

A Computational Model of Syntactic Ambiguity as a Lexical Process¹

Curt Burgess
Department of Psychology
University of Rochester

S.C. Hollbach
Department of Computer Science
University of Rochester

Introduction

The use of semantic information in language comprehension is a matter of controversy. We briefly review Frazier's modular serial parser and Altmann and Steedman's conceptually driven parser. We propose a third model, based on the work of Tanenhaus and Carlson [1988],[Tanenhaus *et al.* 1987], that accounts for the conflicting results of these two models. We have built a connectionist implementation of our model, and, based on the behaviour of the network, make a number of empirical predictions.

The Controversy Defined

The work of Frazier and her colleagues ([Frazier and Rayner 1982]; [Ferreira and Clifton 1986]) is probably best known modular parser. They present considerable evidence that the parser uses a Minimal Attachment strategy. Minimal Attachment is used to explain the Bever sentences, e.g., *The horse raced past the barn fell*. The strategy is straightforward: words are attached to the phrase marker using the fewest possible syntactic nodes permissible by the syntactic rules. Frazier and Rayner [1982] have found that Non-minimal Attachment sentences take longer to read than Minimal Attachment sentences and that eye fixations in the disambiguating region of the sentences were of longer duration than elsewhere in the sentence.

It is possible, however, that a contextual manipulation could override the Minimal Attachment preference. Ferreira and Clifton [1986] consider this possibility. Their first experiment manipulates the animacy of the first noun phrase. They reasoned that the implausibility of an inanimate agent with a verb that is associated with an animate agent would be sufficient to eliminate the garden path in the reduced condition. For example, consider the following two sentences:

- 1a. The defendant examined by the lawyer turned out to be unreliable.
- 1b. The evidence examined by the lawyer turned out to be unreliable.

They found that animacy of the first noun phrase had no effect on reading times at either the *by* phrase or the second verb. However, eye fixations were longer at the first verb when it followed an inanimate NP, suggesting that eye fixations are a valid measure of semantic anomaly.² It seems, then, that the Minimal Attachment preference results in a garden path regardless of the animacy information in the first NP.

Ferreira and Clifton conduct another experiment in which they precede relative clause ambiguities such as *The editor played the tape agreed the story was a big one* with a paragraph that biases the interpretation of the sentence toward either a Minimal Attachment

¹We want to thank Mike Tanenhaus and Bob Peterson for their encouragement, patience and many helpful suggestions. Send correspondence to Curt Burgess, Psychology Dept, University of Rochester, Rochester, NY 14627 [email: RCBU@UORDBV.BITNET]

²Apparently eye fixations reflect semantic as well as syntactic anomaly. It becomes less than completely clear what the linkage is between eye fixations and particular types of psycholinguistic processing.

or Non-minimal Attachment interpretation. The results of this experiment suggest that a preceding contextual paragraph with a Non-minimal Attachment bias is unable to influence the Minimal Attachment preference and that the contextual information is used solely to recompute the structural aspects of the sentences.

While the Frazier and Rayner [1982] and Ferreira and Clifton [1986] results appear quite convincing, this structural approach is not without criticism. Altmann and Steedman [in press] propose that sentences vary not only in structural complexity, but in presuppositional complexity. When a critical sentence in an experiment is encountered, its' relative clauses and prepositional phrases presuppose that a specific set of entities has been introduced by the context. Their claim is that context is ineffectual in the Ferreira and Clifton experiment simply because the context is impoverished and does not set up the necessary presuppositions. When Altmann and Steedman [in press] constructed Non-minimal Attachment supporting paragraphs that explicitly introduce each entity in the test sentence, they did not find the usual garden path effect. Altmann and Steedman's results present a serious challenge to a simple structural explanation such as offered by Frazier and Rayner or Ferreira and Clifton. It appears that lexical or semantic representations may have a substantial role to play in the structural organization of a sentence.

Thematic Roles and Semantic Ambiguity

These apparently conflicting results can be reconciled with a lexical access model of syntactic ambiguity. According to the lexical access model, Ferreira and Clifton's failure to find an animacy effect (e.g., *The evidence examined by the ...*) lies in the lexical representation of the verb. In their first experiment they compare sentences in which the first NP is either animate or inanimate. At the *by* phrase in such a sentence, the verb is shown to be unambiguously a past participle (in the reduced version). The local indeterminacy of the verb (i.e., past tense/past participle) is a morphological ambiguity with each verb form having its own thematic role mapping. The simple past tense tends to be more frequent than the past participle in most verbs. Since the two tenses of the verb are incompatible, the less frequent past participle is inhibited by the more frequent past tense. For example, compare

- 2a. The child bought the dog a leash. 2b. The child bought the dog was pleased.

In 2a, the preferred tense of *bought* is used and the two post-verbal NPs fill thematic roles that are subcategorized by the verb. No garden path is experienced in this sentence. However, in 2b the reader is not aware that the past participle form is necessary for comprehension until the primary verb *was* is encountered. In our lexical model, both the past tense and the past participle are activated in parallel, but the less preferred form is quickly inhibited. As sentence processing continues it becomes increasingly difficult to reactivate the inhibited form. If, later in the sentence, the inhibited form of the verb is required, a garden path will occur.

Now, the results of Altmann and Steedman suggest that a garden path can be avoided under the right contextual circumstances.³ We agree and suggest the following mechanism

³At present, we feel compelled to reject Ferreira and Clifton's argument that an inanimate NP is unable to propose the initial parse. There are at least two problems with their stimuli in experiment 1. Almost half of the inanimate subjects can act as agent (e.g., *The car towed ...*) and a number of the sentences

for the contextual effect: the less preferred past participle form of the verb is selected by the discourse model as it is made available by bottom-up lexical processes. The contextually inappropriate form is then inhibited. This understanding of how context can serve to select the appropriate verb form is important, although we have not yet implemented this in our model.

The model we propose, namely multiple activation of the verb forms followed by contextual selection, parallels the research on lexical ambiguity resolution. In isolation, access of ambiguous word meanings is a frequency coded, multiple access process [Burgess and Simpson 1988a, Simpson and Burgess 1985]. Multiple access obtains even in a sententially biasing context [Seidenberg *et al.* 1982, Burgess *et al.* in press]. The notion of parallel bottom-up activation has proved to be a powerful explanatory tool in lexical processing. Compatible lexical representations remain activated in such a model, while incompatible representations compete with each other based on the entities' resting activation level and previous exposure to related or associated items.⁴ In our model, the garden path occurs as a function of the additional time and processing required to inhibit the inappropriate morphological verb form and to access the appropriate one.

In addition to verb forms, thematic roles play an important role in this model as well. Earlier research has shown that the meaning of a verb can be distinguished from its thematic roles. Tanenhaus, Burgess, D'Zmura and Carlson [1987] showed that a garden path occurs with a less preferred verb sense ambiguity (3a), but not with the preferred sense (3b). In contrast, no garden path occurs with either the less preferred (3c) or more preferred (3d) thematic role ambiguity.

- 3a. Bob passed the test to his friend. 3c. Cathy threw the cat off the bed.
 3b. Bob passed the test to his surprise. 3d. Cathy threw the cat some Friskies.

Thematic roles behave differently than sense ambiguities. All thematic roles associated with a verb are activated in parallel at the point the verb is retrieved and are provisionally assigned to verb arguments as soon as they are activated. Roles that are subcategorized by the verb remain activated even if unfilled and thus allow for rapid integration of concepts into the discourse model. For this reason they are an important entity for rapidly resolving local ambiguity.

The Connectionist Implementation

We have built a connectionist network that embodies the lexical model of syntactic ambiguity on the Rochester Connectionist Simulator [Goddard 1987]. The network is displayed in iconic form on a Sun Workstation, as depicted in Figure 1. The implementation adheres to the localist style of connectionist models, which is to say that the network is highly structured and each node has a precise semantics. Following Cottrell [1985], we adopt a form of exploded case roles, with a distinct node for each element of the cross product of thematic roles and syntactic position. For inner cases such as 'theme' there can be only one syntactic unit that plays this role, hence there are inhibitory links between all competing syntactic

continue the ambiguity beyond the *by* phrase. We've benefitted from discussions with Mike Tanenhaus, Susan Garnsey, Tom Bever and John Trueswell concerning this.

⁴This view can be contrasted with that of Marslen-Wilson and cohorts [Marslen-Wilson and Tyler 1987] in which top-down processing provides for selective access of lexical information.

BURGESS AND HOLLBACH

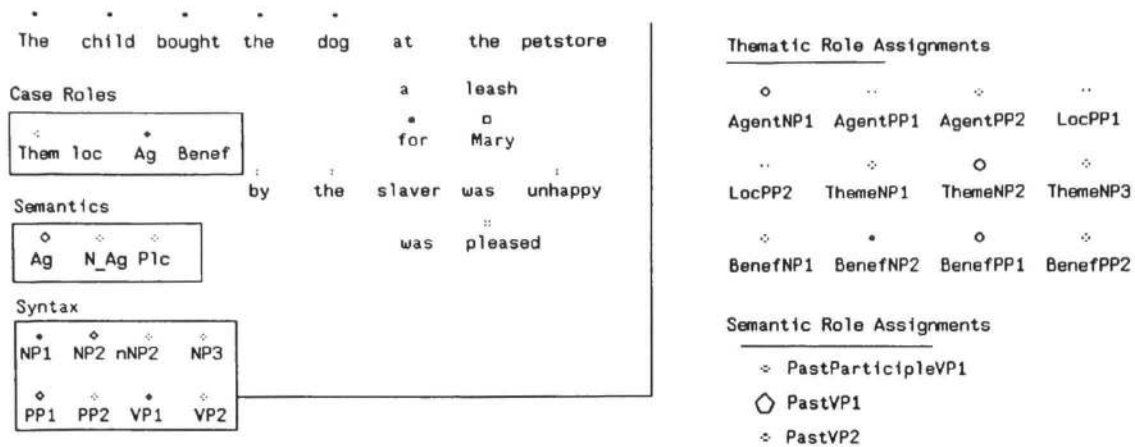


Figure 1: Visual display panel for the network. The input sentence ‘The child bought the dog for Mary’ is shown 10 iterations after sentence completion.

assignments. Moreover, a given syntactic unit can only take on one interpretation, hence there are inhibitory links between all competing thematic interpretations.

There are three distinct node types in the network: input nodes, role assignment nodes, and ‘other’ nodes. The activation function for ‘other’ nodes is simply to set the output of the node to the sum of all inputs:

$$a_t = \Sigma I_t,$$

where a_t is the activation level at time t and I_t is the input vector. Input nodes (or lexical item nodes) ignore their inputs, as the user keys activation in by hand for these nodes. The effect of reading a word in a sentence is simulated by the user setting the activation level of the relevant input node to 1.0. Activation of this node decays thereafter according to the recurrence relation:

$$a_t = a_{t-1} - \frac{1}{a_{t-1}}$$

In order to enter a stimulus sentence, the user keys in each word in turn, running the simulation a few steps between each word. The gradual decay of word activation means that by the time the last word of the sentence has been keyed in, the first word is only marginally active (see Figure 1). Thus any assumptions that have been made early on in sentence processing take a long time to reverse when invalidated by later constraints. The activation function for role assignment nodes incorporates a certain amount of inertia. If the potential of the unit is high, a relatively low weight is accorded to the input signal, while if the potential is low, the input signal has a larger impact on the updated activation level. This is accomplished by dividing the input signal (I_t) by the prior activation before adding the whole to the current activation:

$$a_t = a_{t-1} + \frac{cI_t}{a_{t-1} + k}, \quad c = 5, \quad k = 10.$$

Interpretations are advanced as the sentence progresses. A schematic of the some of the important features of network structure appears in Figure 2. As shown in Figure 2, any number of thematic role assignments are possible after the first noun phrase (NP1), so all are advanced with equal confidence. For the verb ‘bought’ (VP1) there are two possible interpretations: the more likely is the past tense, the less likely the past participle. Thus the past tense receives preferential activation, and, as the two interpretations are in direct

BURGESS AND HOLLBACH

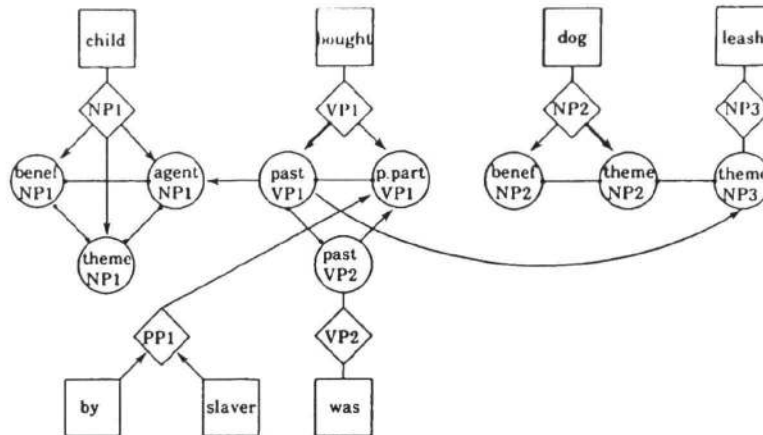


Figure 2: Rough schematic of network structure. Not all nodes and links are shown. Bold arrows indicate preferential weighting on one of multiple outputs. Input nodes appear as squares, role nodes as circles, and other nodes as diamonds.

competition, the second alternative tends to decrease in likelihood as the sentence processing continues.

The first noun phrase encountered after the verb (NP2) is ambiguous itself, but constrains the role of NP1 to either *agent* (if the verb is in the past tense), or *beneficiary* (if the past participle). The ambiguity of NP2 is resolved by the subsequent syntactic unit: if NP3 occurs, then the thematic role assignments are *benefNP2* and *themeNP3*; but if a prepositional phrase (PP1) is encountered, the assignment becomes *themeNP2*, with the role of PP1 depending on its composition.

If a past tense verb (VP2) is now encountered, the only possible interpretation for VP1 becomes the past participle. By this time, however, activation on the *pastVP1* role assignment node will have reached significant proportions, and will only respond slowly to the attempt to inhibit it, so the reversal of the semantic ambiguity is a gradual process. Upon successful reversal, the interpretation of NP1 is altered from *agent* to *beneficiary*.

One syntactic cue that will tend to increase the confidence of the past participle interpretation in mid-sentence is the occurrence of a prepositional phrase immediately following the verb, since *to buy* is transitive.

All of the above constraints are tabulated in Figure 3, and timing results of running the network on the sample stimuli are given. There are three classes of sentence: a straight-forward past tense, understood quickly, an unusual but clearly signalled past participle construction, understood after a moderate delay, and a misleading garden path involving a past participle, understood only after a lengthy reconsideration. Timing results reflect the time elapsed from reaching the end of the sentence until each of the correct thematic and semantic role assignment nodes reaches an activation level of at least 0.5.

Conclusions

We have described a model that accounts for a variety of results without appealing to Minimal Attachment strategies. The parser differs from that of Frazier in several respects. It utilizes parallel activation of morphological verb forms that makes available more than one parse option that can then be selected by context. The model accounts for the Non-minimal Attachment garden path as a function of the frequency asymmetry of the verb forms. This

NP_{agent}	'bought' _{past}	NP_{theme}	{PPs}	timing
The child	bought	a dog.		46
The child	bought	a dog	at the petstore.	47
The child	bought	a dog	for Mary.	48
NP_{agent}	'bought' _{past}	NP_{benef}	NP_{theme}	
The child	bought	the dog	a leash.	46
NP_{theme}	'bought' _{p.part}	PPs	$VP_{past \dots}$	
The child	bought	by the slaver	was sad.	62
NP_{benef}	'bought' _{p.part}	NP_{theme}	$VP_{past \dots}$	
The child	bought	a dog	was pleased.	109

Figure 3: Allowable syntactic forms of sentences using 'bought', and network times for successful comprehension.

would predict that the magnitude of the garden path would be related to this asymmetry, assuming that context is held constant. This model also predicts that preceding context can affect the parse by selecting, after lexical access, the appropriate form. As currently envisioned, context does not selectively affect access but merely functions as a post-access selection mechanism, thus allowing the model to be modular in nature.

The model is closer in spirit to how Altmann and Steedman [in press] have conceptualized sentence comprehension. Their contextual presuppositions serve to select the appropriate morphological verb form, thus reducing the garden path. The current proposal adds to Altmann and Steedman's model in that we have specified the mechanism by which context exerts its influence. The immediate empirical questions to be answered include tracing the timecourse of the activation and inhibition of the morphological verb forms and evaluating the locus of the contextual effect.

Neuropsychological Processes

Placing syntactic processing in the lexicon raises some intriguing questions concerning the locus of processing in garden path recovery. Burgess and Simpson [1988a] found that subordinate associates to ambiguous words are rapidly inhibited in the left hemisphere, but maintain activation over a relatively long period of time in the right hemisphere. Dominant associates maintain activation over time, but appear to decay in the right hemisphere. It is possible that these cerebral asymmetries may aid in semantic garden path recovery [Burgess and Simpson 1988b]. That is, if a sentence was not disambiguated until after the ambiguous word and the subordinate meaning were needed, the language processor would just access this meaning from the right hemisphere. This precludes the need to reactivate the meaning in the left hemisphere. If the syntactic lexical effects discussed above are bilateral in their representation, then one might expect a role for the right hemisphere in syntactic garden path recovery. While this is speculative, particularly since syntax is generally thought to be the domain of left hemisphere processing, Schneiderman and Saddy [in press] have reported that right brain damaged patients have difficulty inserting words into a sentence when the

insertion requires a syntactic reassignment of a word.

Summary

This model proposes a lexically located syntactic processor that makes available multiple morphological verb forms in parallel. It is modular in architecture, but differs from the modular serial parser proposed by Frazier. The model could easily be extended to allow for post-access contextual selection of the appropriate verb form, thus avoiding the Minimal Attachment garden path. The model suggests that the nature of this contextual effect, as detailed by Altmann and Steedman, is lexical selection. Several empirical issues involving frequency asymmetries of the verb forms, the speed of retrieval and inhibition, and interactions with context were discussed. This research is necessary in order to determine the correctness of the present implementation.

References

- [Altman and Steedman in press] Gerry Altman and Mark Steedman, "Interaction with Context during Human Sentence Processing", *Cognition*, in press.
- [Burgess and Simpson 1988a] Curt Burgess and Greg B. Simpson, "Cerebral Hemispheric Mechanisms in the Retrieval of Ambiguous Word Meanings", *Brain and Language*, 33:86-103, 1988.
- [Burgess and Simpson 1988b] Curt Burgess and Greg B. Simpson, "Neuropsychology of Lexical Ambiguity Resolution: Contribution of Divided Visual Field Studies", *Lexical Ambiguity Resolution in the Comprehension of Human Language*, Los Altos, CA: Morgan Kaufmann, 1988.
- [Burgess *et al.* in press] Curt Burgess, Michael Tanenhaus and Mark Seidenberg, "Context and lexical access: Implications of nonword interference for lexical ambiguity resolution", *Journal of Experimental Psychology: Learning, Memory and Cognition*, in press.
- [Cottrell 1985] Garrison W. Cottrell, "A Connectionist Approach to Word Sense Disambiguation", Technical Report 154, Department of Computer Science, University of Rochester, May 1985.
- [Ferreira and Clifton 1986] Fernanda Ferreira and Charles Clifton, "Independence of Syntactic Processing", *Journal of Memory and Language*, (25):348-368, 1986.
- [Frazier and Rayner 1982] Lyn Frazier and Keith Rayner, "Making and Correcting Errors during Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences", *Cognitive Psychology*, (14):178-210, 1982.
- [from Right Hemisphere Damage in press] A Linguistic Deficit Resulting from Right Hemisphere Damage, "E. Schneiderman and J.D. Saddy", *Brain and Language*, in press.
- [Goddard 1987] Nigel Goddard, "The Rochester Connectionist Simulator User Manual", Technical Report, Computer Science Department, University of Rochester, April 1987.
- [Marslen-Wilson and Tyler 1987] W. Marslen-Wilson and L. Tyler, "Against Modularity", *Modularity in Knowledge Representation and Natural Language Understanding*, pages 37-62, Bradford Books, MIT Press, 1987.
- [Seidenberg *et al.* 1982] Mark Seidenberg, Michael Tanenhaus, J.M. Leiman and M. Bienkowski, "Automatic Access of the Meanings of Ambiguous Words in Context: some Limitations of Knowledge-based processing", *Cognitive Psychology*, 14:489-537, 1982.
- [Simpson and Burgess 1985] Greg B. Simpson and Curt Burgess, "Activation and Selection Processes in the Recognition of Ambiguous Words", *Journal of Experimental Psychology: Human Perception and Performance*, 11:28-39, 1985.
- [Tanenhaus and Carlson 1988] Michael K. Tanenhaus and Greg Carlson, "Lexical Structure and Language Comprehension", *Lexical Representation and Process*, Cambridge, MA: MIT Press, 1988.
- [Tanenhaus *et al.* 1987] Michael K. Tanenhaus, Curt Burgess, Susan H. D'Zmura and Greg Carlson, "Thematic Roles in Language Processing", *Proceedings of the 1987 Conference of the Cognitive Science Society*, 1987.