

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

Arguments and Adjuncts: A Computational Explanation of Asymmetries in Attachment Preferences

#### **Permalink**

<https://escholarship.org/uc/item/7jz8d7xm>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

#### **Author**

Stevenson, Suzanne

#### **Publication Date**

1995

Peer reviewed

# Arguments and Adjuncts: A Computational Explanation of Asymmetries in Attachment Preferences

Suzanne Stevenson

Center for Cognitive Science  
and Dept. of Computer Science  
Rutgers University  
New Brunswick, NJ 08903  
suzanne@rucss.rutgers.edu

## Abstract

An explanatory model of ambiguity resolution in human parsing must derive a multitude of preference behaviors from a concise computational framework. One behavior that has been difficult to account for concisely is the preference to interpret an ambiguous phrase as an argument of a predicate, rather than as a modifier that is less integrally related to a phrase (an *adjunct*). Previous accounts of the argument preference have relied on assumptions about adjuncts requiring a more complex structure or entailing a delay in their interpretation. This paper explores a more fundamental distinction between arguments and adjuncts—that the number of potential arguments of a predicate is fixed, while the number of adjuncts for a phrase is unpredictable. This simple difference has important computational consequences within the competitive attachment model of human parsing. The model exhibits a preference for arguments over adjuncts due to the necessary differences in competitive properties of the two types of attachment site. The competitive differences also entail that adjuncts accommodate more easily than arguments to contextual effects. The model thus provides a concise and explanatory account of these argument/adjunct asymmetries, avoiding the unnecessary structural or interpretive assumptions made within other approaches.

## Introduction

In developing a model of human parsing, it is crucial to discover a small set of computational principles that can account for the human ability to effectively resolve linguistic ambiguities. An explanatory model must derive a multitude of observed human preference behaviors from a concise computational framework. One aspect of human syntactic processing that has proven resistant to integration within an explanatory account is the following argument attachment preference: People prefer to interpret an ambiguous phrase as an *argument* of a predicate—for example, as the direct or indirect object of a verb—rather than as an *adjunct*, which is a modifier of a phrase. As an example, consider a sentence beginning *Sara put the boxes on the table. . .*, in which the attachment of the prepositional phrase (PP) *on the table* is ambiguous. The PP can attach to the verb phrase as the location argument of the verb *put* (i.e., the place where Sara put the boxes), or to the noun phrase as an adjunct modifier of the noun *boxes* (i.e., the boxes are those that are on the table). In a choice of this kind between an argument and adjunct attachment for a phrase, the human parser shows a strong preference for interpreting the ambiguous phrase as an argument. Explaining the argument

attachment preference has been an important goal in modeling human parsing, and yet to date, no theory has given an account of this phenomenon that avoids *ad hoc* assumptions about adjuncts.

Some theories assume that an adjunct attachment requires a more complex structure than an argument attachment. The preference for the argument attachment then follows from a more general preference for building simpler syntactic structures (Frazier, 1978, 1990; Gorrell, 1995). However, a strict adherence to this type of structural complexity approach cannot easily accommodate the accumulating evidence that lexical, semantic, and discourse contexts affect initial attachment decisions (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Taraban & McClelland, 1990). Furthermore, the crucial assumption in such an account—that adjunct attachments require the on-line addition of more nodes into the parse tree—is not a computational necessity. For example, in parsing approaches based on pre-allocated  $\bar{X}$  templates (e.g., Lin, 1993; MacDonald et al., 1994; Stevenson, 1994a), the structure-building costs of arguments and adjuncts is equivalent.

Other models posit that the increased difficulty of processing adjuncts arises in the interpretive component of the language processor, rather than in the structure-building component. Some models formulate this as an explicit preference for thematic (argument) attachments (Abney, 1989; Gibson, 1991; Pritchett, 1992); related approaches assume that determining the thematic role of an argument in the discourse is faster than determining that of an adjunct (Crocker, 1992; McRoy & Hirst, 1990; Weinberg, 1991). None of these approaches depend on adjuncts having increased structural complexity. However, the evidence cited above that strongly supports the immediacy of contextual influences makes it less plausible that receiving a thematic role or not from a predicate would lead to a significant difference in complexity or speed of interpretation. Thus, the thematic/non-thematic distinction between arguments and adjuncts, like the structural distinction, appears unable to provide an explanatory account of the argument preference.

In fact, neither structural complexity nor interpretive difficulty is a *necessary* property of the adjunct modifying relationship. The proponents of both the structural and the interpretive approaches are building into these models the ob-

served argument/adjunct distinction, by explicitly formulating a tree-building or interpretive cost to be associated with adjuncts. On the other hand, a purely contextual approach relying entirely on differential frequencies of lexical associations (e.g., MacDonald et al., 1994) appears simply to restate the original problem, since presumably different frequencies are not accidental but arise from some more fundamental distinction.<sup>1</sup> The question then is: Are there any *necessary* properties of adjuncts that differentiate them computationally from arguments? Surprisingly, there is a very simple difference between arguments and adjuncts whose consequences have not previously been explored. Namely, nodes in a parse tree license an exact number of arguments—0, 1, 2, or 3—and each of those individual arguments may be optional or obligatory. By contrast, nodes can be modified by an arbitrary number of adjuncts—0 or more—each of which is always optional. Note that this is essentially the definition of what it means to be an adjunct as opposed to an argument.

In the competitive attachment parser (Stevenson, 1994a), this simple, fundamental difference between arguments and adjuncts leads to an important difference in computational properties between the two types of attachment sites. In the model, attachments are decided by a competitive process of spreading numeric activation through a massively parallel network, which directly represents the possible parse tree structures. A competition for activation at each attachment site focuses the activation within the network onto the preferred set of attachments. The competitive attachment process has been shown to provide an explanatory account of a number of human behaviors in processing syntactic ambiguities that involve argument attachments (Stevenson, 1993a, 1993b, 1994b). This paper discusses the extensions that are required to ensure the activation of the appropriate number of attachments to an argument or adjunct attachment site. The result is a necessary difference in the degree of competition at argument and adjunct sites, which leads to asymmetries in their behavior. One important consequence is that the competitive attachment model mimics the argument attachment preference observed in human parsing. The paper demonstrates additional asymmetries of argument and adjunct attachments, as well as the smooth integration of contextual preferences and recency effects into the relevant attachment competitions. The results extend the explanatory account of the model by showing how independently motivated properties of the competitive attachment mechanism concisely account for differences in the human parsing of arguments and adjuncts.

### The Competitive Attachment Process

The competitive attachment model is implemented within a hybrid connectionist framework, in which a parsing network directly represents syntactic phrases and the potential attach-

<sup>1</sup> Furthermore, a purely contextual account of the data thus far appears insufficient. For example, Hindle & Rooth (1993) found that a statistical model of PP attachment, based on lexical associations derived from a large corpus, achieved only 80% accuracy even when tested on (previously unseen) sentences from the same corpus.

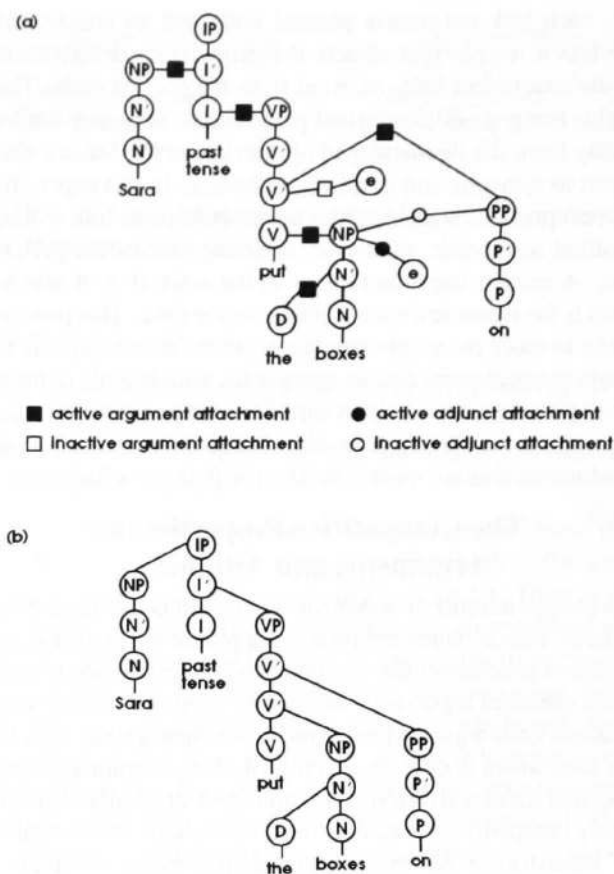


Figure 1: (a) The competitive attachment network parsing *Sara put the boxes on the table*, at the word *on*. Attachment nodes connect potential sisters in the parse tree; attachments to empty (*e*) nodes represent unfilled attachment sites. (b) The partial parse tree represented by (a).

ments among them. The network is created “on the fly” in response to a sequence of words; each word triggers the activation of a phrase and its initialization with the appropriate grammatical features. The current phrase is connected to the existing network with *attachment nodes*, which represent the potential attachment of the current phrase at each of the attachment sites along the right edge of the partial parse tree being developed; see Figure 1. Attachment nodes connect two potential sisters in the parse tree; compare the network of Figure 1(a) with the partial parse tree that it represents in Figure 1(b). After the current phrase has been connected to the developing parse tree structure, the network enters a spreading activation loop in which the attachment nodes compete for the available activation. When the network settles, the set of winning attachments represents the preferred parse tree structure for the input, up to and including the current phrase. The number of iterations required for the network to settle indicates the degree of difficulty in deciding on the current set of attachments.

Two factors that influence the outcome and speed of the attachment competition will be relevant to the results here.

First, each link between a phrasal node and an attachment node has a *weight* that affects the proportion of activation that the attachment node receives from the phrasal node. The weights integrate all contextual preferences; they may derive directly from the thematic grid of a lexical entry, but are also subject to semantic and discourse influence. In this paper, the different possible weight values on an attachment link will be classified as a weak, moderate, or strong contextual preference. A second important factor in the activation of attachments is the numeric *decay* of phrases over time. This process results in more recent phrases having more activation to contribute to attachment nodes, making the attachments to more recent phrases more likely to win. The competitive activation functions of the parsing network focus activation onto a set of attachments that are most consistent with these influences.

### The Competitive Properties of Arguments and Adjuncts

The precise number of arguments for a word is specified in the lexicon. Hence, when the nodes of a phrase are activated, an argument attachment site is established corresponding to each of the specified arguments. In order to satisfy its lexical specification, each argument attachment site must activate exactly one attachment node.<sup>2</sup> To accomplish this, the numeric functions that focus activation onto a preferred attachment must be highly competitive—i.e., they must have sharp “winner-take-all” behavior, so that only a single argument attachment node is activated.<sup>3</sup> Previous work has shown that the argument attachment activation functions achieve this necessary single-winner-take-all behavior (Stevenson, 1994a).

As mentioned above, adjuncts are fundamentally different from arguments in this regard, since an arbitrary number of adjuncts may modify a given phrase. Because the number of adjuncts of a phrase cannot be known ahead of time, a single attachment site for each potential adjunct cannot be allocated along with the phrase, in a manner similar to the allocation of argument attachment sites. The model is thus forced to allow an arbitrary number of attachments to be activated by a single adjunct site. (Whether that site is an *X'* or *XP* is irrelevant to the results here.) Because the competitive activation functions for arguments sharply focus activation onto a single attachment possibility, they are unable to support the simultaneous activation of multiple attachments. Adjunct sites must employ a much less competitive activation function—one which enables multiple adjunct attachments by allowing “multiple-winners-take-all” behavior. Perhaps surprisingly, no additional activation functions had to be added to the model to satisfy the requirements of adjuncts. The parser already incorporated a “multiple-winners-take-all” activation function for the *stack* data structure, which is used for proper input sequencing (Stevenson, 1994a). Thus, the required behavior for

<sup>2</sup>The attachment may be to an “empty” node if the argument is optional (or a *trace*); see Stevenson (1993b).

<sup>3</sup>Specifier attachment sites have the same requirement, since they also license exactly one attachment; thus, they are able to use the same activation functions as argument sites.

adjuncts is achieved through an existing and independently necessary mechanism of the model. Furthermore, the stack and adjunct activation functions are precisely the same as the argument functions, except for a parameter that adjusts the degree of competition at the attachment site.

As an example of multiple adjunct attachments, consider a sentence such as:

- (1) I [<sub>VP</sub> went [<sub>PP</sub> with Sara] [<sub>PP</sub> from London] [<sub>PP</sub> to Madrid] [<sub>PP</sub> on TWA]].

The verb phrase (VP) is modified by four separate prepositional phrases (PPs), necessitating multiple adjunct attachments to the VP. The first PP, *with Sara*, attaches to the VP in 17 iterations of the network. This is quite fast, representing an easy attachment for the parser; in fact, this is the same number of iterations required for the parser to decide on the attachment of a phrase to an obligatory argument site. In contrast to other models (e.g., Frazier, 1978; McRoy & Hirst, 1990), there is no inherent slow-down in the speed of syntactic operations solely due to the attachment of a phrase as an adjunct rather than as an argument, if there are no competing attachment sites. In processing the remainder of sentence (1), the next three PPs successfully activate attachments to the VP, in each case requiring 22 network iterations.<sup>4</sup> Thus, the degree of competitiveness of an adjunct attachment site is indeed mild enough to support attachment of arbitrary numbers of adjuncts, as required.

### Syntactic Ambiguity and Attachment Competitions

As explained above, potential attachments to an adjunct site must not compete too strongly with each other, or modification of a single node by multiple adjuncts would not be supported. The decreased level of competition means that the available activation at an adjunct site is spread more evenly across the potential attachments. During an attachment competition, nodes that connect to an adjunct site can maintain similar activation levels. This contrasts with argument sites, where the available activation must be focused onto a single attachment. These sites force a greater difference between the activation levels of the attachment nodes connected to them. The result is that one attachment node at an argument site typically becomes a much stronger attractor of additional activation, while all the attachment nodes at an adjunct site attract activation less strongly. The crucial consequence for the modeling of human behavior is that the decrease in competitiveness among attachment nodes for activation from the *same* adjunct site has the indirect effect of decreasing the ability of those attachment nodes to compete with attachments to *other* (adjunct or argument) sites. Thus, in situations of syntactic ambiguity, when more than one attachment site is competing to activate an attachment to the current phrase, asymmetries

<sup>4</sup>The attachment of additional adjuncts to the same phrase is slower than attachment of the first because there is an increase in competition for the activation being output from the VP.



in argument and adjunct attachment behaviors arise from their differing competitive properties.

### Competition between Two Adjunct Sites

First consider the following well-known example, in which two different adjunct sites compete for the attachment of a single phrase:

(2) I [<sub>VP</sub> saw [<sub>NP</sub> the man [<sub>PP</sub> with the telescope].

The VP *saw* and the NP *the man* compete for the attachment of the PP as an adjunct. Models of human parsing disagree on the preferred attachment in these cases. Some theories claim that there is a VP-attachment preference (Abney, 1989; Frazier, 1990), while others predict that the most recent (NP) attachment is preferred (Gibson et al., 1995; McRoy & Hirst, 1990). The competitive attachment model provides an account that captures the variability in intuition surrounding this structure. In the parsing network, the decay of activation of phrasal nodes entails a general recency effect (Stevenson, 1994b), leading to a preference for the PP to modify the more recent phrase (the NP). However, the decision to attach to the most recent phrase is affected by only slight changes in lexical or contextual preferences. If the VP shows even a slight preference to have a modifier (or the NP to *not* have a modifier), then the PP will attach as the adjunct of the VP instead of the NP, requiring only one or two more iterations than the attachment to the NP in the neutral context case. Thus, a slight change in preference can cause the parser to change the preferred attachment site, with minimal cost in the time it takes to make the attachment. By smoothly integrating both recency and contextual preferences, the model accounts for the evidence that, in these types of modification structures, appropriate contexts can easily change syntactic attachment preferences, with little or no penalty in processing time (Taraban & McClelland, 1990).

Lexical or contextual preferences can similarly change the preference to attach a phrase from the more recent to the less recent of two argument attachment sites. By contrast with adjunct attachments, however, it is quite costly to do so. The high degree of competitiveness of argument sites causes a slow-down of 25–35% for the network to settle on the less recent of two argument attachments (Stevenson, 1994a). The competitive attachment model thus predicts that adjunct attachments more easily accommodate to contextual influences than argument attachments. This behavior does *not* rely on an assumption regarding a difference in when thematic or semantic interpretation occurs, but rather results from necessary differences in the competitive properties of making adjunct versus argument attachments. Thus, behavioral asymmetries between arguments and adjuncts arise even when they are not in competition with each other.

### Competition between Argument and Adjunct Sites

Now let's turn to cases in which an adjunct attachment site competes with an argument attachment site. This type of competition may show up even within the same phrase, as in the

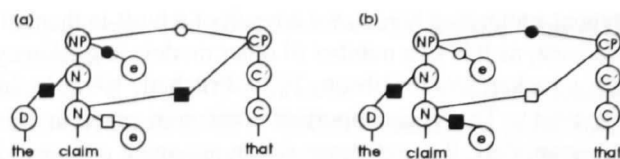


Figure 2: (a) Moderate-to-strong argument preference. (b) Weak argument preference.

NP beginning with *the claim* in:

- (3a) The evidence supported [<sub>NP</sub> the claim [<sub>CP</sub> that Sara is innocent]]
- (3b) The evidence supported [<sub>NP</sub> the claim [<sub>CP</sub> that the prospector disputed *t* ]]

In (3a), the CP (clausal phrase) is a complement clause that attaches as an argument of *claim*, while in (3b) the CP is a relative clause that attaches as a modifier of the NP. Thus, at the word *that*, the parser is choosing between an argument and adjunct attachment site for the CP. The competitive attachment model does not make an adjunct attachment less desirable in direct comparison with an argument attachment; there is no built-in difference in their weights or activation levels. However, as described above, the decrease in competitiveness among attachments at the adjunct site has the indirect effect of making the adjunct site less able to compete with other attachment sites. Thus, the network settles on the argument attachment for the CP, exhibiting an argument attachment preference, as in Figure 2(a).

However, the adjunct attachment possibility is not irrelevant, as can be seen when lexical or contextual preferences are taken into account. In this configuration, if the contextual preference for the argument is moderate to strong, then the current phrase will attach as an argument. On the other hand, if the preference for the argument is weak, the phrase will instead attach as an adjunct; see Figure 2(b). Again we see that lexical or contextual preferences may affect the parser's attachment decision. However, in a choice between an argument and an adjunct attachment, the greater competitiveness of the argument site entails that a significant change in external preference is required in order to overcome the argument attachment preference and select the adjunct attachment instead. The choice of argument or adjunct attachment is thus susceptible to contextual influences, but slight changes in preference are insufficient to shift activation away from the highly competitive argument attachments.

It is important to emphasize here that the preference for the argument attachment arises solely from independently motivated properties of the competitive activation functions, and not from some inherent difference in value assigned to argument and adjunct attachments. In the competitive attachment model, adjuncts are assumed to satisfy grammatical constraints as well as arguments (in contrast to approaches such as in Gibson (1991) or Pritchett (1992)), giving them equally high *a priori* activation levels. The demonstrated argument at-



the obligatory argument attachment for *on the table*, attaching the PP to the higher verb *put*, thereby *not* avoiding the garden path. However, if the lower verb *threw* has even a slight preference to occur with the optional argument, then the first PP is attached to that verb and the garden path is avoided. Thus, in the competitive attachment model, the ability to avoid the garden path relies not on a discrete fact of the lower verb being able to take a PP argument, but rather on the preference of the lower verb to do so. Although I know of no experimental evidence that bears on this, I believe the prediction of the competitive attachment model is more likely to be borne out, given the evidence for lexical and semantic influences on attachment preferences.

### Conclusions

This paper has demonstrated that the competitive attachment model exhibits asymmetries in processing argument and adjunct attachments due to the competitive properties of the activation functions underlying the two types of attachment. The advantage claimed for the explanation given here compared to other approaches is that the competitive attachment model does not rely on building in dubious assumptions regarding the structure and interpretation of adjuncts. An obvious question to ask is why the account is more explanatory, if the difference between arguments and adjuncts is also “built-into” the competitive attachment parser, by providing argument and adjunct attachment sites with different spreading activation functions.<sup>7</sup> The crucial point is that the difference in processing mechanisms is necessary to accommodate a fundamental distinction—that the number of potential arguments of a predicate is fixed, while the number of adjuncts for a phrase is unpredictable. The model thus has an explanatory advantage over approaches in which argument/adjunct asymmetries arise from controversial structural properties of adjuncts, or from unnecessary assumptions regarding differential speeds of interpretation.

### References

Abney, S. (1989). “A Computational Model of Human Parsing.” *Journal of Psycholinguistic Research* 18:1, 129–144.

Crocker, M. (1992). “A Logical Model of Competence and Performance in the Human Sentence Processor.” Doctoral dissertation, University of Edinburgh.

Frazier, L. (1978). “On Comprehending Sentences: Syntactic Parsing Strategies.” Doctoral dissertation, University of Connecticut.

Frazier, L. (1990). “Parsing Modifiers: Special Purpose Routines in the Human Sentence Processing Mechanism.” In D. Balota, G. Flores d’Arcais, & K. Rayner, Eds., *Comprehension Processes in Reading*, Hillsdale, NJ: Lawrence Erlbaum.

Gibson, E. (1991). “A Computational Theory of Human Linguistic Processing: Memory Limitations and Processing

Breakdown.” Doctoral dissertation, Carnegie-Mellon University.

Gibson, E., N. Pearlmutter, E. Canseco-Gonzalez, and G. Hickok (1995). “Recency Preference in the Human Sentence Processing Mechanism.” Manuscript, MIT.

Gorrell, P. (1995). *Syntax and Parsing*. Cambridge University Press.

Hindle, D. and M. Rooth (1993). “Structural Ambiguity and Lexical Relations.” *Computational Linguistics* 19:1, 103–120.

Kimball, J. (1973). “Seven Principles of Surface Structure Parsing in Natural Language.” *Cognition* 2:1, 15–47.

Lin, D. (1993). “Principle-Based Parsing without Overgeneration.” *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, 112–120.

MacDonald, M., N. Pearlmutter, and M. Seidenberg (1994). “Lexical Nature of Syntactic Ambiguity Resolution.” *Psychological Review* 101:4, 676–703.

McRoy, S. and G. Hirst (1990). “Race-Based Parsing and Syntactic Disambiguation.” *Cognitive Science* 14, 313–353.

Pritchett, B. (1992). *Grammatical Competence and Parsing Performance*. University of Chicago Press.

Spivey-Knowlton, M., J. Trueswell, and M. Tanenhaus (1993). “Context Effects in Syntactic Ambiguity Resolution: Discourse and Semantic Influences in Parsing Reduced Relative Clauses.” *Canadian Journal of Experimental Psychology* 47:2, 276–309.

Stevenson, S. (1993a). “Establishing Long-Distance Dependencies in a Hybrid Network Model of Human Parsing.” *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, 982–987.

Stevenson, S. (1993b). “A Competition-Based Explanation of Syntactic Attachment Preferences and Garden Path Phenomena.” *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, 266–273.

Stevenson, S. (1994a). “A Competitive Attachment Model for Resolving Syntactic Ambiguities in Natural Language Parsing.” Doctoral dissertation, University of Maryland. Rutgers Center for Cognitive Science, TR-18.

Stevenson, S. (1994b). “Competition and Recency in a Hybrid Network Model of Syntactic Disambiguation” *Journal of Psycholinguistic Research* 23:4, 295–322.

Taraban, R. and J. McClelland (1990). “Parsing and Comprehension: A Multiple-Constraint View.” In D. Balota, G. Flores d’Arcais, & K. Rayner, Eds., *Comprehension Processes in Reading*, Hillsdale, NJ: Lawrence Erlbaum.

Weinberg, A. (1991). “A Parsing Theory for the Nineties: Minimal Commitment.” Manuscript, University of Maryland.

<sup>7</sup>In fact, as mentioned earlier, the same activation functions are used, with only a parametric difference in the degree of competition induced.