A Constraint-Based Lexicalist Account of the Subject/Object Attachment Preference

Cornell Juliano* and Michael K. Tanenhaus

When a noun phrase could either be the object of the preceding verb or the subject of a new clause or a sentence complement, readers and listeners show a strong preference to parse the noun phrase as the object of the verb. This can result in clear garden paths for sentences such as The student read the book was stolen and While the student read the book was stolen. Even when the verb does not permit a noun phrase complement, some processing difficulty is still found. This has led some theorists to propose models in which initial attachments are lexically blind, with lexical information subsequently used as a filter to evaluate and revise initial analyses. In contrast, we show that these results emerge naturally from constraint-based lexicalist models. We present a modeling experiment with a simple recurrent network that was trained to predict upcoming complements for a sample of verbs taken from the Penn Treebank corpus. The model exhibits an object bias and it also shows effects of verb frequency which are similar to those found in the psycholinguistic literature.

This article is a preliminary report on a modeling experiment we are conducting to evaluate the plausibility of a constraint-based lexicalist account of attachment preferences involving subject/object ambiguities. Consider the examples in 1a and 1b:

1a. The student read the book was stolen.
1b. While the student read the book was stolen.

We wish to thank Michael Spivey-Knowlton for helpful comments. This work was supported by NIH grant HD-27206.

* University of Rochester, Rochester, New York.
\[2 Address all correspondence to Cornell Juliano, Department of Psychology, Meliora Hall, University of Rochester, Rochester, New York 14627-0038.
In both of these sentences, the noun phrase the book is initially taken to be the object of the verb read, resulting in processing complexity, and for many people, a conscious feeling of being garden-pathed when was is encountered (e.g., Frazier & Rayner, 1982). The garden-path effect is clearly much stronger in sentence 1b than it is in sentence 1a.

A surprising result is that readers experience processing difficulty even when the verb does not permit an NP complement, as is illustrated in example 2a:

2a. The student implied the book was stolen.
2b. The student implied that the book was stolen.
2c. The student read the book was stolen.

Under these conditions processing difficulty occurs at the noun phrase (Garney, Lotocky, & McConkie, 1992; Juliano & Tanenhaus, 1993; True- swell, Tanenhaus, & Kello, 1993), though it may spill over onto the verb in the complement (Ferreira & Henderson, 1990). Reading times at the noun phrase the book would be longer compared to either a control sentence like 2b in which there is a complementizer or a sentence like 2c in which the verb is typically used with an object.

The sentences in 3 illustrate the effect demonstrated by Mitchell (1987, 1989; also Adams, Mitchell & Clifton, 1993):

3a. While the student hesitated the book was stolen.
3b. While the student hesitated, the book was stolen.
3c. While the student bought the book her parents waited in the car.

In the absence of a comma, readers have difficulty processing the subject of a main clause when it immediately follows the verb in the subordinate clause, even when the verb does not subcategorize for an NP complement. The comparisons which reveal this difficulty are between sentences like 3a and either 3b, which has a comma after the verb, or 3c in which the noun phrase is the object of the verb. The only possible situation where difficulty does not obtain in the commaless sentences is when the main clause begins with a case marked pronoun (Trueswell, personal communication).

Both the results for sentence complements and subordinate/main clause constructions might appear to present serious problems for any parsing model that makes immediate use of lexical information (e.g., Abney, 1989; Pritchett, 1993; Gibson, 1991). Rather, they seem to provide striking support for two-stage parsing models in which initial attachments are made on the basis of category information (Clifton, Speer, & Abney, 1991; Frazier, 1987;