Right Association in Parsing and Grammar*

Colin Phillips, *Massachusetts Institute of Technology*

1. Introduction

It is straightforward enough to show that sentence parsing and grammaticality judgements are different. There are sentences which are easy to parse but ungrammatical (e.g. *that-trace* effects), and there are sentences which are extremely difficult to parse, but which may be judged grammatical given appropriate time for reflection (e.g. multiply center embedded sentences). This classic argument shows that parsing and grammar are not identical, but it tells us very little about just how much they have in common. The goal of this paper is to point to one principle which may account for a number of effects of ambiguity in both domains.

I have in mind two closely related kinds of ambiguity. Both occur in situations in which structure is underdetermined by meaning. The first kind of ambiguity is familiar from the literature on sentence comprehension: there are points in the comprehension of sentences where the input is consistent with more than one structure. Not knowing the speaker's intended meaning, the comprehender must make a structural choice independent of meaning. This is a typical instance of a *processing ambiguity*. There is, however, an additional kind of situation in which structural choices are undetermined by meaning, one which has received less attention. When I say something, knowing the meaning that I intend to convey and the words that I intend to use may not be sufficient to determine the structure that my grammar assigns to them. There could be more than one way of structuring the same set of elements such that the semantic rules of my grammar assign the intended interpretation to the sentence. This situation I will refer to as a *syntactic ambiguity*. The goal of this paper is to show that the resolution of both of these kinds of ambiguity is governed by a version of the well-known principle of Right Association (Kimball 1973).

Many different structural parsing strategies have been proposed in the sentence processing literature. Most have been quite controversial. But there has

been one remarkably uncontroversial strategy, namely Right Association: the idea that all other things being equal, incoming syntactic material is preferentially construed with recently built structures. The idea has surfaced in various guises under various names, but the basic idea has been preserved throughout. In the parsing sections of this paper my aim is to show that when properly formulated, this most uncontroversial of all parsing strategies has far broader empirical coverage than is generally assumed.

My main empirical claim about processing ambiguities is that all structural parsing preferences can be reduced to the generalization in (1).

(1)　　　　All other things being equal,[1] build the most right-branching structure available.

The paradigm case of a 'right-branching' structure is a tree in which all branching nodes are the right-hand sister of a terminal node, as in (2). I will come shortly to what properties make one structure more right-branching than another.

(2)
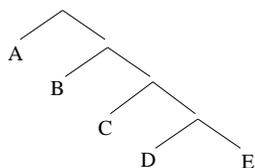
I claim that the generalization in (1) is a consequence of the principle of Right Association, which determines structural choices in processing ambiguities. My version of Right Association, which differs from previous formulations in some important respects, is shown in (3). It requires the last head attached and the incoming head to be as close as possible to one another.

(3)　　　　*Right Association*[2]
　　　　　　For any pair of overt terminal elements X Y, where X is the last terminal attached, reject attachments of Y which require unnecessary steps in the path between X and Y.

---

[1] Factors that might prevent all other things from being equal include biases due to semantics, discourse and frequency.

[2] Fodor and Frazier (1983) propose a closely related principle, which they dub *Minimal Connections*. Fodor and Frazier show how a number of standard attachment preferences (much of the material discussed in sections 4 and 5 of this paper) can be handled by the single principle of *Minimal Connections*.

Right Association in Parsing and Grammar
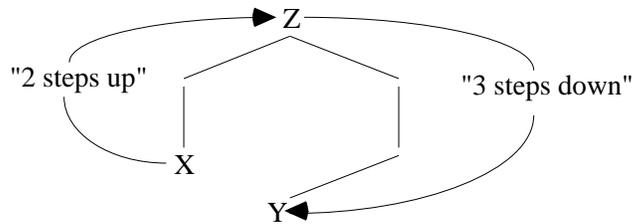
The details of how this principle is implemented in specific cases will unfold as the paper proceeds. If successful, the generalized version of Right Association both subsumes much of the work of more controversial strategies like *Minimal Attachment* (Frazier 1978; Frazier & Fodor 1978) and *Prefer Argument Attachments* (Abney 1987, 1989; Ford, Bresnan & Kaplan 1982; Pritchett 1988, 1992; Gibson 1991) and in a number of instances makes divergent predictions. It also provides a principled account of some asymmetries in gap filling preferences in questions and relative clauses, including contrasts between a number of languages, and some previously unexplained contrasts between local and long-distance extraction.

In the second part of the paper I try to show that the same principle of Right Association may provide an account of some puzzling recent results in the theory of phrase structure, in particular some apparent contradictions pointed out by Pesetsky (1995). Pesetsky shows that the results of some fairly basic syntactic constituency tests indicate that phrase markers are not only more right-branching than has standardly been assumed, but right-branching to such an extent that core assumptions about the structural relationship between heads and their complements must be violated. Meanwhile, the results of other standard tests indicate quite the opposite, that the same phrase markers appear to be radically left-branching. I suggest an account of these contrasting effects which relies heavily on the principle of Right Association, and which avoids postulating multiple syntactic representations for every sentence.

1.1   Two caveats

I should stress at the outset something that this paper is *not* trying to do. I am not trying to show that structural factors dominate pragmatic or probabilistic factors in normal parsing—there is plenty of evidence that these factors do play an important role in language understanding. Pragmatic and probabilistic influences are quite consistent with my system, but they are not everything. I am merely assuming that in situations where these other factors have been properly balanced, the human processor does not just give up, instead it makes a choice based on structural preference. My focus here will be on showing that there are interesting structural commonalities among the choices that the parser makes in a wide variety of situations.

At this point the reader may also be concerned that I am ignoring a fundamental difference between parsing preferences and grammaticality

judgements. Parsing preferences are soft and violable. They just say that one structure is *better* than another, and the preferences can be weakened or even reversed by appropriate contextual biases. Grammaticality judgements, meanwhile, are hard and inviolable. They say that a structure is either *possible* or *impossible*, and do not tolerate in-between cases. Given this distinction, then, is it not misguided to try and attribute grammaticality judgements and parsing preferences to the same principle?

I contend that there is no conflict here, once we focus on truly ambiguous situations. When structure is underdetermined by meaning, whether in sentence generation or sentence comprehension, the principle of Right Association always dictates which structure *must* be chosen. The structure in which precedence and c-command relations most closely match up is the only available option. In syntactic ambiguities this austere character of Right Association is easy to see, because these are situations where complete knowledge of the intended meaning is entirely compatible with two structures. In processing ambiguities, on the other hand, the appearance of 'softness' arises because it is rare to find a situation in which knowledge of meaning genuinely underdetermines the choice between different possible structures. Where meaning is truly underdetermined, however, Right Association should apply just as clearly as in grammaticality judgements. The softness of parsing preferences is observed in situations of *partial* ambiguity, where knowledge of meaning does in fact favour one structure over another. In other words, *it is ambiguities that are soft, and not the parser*.

1.2   Why Right Association?

The principle of Right Association and its descendants have traditionally been viewed as the implementation of working memory constraints favouring association of new material with recent material, and it is probably in large part due to this grounding in well-motivated cognitive constraints that Right Association has been so uncontroversial. My use of Right Association here will be primarily as a condition which minimizes the resources devoted to representing syntactic structures. This version of Right Association will entail the recency favoring property of earlier formulations of the principle, but now as only a subcase of a general principle which simplifies syntactic representations.

I make the (relatively uncontroversial) assumption that when we represent a sentence we need to represent information about both the linear order of elements (i.e., precedence relations)[3] and hierarchical relations among elements (i.e., c-command relations[4]). In a strictly right-branching structure like

---

[3] This is the controversial part of the assumption, in the light of recent attempts by Kayne (1994) and others to show that linear order is entirely predictable from hierarchical structure.

[4] Dominance relations are of course another way of expressing hierarchical relations among the nodes of a phrase marker. However, since we will be mostly concerned

Right Association in Parsing and Grammar

(2) there is complete correspondence between precedence relations and c-command relations: if a terminal X precedes terminal Y, then X also c-commands Y.[5] As an extension of this, I assume that the greater the correspondence between precedence relations and c-command relations, the more right-branching a structure is: mismatches between precedence and c-command make a structure less right-branching. Why should it matter how right-branching a structure is? I suggest that it is crucial because less representational resources are used up when precedence relations and c-command relations match-up — perhaps simultaneous precedence and c-command between two elements can be encoded as a single relation. Therefore the reason why right-branching structures are preferred is because they reduce the resources needed to represent a sentence.

The effect of the principle of Right Association is to choose the most right-branching structure available at any choice point in structure building. The longer the path between a pair of adjacent terminals, the more precedence/c-command mismatches there will be, with one exception. Let me illustrate — in all of the examples in (4) A and B are adjacent terminals, and A precedes B, but the length of the path between A and B (boldface branches) varies.

(4) a.   b.
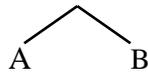
A-B: 1 step up, 1 step down   A-B: 2 steps up, 1 step down
A precedes B   A, C precede B
A c-commands B   A, C *do not* c-command B
  2 steps, 0 mismatches   3 steps, 2 mismatches

---

with hierarchical relations among terminal elements here, I restrict attention to c-command relations here.

[5] The one exception, of course, is the final terminal in the sentence (node E in (2)) which c-commands the preceding terminal, but does not precede it.

c.   d.
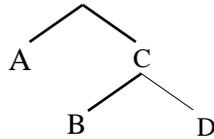
A-B: 1 step up, 2 steps down   A-B: 1 step up, 3 steps down
A precedes B   A, B, E precede F
A c-commands B   A c-commands F
    3 steps, 0 mismatches   B, E *do not* c-command F
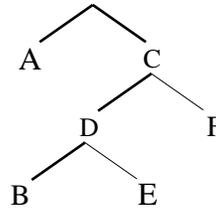   4 steps, 2 mismatches

Case (4c) is the one situation where lengthening the path from A to B does not entail a mismatch of precedence and c-command relations. However, it will become apparent later in the paper that when paths from A-B like (4c) are required at an intermediate stage in structure building, this is generally a stage in the construction of a tree like (4d), in which there are precedence/c-command mismatches: node B in (4c) is almost always followed by additional material which forces B to be expanded to form a complex specifier, in which the path between A and B is as long as in (4d) or longer. Therefore, since longer paths between adjacent terminals almost always entail precedence/c-command mismatches, a structure building principle which attempts to minimize the lengths of paths between adjacent terminals will be a fairly effective way of ensuring that a maximally right-branching structure is built, and representational resources are spared. This is what I suggest that Right Association achieves.

The paper is organized as follows. Section 2 shows why it was thought in the 1970s that Right Association applies only in a narrow range of situations in processing, and why as a consequence additional parsing strategies must be introduced. Section 3 motivates the syntactic analyses which are crucial for section 4. In section 4 I show how my version of the Right Association gets around the limitations noticed in the 1970s and how it accounts for both standard Right Association preferences and preferences commonly attributed to *Minimal Attachment* or argument attachment strategies. Section 5 discusses a number of ambiguities where the predictions of Right Association differ from the predictions of standard structural parsing theories, apparently in favour of Right Association. Section 6 shows how Right Association accounts for a range of gap-filling preferences in English, Dutch, German and Italian, including dependencies of different lengths and types. In section 7 I turn to the role of Right Association in grammar, and shows how it may explain both the apparent coexistence of left-branching and right-branching structures, and why different syntactic constituency tests give the results they do. Section 8 concludes.

2.   How general is Right Association: "Kimball's Problem"

Right Association is the second of Kimball's (1973) "Seven Principles of surface structure parsing" (5).

(5)   *Right Association* (Kimball 1973)
         Terminal symbols optimally associate to the lowest non-terminal node.

Kimball's Right Association principle favors attachment of new material to recently built structure, by requiring that the lowest node dominating both the new terminal and the immediately preceding terminal be as low as possible. In terms of the path between these two terminals, then, it requires that there be as few 'steps up' as possible, but does not care about the number of 'steps down'.[6] This principle accounts for a number of fairly robust biases in construing new words or phrases with partially constructed sentences, as the examples in (6–8) show. In (6) the particle *out* can be construed with either the matrix verb *figured* or the embedded verb *take*, and there is a strong preferencefor the embedded construal, even when this leads to an extremely implausible interpretation, as in (6b). In (7–8) the phrases *'s friend* and *that was fluffy* can be construed with one of two NPs, and in each case the lower and more recent choice is taken, again overriding plausibility.

(6)   a.   Joe figured that Sue wanted to take the cat out.
       b.   Joe figured that Sue wanted to take the train to New York out.

(7)   a.   I met the boy who Sam introduced to Mary's friend.
       b.   I met the boy who Sam took to the park's friend.

(8)   a.   The cat slept on the cushion that was fluffy.
       b.   The cat let out a huge meow that was fluffy.

However, there are some cases where the principle of Right Association seems to make the wrong predictions. In the examples in (9) the phrases *for Susan*, *to Mary*, and *that Bill knew* can all be either an argument of the verb, or a modifier of the NP following the verb. In each case most speakers' intuitive preference is to construe the ambiguous phrase as an argument of the verb, which under Kimball's syntactic assumptions meant that the ambiguous phrase was attached higher than it would be if it was modifying the NP.

(9)   a.   Joe bought the book for Susan.
       b.   John read the letter to Mary.
       c.   John told the girl that Bill knew to leave.

---

[6] In terms of the diagram in (3), 'steps up' corresponds to the path from node X to node Z, and 'steps down' corresponds to the path from node Z to node Y.

Kimball was aware of this, and his interpretation of the problem was that the Right Association preference holds only when the competing attachment sites are some distance apart, as in (6–8). Fodor and Frazier (1978) essentially follow this interpretation of the problem, and propose the principle of *Minimal Attachment* to account for the verb attachment preferences in the examples in (9). Minimal Attachment prefers attachments which minimize the number of new nodes built, and therefore attachment as an argument of the verb is preferred over attachment as a modifier of the verb in (9), since Chomsky-adjunction of the modifier to the NP requires the insertion of a new NP node, whereas an attachment site for VP-arguments is already available, without creation of new nodes.

Many other structural parsing theories incorporate a principle which does more or less the same work as Kimball's Right Association: eg. *Late Closure* (Frazier 1978, 1987), *Recency* (Gibson 1991), *Connect Bottom-up* (Stabler 1994), or *Low Attachment* (Abney 1987, 1989). Therefore, it is relatively uncontroversial that something along the lines of Right Association plays a role in determining structural parsing preferences.[7]

The aim of sections 2 and 3 is to show that the verb attachment preferences in (6) may not present a problem for the Right Association strategy after all. The first step is to review some syntactic evidence that multiple complement constructions have a rather different structure from what has generally been assumed in the sentence processing literature.

3. Right-branching structures in syntax

Traditional analyses of VPs containing multiple complements have assumed either (i) that the two arguments of the verb are sisters in a ternary branching structure like (10a), or (ii) that the second argument asymmetrically c-commands the first argument, as in (10b). A number of results from the past 10 years show that by applying quite basic constituency tests to these structures, we are led to a rather different picture of the structure of VP.

---

[7] Cuetos and Mitchell (1988) question the generality of the Right Association strategy, based on differing preferences between English and Spanish speakers in processing temporary ambiguities like *the daughter of the colonel who . . .* English speakers have a slight low attachment preference (*who* modifies *colonel*), whereas Spanish speakers have an equally slight high attachment preference (*who* modifies *daughter*). Some recent papers have corroborated this basic finding, but shown that it would be an oversimplification to claim that the Spanish speakers are showing an anti-recency effect (Gilboy et al. 1995; Gibson et al. 1995). These studies show that Spanish speakers do show recency preferences, but that these effects are sometimes obscured by other factors. I have nothing new to contribute to this puzzle here.

(10) a.   b.
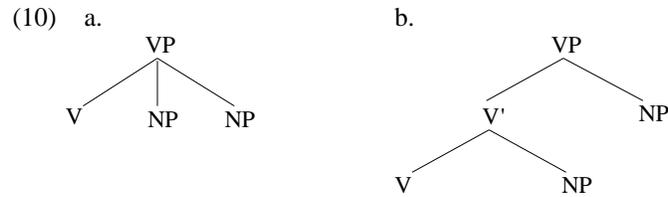
First, Barss and Lasnik (1986) showed that a series of well-known tests for c-command relations all indicate that the first object in a double object construction c-commands the second object, but not vice versa. Their tests included reflexive binding (11), bound variable anaphora (12), superiority (13) and negative polarity item licensing (14). These results are quite unexpected under the traditional analyses of double object constructions, which predict either mutual c-command (10a), or that the second object c-commands the first (10b).

(11) a.   I showed John himself in the mirror.
      b.   *I showed himself John in the mirror.

(12) a.   I denied each worker his paycheck.
      b.   *I showed its trainer every lion.

(13) a.   Who did you give which book?
      b.   *Which book did you give who?

(14) a.   I gave no one anything.
      b.   *I gave anyone nothing.

Second, Larson (1988) showed that the two objects in a double object construction can be coordinated to the exclusion of the verb (15). Given the standard assumption that only constituents can be coordinated, this implies that the two objects of the verb form a constituent, which neither of the structures in (10) predict.

(15) a.   John sent [a letter to Mary] and [a book to Sue]
      b.   John sent [Mary a letter] and [Sue a book]

The binding and coordination facts imply a structure for double objects like (16), where the two objects are both contained in a single constituent which is sister of the verb—this captures the coordination facts—and in which the second object is sister of a further constituent — this captures the findings about asymmetric c-command. A widespread assumption is that this further constituent is the intermediate level projection (X') of the complement of the verb, and that the two objects are its specifier and complement respectively. There is less consensus about what the category of the head of this constituent is, and its precise identity will not be crucial at this point. For the sake of exposition,

9

though, I will follow Larson (1988) in assuming that it is the trace of the moved verb.

(16)
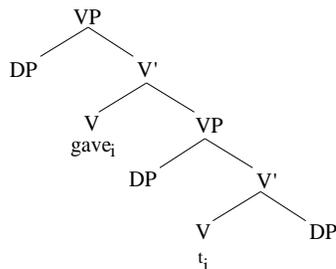
As Larson (1988) and Pesetsky (1995) observe, the binding and coordination tests don't only point to a revised structure for predicates with multiple arguments. They also indicate that a new structure is required for sets of non-arguments following a verb. When the same tests in (11–15) are applied to either multiple adjuncts, or to combinations of arguments and adjuncts, they yield the same result: the linear ordering of phrases predicts the c-command relations among them. Phrases on the left c-command phrases to their right, and groups of arguments and adjuncts at the right end of VP always behave like a constituent for coordination (17–18, from Pesetsky 1995).

(17) a. Sue gave books to these people$_i$ on each others$_i$ birthdays.
      b. Sue spoke to Mary about these people$_i$ in each others$_i$ houses on Tuesday.

(18) a. Sue will speak to Mary [about linguistics on Friday] and [about philosophy on Tuesday]
      b. Kremer will perform this concerto [in Rome on Tuesday] and [in Somerville on Thursday]

These facts imply that the right-branching structure for double object constructions in (16) should be extended to verb-argument-adjunct and verb-adjunct-adjunct sequences. This observation will become important in later sections.

The next step is to show how these syntactic results affect our understanding of processing strategies.

## 4.   Generalizing Right Association

4.1   A single strategy

Kimball's Right Association principle requires that incoming material be attached at the lowest available site. As pointed out above, this has the effect

of minimizing the length of the path from the last terminal attached to the lowest node dominating both that terminal and the new terminal. I propose a slight modification of Right Association, according to which incoming terminals are required to be as close as possible to the most recently attatched terminal, and the length of the *entire path* between the two terminals matters not only the 'steps up' as in Kimball's formulation. In ambiguous situations, then, the preferred structure is the one in which the path between the last word heard and the word currently being processed is as short as possible. The principle is given in (3), repeated as (19).

(19)   *Right Association*
           For any pair of overt terminal elements X Y, where X is the last terminal attached, reject attachments of Y which require unnecessary steps in the path(s) between X and Y.
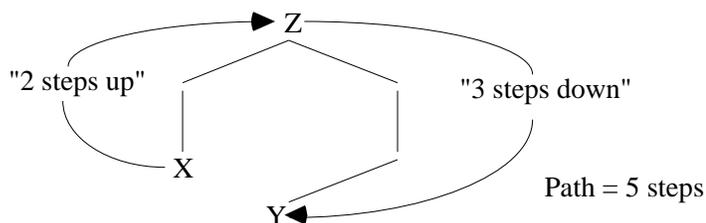
X and Y are phonetically adjacent terminals. In many cases X and Y will also be syntactically adjacent terminals, but this will not always be the case, because phonetically null terminals will sometimes need to be postulated between X and Y. In these cases, Right Association will need to evaluate multiple paths between the syntactically adjacent terminals between X and Y. Such cases will be discussed shortly.

4.2   Traditional Right Association preferences

Not surprisingly, the modified version of Right Association in (19) makes correct predictions for preferences traditionally attributed to Right Association, like those in (6–8) above, and (20) below. In each case a particle or modifier can potentially be attached at one of two sites. In both alternatives the ambiguous phrase is attached to the same category and has the same status as argument/adjunct/predicator. However, the path between the ambiguous word and the preceding word is much shorter in the preferred parse. (21) illustrates the

alternative structures for (20a),[8] and the length of the paths involved.[9] In the list of path lengths following (21) I adopt the convention used throughout this paper of marking fatal paths in boldface text.

(20) a.   John said Bill left yesterday.
      b.   The reporter interviewed the daughter of the colonel who had had the accident.

(21) a.   b.
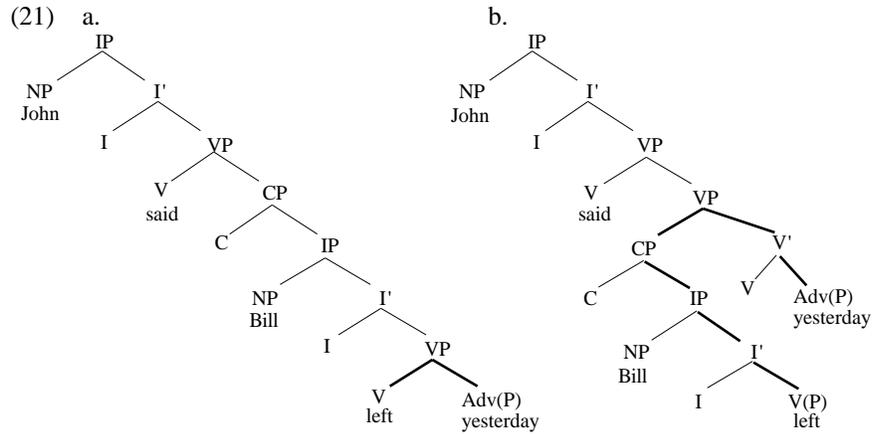
low attachment: high attachment:
2: V-VP-Adv(P)   **6: V(P)-I'-IP-CP-VP-V'-Adv**

        Examples (22–23) show further temporary ambiguities for which preferences have standardly been attributed to Right Association or one of its variants. The sentences in (22) involve ambiguous scope of the conjunction *and*, for which the intuitive preference is always to favour the low attachment, narrow scope for *and*. In these cases the difference between the alternative structures in *steps up* means that low attachment of *and* entails a shorter path from the preceding word to *and*.

(22) a.   John kissed <u>Mary and</u> her sister laughed.
      b.   The <u>janitors and</u> doctors who supported the strike picketed in large numbers.[10]

---

[8] I am assuming throughout that phrase markers contain only terminal nodes and branching nodes, i.e. that there are no vacuous projections, along the lines of Chomsky (1994). This is why the verb *left* in (21b) is labelled as a V(P) node: it is both a terminal node and the maximal projection of V.

[9] For expository purposes I am ignoring the phonetically null V node that is sister of *yesterday* in (21b), since it does not play a crucial role in these examples. The role of such nodes will become clear in the section 4.3.

[10] If *and* coordinates at the level of NP the relative clause restricts both the janitors and the doctors; if *and* attaches at the level of DP, though, the relative clause only

In (23) the determiner *the* attaches to a different category with a different role in the two alternative structures to be evaluated. In the first (24a), *the* heads the subject of the matrix clause, and in the other *the* heads the direct object of the verb *mending*. Therefore there is a small difference in *steps down*, slightly favouring the direct object attachment. But much more important than this is an enormous difference in *steps up*, again favouring the direct object attachment. Therefore Right Association correctly predicts (23) to be a garden path. The alternative structures are shown in (24).

(23)   While Mary was <u>mending the</u> sock fell off her lap.

(24)   a.   b.
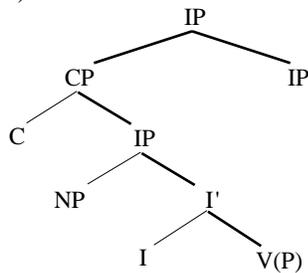
high attachment: low attachment:
6: V(P)-I'-IP-CP-IP-IP-D(P)   2: V-VP-D(P)
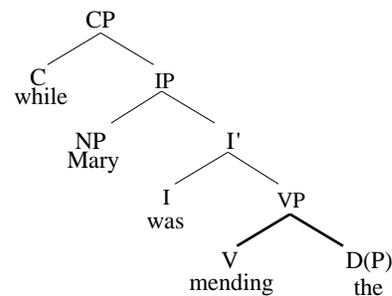
4.3   Null heads

So far we have only considered cases in which the last word to be attached and the incoming word are syntactically as well as phonetically adjacent. But if there are some syntactic terminal elements which are not pronounced, as is almost universally assumed,[11] then the parser will encounter situations in which a null head must be postulated between the last overt word and the incoming overt word. In other words, two syntactic heads must be introduced in order to attach a single overt head. The question then arises of how these cases are handled by Right Association.

I assume that when null heads need to be postulated, Right Association breaks the path between the overt heads X and Y in (3/19) into paths between each pair of syntactically adjacent heads, and measures the paths between each

---

restricts the bare plural DP *doctors*. I assume that relative clauses have the status of right-hand specifiers of DP.

[11] Of course the existence of certain phonetically null categories has been extremely controversial, eg. PRO, or NP-trace. However, many fewer people are squeamish about the possibility of elements like a null complementizer that alternates with *that*.

pair of *syntactically* adjacent heads. In other words, Right Association measures the paths between null heads in just the same way as it does for overt heads. However, since the parser is trying to determine the optimal attachment of an *overt* head, it evaluates the *sets of paths* between X and Y, the two *phonetically adjacent* heads. Right Association rejects the attachment that requires the longest individual path between adjacent heads. To illustrate, consider the hypothetical situation, in (25). X is the last overt head attached, and Y is the incoming overt head, as in (3/19), and each of two possible attachments of Y requires postulating intermediate null heads.

(25) a.   X e_1 Y

                    2 steps   4 steps

       b.   X e_1 e_2 Y

                  2 steps   3 steps   3 steps
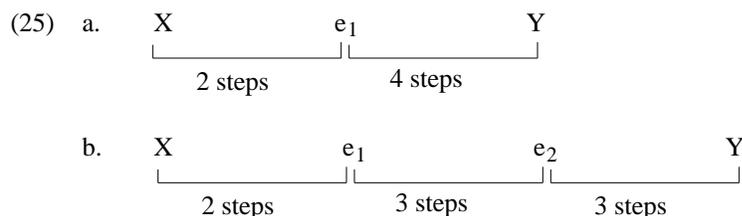
Although the attachment of Y in (25a) requires a smaller total number of heads to be introduced, and the sum of the lengths of the paths between adjacent pairs of heads is smaller than in (25b), (25a) is rejected by Right Association because it requires the longest path between any pair of heads—the 4-step path between e_1 and Y. Where the two alternatives contain a different number of null heads, and hence a different number of sub-paths, I assume that null paths can be freely generated for the purposes of comparing candidate structures, as in (26), which is generated from (25).[12]

(26) a.   2: $X...e_1$ b.   2: $X...e_1$

             4: $e_1...Y$ 3: $e_1...e_2$

             0: --   3: $e_2...Y$

(27) gives an explicit set of instructions for implementing this evaluation procedure.

(27) a.   Construct pairs of adjacent heads, overt or null, in the possible paths from X to Y.

       b.   Find the length of the path between each pair of syntactically adjacent heads.

---

[12] The null path plays no special role in (26), where the 4 step path automatically rules out parse (26a). However, if all of the paths in (26a,b) were of length 2 steps, as in 26', then the null path would become relevant.

(26') a.   2: $X...e_1$ b.   2: $X...e_1$

             2: $e_1...Y$ 2: $e_1...e_2$

             0: --   2: $e_2...Y$

The only difference between the alternative parses in (26') is that (26'a) contains a null path where (26'b) contains a path of length 2 steps. In this case obviously (26a) wins.

c.   Reject the set containing the worst path.
[d.   If more than one alternative remains, iterate step (c).]

       Now we are ready to return to Kimball's problem from section 2, and show how it does not arise the new version of Right Association.

4.4   A solution to Kimball's Problem

       (28) shows an example of the kind of ambiguity that led Kimball to question the generality of his Right Association principle.

(28)   *Clausal complement vs. relative clause*
           John told the girl that Bill saw to leave.

       The clause introduced by *that* can be attached either as an argument of the verb, or as a modifier of the NP. Subjects prefer the attachment as an argument of the verb (cf. Mitchell, Corley & Garnham 1992), but under Kimball's syntactic assumptions Right Association should prefer attachment as an NP modifier.

       However, notice what happens when we apply the modified Right Association principle to this example, using the syntactic analysis of multiple complement constructions motivated in section 3. The alternative structures are shown below in (29a-b): the final structures are shown first, with the individual paths required for the attachment of each head below. In the argument attachment structure (29a) *girl* and *that* are separated by a null V head, and therefore the paths between *girl* and *that* are of length 3 steps (NP-DP-VP-V) and 2 steps (V-V'-C).[13] In the adjunct attachment structure (29b) *girl* and *that* are separated by a null operator corresponding to the relativized argument in the relative clause. The paths between *girl* and *that*, then, have lengths of 4 steps (NP-DP-DP-CP-Op) and 2 steps (Op-CP-C). Therefore structure (29b) is rejected because it contains the worst single path (cf. 27c).

---

[13] Example (29) illustrates an important feature of how attachments involving multiple subpaths are evaluated. Right Association evaluates the length of the path required to attach the last head to the new head *at the point when the new head is attached*. Any subsequent projection (and hence lowering) of the new head in order to attach later material is ignored. This is why the first subpath under (29a) is of length 3 steps whereas the first subpath under (29b) is of length 4 steps, although the paths between the same pairs of heads are of identical length after *that* has been attached.

(29) a.   b.

VP
V VP
told
DP V'
D N(P)   V   C(P)
the girl   that

VP   V'
DP V   V   C(P)
NP

VP
V DP
told
DP CP
D N(P) Op   C
the girl   that

DP   CP
DP CP Op C
NP   Op

argument attachment: modifier attachment:
3: N(P)-DP-VP-V   **4: N(P)-DP-DP-CP-Op**
2: V-V'-C(P)   2: Op-CP-C
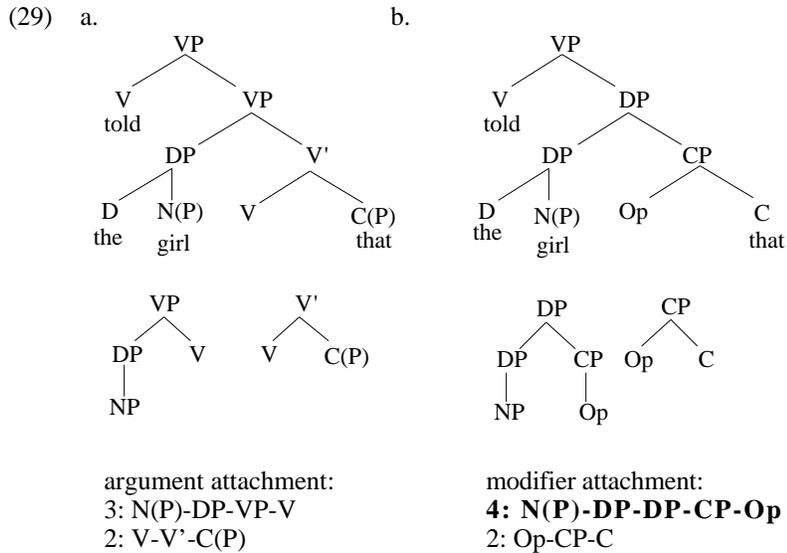
Therefore Right Association correctly predicts subjects' preferences for argument attachments in ambiguities like (29), thereby avoiding the problem with Kimball's version of Right Association.[14]

4.5   "Minimal Attachment" preferences

I next turn to a series of parsing preferences which have standardly been assumed to follow from either Frazier's *Minimal Attachment* strategy, or from a principle which favours satisfaction of predicate argument relations (Ford, Bresnan & Kaplan 1982; Pritchett 1988, 1992; Gibson 1991; Abney 1987, 1989; Tanenhaus, Carlson & Trueswell 1989). I show that these predictions are all matched by the new version of Right Association.

In (30) the NP *the dog* can be attached either as the second argument of *give*, or as the subject of a relative clause modifying *the man*. The alternative structures, and the individual paths required in each alternative, are shown in (31a-b). The adjunct attachment in (31b) is rejected by Right Association because it contains the worst individual path (4 steps: NP-DP-DP-CP-Op). The argument attachment preference that Right Association predicts conforms to

---

[14] The other ambiguities in (9) above, repeated in (i), that created problems for Kimball's Right Association are resolved in very much the same way as (28).

(i)   a.   Joe bought the book for Susan.
   b.   John read the letter to Mary.

The only difference between these examples and (28) is that (28) involved a CP which could be either an argument or an adjunct, whereas (ia–b) contain ambiguous PPs. I assume that PP modifiers of NPs are structurally parallel to full relative clauses in that they contain a null specifier PRO, coindexed with the head of the NP.

subjects' on-line preferences (Frazier & Rayner 1982; Ferreira & Henderson 1990; but cf. Holmes 1987, Kennedy et al. 1989).

(30)   *NP complement vs. subject of relative clause*
         Fred gave the man the dog bit a package.
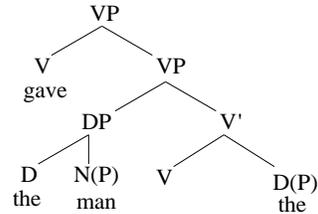
(31)   a.   b.
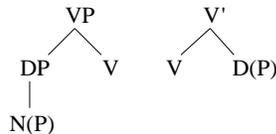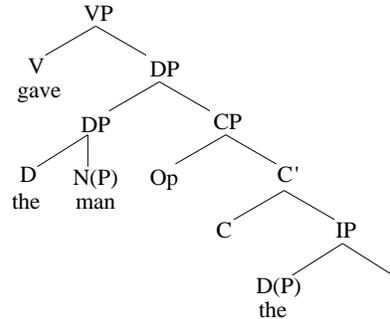
argument attachment:
3: N(P)-DP-VP-V
2: V-V'-D(P)
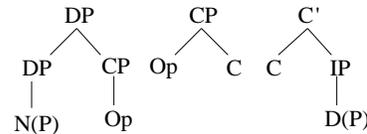
relative clause attachment:
4: N(P)-DP-DP-CP-Op
2: Op-CP-C
3: C-C'-IP-D(P)

The sentences in (32) illustrate the intensely studied main verb/reduced relative clause ambiguity in English. Subjects' on-line preference is to choose the main verb interpretation of the ambiguous verb (Bever 1970; Rayner, Carlson & Frazier 1983). (33a-b) show the alternative structures, with the individual paths beneath each one. Right Association prefers the main verb attachment in (33a) with a worst path of 3 steps over the reduced relative attachment in (33b) with its worst path of 4 steps, correctly modelling speakers' preferences.[15,16]

---

[15] At this point I should stress once again that I am not trying to claim that Right Association is the only factor involved in resolving ambiguities like (32). At this point there is an immense amount of literature showing that other factors, such as plausibility, and frequency, modulate the strength of the main clause attachment preference, although I am unaware of any evidence that they can reverse the standard preference. (Cf. MacDonald et al. 1994, Truewell & Tanenhaus 1994 and references cited therein for extensive discussion). These factors are quite consistent with the structural component of the parser which I am focussing on here.

[16] The argument structure of the verb used in this ambiguity seems to play an important role in the strength of the main verb preference. The verbs used in (32)

(32)   *Main verb vs. reduced relative clause*
  a.   The boat floated down the river sank.
  b.   The horse raced past the barn fell.
  c.   The dog walked to the park chewed the bone.

(33) a.   b.

main verb attachment: reduced relative clause attachment:
3: N(P)-DP-IP-I   **4: N(P)-DP-DP-IP-PRO**
2: I-I'-V(P)   2: PRO-IP-I
                                         2: I-I'-V(P)

The last 3 ambiguities considered all share the property that they involve choices between one attachment that satisfies a predicate-argument relation and an alternative that introduces a relative clause. In every case the relative clause structure is clearly dispreferred. The next standard ambiguity we turn to is slightly different in that it involves a choice between a pair of structures that both satisfy a predicate-argument relation. In (34a–b) the underlined determiner can head either the direct object of the matrix verb, or it can head the subject of a clausal complement. There has been heated controversy in the recent literature over whether subjects show any reliable general preference for one attachment over the other (cf. Frazier & Rayner 1982; Ferreira & Henderson 1990; Juliano & Tanenhaus 1993; Trueswell, Tanenhaus & Kello 1993), but to the extent that there is a preference, it is clearly a preference for the direct object attachment (35a).[17]

---

float, *race* and *walk* all alternate between an intransitive and a transitive (causative) frame. Because the second argument is not obligatory, the sentences are disambiguated only when the sentence final verb is encountered. If, on the other hand, obligatorily transitive verbs like *find* are used, as in *The bird found in the bush was rare*, disambiguation is reached much earlier, and the initial mis-parse is much easier to recover from.

[17] As in the case of the main verb/reduced relative ambiguity, manipulation of plausibility and frequency factors have succeeded in reducing the strength of the direct object attachment preference, possibly even to such an extent that the preference disappears, but again I am unaware of any reports that the preference can be reversed. The most important point is that in properly ambiguous cases, in which biases due to context and frequency have been controlled for, the structural preference for direct object attachment seems to hold.

(34)   a.   I <u>know the</u> answer to the physics problem is wrong
        b.   Bill <u>heard the</u> gossip wasn't true

(35)   a.   b.

```
        VP                          VP
       /  \                        /  \
      V    D(P)                   V    CP
    know   the               know    /  \
                                    C    IP
                                         |
                                       D(P)
                                        the
```
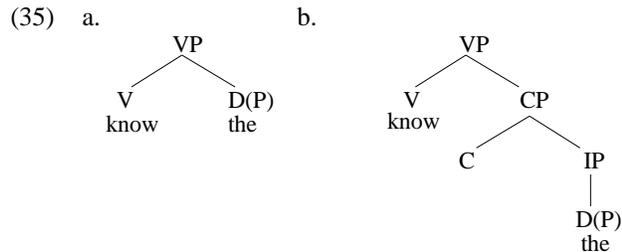
direct object attachment: embedded subject attachment:
2: V-VP-D(P)   2: V-VP-C(P)
                                     3: C-CP-IP-D(P)

Right Association prefers the direct object attachment of *the* in (35a) because its worst path of 2 steps is shorter than the 3 step path required by the alternative parse in (35b).

5.   Some divergent predictions

In most of the examples that we have considered so far Right Association prefers argument attachments over non-argument attachments, or chooses among different potential argument attachments. Most of these predictions are matched by a variety of other accounts of structural parsing. A better test of Right Association is cases in which it forces an attachment which is either the 'non-minimal' attachment, or involves an adjunct attachment in preference over an argument attachment. These would be cases to distinguish Right Association from other standard structural parsing strategies, like *Minimal Attachment* or an argument attachment strategy. This section considers such ambiguities.

5.1   Subject-subject ambiguities: A 'non minimal' preference

A number of verbs can either be intransitive, or select a complement clause (without *that*), and do not allow an direct object NP complement. This means that it is possible to construct local ambiguities in which the NP in a V-NP sequence is either the subject of the verb's clausal complement or the subject of an independent higher clause. Some examples are given in (36–38) and the alternative structures at the choice point are shown in (39).

(36)   a.   Whenever Mary started to <u>think the class</u> had finished there was still an hour remaining.
        b.   Whenever Mary started to <u>think the class</u> had finished and everybody had left.

(37)   a.   When Bill <u>complained the class</u> was boring it usually was.
        b.   When Bill <u>complained the class</u> was annoying as a result.

(38)   a.   If Tom <u>confessed the gun</u> was his he would be in trouble.
        b.   If Tom <u>confessed the gun</u> was his.
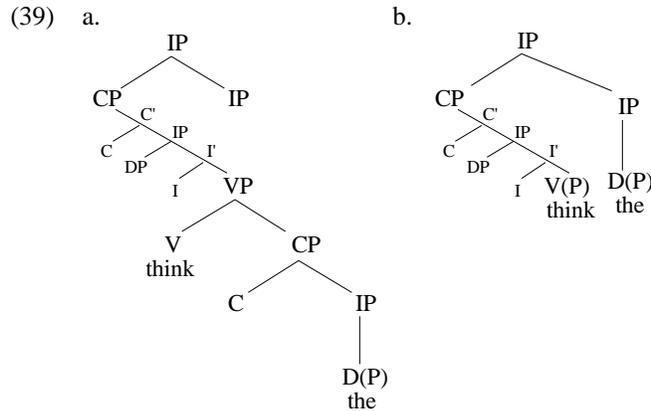
(39)   a.   b.

embedded subject attachment: matrix subject attachment:
2: V-VP-C(P)   **7: VP-I'-IP-C'-CP-IP-IP-DP**
3: C-CP-IP-D(P)

Notice that in the continuation in which the determiner heads the subject of the complement of *think*, many new nodes must be built, but the paths between adjacent heads are kept short. In the analysis in (39b), on the other hand, fewer new nodes must be built than in (39a), but the path between *think* and *the* is extremely long. Therefore, the structure in (39b) contains the more 'minimal' attachment of *the*, but the structure in (39a) is favoured by Right Association, a clear divergence of predictions.

Similar examples can be constructed with a handful of other verbs (e.g., *agree*, *dissent*, *concur*, *joke*, *cry*), and speakers generally prefer the embedded subject continuation (the (a) examples in 36–38). The preferences have not been experimentally confirmed, but if the intuitive preference turns out to be correct, then these subject-subject ambiguities favour the Right Association theory over the theory based upon Minimal Attachment.[18]

---

[18] A concern that arises immediately is whether the intuitive preference for low attachment in these examples is due to a strong frequency bias for the transitive frames of the verbs involved. I have checked for this possible confound by counting frequencies of transitive and intransitive frames in the parsed Brown Corpus (available in the Penn Treebank) for verbs displaying the subject-subject ambiguity. It turns out that the frequency of a given subcategorization is heavily dependent on

5.2   A preference *for* a relative clause, for a change

The ambiguity in the examples in (40–41) is a variant on the same theme: at the underlined pronoun a choice must be made between continuation of the embedded clause by beginning a relative clause (40a, 41a) and beginning the matrix clause (40b, 41b). An important property of this ambiguity is that the lower, more right-branching attachment is a non-argument attachment, whereas the higher attachment is an argument attachment.[19]

(40)   a.   Although the dissident eventually destroyed the letters <u>he</u> had kept beneath the floorboards it pained him to do so.
       b.   Although the dissident eventually destroyed the letters <u>he</u> had kept them beneath the floorboards for months.

(41)   a.   Because the conference organizers liked the caterers <u>they</u> hired for the banquet they asked them back the following year.
       b.   Because the conference organizers liked the caterers <u>they</u> hired them for the banquet in honor of the former president.

---

the tense and aspect of the verb. Some tenses/aspects show strong frequency biases, while other tenses/aspects of the same verb are roughly equibiased. For example, present tense *think(s)* is strongly biased in favour of a sentential complement, but past tense *thought* occurs with approximately equal frequency with a complementizerless sentential complement or with no complement at all. Since it it not yet known how frequency biases carry over from one part of a verbal paradigm to another, it is hard to say whether it would be possible to experimentally verify this preference using balanced materials.

Additionally, the principle of *Minimal Attachment* might be made immune to the consequences of the low attachment preference on these examples, by adopting an extreme version of the claim that semantic and argument structure information is delayed in processing. It might be claimed that the NP following verbs like *think* and *confess* is initially attached as the direct object of the verb, even though this is not one of the possible frames for the verb used. This satisfies *Minimal Attachment* optimally and relegates the process of deciding between the two *grammatical* argument structures to separate and poorly defined processes of reanalysis which come into effect later in processing.

While it is of course difficult to refute such a sidestep, I should note that it crucially relies on the assumption that structures can be built in processing which violate the lexical requirements of verbs. This assumption has been argued for by Mitchell (1987), but called into question by Boland and Tanenhaus (1991).

[19] Thanks to Ted Gibson for brining this ambiguity to my attention (see Gibson and Broihier (to appear)). We are currently conducting an experiment to test the intuitive preference for low attachment in these examples.
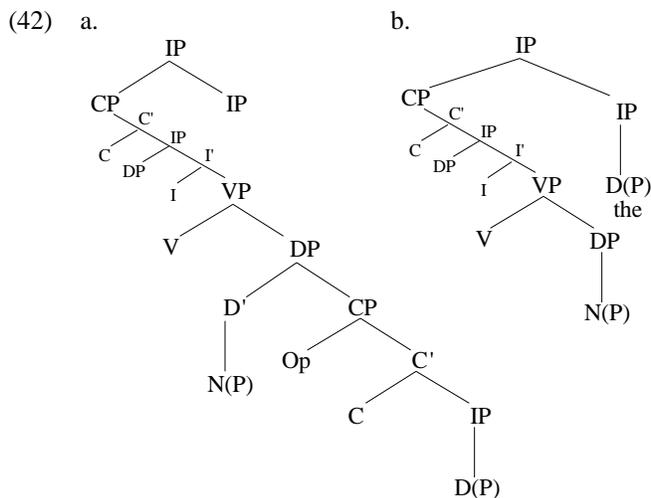
(42)   a.   b.

[Syntax tree diagrams for (42a) and (42b)]

relative clause attachment: matrix subject attachment:
4: N(P)-D'-DP-CP-Op   **9: N(P)-DP-VP-I'-IP-C'-CP-**
2: Op-CP-C   **-IP-IP-D(P)**
3: C-C'-IP-D(P)

Speakers' intuitions favor the relative clause attachment in the (a) examples. The fact that the lower attachment is preferred is striking, since it is a rare case in which preferences are for *avoiding* an argument attachment in favor of a relative clause. This is the opposite scenario to the examples from Kimball in (9) which we have discussed at length, but again this preference is exactly what Right Association predicts. Although relative clause attachments are usually dispreferred when an argument attachment is available (see section 4 above for many examples), the potential argument attachment in (42b) (cf. 40b, 41b) requires such a long path that Right Association opts for the relative clause continuation (42a) in this case.

5.3   Possessor-argument   ambiguity: A   further   non-argument preference

The examples in (40–42) show a preference for building an adjunct rather than a subject of a main clause. The examples in (43) show another advantage of not explicitly building into the theory an argument attachment preference. In these examples, the underlined words following the ditransitive verb may be interpreted either as a single argument containing a possessor, or as both arguments of a double object contstruction. To the extent that speakers' intuitions show a preference for one of these two interpretations of *her hamsters*,

they slightly favour the single argument interpretation in (43b) (cf. Kennison & Clifton 1994).[20]

(43) a.   I gave <u>her hamsters</u> on her birthday.
      b.   I gave <u>her hamsters</u> to Mary.

(44) a.   b.

2: DP-VP-V   2: A(P)-DP-D
2: V-V'-D(P)   2: D-D'-N(P)
2: D-DP-N(P)   0: --
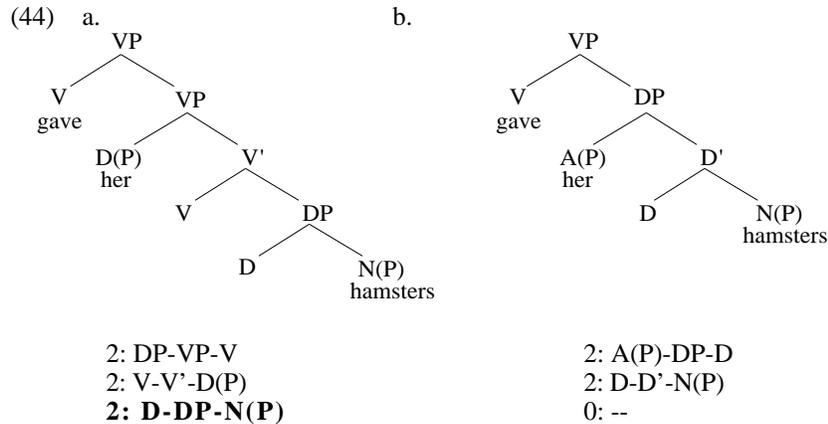
This is the structure which Right Association predicts to be preferred. The worst paths in each structure are identical, so neither attachment is favoured based on the very worst path. In this situation, Right Association must look to the next worst path in each parse for a difference which would favour one alternative over the other. This process iterates until a winner is found. In (44) the second worst path in each parse is also identical in the two parses. The only difference between the two parses is that the third longest path in (44a) is of length 2 steps, whereas the corresponding path in (44b) is a null path of length 0. Right Association therefore rejects (44a) and opts for the single object continuation in (44b). A theory in which argument attachments are preferred incorrectly predicts the double object attachment in (44a) to be favored.[21]

---

[20] There is a clearer preference for the single argument interpretation when *give* in (43) is replaced by an optionally ditransitive verb like *find* or *carry*, as Frazier (1994) points out.

(i)   I found her hamsters.
(ii)   I carried her shopping.

This increased preference for the single argument reading may be due to the parser's avoidance of benefactive roles, or to the fact that in the double object construction reading of (i–ii) an optional argument precedes an obligatory argument.

[21] The ambiguity in (43) does not lead to garden path effects, unlike some of the ambiguities discussed in section 3. The contrast between hard and easy ambiguities is a topic which I will have nothing to offer on here (cf. Pritchett 1988, 1992; Gibson 1991; Inoue & Fodor 1994; Frazier 1994; Gorrell 1995).

5.4 PP attachment: argument/adjunct contrast for nouns, not for verbs

If argument attachment preferences, such as we have observed in section 4, are directly encoded in the parser, then we predict that adjuncts should exact a greater processing load than arguments. This increased cost should hold regardless of the category that that is being attached to, and regardless of whatever structural contrast or parallel there is between arguments and adjuncts of that category. On the other hand, a purely nodular theory like Right Association predicts that if some categories structurally differentiate arguments and adjuncts, whereas other categories do not, this difference should be reflected in processing costs. In this section I argue that in intensively studied PP attachment ambiguities like (45), an argument advantage is found for nouns, but not for verbs, and that this difference is precisely what Right Association predicts.

(45)   The spy saw the cop with the revolver.

In V-NP-PP sequences at least two choices must be made in attaching the PP. First, the choice of whether to attach the PP to a projection of N or V; second, whether to attach the PP as an argument or as a modifier of the chosen category. Some typical examples are shown in (46–47) (taken from Clifton, Speer & Abney 1991). In (46) the PP can be either an argument of N or a modifier of V, and in (47) the PP can be either an argument of V (47a) or a modifier of N (47b).[22]

(46) a.   The man expressed his interest in the Volvo at the sale.
      b.   The man expressed his interest in the Volvo on the drive home.

(47) a.   I put the candy on the table into my mouth.
      b.   The cook placed the cake in the oven on the table.

If we apply the now familiar constituency tests to examples like these, we find that arguments and adjuncts of V behave alike, whereas arguments and adjuncts of N behave differently, implying that the argument/adjunct contrast does not correspond to consistent structural contrast.

PPs which are arguments and modifiers of V behave alike on coordination and reciprocal binding tests (48–49).
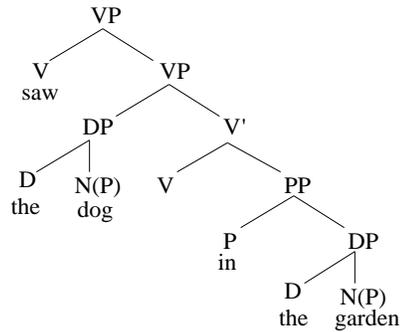
(48) a.   I saw [the dog in the garden] and [the cat on the roof]
      b.   The baker put [the cake in the oven] and [the pastry on the shelf]

(49) a.   I saw the dogs in each other's kennels.
      b.   He put the children in each others beds.

---

[22] See Schütze (this volume) for a comprehensive survey of diagnostics for argument and adjuncthood.

24

These parallels were noted by Larson (1990), who concluded that PP modifiers of V must be structurally identical to PP arguments of V, and therefore attached as complements of the lower VP in a nested VP structure like (50).

(50)

```
                        VP
                      /    \
                    V        VP
                   saw      /  \
                          DP      V'
                         /  \    /  \
                        D   N(P) V    PP
                       the  dog       /  \
                                     P     DP
                                     in   /  \
                                         D    N(P)
                                        the  garden
```

In the absence of any structural difference between arguments and adjuncts of V, then, we do not expect there to be any preference for PPs to attach as V-arguments rather than V-adjuncts, or vice versa.

In contrast to the structural parallel between PP complements and modifiers of V, coordination facts show an asymmetry between arguments and modifiers of nouns. N-PP sequences can be coordinated to the exclusion of a determiner if PP is an argument of N (51a), but N-PP cannot be coordinated if PP is a modifier of N (51b). I take this as evidence for the two structures in (52).

(51) a.   I will never forget the destruction of the city and pillaging of all its riches.
      b.   *I most liked the dog in the garden and cat on the roof.

(52) a.   b.

```
              DP                                    DP
             /  \                                  /  \
            D    NP                               DP    PP
                /  \                             /  \
               N    PP                          D    NP
                                                      |
                                                      N
         N-complement                          N-modifier
```
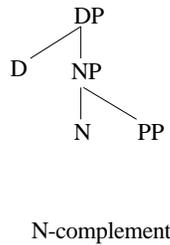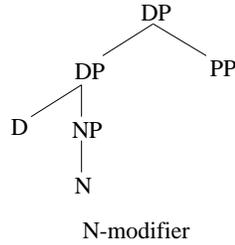
As a result, PP attachment as an argument of N should be preferred over attachment as a modifier of N, because the path from N to P is longer in the adjunct attachment. Applying the results of these tests, the Right Association principle predicts the following ranking of potential PP attachments in V-NP-PP sequences (53). Arguments of N are predicted to be most preferred, followed by any attachment to V, with attachment as a modifier of N predicted to be worst.

(53)   *Predicted preferences in V-NP-PP sequences:*
      a.   N-arguments
      b.   V-arguments and V-adjuncts
      c.   NP-adjuncts

These predictions are mostly consistent with the findings of Clifton, Speer, and Abney (1991), which was to my knowledge the first study to independently manipulate argumenthood and the category of the attachment site for PPs. They found that NP-adjuncts are clearly the most dispreferred attachments,[23] and they also found no contrast between arguments and modifiers of verbs in first pass reading times. But Clifton et al. failed to find any special preference for argument attachments to nouns, contrary to what Right Association predicts.

However, Schütze (1995a, 1995b) presents both an explanation for why Clifton et al. (1991) failed to find the N-argument preference in their experiment, and experimental evidence for a preference for N-argument attachments over V-adjunct attachments. Schütze points out that many of Clifton et al.'s intuitively chosen argument and adjunct PPs fail to pass a number of syntactic tests of argumenthood or adjuncthood. Given that the experimental materials do not properly contrast arguments and adjuncts, it is therefore not surprising that Clifton et al.'s results failed to find a preference for N-argument attachment. Schütze also reports the results of a new self-paced reading study which he conducted with Ted Gibson using materials which pass the appropriate syntactic tests of argument or adjunct status, and found a significant reading time advantage for N-argument over V-adjunct attachments.[24] Thus, the unusual predictions of Right Association for preferences in PP attachment ambiguities appear to be supported by experimental results: argument PPs have an advantage in attaching to noun phrases, but not in attaching to verb phrases.

6.   Gap-filling: Effects of length and language

Sections 3 and 4 have laid out an account of how the principle of Right Association alone predicts a wide range of structural preferences in attaching new words into phrase structure trees. This section shows how Right Association also accounts for a range of differing preferences for associating extracted phrases with argument positions. This includes both cross-linguistic differences between Dutch, German, English and Italian, and within-language differences between long and short extraction.

---

[23] In Clifton et al.'s eye-tracking study, there was no difference between arguments and modifiers of nouns in first pass reading times for the PP itself, but a very large difference in first pass reading times for material immediately following the PP.

[24] In Schütze and Gibson's study, the reading time advantage for N-argument attachments was significant by subjects, but not by items. Schütze (1995a) gives an explanation for the differences they found between items.

Right Association in Parsing and Grammar

6.1   Local *wh*-movement in English, Dutch and German

The preferences for gap-filling in English and Dutch/German have been reviewed in detail elsewhere (Frazier 1993; Gibson, Hickok & Schütze 1994), so I will restrict myself to a brief summary here. I assume that the basic cross-linguistic contrast to be captured is that local extractions are preferentially taken to be object extractions in English and subject extractions in Dutch/German.

The English object gap preference is motivated by studies showing so-called "filled gap" effects for object gaps but not for subject gaps. The filled-gap effect is a slow-down in reading time for a phrase occupying a potential gap position, relative to the reading time for that same phrase when its position is not a potential gap position (Crain & Fodor 1985; Stowe 1986). For example, the underlined word *us* is read more slowly in (54b: Stowe 1986), where it occupies a potential gap site for the filler *who*, than in (54a), in which it is not a potential gap site.

(54)   a.   My brother wanted to know if Ruth will bring <u>us</u> home to Mom at Christmas.
       b.   My brother wanted to know who Ruth will bring <u>us</u> home to at Christmas.

This kind of contrast is standardly interpreted as follows. At some point before *us* is encountered in (54b) the trace of extraction is posited in object position, and therefore reanalysis is required when the overt object "us" is encountered. It is this reanalysis that is responsible for the slowdown.

In contrast to the filled-gap effects that Stowe found for object positions in English, Stowe found no filled gap effect in subject position. In other words, the subject NP *Ruth* was read just as quickly in (54a) in which it is a potential gap position, as in (54b) in which is it not a potential gap position. I interpret this subject-object contrast to reflect a preference for object gaps in English.[25]

In Dutch, on the other hand, a series of experiments have pointed to a preference for associating extracted arguments with subject gaps.

---

[25] Others have given rather different interpretations to Stowe's findings. Gibson et al. (1994) assume that in subject position there is no pressure on the parser to posit a gap, but also no pressure not to. On the other hand Frazier argues that English does in fact have a subject gap preference, but that this preference is obscured in Stowe's experiment because of the fact that the filler and the potential subject gap are linearly adjacent.

Pickering and Shillcock (1992) argue that subject extractions are easier to process than object extractions, in both local and long-distance extraction. Their conclusions are based on comparing reading times for subject and object extractions, with the result that reading times for verb + gap are compared with reading times for an NP. This makes it more difficult to interpret their results than Stowe's.

Because Dutch is verb final, more tests are available for investigating preferred extraction sites than in English. The sequence *wh*-NP NP V can be interpreted as containing either a subject gap or an object gap, and the case marking and number agreement on the verb can be manipulated so that the clause is either entirely ambiguous, or ambiguous until the verb disambiguates based on subject number agreement.

Frazier (1987) conducted both on-line and off-line experiments of preferences in Dutch. In an off-line study, subjects were asked for their preferred interpretation for fully ambiguous relative clauses like (55). The results showed a strong preference for subject gap interpretations.

(55)   Karl hielp de mijnwerker die de boswachter vond.
            'Karl helped the mineworker who __ found the forester.'
            'Karl helped the mineworker who the forester found __.'

Materials for the on-line study were just like (55), except that disambiguating agreement was encountered on the clause-final verb (56). Reading times on the verb were significantly longer when agreement disambiguated in favour of an object gap reading (56b), implying that the relative clause is initially interpreted as containing a subject gap.

(56)   a.   Karl hielp de mijnwerkers die de boswachter vonden
             Karl helped the mineworkers who the forester found-pl
             'Karl helped the mineworkers who found the forester.'

       b.   Karl hielp de mijnwerkers die de boswachter vond
             Karl helped the mineworkers who the forester found-sg
             'Karl helped the mineworkers who the forester found.'

Both of these results point to a subject gap preference for extracted arguments in Dutch,[26, 27] contrasting with the object gap preference found in English.

There has been much discussion in the literature on processing *wh*-questions about whether special parsing strategies are employed in determining the appropriate trace site for extracted arguments. I will assume that there are no special strategies for gap filling. Instead, as is generally assumed in the syntactic literature, I take filler-gap chains to be syntactic units like any other, with the sole difference that they are spread across more than one position. Therefore I predict that *wh*-chains should be constructed on-line according to the same

---

[26] The subject gap preference appears to hold only for local extractions in Dutch. Haverkort (1986) and Frazier (1993) report that off-line judgements strongly favour object gap interpretations for long-distance extractions. See section 6.2 below for an account of this contrast.

[27] Frazier and Flores d'Arcais (1989) present evidence for similar preferences in Dutch *wh*-questions.

principle of Right Association that guides attachment of single member chains. When a *wh*-phrase is encountered, then, it is not sufficient to just attach the overt *wh*-phrase in the way that best satisfies Right Association. An attachment site for the tail of the chain must also be created, and this site must also be chosen in a way that best satisfies Right Association.

Consider first what happens when a *wh*-phrase is encountered in English, as in the sentence fragment in (57).

(57)   I know [who...t]

There is only one possible attachment site for the head of the *wh*-chain *who*—namely [Spec,CP]—but there are multiple possible attachment sites for the tail of the chain. For now let us consider just two possibilities, that the tail of the chain is in either subject or object position of the complement clause of *know*. If the trace is in subject position, then the paths in (58a) are required in order to create an attachment site for the trace. If, on the other hand, the trace is in object position, then the paths in (58b) must be built.[28]

(58)   a.   *subject gap*   b.   *object gap*
            2: wh-CP-C   2: wh-CP-C
            3: C-C'-IP-DP   2: C-C'-I(P)
                                                        2: I-IP-V(P)
                                                        2: V-VP-D(P)

Applying the same evaluation procedure employed in sections 4–5 above, the object gap analysis is chosen (58b), because the subject gap analysis (58a) requires a 3-step path, which is longer than any of the paths required under the object gap analysis.[29]

Turning now to a situation in Dutch that is equivalent to (57), I assume that the *wh*-phrase is attached in just the same way as in English, in [Spec,CP]. I also assume that the position of the subject in Dutch is identical to English, as reflected by the identical cost of a subject gap shown in (59a). The only difference between Dutch and English is in the relative ordering of object

---

[28] Notice that under the object gap analysis, only those heads are attached that are necessary in order to create an attachment site for the object gap. The subject position, for example, is not built. This differs from the attachment of overt heads in sections 4–5, which required attaching of all intervening null heads, even if they were not essential for creating an attachment site for the overt head. This difference is found because when the object gap is postulated the parser is in effect "looking ahead," and will still get to hear an overt subject, whereas in the examples discussed in sections 3–4 the parser is not looking ahead.

[29] Notice that when an attachment site for an object gap is constructed, some positions which will eventually be necessary are ignored, e.g., subject position in [Spec,IP]. This is because only those positions are created which are required in order to create an attachment site for the trace.

position and the verb: in Dutch, object position precedes the verb, and therefore the series of steps between the head and tail of the *wh*-chain that Right Association evaluates is different from English (506b).

(59) a. *subject gap*   b. *object gap*
          2: wh-CP-C   2: wh-CP-C
          3: C-C'-IP-D(P)   2: C-C'-I(P)
          0: —   3: I-IP-VP-D(P)

In this case, then, the worst path in both series contains 3 steps. Therefore, the subject gap analysis is not immediately excluded as it was in English. The second worst path in both series is also equivalent, having 2 steps. The only difference between the two analyses is that the object gap analysis contains an extra 2 step path. Therefore, the subject gap analysis is predicted to be the preferred analysis in Dutch.[30]

Therefore, Right Association accounts for the contrasting gap-filling preferences between English and Dutch local extractions.

6.2   Differences between local and long-distance extraction

Cross-linguistically, preferences for gap filling in long-distance extraction constructions present two rather intriguing contrasts. First, the preference for a subject gap in short-distance extraction in Dutch and German disappears in long-distance extractions, and is replaced by a strong object gap preference. Thus, the ambiguous Dutch long-distance question in (60) is preferentially understood with the object-gap reading in (60b), although the subject gap reading is entirely grammatical (cf. Haverkort 1986; Frazier 1993).[31] So the first puzzle is that there is a reversal of preferences internal to Dutch and German, depending on whether extraction crosses a clause boundary or not.

(60)   Welke patienten meende Jan dat de dokters bezochten?
       a.   Which patients did Jan mean that __ are visiting the doctors?
       b.   Which patients did Jan mean that the doctors are visiting __?

_____

[30] I am assuming here for purposes of exposition that the differing ordering of verb and object in Dutch and English is due to whether or not the verb selects a complement to its right (English) or to its left (Dutch). However, the analysis works in exactly the same way if we instead adopt the proposal of Zwart (1993) that both Dutch and English are uniformly head initial languages. Under Zwart's analysis the fact that the object precedes the verb in Dutch is the result of obligatory movement to the specifier of a higher projection. This analysis of Dutch fails to affect the predictions about gap-filling preferences because extraction from specifier position still requires a 3-step path, and there will still be more 2-step paths for object gaps than subject gaps.

[31] The fact that preferences reverse between short and long distance extractions apparently cannot be explained away in terms of grammaticality. Dutch lacks *that-t* effects, so both the subject gap and the object gap analysis of (60) are entirely grammatical.

Meanwhile, Italian differs from Dutch and German in that the crossing of a clause boundary typically doesn't affect gap filling preferences. In both short and long-distance extraction of *wh*-phrases like *chi* 'who' Italian speakers show a subject gap preference. So, for example, the ambiguous long-distance extraction in (61) is preferentially interpreted as in (61a) (de Vincenzi 1995).

(61)   Chi pensi abbia licenziato il metalmeccanico sensa preavviso?
    a.   Who do you think __ fired the steel-worker without notice?
    b.   Who do you think the steel-worker fired __ without notice?

Therefore the second puzzle is: why does Italian lack the contrast between long and short extraction found in Dutch and German?

By the end of this section, I hope to have shown how both of these puzzles can be explained as a consequence of the principle of Right Association, in basically the same way as Right Association accounted for the contrast between English and Dutch/German local extraction in section 6.1. I first review some evidence which suggests that the parser is able to associate *wh*-phrases with gap positions inside syntactic islands—this evidence comes from the processing of grammatical parasitic gap constructions and ungrammatical subjacency violations. Then I show that once we assume the limited island-crossing capacity necessary for the processing of parasitic gaps is available to the parser in general, we predict that the parser will be able to build long filler-gap dependencies across clause boundaries when the gap is in some positions, but not in others, and that these contrasts will vary across languages. In combination with these independently motivated properties of gap-processing, the principle of Right Association accounts for the two puzzles just outlined.

6.2.1   Gap-filling across islands

6.2.1.1 Parasitic Gap constructions

The property of parasitic gap constructions like those in (62–63) that has most interested linguists is that they contain a gap in a position that is normally not allowed in the language, located inside a syntactic 'island' (Taraldsen 1981; Chomsky 1982; Engdahl 1983). Informally speaking, the illegal gap is 'saved' by the presence in the same sentence of another gap which is in a perfectly legal position. (62) shows the contrast between an illegal extraction from inside a subject when there is no other gap (62a) and when there is another legal gap (62b). (63) contrasts extraction from an adjunct clause with and without a second gap.

(62)   a.   *Who did your attempt to persuade annoy everybody?
    b.   Who did your attempt to persuade annoy?

(63)   b.   *Which reports did the police chief file complaints without reading?
    a.   Which reports did the police chief file without reading?

Parasitic gap constructions involving a gap inside a subject are particularly interesting from a parsing perspective, because the object gap that is responsible for rescuing the grammaticality of the construction does not occur until after the illegal gap inside a subject. There are two ways in which a parser could handle this kind of situation. The first possibility is that it could avoid postulating a gap inside an island until the rescuing object gap is encountered. The alternative is that the parser could liberally postulate gaps inside islands, regardless of the fact that they may turn out to be ungrammatical, and then rely on the later appearance of a gap position that is not inside an island to rescue the first gap.

Kurtzman and Crawford (1990) provide interesting experimental evidence for the second of these alternatives, the liberal gap-filler. They used a speeded on-line grammaticality judgement task to test the acceptability of parasitic gaps at different points in the processing of a parasitic gap construction. Subjects read sentences presented word-by-word on a computer screen, until at some unexpected point a beep sounded, instructing the subject to make a judgement on whether the sentence was "grammatical so far". In constructions with a subject parasitic gap they compared acceptability levels before and after the legal gap has been encountered. Surprisingly, they found that subjects are just as likely to respond "grammatical" when they have only encountered the parasitic gap (64a) as when they have encountered both the parasitic gap and the legal gap (64b). Moreover, these acceptability rates are just as high as for (65), which contains a garden-variety matrix object gap and no other gaps.

(64) a.   Who did your attempt to instruct __   68%
       b.   Who did your attempt to instruct __ confuse __   66%

(65)   Which antiques did Joan lift __ up in order to examine them   67%

I take Kurtzman and Crawford's finding to mean that the parser has no difficulty in initially positing a gap inside a subject island, as in (64a). They also found high acceptability ratings for parasitic gaps inside adjunct clauses like (66). But since the parasitic gap follows the object gap in these examples, the high degree of acceptability would also be expected if the parser avoided positing gaps inside islands until a legal gap has been encountered.

(66)   Which antiques did Joan buy after examining __   73%

6.2.1.2 Subjacency violations

The smooth processing of parasitic gap constructions implies that the parser is able to postulate gaps inside islands. A number of studies have looked at this issue using other techniques, with varying results. I believe that some of the conflicting findings about the parser's ability to look inside islands is due to the fact that some studies have looked at more impenetrable islands than others.

Based on evidence from the sentence matching task, Freedman and Forster (1985) argue that subjects construct representations containing the ungrammatical gaps in (67a-b). In (67a) the gap is again inside a complex NP (relative clause) island; in (67b) it is inside a definiteness island ('specified subject island').[32]

(67) a.   *Who do the police believe a claim that John shot __?
      b.   *What did the scientist criticize Max's proof of __?

On the other hand, some studies which have looked for filled-gapeffects inside islands have failed to find any good evidence that the parser tries to postulate gaps inside islands.

Pickering et al. (1994) report some experiments on constructions like (68–69), which include a subject containing a relative clause (68) or a subject containing an adjunct (69). They found no significant differences in reading-times for the underlined phrase in (68b) relative to the same phrase in (68a). A slow-down would be expected if the parser had tried to complete the *wh*-chainheaded by *what* inside the relative clause.[33] Similarly, no differences were found in reading times for the underlined region in (69).

(68) b.   I realize that the artist who painted <u>the large mural</u> ate cakes.
      a.   I realize what the artist who painted <u>the large mural</u> ate today.

(69) a.   I know that a book about <u>the local election</u> discussed the poll.
      b.   I know what a book about <u>the local election</u> discussed the most.

---

[32] The sentence matching technique is an extension of classic results which found that subjects are faster to identify a string of letters as matching if it forms a possible word than if it does not. For example, the matching task is performed much faster with strings like HOUSE than with strings like HSEUO. Freedman and Forster (1985) showed that certain kinds of ungrammaticality (e.g., agreement violations) led to slow-down on sentence matching compared to relevant controls, whereas other kinds of ungrammaticality, like the subjacency violations in (67) did not cause slow-down. Freedman and Forster interpret the contrast in matching times as reflecting the differences between sentences which the grammar cannot generate (slow-down), and those sentences which the grammar can generate, but are filtered out by grammatical constraints (no slow-down). However, Freedman and Forster's interpretation of their findings have been challenged. See Crain and Fodor (1985a, 1985b), Stowe (1992) for alternative explanations, and Forster and Steedman (1987) for a defense of Freedman and Forster's account.

[33] Pickering et al. (1994) did find a small but reliable differences in reading times for the verb *painted* between (68a) and (68b). However, this is not comparable to the filled-gap effect that Stowe found in non-island contexts, and the effect could just be due increased complexity in (68b): *painted* immediately follows the word *who*, which may cause mild processing overload when there is already an incomplete filler-gap dependency in the sentence. Pickering et al. found no related effects in (69).

Stowe (1986, Experiment 2) also tested to see whether filled-gap effects would be found inside complex NP islands, and also found no evidence for the postulation of gaps inside islands. Stowe compared reading times for the word *Greg's* in (70a) and (70b), and reasoned that if the processor was able to posit gaps inside islands, then there would be a slow-down for reading the word *Greg's* in (70b) relative to (70a), because (70b) contains a *wh*-filler. Stowe found no slowdown, and therefore argued that the parser is unable to look inside islands when searching for a gap to associate with a *wh*-filler.

(70) a   The teacher asked if the silly story about <u>Greg's</u> older brother was supposed to mean anything.
  b.   The teacher asked what the silly story about <u>Greg's</u> older brother was supposed to mean.

However, I think that Stowe's conclusion is too strong. The various studies reviewed in this section differ in whether they looked at constructions involving a single island or constructions involving a more complex 'double' island. Not surprisingly, it was the studies that looked at single islands that found evidence for gap-filling inside islands, whereas the studies that looked at double islands found no evidence for gap-filling inside islands.

The position of *Greg's* in (70b) is inside *two* islands—both a subject island and an adjunct island. This is confirmed by the impossibility of a parasitic gap in a comparable position in (71). A number of authors have observed that parasitic gaps inside two islands are much less acceptable than parasitic gaps inside one island (cf. Kayne 1983; Bennis & Hoekstra 1984; Longobardi 1984; Koster 1986).

(71)   *The principal asked who the silly story about upset.

Along similar lines, Pickering et al. (1994) used sentences with subjects containing relative clauses. These are also double islands, and do not allow parasitic gaps.[34]

(72)   *I realize what the decorator who painted ruined.

I therefore assume that the parser is also able to posit *wh*-chains that cross one island, but no more than one island. This is why parasitic gaps can be incrementally processed without backtracking. However, for some reason which we will not look into more closely here, two islands are more than the parser is prepared to look inside. Having noted that this appears to explain some

---

[34] The ungrammaticality of (72) is not just due to the presence of a relative clause. Norvin Richards (p. c.) points out that (i), in which the parasitic gap is inside a relative clause attached to an object NP, is significantly better than (72), although it is not entirely well-formed.
(i)   ?I know what John persuaded the men who bought to sell.

conflicting findings in the literature on processing *wh*-clauses, we proceed to apply this generalization to the processing of less exotic *wh*-clauses.

6.2.2   Reversal of preferences in Dutch and German embedded clauses

We are now in a position to return to the puzzles of long-distance gap-filling preferences across languages that we started this section with, and the relevance of the principle of Right Association. I turn first to the contrast between clause-internal and long-distance extraction in Dutch and German: recall that in these languages there is a subject-gap preference in clause internal extraction, but that preferences switch to an object-gap preference in long-distance extraction.

Consider the possible long-distance extractions available to a parser which can build *wh*-chains that span one island. I assume that syntactic islands for extraction include subjects, adjuncts, the CP/IP complex[35] among others. If the parser posits an embedded object gap it will be able to construct a single-link chain, as in (73b). In failing to postulate an intermediate trace in [Spec,CP] one island category is crossed—the CP/IP complex—but no other islands are crossed.[36] The parser will not be able to construct a similar one-link chain from embedded subject position, because this will entail the crossing of two islands, the subject position and the CP/IP complex. Instead, the parser will have to construct a two-link chain, with an intermediate trace in [Spec,CP] (73c).

(73)   a.   Welke patienten meende Jan dat de dokters bezochten?
        b.   *Which patients did Jan mean that the doctors are visiting __?*

        c.   *Which patients did Jan mean t that __ are visiting the doctors?*

Consider how Right Association evaluates these two alternatives. Measuring from the point at which the word *dat* is encountered in processing (73a), at which point the speaker knows that the gap cannot be in the matrix clause, the paths required to create an attachment site for an embedded subject gap are given in (74a), and for an embedded object gap in (74b).

(74)	a.	*subject gap*	b.	*object gap*
		3: V-VP-CP-t		2: V-VP-C(P)
		2: t-CP-C		2: C-CP-I(P)
		3: C-C'-IP-t		**3: I-IP-VP-t**

---

[35] I use this term to capture the fact that crossing CP or IP individually is licit, but crossing both CP and IP in one step generally leads to ungrammaticality. See Chomsky 1986 for discussion.

[36] For evidence that some arguments are grammatically allowed to move across clause boundaries, rather than move successive cyclically, see Pesetsky 1987, Cinque 1990, Chung 1994.

In both alternatives the worst path contains 3 steps, but in the subject gap analysis the next worst path also contains 3 steps, whereas all other paths in the object gap analysis contain 2 steps. Therefore Right Association predicts that the embedded object gap analysis is preferred over the embedded subject gap analysis.

This explains why Dutch and German show a reversal of preferences between long and short extraction. Although Right Association makes no explicit reference to any specific properties of gap-filling, and in particular makes no demands about the length or complexity of *wh*-chains, it favors the object gap dependency in (73b) because it avoids the unnecessarily long subpath forced by the intermediate trace in [Spec,CP] in (73c).[37] The next task is to figure out why Italian fails to show this reversal in gap-filling preferences.

6.2.3   Italian

De Vincenzi (1995) shows that speakers of Italian show a subject gap preference in ambiguous long and short-distance extractions alike (75–76). Italian questions are ambiguous not because objects can front over subjects, as in German and Dutch, but because Italian allows postverbal subjects, which are standardly assumed to be linked to a null subject in the normal preverbal subject position. This means that WH V NP sequences may be interpreted as containing a subject gap and an overt object NP (77a), or an object gap and a postverbal subject (77b).

(75)   Chi ha licenziato il metalmeccanico/il proprietario senza dare il
          preavviso?
          'Who fired the steel-worker/the owner without giving notice?'
          'Who did the steel-worker/the owner fire without giving notice?'

(76)   Chi pensi abbia licenziato il metalmeccanico/il proprietario sensa
          preavviso?
          'Who do you think fired...'

(77) a.   Subject gap: WH __ V NP
      b.   Object gap: WH *pro* V __ NP

Why should Italian differ from Dutch and German in gap filling preferences, when the three languages do not differ in the grammaticality of extracting arguments from embedded clauses?

I suggest that the crucial factor here comes from Rizzi's observation that Italian differs from English in the kinds of contexts that yield subjacency violations (Rizzi 1982). Rizzi shows that while both of the English sentences in (78) are ungrammatical, only one of their Italian counterparts in (79) is ungrammatical. Italian shows relative clause islands (79b), but not *wh*-islands.

---

[37] See Frazier 1993 and Kaan 1995 for alternative accounts of the contrasting gap-filling preferences between local and long-distance extraction in Dutch.

(78) a.   *your brother, to whom I wonder which stories they told
     b.   *your brother, to whom I fear the possibility that they have told everything

(79) a.   tuo fratello, a cui mi domando che storie abbiano raccontato
          your brother, to whom myself I-ask which stories they have told
     b.   *tuo fratello, a cui temo la posibilità che abbiano raccontato tutto
          your brother, to whom I fear the possibility that they have told everything

Rizzi's explanation of this was that the Subjacency parameter must be parameterized, such that the node S' is a bounding node in English (and German and Dutch, which also show *wh*-island effects) but not in Italian. In the terms we have been using here, the CP/IP complex is not an island in Italian.

Recall the reason why Right Association favors the embedded object gap in Dutch and German long-distance extractions. The subject gap analysis requires the postulation of an intermediate trace, which entails the presence of a long path which is not needed under the object gap analysis. It is impossible for the *wh*-subject to move to the matrix [Spec,CP] in one step because this would entail the crossing of two islands, which the parser does not allow. In Italian, however, since the CP/IP complex is not an island it should be possible to create a single-link filler-gap dependency from the matrix [Spec,CP] to an embedded subject position, as this involves crossing only one island—the subject position itself (80b).

(80) a.   Chi pensi abbia licenziato il metalmeccanico sensa preavviso?
     b.   Who do you think __ fired the steel-worker without warning

     c.   Who do you think pro¡ fired __ the steel-worker¡ without warning

Not surprisingly, a single-link dependency with an embedded object position is also possible (80c). I assume that this chain crosses no barriers in Italian, since CP/IP is not an island. Notice that this means that a single-link chain to the embedded object position still crosses one fewer island categories than a single-link chain to the embedded subject position. However, the parser does not care about whether or not an island is crossed in selecting a gap position, as long as two islands are not crossed, so this difference between embedded subject and object gaps in Italian is irrelevant to the parser. Now consider how Right Association evaluates potential subject and object gaps in Italian, measuring from the point at which it becomes clear that an embedded gap is required.

(81) a.   *subject gap*   b.   *object gap*
          2: V-VP-C(P)   2: V-VP-C(P)
          3: C-C'-IP-t   2: C-CP-I(P)
                                     3: I-IP-VP-t

Here the difference between a subject and an object gap is just the same as in matrix extractions in Dutch and German—the subject gap requires less the construction of fewer long paths than the object gap, and therefore Right Association opts for the subject gap. This explains the second puzzle.

6.3   Conclusion

This section has extended the coverage of the principle of Right Association to temporary ambiguities in the location of *wh*-traces. If we take the task of finding an appropriate location for a *wh*-gap to be just the second part of the task of connecting a *wh*-chain to a phrase marker, then there is no reason to expect the parser to include special purpose routines for gap-finding. I hope to have shown in this section not only that it is unnecessary to invoke special gap-finding routines in the parser, but also that Right Association is sufficient to account for a surprisingly broad range of gap-filling phenomena across constructions and languages.

The analysis presented here also depended on the assumption that the parser is freely able to construct filler-gap dependencies which cross an island. This assumption is independently motivated, in that it is necessary in order to account for why the processing of parasitic gap constructions unfolds as smoothly as it apparently does (cf. section 6.2.1.1 above). However, it is worthwhile to consider whether the empirical advantages I have claimed for Right Association here are mostly due to the assumption that islands can be crossed, and whether the cross-linguistic account I have given here could easily be translated into other accounts of gap-filling preferences.

First, the results of this section are due to Right Association at least as much as to the allowing of island-crossing dependencies. The important effect of allowing one island but no more than one island to be crossed is to determine when intermediate traces are and are not required in extraction from embedded clauses. Intermediate traces make a difference to Right Association not because they create extra chain links—Right Association does not care about this—but because intermediate traces are specifiers, and introducing specifier positions typically entails long paths between adjacent terminals, which is what Right Association objects to. However, this is only one kind of situation in which a gap in a specifier position is required—subject gaps occupy a specifier position in all of the languages discussed here, and so do object gaps in some but not all of the languages. The varying preferences for gap filling that we have reviewed in this section are determined by the combination of these factors. The preferences cannot be explained in terms of a tax on gaps in preverbal object position—this would fail to account for the variation between Dutch and German local and long distance extraction.

Nor could the preferences be explained in terms of a tax on intermediate traces. Intermediate traces do incur a penalty in Dutch and German long distance extraction ambiguities. However, if we focused on intermediate traces alone, we would be lacking an account of why subject gaps are preferred in Italian long-

distance extractions, where embedded subject and object gaps are alike in not requiring an intermediate trace.[38]

As for the second question, whether my account here could be readily translated into the terms of other structural accounts of gap-filling preferences, it is always difficult to answer this kind of question with confidence, but I can at least point to some crucial features of my account which depart from the basic assumptions of other accounts of gap-filling. The presence or absence of specifier positions plays an important role in the Right Association account of gap-filling. This hierarchical factor is independent of the relative linear order of alternative gap positions, and of the relative linear order of heads and their complements. Therefore accounts like de Vincenzi's (1991) *Minimal Chain Principle* and Frazier's (1987) *Active Filler Strategy*, which take the relative linear ordering of different alternative chains to be crucial, would need to be augmented to take account of hierarchical factors.[39]

Similar augmentation would be required of an account like the one proposed by Gibson, Hickok and Schütze (1994). Their account focuses on the satisfaction of predicate-argument relations, and attributes the contrast between English and Dutch/German local extraction preferences to the fact that English has VO order whereas Dutch and German have OV order. Since the relative ordering of verbs and objects remains constant across matrix and embedded clauses in all of these languages, no reversal of gap-filling preferences is expected under Gibson et al.'s theory. It would be possible to extend the coverage of their theory enrich it with factors which refer to islandhood

---

[38] Right Association makes a related prediction about English, which to my knowledge has not been tested. In English long-distance extractions I assume that a single link dependency with an object gap can be built, crossing only one island, but a single link dependency with a subject gap cannot be built, as it would cross two islands. Therefore we expect filled gap effects for embedded object gaps in English, as Clifton and Frazier (1989) found, but not for embedded subject gaps.

If we turn to extraction from a non-finite embedded clause, however, it should be possible to form single link dependencies with both subject and object positions, assuming that either there is no CP/IP complex in non-finite clauses, or that for some reason CP/IP is not an island in these environments. We predict this situation to be like Italian, then, in that the dependencies to subject and object position have the same number of chain links. Importantly, though, Right Association predicts that English infinitival clauses will not parallel Italian finite embedded clauses in gap-filling preferences. Given our assumption that objects are extracted from a complement position to the right of the verb in English, but not in Italian, the object gap preference should is predicted to persist in extraction from English infinitivals. If this contrast between English and Italian is verified, it would clearly not be explainable in terms of a tax on intermediate traces.

[39] De Vincenzi's *Minimal Chain Principle* does include an additional statement about avoiding extra chain links, which could be used in combination with my assumptions about island crossing to account for the reversal of preferences between long and short extraction in German and Dutch. However, this still leaves no account for why in local extraction English shows an object gap preference but Italian/Duch/German show a subject gap preference.

parameters and to the presence or absence of intermediate traces, but this would mark a departure from Gibson et al.'s otherwise semantically based account.

The point of these remarks has not been to claim that Right Association provides the only possible account of cross-linguistic contrasts in gap-filling preferences. Rather, I hope to have shown that the reason for the success of Right Association this domain lies in its attention to properties of hierarchical structure. Other theories may incorporate be extended to cover the range of facts covered here, but I suspect that in order to do so they will be forced to add mechanisms which pay attention to hierarchical structure. The advantage of a theory based on the principle of Right Association is that hierarchical structure is what it most naturally pays attention to.

7. Right Association and right branching syntax

Having discussed at length the promise of the Right Association principle in (3) in the resolution of processing ambiguities, the traditional domain of parsing theories, I now turn to some recent problems in the theory of phrase structure. The issues come out of a quite different tradition from ambiguity resolution in sentence comprehension, but I hope to show that the phrase structure construction algorithm presented here can solve some problems in both areas.

I will be focussing on a pair of problems for the theory of phrase structure that are brought out quite clearly in Pesetsky (1995). Both are the result of applying straightforward tests of constituency like the ones we used in section 3 above. The first problem challenges one of the most basic assumptions about the syntax-semantics mapping: namely, that the internal argument of a head is the syntactic complement of that head. The second problem is both more syntax-internal and more puzzling: it concerns conflicting evidence that shows that the structure of VP is both chronically left-branching and chronically right-branching. In both cases I aim to show that when we take into consideration how phrase structures are constructed, in particular taking into account the role of the principle of Right Association, both problems can be avoided.

7.1 Complements that are not arguments, arguments that are not complements

In section 3 above we saw how constituency tests based on binding and coordination indicate that the structure of a VP with multiple complements and/or adjuncts is more right-branching than has traditionally been assumed. To a rough approximation, rightward equals downward (see 17–18 above). The structures may take a while to get used to, but they do at least preserve some of our basic assumptions about the structural relationship between heads and their arguments. The arguments of a head are either specifiers or complements of that head at some level.

However, the reader may have noticed that the sentences in (17a-b), which I passed over with great haste, actually present a problem, either for the c-command tests which I have been assuming, or for standard assumptions about the relationship of heads and complements. In both (17a) and (17b), which are repeated below, the NP inside the first PP binds an element in the second PP. Given the completely uncontroversial assumption that PPs like *about these people* are constituents, in which PP dominates NP, this binding ought to be impossible—the antecedent should fail to c-command the anaphor, given that it is embedded inside a PP.

(17) a. Sue gave books to these people$_i$ on each other's$_i$ birthdays.
      b. Sue spoke to Mary about these people$_i$ in each other's$_i$ houses on Tuesday.

The problem of PP complements appearing to c-command things they shouldn't has been around for some time (cf. Reinhart 1981, pp. 631–632; 1983, pp. 175–176; Jackendoff 1990, pp. 430ff.), and has generally been either left as an unsolved problem, or posed as a problem for the notion of c-command. Pesetsky (1995) chooses the other course of action, and argues that examples like these force us to rethink the structure of complementation: Pesetsky suggests that we should retain the standard definition of c-command, according to which a node c-commands its sister and all subparts of its sister, and instead question the assumption that *about these people* is a constituent. He proposes that in sentences like (17a) the preposition *to* is sister to a constituent containing *these people on each other's birthdays*, and therefore *these people* is the specifier of the complement of *to*, as shown in (83). This constituency is lent further support by the coordinations in (82a-b), which are variants of (18a-b) above.

(82) a. Sue will speak to Mary about [linguistics on Friday] and [philosophy on Tuesday]
      b. Kremer will perform this concerto in [Rome on Tuesday] and [Somerville on Thursday]

(83)
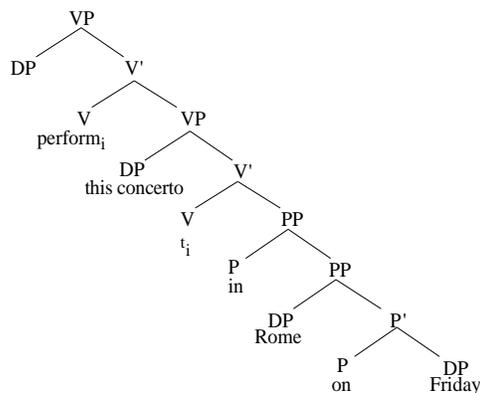
This approach therefore entails that the internal arguments of a head need not be the syntactic complement of the head at any level. This is somewhat of a radical assumption, but it is the logical consequence of using standard syntactic argumentation and fairly simple examples. This is the first problem.

7.2   VPs that are both left-branching and right-branching?

By applying binding and coordination tests to our heart's content we are led to postulate VP structures something like (83). But as Pesetsky (1995) points out, if we apply another kind of constituency test, involving movement, we are led to a rather different view of VP.

As we have already seen, binding and coordination tests indicate that any subpart of VP tht includes the *right edge* of VP is a constituent. Tests of movement, on the other hand, indicate quite the opposite. The contrast between (610) and (611) indicates that any subpart of VP that includes the *left* edge of VP is a constituent.

(84)   a.   *[To John about himself] Mary spoke __.
       b.   *I wonder [to whom about physics] Mary spoke __.
       c.   *[In the garden on Tuesday] Bill said [he gave the book to Sue __].

(85)   a.   ...and give the book to them in the garden on Tuesday he did.
       b.   ...and give the book to them in the garden he did on each other's birthdays.
       c.   ...and give the book to them he did in the garden on each other's birthdays.

These movement tests motivate a structure for VP like (86), in which rightward roughly equals upward.[40]

(86)
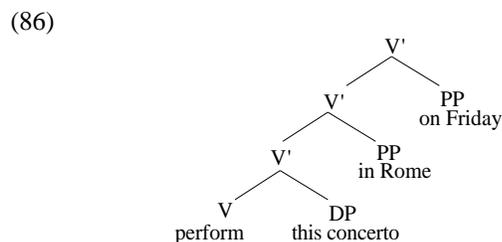
Structures like this are of course not novel in the literature. The problem is how to reconcile the results of the binding and coordination tests,

---

[40] The generalization that subparts of VP that include the left edge of VP can always be fronted has one important exception. The verb and the first object of a ditransitive verb cannot strand the second complement, as in (i) and (ii).

(i)   *...and put the book he did on the table.
(ii)   *...and give John he did the book on his birthday.

which yield the structure in (83), with the results of movement tests, which yield the structure in (86). One response would be to try to deny the existence of one of the structures. Pesetsky's response is to bite the bullet and assume that *both* structures are genuine, and that speakers represent both a right-branching *cascade structure* and a left-branching *layeredstructure* in parallel. Of course this is not the simplest theory of grammar imaginable, but it is a conclusion that does seem to be forced upon us by the facts.

In the next section I will try to show that it may be possible to have our cake and eat it, if we assume that the Right Association principle is a condition on grammatical derivations. I'll show how the effects of both left-branching and right-branching structures can emerge without assuming multiple parallel representations. This will have the additional consequence that we can maintain the assumption that semantic complements are syntactic complements.

An account of the apparently left-branching and right-branching effects that we have seen above will need to answer the questions in (87).

(87) a. Why do complex VPs show right-branching (cascade) structures?
     b. Why the mismatch between semantic and syntactic complementation in multiple PP sequences?
     c. Why is it possible to move some cascade non-constituents?
     d. Why is impossible to move some cascade constituents?

In government-binding theory (e.g., Chomsky 1981, 1986) it was standardly assumed that D-structures are subject to various well-formedness constraints, but that D-structures are not derived by any deterministic concatenation process. In other words, D-structures could be built in any order or manner, but once constructed they had to satisfy a number of filters: the best known of these filters is the *theta-criterion*. There have been more recent proposals to the effect that global conditions on D-structures should be replaced by a combination of LF constraints and local well-formedness conditions on concatenation of items (Chomsky 1994), but still no deterministic procedure for concatenation is assumed.

Let us suppose that the structure of clauses, and of VP in particular, is generated in a more derivational manner than has generally been assumed in government-binding theory and its descendants. I assume that VP is derived by concatenation of items from left to right, i.e., in the same order that sentences are produced and heard, and that there are conditions on concatenation that make the process relatively deterministic.

A given ordering of elements leaves open many possible ways of hierarchically structuring those items. If we exclude those hierarchical structures that are uninterpretable, or that lead to the wrong interpretation, then the range of possible structures is significantly reduced. However, considerations of interpretation will not the range of possible structures to one candidate in all cases. For example, interpretability is a fairly weak constraint on the structure of

a complex VP: it is commonly assumed that PP modifiers like *on Tuesday* or *in the garden* are interpreted as conjoined sentences, in which the PP is predicatedof an event argument. Some semantic proposals in the neo-Davidsonian tradition extend this analysis to arguments. Under this kind of approach, the interpretation of a sentence like *John saw Mary on Tuesday* would amount to something like *There was a seeing event, John was the agent, Mary was the theme, and it happened on Tuesday*.[41] If this is all that the interpretive system has to extract from a sentence, then its main task will be to access the thematic roles of the arguments.

The main lesson of this example is that VP structure is often underdetermined by the demands of interpretation. I suggest that these instances of structural underdetermination by semantics should be treated as parallels to the more familiar on-line parsing ambiguities discussed in sections 2–6, where structure is underdetermined because the intended interpretation is unknown to the listener. In both types of situation I contend that the principle of Right Association determines the structural choices made.

First I will show how Right Association can explain both cascade and layered constituency effects, without recourse to multiple constituency. Next I show how my account for why specific constituency tests yield the results they do. Finally, I show a novel prediction of my account of why right-branching cascade structures exist.

7.3   Multiple constituency without multiple structures

7.3.1   Cascade structures

Consider now how a complex VP *give candy to the children on their birthday* will be constructed, if it is to maximally satisfy Right Association. Let's assume that we start with just a verb, such as *give*. This both has the category label V and is the maximal projection of that category, as in (88).

(88)

V(P)
give

Next an object NP *candy* is added. This must be concatenated with the verb in a way that best satisfies Right Association. As usual, the best way to satisfy Right Association is to be a sister of the preceding head. In order to achieve this, the V(P) node must be extended, so that VP may branch, as in (90). I assume that new nodes may be freely generated above or below a node, subject

---

[41] cf. Vlach 1983; Higginbotham 1985; Parsons 1990.

to the constraint in (89), which simply ensures that dominance relations are preserved from one step in the derivation to the next.[42]

(89)   *Preserve Dominance*
    a.   If X is immediately dominated by Y in the input structure, then X is immediately dominated by Y in the output structure.
    b.   If X immediately dominates Y in the input structure, then X immediately dominates Y in the output structure.
        [For both of these conditions, X, X', XP are taken to be identical, as in Chomsky 1994.]

(90)

```
        VP
       /  \
      V     NP
     give  candy
```
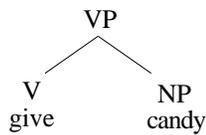
At this stage in the derivation, then, [V NP] is a constituent. Also, NP is both a syntactic and semantic complement of V.

Next the P head *to* is concatenated with the structure in (90). It must be attached to [V NP] in the manner that best satisfies Right Association. P could be attached by further projecting VP and making the P the sister of [V NP], as in (91), but there is an attachment which better satisfies Right Association, shown in (92). First a copy of V is inserted to the right of NP: this projects such that it is dominated by a VP node: this is one of the options that is freely available, given the condition in (89). This new VP node is attached such that it dominates NP and is sister of the original V head. This "lowering" of NP is licit, because it is still immediately dominated by a projection of the same head. (I assume that the lower V is identical to the higher V, following Larson 1988). In addition, NP and the new V head are sisters, thereby optimally satisfying Right Association. This intermediate structure is shown in (92a). The P is now straightforwardly attached by projecting the rightmost V, and attaching P as sister of V, again optimally satisfying Right Association, as can be seen in (92b).

(91)
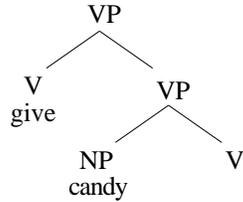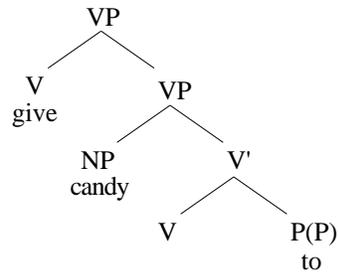
```
          VP
        /    \
      V'      P(P)
     /  \      to
    V    NP
  give  candy
```

_____
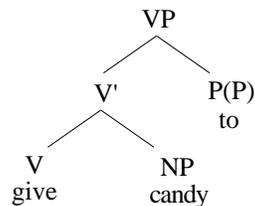
[42] See Gorrell (1995) for an approach to structural parsing which applies related constraints to different problems in on-line parsing.

(92)   a.   b.

```
            VP                                        VP
          /    \                                    /    \
        V       VP                                V       VP
       give    /  \                              give    /   \
            NP      V                                 NP       V'
           candy                                    candy     /  \
                                                            V      P(P)
                                                                    to
```
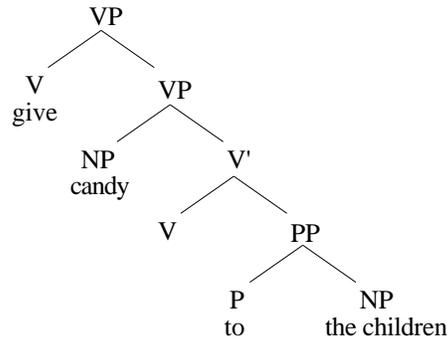
Next the NP complement of P, *the children,* is attached. This is achieved in the now familiar manner, by projecting P and attaching NP as sister of P, as in (93). Right Association is maximally satisfied.

(93)

```
              VP
            /    \
          V       VP
         give    /   \
              NP       V'
             candy    /   \
                    V       PP
                          /    \
                        P        NP
                        to    the children
```
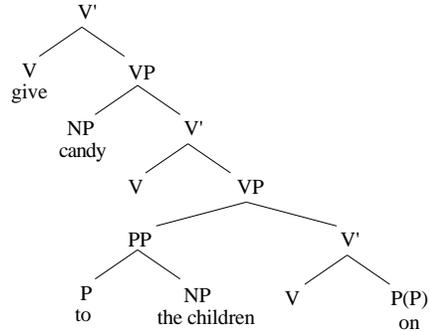
Notice two important features of this derivation. First, a right-branching cascade structure is emerging. Second, certain constituents are changing from one step in the derivation to the next. At this stage in the derivation, [V NP PP] is a constituent, and the semantic complement of P is also its syntactic complement. However, [V NP] is no longer a syntactic constituent in (93) as it was in (90).

Next the PP modifier *on their birthday* is attached. The P could be attached as sister of V, as in (94a), which parallels (92b). However, since PP is not a selected phrase it does not need to be sister of a selecting head at any point in the derivation. Consequently, there is an alternative structure available which better satisfies Right Association. NP can be projected to allow P to be its sister, and then P can be projected to allow its NP complement to be attached as its sister (94b). Both of these steps maximally satisfy Right Association.[43]
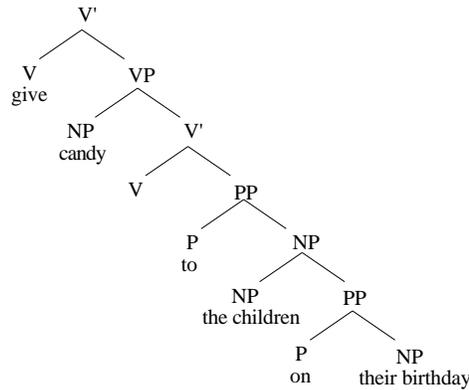
---

[43] I assume that structures like (94b) in which PP is adjoined to NP are not necessarily interpreted as modifying NP. Modifiers are predicated either of (i) a head which is coindexed with an operator inside the modifier, as is standardly assumed for restrictive relative clauses, or (ii) a locally c-commanding event-denoting head,

(94) a.

```
                    V'
                  /    \
                 V      VP
               give    /  \
                     NP     V'
                   candy   /  \
                          V    VP
                             /      \
                           PP        V'
                          /  \      /  \
                         P    NP   V    P(P)
                         to  the children  on
```

b.

```
                    V'
                  /    \
                 V      VP
               give    /  \
                     NP     V'
                   candy   /  \
                          V    PP
                              /  \
                             P    NP
                             to  /  \
                               NP    PP
                          the children /  \
                                      P    NP
                                      on  their birthday
```

Now [V NP PP PP] is a constituent, as are [NP PP PP], [PP PP], and
even the final [NP PP]. Meanwhile, the PP *to the children* is no longer a
constituent. This is precisely the constituency predicted by the binding and
coordination tests above.

At this point, then, a complex VP has been built in a way that
maximally satisfies Right Association at each step of the derivation, and the
resulting structure is a radically right-branching *cascade* structure. This provides
an answer to the first question in (87): the principle of Right Association is to
blame for cascade structures. It also suggests an answer to (87b): why are there
apparent mismatches between semantic and syntactic complementation? If the
derivation in (88–94) is correct, then semantic complements do enter the syntax
as syntactic complements. This may be sufficient for theta-marking. It is only
later in the derivation that mismatches between syntactic and semantic
complementation emerge.

---

where the appropriate locality relation remains to be defined. I assume that (94b) is
an appropriate structure for the PP to be predicated of V.

7.3.2   *Layered stuctures*

From section 7.3.1 we now have a possible explanation for how cascade structures arise, and we may be able to avoid one of their more unfortunate consequences for the theory of complementation. But the more difficult question, which still remains, is how to make sense of the evidence for left-branching layered structures, without actually postulating parallel structures for every sentence.

The first important point to notice is that the grammaticality effects motivating layered structures are different from the effects motivating cascade structures. Cascade structures are motivated both by constituency tests (coordination) and by c-command tests (e.g., binding). The movement tests motivating layered structures, on the other hand, only show that such and such a part of a complex VP can be treated as a constituent. The evidence for the c-command relations predicted by layered structures is much less direct. The movement tests show that strings including the left edge of VP are constituents, paralleling what coordination tests tell us about cascade structures. But they do not directly diagnose c-command relations from right to left—we merely infer right-to-left c-command in order to make sense of the constituency results.[44] This places much looser constraints on an account of the layered structure effects than on an account of the cascade structure effects.

The second important observation was foreshadowed in the derivation of the cascade VP structure in (88–94). All of the constituents of a layered VP structure are constituents at some point of the derivation of the cascade structure; it is only later in the derivation that the layered constituents cease to be cascade constituents.

In order to capture the layered-syntax movement effects, I assume the copy theory of movement, which assumes that links in a syntactic chain are copies of one another, with only one of the copies being pronounced (Chomsky 1993). The constituents of layered syntax can be freely generated in A' positions and copies generated in VP (95). Both copies of this partial VP will have an internally *right*-branching structure.[45] Subsequent to this, additional VP material can be added to the lower copy (96).

(95)   ...and [give [candy [to the children]]] he did [~~give~~ ~~[candy~~ ~~[to the~~
            ~~children~~]]]

(96)   ...and [give [candy [to the children]]] he did [~~give~~ ~~[candy~~ ~~[to~~ ~~[the~~
            ~~children~~ [in the garden]]]]]

---

[44] Scopal relations among adverbial modifiers have been used as a c-command test. However, I argue below that this kind of test may be misleading.

[45] This analysis is compatible with both a derivational (movement) and non-derivational (chain-formation: Rizzi 1986; Brody 1992) account of A' chains.

This provides an account of the question in (87c): why can cascade non-constituents move? They can move because they are constituents at some point in the derivation.

This account of movement has the effect of making the question in (87d) more acute. We already know that certain strings which are constituents in final cascade structures cannot undergo movement. In Pesetsky's dual system this is explained by layered structures, which severely restrict the available structures for movement. However, the account that I have suggested, according to which it is sufficient for a string to be a cascade constituent at *some* point in the derivation in order for it to move, now overgenerates possible movement structures. For example, in a sequence V NP P NP P NP P NP I predict that *any* contiguous string of elements should be able to undergo movement, a prediction that is clearly false, as can be seen from (97). In the absence of layered structures I need an alternative account of the restrictions on movement.

(97)   a.   *...and [give candy to] he did the children in the garden on Tuesday.
        b.   *...and [give candy to the children in] he did the garden on Tuesday.

The appropriate generalization about possible movements from complex VPs seems to be the following: individual arguments (NP or CP) or modifiers (PP) can be fronted, and sequences of these categories can be fronted too, but only if the sequence includes the verb. Therefore, the most important type of movement to exclude is cases like (98).

(98)   *Sue asked me [in which capital city on Tuesday] Bill said [he gave the book to Mary __]

Although I do not have a complete story to tell about this restriction at this point, I suggest that the restriction follows from the same interpretive factor that makes cascade structures for multiple PPs *possible*. Recall from (94) and footnote 43 the claim that modification relations are permitted when the modifier is locally c-commanded by the event-denoting head that is it predicated of. I assume that this weak interpretive condition constrains the structure of multiple PP modifiers in complex VPs. However, if a [PP PP] structure is fronted without a verb, the only possible construal of the second PP will be as a modifier of the NP in the first PP, since there is no locally c-commanding event-denoting head: notice that this is a possible, albeit rather bizarre, interpretation of sentences like (98). Only when the multiple PP sequence is fronted together with the verb is it possible for PPs other than the first to be construed as modifiers of the verb.

Therefore my account of why certain sequences which *are* cascade constituents cannot be fronted assumes that the relevant restrictions follow from independent conditions. This account is clearly rather speculative at the moment.

7.4 Why the constituency tests show what they show

The account of cascade and layered constituency effects just presented has the additional advantage of explaining why it is that each specific constituency test yields the results it does. In other words, we can account for why VP-fronting tests point to left-branching constituency, while binding and coordination tests point to right-branching constituency, and not the reverse.

Since the apparent constituents in the left-branching structure are only constituents at the stage in the construction of the phrase marker where they are the *rightmost* phrase in the sentence, it follows that these constituents can only ever enter into dependencies with phrases to their left: once material is added to their right, they cease to be constituents. This accounts for why leftward movement tests diagnose left-branching structures.

Meanwhile, the tests which diagnose the surprising properties of right-branching, cascade structures (eg., NPs that are not sisters of the Ps that select them) are tests involving syntactic relations with elements to the right. For example, the unexpected position of the NP *these people* in (17) is diagnosed by the fact that it can bind an anaphor to its right; similarly the position of the NP *linguistics* in (82) is diagnosed by the fact that it can coordinate with the PP to its right *on Friday*, to the exclusion of the P *about.* Given our account of how phrase markers are constructed from left-to-right, and of where in a derivation arguments can cease to be the sister of the head selecting them, we predict that it will be the tests involving rightward looking relations that diagnose right-branching structures.

To my knowledge, this correspondence between the linear properties of a given constituency test and the direction of branching that the test predicts goes unexplained in theories which assume multiple parallel constituency in order to account for the conflicting results of the different diagnostics.

This leaves at least one loose end still to deal with. There is one further kind of evidence for the c-command relations predicted by layered structures, which I have ignored up to this point. This evidence comes from the relative scope of multiple postverbal modifiers. It has been claimed that modifiers closer to the verb take narrower scope than modifiers to their right (Andrews 1982, 1983; Pesetsky 1995), as seen in examples like (99), taken from Pesetsky 1995, chapter 7. (99a) most naturally restricts quartet playing in foreign countries to weekends, while (99b) most naturally restricts quartet playing on weekends to foreign countries.

(99)   a.   John plays quartets in foreign countries on weekends.
       b.   John plays quartets on weekends in foreign countries.

In the right-branching structures which I am claiming to be the only structures available for complex VPs, the correct scopal relations are not predicted by c-command relations.[46]

I suspect that it may be an oversimplification to equate preferences for what modifiers restrict with c-command relations. Consider the variants of (99) in (100).

(100) a.   John plays quartets in foreign countries$_i$ on their$_i$ national holidays.
      b.   John plays quartets on federal holidays$_i$ in their$_i$ first 5 years of existence.

My intuition in these sentences is that each of the two PP modifiers is dependent on the other, although in slightly different ways. The rightmost PP in each case contains a pronoun bound by the NP inside the first PP. This implies that the first PP takes scope over the second. At the same time, though, the preference on restriction relations reported for (99) remains: in (100a) quartet playing in foreign countries is restricted to national holidays, and in (100b) quartet playing on federal holidays is restricted to the first 5 years in which that holiday is celebrated. If these restriction preferences entail wide scope for the second PP, we are left with a scopal contradiction. Alternatively, we could assume that the restriction relation corresponds to a topic-focus partitioning of the VP which is either not directly read off c-command relations or is determined by LF c-command after focussed material has undergone A' movement.

7.5   A constraint on radically right-branching structures

Throughout this paper I have been arguing for the view that right-branching structure are the 'default' syntactic arrangement of strings of items. In the last few sections I have presented this as an explanation of why complex VPs appear to be far more radically right-branching than has traditionally been assumed: they are right-branching first because no interpretive constraint prevents them from being right-branching, and second, because the principle of Right Association prefers them to be right-branching.

Support for this view of right-branching structures as default structures, constrained by interpretation comes from the fact that the standard diagnostics of right-branching-ness that we have been using throughout this paper fail in precisely those contexts in which interpretation does place tighter demands on constituency.

---

[46] Notice, however, that the c-command relations in the cascade structures are not in direct conflict with the intuitions on (99). They just tell us nothing about the relative scope of the PP modifiers, since the PPs are not even constituents in the cascade structure.

We have already seen that NPs which are selected by a preceding P can behave syntactically as specifiers of a following adjunct—provided that the following adjunct lacks internal predicational structure of its own (101).

(101) a.   Kremer will perform this piece in [PP Rome [P' on Tuesday]] and [PP Chicago [P' on Friday]]
   b.   John gave books to [PP every child$_i$ on his$_i$ birthday]

However, if the adjunct has its own internal predicational structure, as for example in a small clause or secondary predicate, it no longer allows a preceding NP argument of P to behave as its specifier, as shown by the examples in (102), pointed out to me by Alec Marantz. Nevertheless, there is evidence that the adjunct does attach fairly low in the VP (103).

(102) a.   *John mailed parcels to [SC Milan [PRO naked]] but [SC Pisa [PRO fully dressed]]
      b.   ??The teacher sent a thank-you letter to [SC every child$_i$ dressed in his$_i$ gift]

(103) a.   John mailed parcels [PP to Milan naked] but [PP to Pisa fully dressed]
      b.   The teacher sent [VP every child$_i$ a thank-you letter dressed in his$_i$ gift]

These examples imply that the secondary predicates form a constituent with the preceding PP, as in (104), but do not form a constituent with the NP of the PP. This structure is the 'next best' alternative for Right Association, and it closely resembles the structure that was rejected in (94a).

(104)
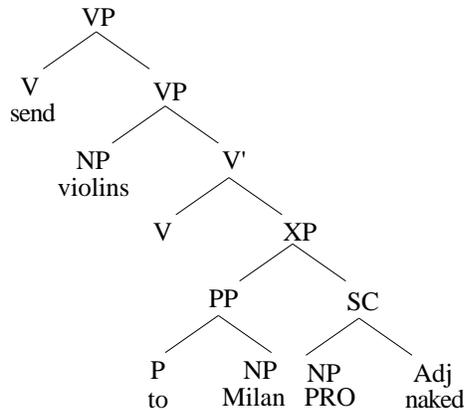
This concludes the discussion of Right Association and VP structure. To summarize, I have suggested that the generation of grammatical structures and the process of comprehending utterances are more similar than is standardly

assumed. Both domains contain situations in which the appropriate structure building is underdetermined. In these situations the principle of Right Association determines the choice among otherwise equivalent structures. In grammatical derivations, this provides a possible account of apparently contradictory conclusions from different constituency tests. As shown in section 7.4, a particular attraction of this account of the conflicting constituency results is that it actually explains why the different constituency tests produce the results they do.

8.   Conclusion

Right Association may have much broader consequences than has traditionally been thought. I have argued here that this principle alone accounts for a wide range of structural preferences in parsing, and also explains some puzzles in the theory of phrase structure. Given that most theories of parsing assume that some analog of Right Association is needed, this is an attractive result.

Of course, these remarks are very preliminary at present, and numerous questions arise, involving topics such as the interaction of Right Association with non-structural factors in parsing, and why some structural preferences are easy to recover from while others are not. Additionally, the assumption that syntactic derivations involve left-to-right concatenation of items demands a rethinking of many basic issues in syntax. For now I leave these problems as acknowledged, but unsolved.

References

Abney, S. (1987). Licensing and parsing. *Proceedings of NELS 17*, Amherst, Mass.: GLSA.

Abney, S. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research* 18, 129–144.

Andrews, A. (1982). A note on the constituent structure of adverbials and auxiliaries. *Linguistic Inquiry* 13, 313–317.

——— (1983). A note on the constituent structure of modifiers. *Linguistic Inquiry* 14, 695–7.

Barss, A. & H. Lasnik (1986). A note on anaphora and double objects. *Linguistic Inquiry* 17, 347–354.

Bennis, H. & T. Hoekstra (1984). Gaps and parasitic gaps. *The Linguistic Review* 4, 29–87.

Bever, T.G. (1970). The cognitive basis for linguistic structures. In: J.R. Hayes (ed), *Cognition and the Development of Language*, New York: Wiley, 279–352

Boland, J. & M. Tanenhaus (1991). The role of lexical representations in sentence processing. In: G.B. Simpson (ed), *Understanding Word and Sentence*, North-Holland: Elsevier, 331–366.

Brody, M. (1992). Theta theory and arguments. *Linguistic Inquiry* 23, 1–25.

Chomsky, N. (1973). Conditions on transformations. In: *A festschrift for Morris Halle*, ed. S. Anderson & P. Kiparsky, New York: Holt, Reinhart & Winston, 232–286

——— (1981). *Lectures on government and binding*. Dordrecht: Foris.

——— (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

——— (1986). *Barriers.* Cambridge, MA: MIT Press.

——— (1993). A minimalist program for linguistic theory. In: K. Hale & S.J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press.

——— (1994). *Bare phrase structure*. MIT Occasional Papers in Linguistics #5. MITWPL.

——— (1994). *Wh*-agreement and "referentiality" in Chamorro. *Linguistic Inquiry* 25, 1–44.

Cinque, G. (1990). *Types of A' dependencies*. Cambridge, MA: MIT Press.

Clifton, C., S. Speer & S. Abney (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language* 30, 251–271.

Crain, S. & J.D. Fodor (1985a). How can grammars help parsers? In D.R. Dowty, L. Karttunen & A. Zwicky (eds.), *Natural language parsing: Psychological, computational and theoretical perspectives*, Cambridge, UK: Cambridge University Press, 94–128.

——— (1985b). Rules and constraints in sentence processing. *NELS* 15, Amherst, MA: GLSA, 87–104.

——— (1987). Sentence matching and overgeneration. *Cognition* 26, 123–169.

Cuetos, F. & D. Mitchell (1988). Cross-linguistic differences in parsing: restrictions on the use of the Late Closure strategy in Spanish. *Cognition* 30, 73–105.

de Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.

——— (1995). Syntactic analysis in sentence comprehension: Effects of dependency types and grammatical constraints. ms: CNR, Rome.

Engdahl, E. (1983). Parasitic gaps. *Linguistics and Philosophy* 6, 5–34.

Ferreira, F. & J.M. Henderson (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory and Cognition* 16, 555–568.

Ford, M., J. Bresnan & R. Kaplan (1982). A competence based theory of syntactic closure. In J. Bresnan (ed.), *The Mental Representation of Grammatical Relations*, Cambridge, MA: MIT Press, 727–796.

Frazier, L. (1978). On comprehending sentences: Syntactic parsing strategies. PhD thesis, University of Connecticut.

——— (1987a). Sentence processing: A tutorial review. In: M. Coltheart (ed) *Attention and Performance XII*. Hillsdale, N.J.: Erlbaum.

——— (1987b). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5, 519–560.

——— (1993). Processing Dutch sentence structures. *Journal of Psycholinguistic Research* 22, 85–108.

——— (1994). Sentence (re-)analysis. ms, UMass, Amherst.

Frazier, L. & C. Clifton (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4, 93–126.

Frazier, L. & G. Flores d'Arcais (1989). Filler-driven parsing: a study of gap filling in Dutch. *Journal of Memory and Language* 28, 331–344.

Frazier, L. & J.D. Fodor (1978). The sausage machine: a new two-stage parsing model. *Cognition* 6, 291–326.

Fodor, J.D. & L. Frazier (1983). Local attachment in a one-stage parser. ms.

Frazier, L. & K. Rayner (1982). Making and correcting errors during sentence processing: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14, 178–210.

Freedman, S. & K. Forster (1985). The psychological status of overgenerated sentences. *Cognition* 19, 101–131.

Gibson, E. (1991). A computational theory of human language processing: Memory limitations and processing breakdown. PhD thesis, Carnegie Mellon University.

Gibson, E. & K. Broihier (to appear). Optimality theory and human sentence processing. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky (eds), *Is the best good enough?*, Cambridge, MA: MIT Press.

Gibson, E., G. Hickok & C.T. Schütze (1994). Processing empty categories in a parallel parsing framework. *Journal of Psycholinguistic Research* 23, 381–405.

Gibson, E., N. Pearlmutter, G. Hickok & E. Canseco-Gonzalez (1994). Recency preferences in the human sentence processing mechanism. ms, MIT, UIUC, UCSD & Reed College. to appear in *Cognition*.

Gilboy, E., J.-M. Sopena, C. Clifton & L. Frazier (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition* 54, 131–167.

Gorrell, P. (1995). *Syntax and parsing*. Cambridge, UK: Cambridge University Press.

Haverkort, M. (1986). Parasitic gaps: multiple variable binding, connectedness, ATB or chain composition. Undergraduate dissertation, University of Nijmegen.

Higginbotham, J. (1985). On semantics. *Linguistic Inquiry* 16, 547–93.

Holmes, V. (1987). Syntactic parsing: In search of the garden path. In: M. Coltheart (ed.), *Attention and Performance XII*, Hillsdale, NJ: Erlbaum, 587–599.

Inoue, Y. & J.D. Fodor (1994). The diagnosis and cure of garden paths. ms: CUNY Graduate Center.

Juliano, C. & M.K. Tanenhaus (1993). Contingent frequency effects in syntactic ambiguity resolution. *Proceedings of the 15th annual conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum, 593–598.

Kaan, E. (1995). Subject-object ambiguities in Dutch: A constraint-based approach. ms, University of Groningen.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry* 14, 223–249.

Kayne, R. (1994). *Antisymmetry of syntax*, Cambridge, MA: MIT Press.

Kennedy, A., W. Murray, F. Jennings & C. Reid (1989). Parsing complements: Comments on the generality of the principle of minimal attachment. *Language and Cognitive Processes* 4, 51–76.

Kennison, S. & C. Clifton (1994). Resolving the ambiguity of the word *her*. Paper presented at the Psychonomics Society, St. Louis, MO.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition* 2, 15–47.

Koster, J. (1986). *Domains and dynasties*. Dordrecht: Foris.

Kurtzman, H. & L. Crawford (1990). Processing parasitic gaps. In *Proceedings of NELS 21*, Amherst: GLSA, 217–231.

Larson, R. (1988). On the double object construction. *Linguistic Inquiry* 19, 335–391.

——— (1990). Double objects revisited: reply to Jackendoff. *Linguistic Inquiry* 21, 589–632.

Longobardi, G. (1984). Connectedness and island constraints. In J. Guéron, H.-G. Obenauer & J.-Y. Pollock (eds), *Grammatical Representation*. Dordrecht: Foris, 169–185.

McDonald, M.C., N. Pearlmutter & M. Seidenberg (1994). The lexical nature of syntactic ambiguity resolution. to appear in *Psychological Review*.

Mitchell, D. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In: M. Coltheart (ed.), *Attention and Performance XII*. Hillsdale, N.J.: Erlbaum.

Mitchell, D., M. Corley & A. Garnham (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Language, Memory and Cognition*, 18, 69–88.

Parsons, T. (1990). *Events in the subatomic semantics of English*. Cambridge, MA: MIT Press.

Pesetsky, D. (1987). Wh-in-situ: movement and unselective binding. In: E. Reuland & A. ter Meulen (eds), *The Representation of Indefiniteness*, Cambridge, MA: MIT Press, 98–129.

——— (1995). *Zero syntax*. Cambridge, Mass.: MIT Press.

Pickering, M., S. Barton & R. Shillcock (1994). Unbounded dependencies, island constraints, and processing complexity. In C. Clifton, L. Frazier & K. Rayner (eds), *Perspectives on Sentence Processing*, Hillsdale, NJ: Erlbaum, 199–224.

——— & R. Shillcock (1992). Processing subject extractions. In: H. Goodluck & M. Rochemont (eds), *Island Constraints*, Dordrecht: Kluwer, 295–320.

Pritchett, B. (1988). Garden-path phenomena and the grammatical basis of language processing. *Language* 64, 539–576.

——— (1992). *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.

Rayner, K., M. Carlson & 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.

Reinhart, T. (1981). Definite NP anaphora andc-command domains. *Linguistic Inquiry* 12, 605–636.

——— (1983). *Anaphora and semantic interpretation*. London: Croom Helm.

Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht: Foris.

——— (1986). On chain formation. In: H. Borer (ed.), *The Grammar of Pronominal Clitics, Syntax and Semantics 19*. New York: Academic Press.

Schütze, C. (1995a). Another look at minimal attachment: Thematic factors in PP attachment. Paper presented at the Canadian Linguistics Association meeting, Montreal.

——— (1995b). Argumenthood and PP attachment.

Stabler, E. (1994). Syntactic preferences in parsing for incremental interpretation. ms, UCLA.

Stowe, L. (1986). Evidence for on-line gap location. *Language and Cognitive Processes* 1, 227–245.

——— (1992). The processing implementation of syntactic constraints: The sentence matching debate. In: H. Goodluck & M. Rochemont (eds.), *Island Constraints*, Dordrecht: Kluwer, 419–443.

Tanenhaus, M., G. Carlson & J. Trueswell (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes* 4, 211–234.

Taraldsen, T. (1981). On the theoretical interpretation of a class of marked extractions. In A. Belletti, L. Brandi & L. Rizzi (eds), *Theory of markedness in generative grammar*, Pisa: Scuola Normale Superiore, 475–516.

Trueswell, J. & M.K. Tanenhaus (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier & K. Rayner (eds), *Perspectives on Sentence Processing*, Hillsdale, NJ: Erlbaum, 199–224.

Trueswell, J., M.K. Tanenhaus & C. Kello (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19, 528–553.

Vlach, F. (1983). On situation semantics for perception. *Synthese* 54, 129–52.

Zwart, J.-W. (1993). Dutch syntax: A minimalist approach. PhD thesis, University of Groningen.

Department of Linguistics & Philosophy
20D-219
MIT
Cambridge, Mass. 02139
USA

cphill@mit.edu