory does not predict the existence of a context-independent parsing preference, as exhibited in these experiments. The general theory of parsing in which it is situated is not, however, incompatible with the notion that multiple constraints may act together to pick out a favoured analysis out of a set of candidate structures. Minimal Attachment, on the other hand, is places within a set of theoretical assumptions which preclude the interaction of syntactic and discourse-based constraints. As mentioned earlier, one could argue that the referential effect in these experiments is a reanalysis effect, rather than an effect of initial structural assignment. However, the earliness of this effect in real time, and the persistence of the general VP-attachment preference beyond the point at which referential influences are evident mitigate against this interpretation. Similarly, the Thematic Expectations account cannot explain the difference in the magnitude of the garden path effect for items that are identical in all respects except for the definiteness of the NP complement to the verb.

Consequently, we believe that a plausible way to account for these data is within a framework of multiple constraints acting simultaneously to influence the initial syntactic commitments that are made by the parser. This model would involve at least two components in the parsing process: a discourse-driven component that is sensitive to the referential linkings between an utterance and its discourse context, and a locally driven, context-independent component. The nature of the local component is not clear at this time. It could well turn out to reflect a purely structural strategy which favours the building of one type of structure over another. On the other hand, a frequency-based bias towards VP-attachment might plausibly account for the local component of the on-line parsing preference. Another alternative is that the local component itself is comprised of multiple constraints acting together.

The results of the corpus analysis, however, suggest that before the interpretation we have proposed above can be confidently endorsed, it is necessary to understand the extent and nature of the effect of frequency-based information. In principle, a frequency effect could account for either the effect that definiteness has in modulating the size of the VP-attachment preference, or for a general bias towards VP-attachment (perhaps conditionalized over specific lexical entries for prepositions or the head of the VP preceding the ambiguous phrase) that is then modulated by discourse-related factors. It could not, however, account for the reading time data on its own, as the pattern of the frequency counts does not map neatly onto the pattern of reading times found in this study. If knowledge of frequency counts does play a role in on-line parsing, it is important to resolve the issue of the type of information that is encoded by the frequency-based mechanism At present, the possibility of a frequency-based account for some of the on-line data presents a confound for either the use of referential factors, or for a context-independent parsing principle such as Minimal Attachment. We leave it to further research to resolve these issues.

460   JULIE C. SEDIVY AND MICHAEL J. SPIVEY-KNOWLTON

Notes:

1.   This work was partially supported by grant HD 27206 from NIH. In addition, the first author is a recipient of a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada, and the second author is supported by a Graduate Research Fellowship from the National Science Foundation. We would like to thank the participants of NELS 23 and Mike Tanenhaus for helpful comments and questions, and Cornell Juliano for supplying the corpus extraction programs.

2.   Accepting this view, requires, however, accepting that although frequency must have its effect at a point later than Minimal Attachment, there is no measurable difference in the onset of these effects. In addition, it remains problematic when confronted with evidence from context studies that context-based information, whatever its specific nature may be, *does* exert an influence in on-line parsing.

References:

Altmann, G., A. Garnham & Y. Denis. (1992). Avoiding the garden-path: Eye movements in context. *Journal of Memory and Language*, **31**, 685-712.

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

Britt, M., C. Perfetti, S. Garrod, & K. Rayner. (1992). Parsing and discourse: Context effects and their limits. *Journal of Memory and Language*, **31**, 293-314.

Crain, S. & M. Steedman. (1985). On not being led up the garden path: The use of context by the psychological parser. In Dowty, Karttunen, & Zwicky (eds.). *Natural Language Parsing*. Cambridge: Cambridge U. Press.

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

Frazier, L. (1987). Processing syntactic structures: Evidence from Dutch. *Natural Language and Linguistic Theory*, **5**, 519-559.

Frazier, L. (1978). *On comprehending sentences: syntactic parsing strategies.* Ph.D. dissertation, University of Connecticut.

Hindle, A. & M. Rooth. (1990). Structural ambiguity and lexical relations. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics.*

Heim, I. (1982). *The semantics of definite and indefinite nounphrases.* Ph.D. dissertation, University of Massachusetts, Amherst.

Kucera, H. & Francis, W. (1967). *Frequency Analysis of English Usage.* Boston: Houghton Mifflin.

THE EFFECT OF DEFINITENESS ON PARSING AMBIGUITIES   461

Mitchell, D. (1992). Talk presented at the CUNY Sentence Processing Conference, March, 1992, New York.

Ni, W. & S. Crain. (1990). How to resolve structural ambiguities. *Proceedings of the 20th meeting of the North Eastern Linguistic Society.*

Spivey-Knowlton, M. (1992). Another context effect in sentence processing: Implications for the principle of referential support. *Proceedings of the 14th annual meeting of the Cognitive Science Society.*

Rayner, K., M. Carlson and L. Frazier. (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.

Taraban, R. & J. McClelland. (1988). Constituent attachment and thematic role expectations. *Journal of Memory and Language*, **27**, 597-632.