Kayne
Kap. 1 – 4.4

Preface

It is difficult to attain a restrictive theory of syntax. One way of making progress toward that goal is to restrict the space of available syntactic representations, for example, by imposing a binary branching requirement, as I suggested in earlier work. The present monograph proposes further severe limitations on the range of syntactic representations accessible to the human language faculty.

The primary locus of inquiry is the relation between hierarchical structure and linear order. It is standardly assumed that that relation is a flexible one, that is, that linear order can be associated with hierarchical structure quite freely. A head (H) and its complement (C) can be associated in some languages with the order H–C, in others with the order C–H. There may also be languages in which the order varies depending on the category of the head, for example, H–C when H is N, but C–H when H is V. Furthermore, adjunctions can be either to the left or to the right, again depending sometimes on the particular language, sometimes on the particular construction within a given language.

I will argue in what follows that this picture of the human language faculty is incorrect and that the human language faculty is in fact rigidly inflexible when it comes to the relation between hierarchical structure and linear order. Heads must always precede their associated complement position. Adjunctions must always be to the left, never to the right. That is true of adjunctions to phrases and it is true of adjunctions to heads.

This inflexibility extends to specifiers, too, which I argue to be an instance of adjunction. Hence, specifier positions must invariably appear to the left of their associated head, never to the right.

The implications of this new picture of the human language faculty are widespread. For languages like English, right adjunction has standardly been assumed in the characterization of various constructions. Every one of these constructions must be rethought in a way compatible with the unavailability of right adjunction. The range is substantial: right dislocation, right node raising, relative clause extraposition, comparative and result clause extraposition, heavy NP shift, coordination, multiple complements and multiple adjuncts (here the work of Richard Larson has been extremely important), possessives like *a friend of John's*, partitives, and also relative clauses, which must now be reanalyzed in the spirit of the raising/promotion analysis that dates back to the early seventies.

For languages like Japanese, complement positions can no longer be taken to be to the left of their head. The fact that complements do precede their associated head must be reinterpreted as indicating that in Japanese complements necessarily appear in specifier/adjoined positions that are hierarchically higher than the position of the head. A direct object in Japanese will asymmetrically c-command its verbal head, the object of a postposition will asymmetrically c-command that postposition, and the IP complement of a complementizer will asymmetrically c-command that complementizer.

It is legitimate and necessary to ask why the human language faculty displays the particular linear ordering that it does. Why do heads always precede complements and why do specifiers and adjoined phrases always precede heads? I provide a partial answer to this question, starting from the assumption that there exists a mapping between hierarchical structure and observed linear order that is rigid.

The formulation of this idea that I adopt, in terms of what I call the Linear Correspondence Axiom, has implications beyond those concerning linear order itself. It implies not only that specifiers are an instance of adjunction, but also that no phrase can have more than one other phrase adjoined to it. Similarly for heads: no head can have more than one other head adjoined to it.

The conclusion that adjunction is drastically limited in this way implies in turn that languages like Japanese cannot be uniformly head-final. Of course, the term *head-final* has standardly had a meaning that must now be dropped, since no head can follow its associated complement position. However, the term can conveniently be retained with a different meaning. We can say that a head X^0 is final if its complement comes to precede it (by moving to some higher specifier/adjoined position). In this new sense of the term, Japanese verbs still have the property of being heads that are final. Put another way, Japanese VPs are still head-final, in this altered sense.

Consider now a typical SOV sentence in Japanese. The subject must occupy some specifier position. That specifier position is the specifier position of some head Y^0. However, the complement of Y^0 cannot have moved into its specifier position, since that is filled by the subject, by assumption. Therefore, Y^0 is not final in its phrase.

Strictly speaking, this conclusion is not necessarily valid for Y^0. Perhaps the complement of Y^0 has moved into the specifier position of a different head Z^0 higher than Y^0. But then the complement of Z^0 remains to the right of Z^0, in which case Z^0 is not final. One could pursue this further, but given the finiteness of syntactic representations, the conclusion will clearly be that in every representation, at least one head must be initial, in the sense that its complement must have remained in situ.

Thus, Japanese cannot be uniformly head-final, although it could be that all its visible heads are head-final. (I am actually led below to question even this.) Since by this reasoning no language can be uniformly head-final, the conclusion must be that mixed headedness is by far more common than standard typological descriptions would lead one to believe. (Note that in this new sense of *head-final* English is head-final in certain constructions, too—for example, those involving preposition stranding.) From this perspective, the fact that many languages (e.g., Dutch, Hungarian) are visibly of mixed headedness is to be expected.

The Linear Correspondence Axiom has additional consequences of a different sort. It explains certain basic properties of phrase structure that standard X-bar theory has not, for example, the fact that every phrase must have at least one and at most one head. It does so by in essence attributing certain properties of linear order to hierarchical structure, in effect taking linear order to be of more fundamental importance to the human language faculty than is generally assumed. One of these properties is *antisymmetry*, whence the title of this monograph.

Chapter 1

Introduction and Proposal

1.1 Introduction

It is standardly assumed that Universal Grammar (UG) allows a given hierarchical representation to be associated with more than one linear order. For example, postpositional phrases and prepositional phrases are generally taken to be hierarchically identical, differing only in linear order. Similarly, English and Japanese phrases consisting of a verb and its complement are thought of as symmetric to one another, also differing only in linear order.

In this monograph I will propose a restrictive theory of word order and phrase structure that denies this standard assumption. I will argue that phrase structure in fact always completely determines linear order and consequently that if two phrases differ in linear order, they must also differ in hierarchical structure.

More specifically, I will propose that asymmetric c-command invariably maps into linear precedence. I will offer a particular formulation of this simple idea that will yield two major consequences. First, there will follow with few further hypotheses a highly specific theory of word order, essentially that complements must always follow their associated head and that specifiers and adjoined elements must always precede the phrase that they are sister to. I will try to show that this then leads to a series of favorable empirical results.

Second, the requirement that hierarchical structure map uniquely to linear order will turn out to yield a derivation of the essentials of X-bar theory. Put another way, I will argue that X-bar theory is not a primitive component of UG. Rather, X-bar theory in essence expresses a set of antisymmetric properties of phrase structure. This antisymmetry of phrase structure will be seen to be inherited, in effect, from the more basic antisymmetry of linear order.

Let us start from the familiar notion of phrase marker, with the usual distinction between terminal symbols and nonterminal symbols. At least in the PF wing of the grammar, the terminal symbols must be linearly ordered. A linear ordering has three defining properties.

(1) a. It is transitive; that is, xLy & yLz → xLz.
 b. It is total; that is, it must cover all the members of the set: for all distinct x, y, either xLy or yLx.
 c. It is antisymmetric, that is, not(xLy & yLx).

The familiar dominance relation on nonterminals is not a linear ordering. Although it is both transitive and antisymmetric, the dominance relation is not total; that is, there can be two nodes in a given phrase marker such that neither dominates the other.

However, the dominance relation has something significant in common with a linear ordering, beyond being transitive and antisymmetric. Consider a given nonterminal X in a phrase marker, and then consider the set of nonterminals that dominate X. For all X, that set is linearly ordered by the dominance relation, that is, for all X, Y dominates X & Z dominates X → either Y dominates Z or Z dominates Y. Although the dominance relation itself is not total, it becomes total when restricted to the set of nodes dominating a given node. Let us say that it is *locally total*, in this sense.[1] Let us further say that, although the dominance relation is not a linear ordering, it is, by virtue of being locally total, a *locally linear* ordering (in the sense that it becomes linear if one restricts oneself to the local environment of a given node).

The familiar relation of c-command is transitive, but unlike the dominance relation it is not even antisymmetric, since two sister nodes can c-command each other. However, we can add antisymmetry to c-command by simply taking the relation of asymmetric c-command:

(2) X asymmetrically c-commands Y iff X c-commands Y and Y does not c-command X.

This relation is now both transitive and antisymmetric. It is not total, since in a given phrase marker there can be two nodes neither of which (asymmetrically) c-commands the other. But if we restrict ourselves henceforth to binary-branching phrase markers,[2] it is locally total, and hence locally linear, in the same sense as the dominance relation. This is so, since in a binary branching tree, if Y asymmetrically c-commands X and Z (distinct from Y) also asymmetrically c-commands X, then it must be the case that either Y asymmetrically c-commands Z or Z asymmetrically c-commands Y.

We now have two locally linear relations on nonterminals, dominance and asymmetric c-command. The intuition that I would like to pursue is that there should be a very close match between the linear ordering relation on the set of terminals and some comparable relation on nonterminals. By *comparable*, I now mean locally linear. Of the two locally linear relations at issue, it is natural to take asymmetric c-command to be the one that is closely matched to the linear ordering of the set of terminals.

This matching will have to be mediated by the familiar dominance relation that holds between nonterminals and terminals. To keep this relation separate from the above-discussed dominance relation between nonterminals, which I will think of as *D*, I will refer to the nonterminal-to-terminal dominance relation as *d*. This relation *d* is a many-to-many mapping from nonterminals to terminals. For a given nonterminal X, let us call d(X) the set of terminals that X dominates. d(X) can be said to be the "image" under *d* of X.

Just as we can speak of the image under *d* of a particular nonterminal, so we can speak of the image under *d* of an ordered pair of nonterminals ⟨X, Y⟩. What we want to say is that the image under *d* of ⟨X, Y⟩ will be based on d(X) and d(Y), specifically by taking the image to be the Cartesian product of d(X) and d(Y). Put somewhat more formally, d⟨X, Y⟩ (= the image under *d* of ⟨X, Y⟩) is the set of ordered pairs {⟨a, b⟩} such that a is a member of d(X) and b is a member of d(Y).

If instead of simply looking at one ordered pair ⟨X, Y⟩ and its image, we look at a set of ordered pairs and their images under *d*, we can introduce the natural notion that the image of a set of ordered pairs is just the set formed by taking the union of the images of each ordered pair in the original set. For example, let *S* be a set of ordered pairs $\{\langle X_i, Y_i\rangle\}$ for $0 < i < n$. Then d(*S*) = the union for all *i*, $0 < i < n$ of $d\langle X_i, Y_i\rangle$.

1.2 Proposal

To express the intuition that asymmetric c-command is closely matched to the linear order of terminals, let us, for a given phrase marker, consider the set *A* of ordered pairs $\langle X_j, Y_j\rangle$ such that for each j, X_j asymmetrically

c-commands Y_j. Let us further take A to be the maximal such set; that is, A contains all pairs of nonterminals such that the first asymmetrically c-commands the second. Then the central proposal I would like to make is the following (for a given phrase marker P, with T the set of terminals and A as just given):

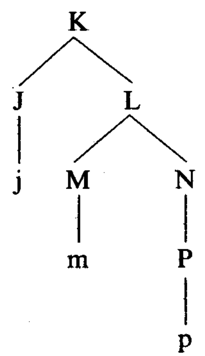
(3) *Linear Correspondence Axiom*
d(A) is a linear ordering of T.

Chapter 2

Deriving X-Bar Theory

To see how the Linear Correspondence Axiom (LCA) works in practice, let us begin with the simple phrase marker in (1).

(1)

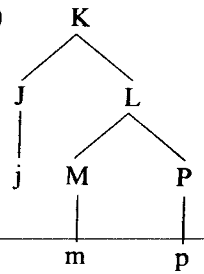


In this phrase marker the pairs that constitute the set A (i.e., the pairs of nonterminal nodes such that the first asymmetrically c-commands the second) are the following: ⟨J, M⟩, ⟨J, N⟩, ⟨J, P⟩, ⟨M, P⟩. Since in this simple case J, M, N and P all dominate just one terminal element, d(A) is easy to exhibit fully: namely, ⟨j, m⟩, ⟨j, p⟩, ⟨m, p⟩.[1] These three ordered pairs do constitute a linear ordering of the set {j, m, p}, given that (1) transitivity holds, (2) antisymmetry is respected, and (3) the ordering is total, in that for every pair of terminals an ordering is specified.

It should be noted that I am crucially taking c-command to be properly defined in terms of "first node up" and not in terms of "first branching node up." Under the latter type of definition the node P in (1) would c-command M, so that M would no longer asymmetrically c-command P, in which case no ordering between the terminals m and p would be specified at all, incorrectly.

The importance of this point can be seen further by considering the phrase marker (2), which is similar to (1) in all respects except that it lacks the node N.

(2)



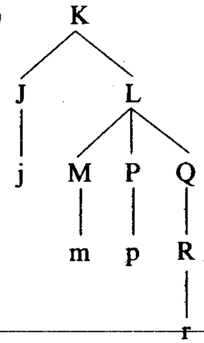
In (2) the set *A* of pairs such that the first nonterminal asymmetrically c-commands the second is as follows: ⟨J, M⟩, ⟨J, P⟩. Consequently, d(*A*) for (2) is composed of the pairs ⟨j, m⟩ and ⟨j, p⟩. Although the set d(*A*) consisting of these two pairs of terminals respects both transitivity (vacuously) and antisymmetry, it does not constitute a linear ordering of the set {j, m, p}, since it specifies no order at all between the two terminals m and p; that is, it fails to be total in the sense of (1b) of chapter 1.

In other words, (2) fails to meet the requirement imposed by the LCA and is therefore not an admissible phrase marker. This has at least two desirable consequences. First, consider whether the complement of a head can itself be a head. The usual assumption, within the context of X-bar theory, is that it cannot be. One could take that to be a basic fact of X-bar theory, but X-bar theory itself clearly provides no account of why it should hold. The LCA given in (3) of chapter 1 does, since having a head whose complement was itself a head would yield precisely the configuration of M, P (and L) in (2), which is inadmissible.

The second desirable consequence related to (2) lies in the even more basic question of why a phrase cannot have more than one head. X-bar theory treats this as a basic fact about phrase structure but does not attempt to provide an explanation for it. The LCA does, since a phrase with two heads would again look like [$_L$ M P] in (2) and would again be excluded. Put another way, the LCA derives both the fact that a head cannot take a complement that is itself a head and the basic X-bar fact that a phrase cannot have two heads.[2]

The exclusion of (2) would not be affected if we added a nonhead sister node to M and P, as in (3).

(3)



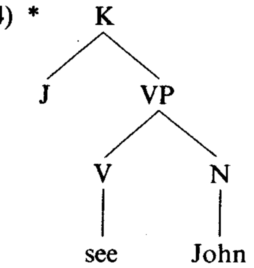
In (3) again neither M nor P asymmetrically c-commands the other (nor are m and p dominated by any other node that is in an asymmetric c-command relation). *A* for (3) is ⟨J, M⟩, ⟨J, P⟩, ⟨J, Q⟩, ⟨J, R⟩, ⟨M, R⟩, ⟨P, R⟩. d(*A*) is then ⟨j, m⟩, ⟨j, p⟩, ⟨j, r⟩, ⟨m, r⟩, ⟨p, r⟩. But again it lacks any pair involving m and p and so does not meet the totality requirement.

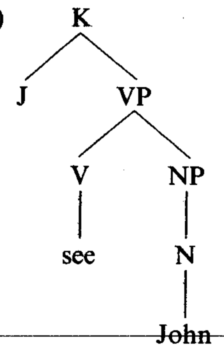
From this perspective, (2) and (3) are excluded essentially because the terminals m and p (and the nonterminals M and P that exhaustively dominate them) are in too symmetric a relation to one another. For that reason, they are not "seen" by the relation of asymmetric c-command and so fail to be incorporated into the required linear ordering. Another informal way to put this, reversing the vantage point, is, to say that the LCA, by virtue of requiring *d* (the dominance relation between nonterminals and terminals) to map *A* into a linear ordering, has forced the set of nonterminals to inherit the antisymmetry of the linear ordering of the terminals.

If we think of L in (2) as a VP = *see John*, with M = *see*, then the preceding discussion tells us that the complement *John* cannot be dominated (apart from VP and higher nodes) solely by N(oun), as in (4), but must also be dominated by (at least) another node NP, as in (5), in order for the phrase marker to be well formed.

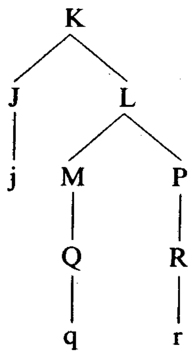
(4) is not an admissible phrase marker, but (5) is (setting aside questions such as the choice between DP and NP). In (4) no linear ordering would be assigned to *see* and *John*. In (5), on the other hand, *see* correctly is ordered with respect to (before) *John*, since V in (5) asymmetrically c-commands N.[3]

Comparing (2) with (1), we see that replacing one of the two symmetric nodes by a more complex substructure breaks the symmetry and renders

(4) *


(5)


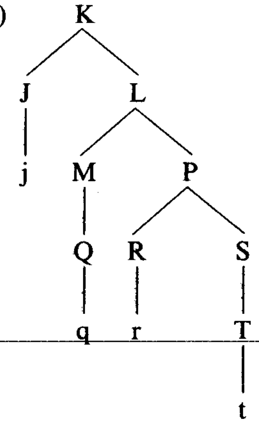
the phrase marker admissible. Now consider the result of adding structure under both M and P in (2), as in (6).

(6)


The asymmetric c-command set *A* for (6) is ⟨J, M⟩, ⟨J, Q⟩, ⟨J, P⟩, ⟨J, R⟩, ⟨M, R⟩, ⟨P, Q⟩. The corresponding d(*A*) is ⟨j, q⟩, ⟨j, r⟩, ⟨q, r⟩, ⟨r, q⟩. (⟨q, r⟩ is in d(*A*) since ⟨q, r⟩ = d⟨M, R⟩; ⟨r, q⟩ is in d(*A*) since ⟨r, q⟩ = d⟨P, Q⟩.) This d(*A*) for (6) is total, but it is not antisymmetric. Therefore, (6) is not an admissible phrase marker.

The problem with (6) is not exactly that M and P are symmetric in that each dominates one other nonterminal. This can be seen by adding more substructure to, for example, P, as in (7).

Concentrating just on the sub–phrase marker whose root node is L, we find that *A* there is ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩, ⟨R, T⟩, ⟨P, Q⟩. But we now see that the addition of S and T, although resulting in a larger *A*, has not changed the heart of the problem in (6), which was the cooccurrence in *A* of ⟨M, R⟩ and ⟨P, Q⟩, which led to both ⟨q, r⟩ and ⟨r, q⟩ being in d(*A*), violating antisymmetry. Exactly the same problem arises in (7).[4]

(7)


Let us call a nonterminal that dominates no other nonterminal a *head.* A nonterminal that does dominate at least one other nonterminal will be a *nonhead.* Then we can sum up the results of this chapter ((1)–(7)) in the following terms: if two nonterminals are sisters and if one of them is a head and the other a nonhead, the phrase marker is admissible ((1) and (5)). If both are heads, the phrase marker is not admissible ((2), (3), and (4)). If both are nonheads ((6) and (7)), the phrase marker is again not admissible (whether or not the number of nonterminals dominated by each of those two nonheads is the same).

The prohibition against nonhead sisters has one clearly desirable and important consequence, and another consequence that will require a (familiar) refinement of the notion "nonterminal." Let me begin with the first. A basic tenet, perhaps the basic tenet, of X-bar theory is that all phrases must be headed. Thus, X-bar theory disallows a phrasal node immediately dominating two maximal projections and nothing else. X-bar theory does not, however, explain why every phrase must have a head. The LCA does. The reason that a phrasal node cannot dominate two maximal projections (and nothing else) is that if it did, there would be a failure of antisymmetry, exactly as discussed above for (6) and (7).

This explanation for the pervasiveness of heads in syntactic structure has a particularly striking subcase in the realm of coordination. Why is it not possible to have sentences such as these with a coordinate interpretation?

(8) a. *I saw the boy the girl.
b. *The girl the boy were discussing linguistics.

Again, the answer is straightforward. A phrase such as '[[the boy] [the girl]]' is not adequately antisymmetric and leads to a violation exactly as described for (6) and (7). The required presence of a word like *and* is now understandable: coordinating conjunctions are heads that serve to bring coordinate structures in line with the antisymmetry requirement imposed by the LCA. Consequently, the constituent structure of *the girl and the boy* must be '[the girl [and [the boy]]]'.[5]

PART II

Chapter 3
Adjunction

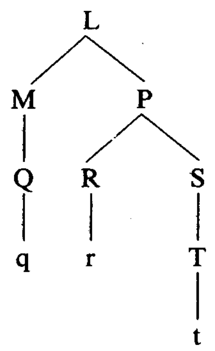
3.1 Segments and Categories

The preceding discussion appears to rule out sentences such as (1), in which the subject clearly must have a sister constituent that is not a head.

(1) The girl saw John.

Put more generally, specifiers and adjoined phrases appear to have no place in the theory being elaborated here. To allow for specifiers or adjoined phrases, I need to add a refinement to the theory of phrase structure presented so far. I will adopt the notion of segment, that is, the distinction between segment and category that was introduced by May (1985) and adopted by Chomsky (1986a). Let us return to the substructure of (7) from chapter 2, repeated here, that was earlier argued to be inadmissible.

(2)



The problem arose with respect to both r and t in their relation to q. Let us look just at r. Since M asymmetrically c-commands R (i.e., A contains ⟨M, R⟩), it follows that d(A) contains ⟨q, r⟩. But d(A) also contains

⟨r, q⟩ by virtue of P asymmetrically c-commanding Q. As a result, d(*A*) violates antisymmetry; that is, it fails to be a linear ordering of the terminals.

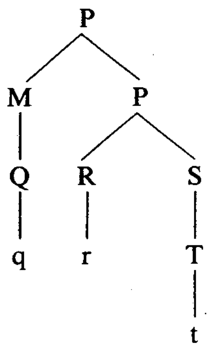
This result is correct for the case in which M and P are both maximal projections dominated by another node L, and we want to maintain it. At the same time we want to allow for the case in which M is adjoined to P. The segment/category distinction leads to the statement that under adjunction L and P are two segments of one category. The question is how that makes (2) compatible with the LCA.

The solution I would like to propose is to restrict c-command to categories—that is, to say that a segment cannot enter into a c-command relation.

(3) X c-commands Y iff *X and Y are categories* and X excludes[1] Y and every category that dominates X dominates Y.

In this light consider (4), the counterpart of (2) in which L is replaced by P, to indicate clearly the adjunction structure.

(4)



In (2) the problem was that d(*A*) contained both ⟨q, r⟩ and ⟨r, q⟩, the former as the image of ⟨M, R⟩, the latter as the image of ⟨P, Q⟩. The adjunction structure of (4), combined with the italicized part of (3), has the effect of eliminating ⟨P, Q⟩ from the *A* of (4), since the lower P is a segment and not a category.[2] Therefore, ⟨r, q⟩ does not belong to the d(*A*) of (4), and the potential violation of antisymmetry is eliminated.

The *A* associated with (4) is thus ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩, ⟨R, T⟩.[3] The corresponding d(*A*) is ⟨q, r⟩, ⟨q, t⟩, ⟨r, t⟩, which constitutes a linear ordering of the set of terminals, as desired. We can now think of the segment/category distinction as being forced upon UG by the need to permit specifiers and adjoined phrases.[4]

Unless there turns out to be another natural way to permit specifiers within the theory developed here, the conclusion must be that a specifier is necessarily to be taken as an adjoined phrase, involving crucial use of the segment/category distinction.[5]

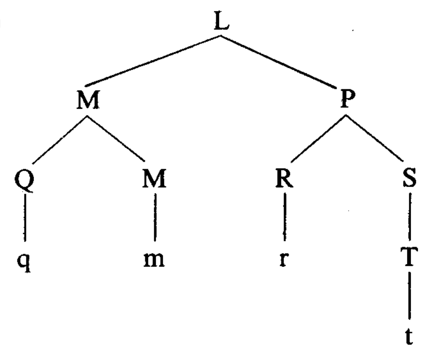
Returning to (4), note that what makes it compatible with the LCA is that (the lower) P does not c-command Q, as a result of the phrase added to the definition in (3). Strictly speaking, though, this property of (4) depends only on X in (3) being restricted to categories, the status of Y is not directly relevant. The idea that Y, too, must be a category (and not a segment) does have potential significance. If a segment cannot be c-commanded, and if antecedent government strictly has c-command as a necessary component, then a segment cannot be antecedent-governed and thus cannot be moved. In other words, a phrase that has something adjoined to it cannot be moved out by itself.

This derives the fact that a head to which a clitic (or other element) has adjoined cannot move up in such a way as to strand the clitic.[6] With respect to adjunction to a nonhead, recall that I have been led to analyze specifiers as involving adjunction. We consequently derive the prediction that the sister node of a specifier cannot be moved. This corresponds to a fairly standard assumption.[7]

3.2 Adjunction to a Head

Adjunction to a head, as in the case of a clitic, is illustrated in the phrase marker (5).

(5)



Here Q is adjoined to M. Since M does not dominate Q (only one of its segments does), the fact that M does not dominate R and S is irrelevant to Q's relation to R and S. In other words, the definition (3) of c-command has the effect that Q in (5) (asymmetrically) c-commands both R and S. The *A* for (5) is therefore ⟨Q, R⟩, ⟨Q, S⟩, ⟨Q, T⟩, ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩, ⟨R, T⟩. (Note that although P c-commands both M and Q, it does not asymmetrically c-command either one, since M and Q both c-command P.) This yields d(*A*) = ⟨q, r⟩, ⟨q, t⟩, ⟨m, r⟩, ⟨m, t⟩, ⟨r, t⟩, which respects transitivity and antisymmetry, but appears to fail the requirement of totality, since no order is yet specified for q relative to m.

Again consider definition (3), repeated here.

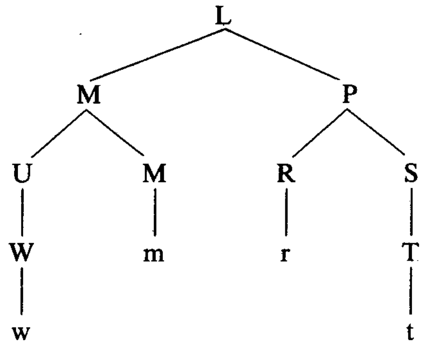
(6) X c-commands Y iff *X and Y are categories* and X excludes Y and every category that dominates X dominates Y.

By (6), M in (5) does not c-command Q because it does not exclude it. Q cannot c-command a segment of M alone, by assumption. However, Q in (5) does c-command the category M. This is so because Q excludes M and every category that dominates Q dominates M.[8]

The fact that Q in (5) c-commands, and hence asymmetrically c-commands, M means that the pair ⟨Q, M⟩ must be added to *A*: ⟨Q, R⟩, ⟨Q, S⟩, ⟨Q, T⟩, ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩, ⟨R, T⟩, ⟨Q, M⟩. This results in the addition of ⟨q, m⟩ to d(*A*), yielding d(*A*) = ⟨q, r⟩, ⟨q, t⟩, ⟨m, r⟩, ⟨m, t⟩, ⟨r, t⟩, ⟨q, m⟩, which is a linear ordering of the set of terminals, as desired.[9]

The fact that Q also asymmetrically c-commands R and S as discussed three paragraphs back is an instance of a more general property of adjoined phrases, namely, that they always c-command "out of" the phrase they are adjoined to. Let us replace Q in (5) by a nonhead, as in (7).

(7)



Here, U, a nonhead, has been adjoined to the head M. As before, M does not dominate U, so that U c-commands P and everything dominated by

P. P itself c-commands U, so that U and P enter into no asymmetric c-command relation (as was true for Q and P in (5)).

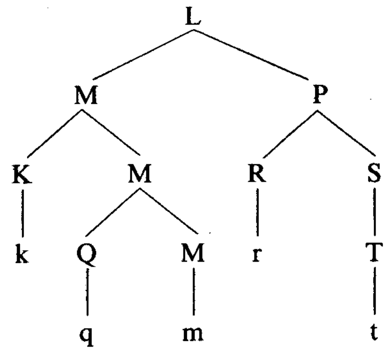
Now consider W. W does not c-command P, because of the intervening presence of U. Therefore, P asymmetrically c-commands W. Hence, ⟨P, W⟩ is in *A* in (7) and ⟨r, w⟩ and ⟨t, w⟩ are in d(*A*). But U asymmetrically c-commands R, S, and T, so that ⟨U, R⟩, ⟨U, S⟩, and ⟨U, T⟩ are also in *A*, and correspondingly ⟨w, r⟩ and ⟨w, t⟩ are also in d(*A*). Thus, d(*A*) for (7) consists at least of ⟨r, w⟩, ⟨t, w⟩, ⟨w, r⟩, and ⟨w, t⟩, which violates antisymmetry, so that d(*A*) is not a linear ordering and (7) is excluded as a violation of the LCA.

Put another way, we have just derived without stipulation the fact that a nonhead cannot be adjoined to a head, in all probability a correct result.[10]

3.3 Multiple Adjunction: Clitics

The phrase marker (5) represents the case of a clitic (Q) adjoined to a head (M). Now consider what happens if a second clitic (K) is adjoined to the same head (M), as in (8).

(8)



As before, Q c-commands P and everything dominated by P, and K does the same. The problem that arises for (8) instead concerns the relation between K and Q, neither of which is dominated by M. Consequently, K and Q c-command each other; that is, neither asymmetrically c-commands the other. Therefore, no linear order is specified for k and q (neither ⟨k, q⟩ nor ⟨q, k⟩ is contained in the d(*A*) of (8)), so that (8) is excluded by the LCA.

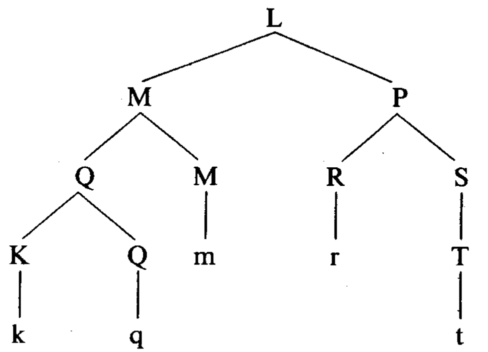
What we see here is that two (or more) clitics adjoined to the same head find themselves in too symmetric a relation: both clitics are dominated by segments of the same head and neither is dominated by that head as category. The required antisymmetry does not hold. The conclusion is inescapable: it is not possible to adjoin two (or more) clitics to the same head.

Run-of-the-mill French sentences like (9) appear to pose a problem.

(9) Jean vous le donnera.
Jean you_{DAT} it will-give
'Jean will give it to you.'

However, two other structures are available for multiple clitics, both of which are compatible with the LCA. The one most likely to be appropriate for (9) is (10).

(10)



Here the clitic Q has adjoined to the head M, and the clitic K has in turn adjoined to the clitic Q. Q c-commands P and everything P dominates, as before. Since K is not dominated by Q (or by M), K c-commands P and everything P dominates, too. In other words, both K and Q asymmetrically c-command R, S, and T (but not P, since P c-commands both K and Q). The *A* for (10) is therefore ⟨K, Q⟩,[11] ⟨K, M⟩, ⟨Q, M⟩, ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩, ⟨K, R⟩, ⟨K, S⟩, ⟨K, T⟩, ⟨Q, R⟩, ⟨Q, S⟩, ⟨Q, T⟩, ⟨R, T⟩. The corresponding d(*A*) is ⟨k, q⟩, ⟨k, m⟩, ⟨q, m⟩, ⟨m, r⟩, ⟨m, t⟩, ⟨k, r⟩, ⟨k, t⟩, ⟨q, r⟩, ⟨q, t⟩, ⟨r, t⟩, which is a linear ordering of the set of terminals, as desired.

The two clitics of (9) could thus be taken to form a constituent '[vous le]',[12] with *vous* adjoined to *le*. From this perspective, clitic ordering and cooccurrence restrictions could be looked at as follows: the impossibility of **Jean le vous donnera* could be due to the impossibility of adjoining another clitic to *vous* (and similarly *nous, me, te, se*)[13] in French.[14] I have no proposal to make concerning the impossibility of expressing the French counterpart to *They will introduce me to him* with two clitics.

(11) a. *Ils me lui présenteront.
they me him_{DAT} will-introduce
b. *Ils lui me présenteront.

Although (11b) may well be excludable on the grounds suggested above (that *me* cannot be adjoined to), (11a) is unexpected. Of interest, nonetheless, is that for many speakers, the clitic sequence *me lui* is better in (12) than in (11a).[15]

(12) ?Elle me lui semble infidèle.
she me_{DAT} him_{DAT} seems unfaithful
'She seems to me unfaithful to him.'

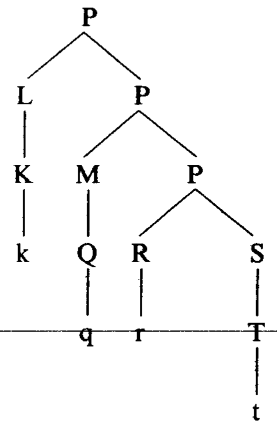
This kind of contrast might be analyzed by attributing to *me lui* in (12) a different constituent structure from that holding in (11a), which probably involves (unsuccessful) adjunction of *me* to *lui*. Although no more than one clitic can be adjoined to a given head, the possibility still remains open that two adjacent clitics are to be analyzed as being adjoined to two distinct (nonclitic) heads. In other words, it might be that *me* in (12) is adjoined to one functional head, and *lui* to the next functional head below that. Whether or not this is correct for (12), it is almost certain to be correct for some instances of adjacent clitics. This amounts to saying that the phenomenon of *split clitics* discussed in Kayne 1991, pp. 660ff., is found not only when the clitics in question are separated by overt material but also, as one would expect, when they are not.[16]

In summary, from the perspective of the LCA, sequences of clitics must not be analyzed as successive adjunctions to the same head but instead should be analyzed as involving either adjunctions to distinct functional heads (e.g., one clitic to Tense, one to Agr) or adjunctions of one clitic to another, or some combination thereof.

3.4 Multiple Adjunctions: Nonheads

The antisymmetry requirement induced by the LCA has the same consequence for adjunctions of nonheads to nonheads as it does for adjunctions of heads to heads, as discussed for clitics in the previous section. The relevant phrase marker has the form shown in (13).

(13)



In (13) the nonhead M has been adjoined to the nonhead P, and the nonhead L has been further adjoined to the two-segment category P that was the output of the first adjunction.

In a way partially parallel to the discussion of (8), there is a problem with (13) concerning the relation between k and q. The problem here is specifically that L asymmetrically c-commands Q and at the same time M asymmetrically c-commands K. Thus, ⟨L, Q⟩ and ⟨M, K⟩ are both in the *A* of (13), so that ⟨k, q⟩ and ⟨q, k⟩ are both in d(*A*), with a consequent violation of antisymmetry.

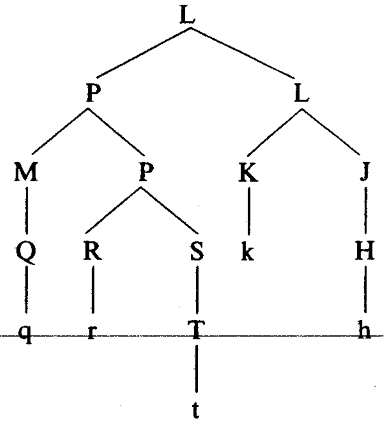
I conclude that the adjunction of more than one nonhead to a given nonhead is impossible. Since in this theory specifiers are a case of adjunction, we derive the fact (stated by X-bar theory) that a given phrase can have only one specifier.

This limitation on specifiers is not controversial (so that its derivation is clearly desirable), but the more general limitation on adjoined phrases is potentially controversial, since it is usually assumed that more than one phrase can be adjoined to a given projection (nonhead) and also that a phrase can be adjoined to a phrase that already has a specifier.

3.5 Specifiers

Let me begin indirectly by pointing out that the present theory does allow a certain kind of multiple adjunction, parallel to that seen above in the case of clitics. More specifically, it is permissible to adjoin Y to X and Z to Y. The relevant phrase marker looks like (14).

(14)



Here, the nonhead P has been adjoined to the nonhead L, and the nonhead M has been adjoined to P.

In (14) M is dominated neither by P nor by L. Consequently, M asymmetrically c-commands K, J, and H (see note 3), so that the asymmetric c-command set *A* contains ⟨M, K⟩, ⟨M, J⟩, and ⟨M, H⟩. Since M dominates q, d(*A*) contains ⟨q, k⟩ and ⟨q, h⟩. Similarly for ⟨M, R⟩, ⟨M, S⟩, ⟨M, T⟩ and ⟨q, r⟩, ⟨q, t⟩; also ⟨R, T⟩, ⟨K, H⟩ and ⟨r, t⟩, ⟨k, h⟩. The remaining pairs ⟨r, k⟩, ⟨r, h⟩, ⟨t, k⟩, ⟨t, h⟩ come into d(*A*) by virtue of P asymmetrically c-commanding K and J. (K and J do not c-command P, since L, which dominates K and J, does not dominate P.) (14) is thus compatible with the LCA.

This type of phrase marker takes on particular interest when we recall that in the theory being developed here specifiers are an instance of adjunction. Therefore, M in (14) could just as well be a specifier of P and P a specifier of L—in which case the specifier of the specifier of L would asymmetrically c-command K and J and everything dominated by K and J. Taking L = IP, K = I, and J = VP, we reach the conclusion that the specifier of the subject of IP asymmetrically c-commands I and VP and everything within VP.

This conclusion has some favorable consequences. First, it brings back into the fold the recalcitrant cases of pronoun binding by a quantifier phrase that are discussed by Reinhart (1983, p. 177). For example:

(15) Every girl's father thinks she's a genius.

From the present perspective, the fact that *every girl* is in the specifier of the subject DP does not interfere with its binding the pronoun. Since specifiers are adjoined phrases, the definition of c-command adopted above, and repeated here, has the effect that *every girl* in (15) does in fact c-command *she*.

(16) X c-commands Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y.

Second, we now have an account of contrasts like the following:

(17) Nobody's articles ever get published fast enough.

(18) *Articles by nobody ever get published fast enough.

(19) *The articles that nobody writes ever get published fast enough.

(20) Nobody ever gets their articles published fast enough.

The polarity item *ever* needs to be c-commanded by an appropriate licenser, in these examples *nobody*. In agreement with Larson (1988, p. 337), I take c-command of the polarity item by the negative licenser to be a necessary condition for well-formedness. (18) and (19) are excluded because c-command between *nobody* and *ever* does not hold. In the present framework, *nobody* in (17) does c-command *ever*, as desired.

Third, Webelhuth (1992, chap. 4) has argued that pied-piping in (nonecho) interrogatives is licensed only by specifiers (as opposed to modifiers and complements). This is particularly clear in embedded interrogatives, where interference from echo constructions is not at issue. For example:

(21) We know whose articles those are.

(22) *We know articles by who(m) those are.

(23) *I wonder people from what city the game is likely to attract.

This contrast can be partially unified with those discussed in the preceding paragraphs on negative polarity items and pronoun binding as follows. Assume that these (nonecho) interrogatives have a [+wh] head (C^0) that the *wh*-phrase (*who, what city*) must be paired with.[17] Then (21) can be distinguished from (22)–(23) if (24) holds as a necessary condition.

(24) The *wh*-phrase in interrogatives must asymmetrically c-command the [+wh] head.

Taking *who* in (21) to be the specifier of (the highest head of) *whose articles*, it follows from (16) (plus the specifiers-as-adjoined-phrases hypothesis) that *who* in (21) meets (24), whereas the *who(m)* and *what* of (22) and (23) clearly do not.[18]

In (my) colloquial English, the pied-piping of a prepositional phrase in interrogatives (and relatives) is not possible.[19]

(25) a. *We want to know about what you're thinking.
b. *Tell me at who(m) you were looking.
c. *Nobody knows to what school he goes.
d. *I can't figure out for which kid you bought it.
e. *Tell us from where you got that.

This follows directly from (24) since the complement of a preposition internal to Spec,CP does not c-command C^0.

The many languages that do allow the equivalent of (25) must permit the object of the preposition to move to Spec,PP at LF (cf. Chomsky 1993, p. 35).

(26) ... $[_{CP}[_{PP}$ what$_i$ $[_{PP}$ about $[e_i]]]$ $[_{IP}$ ...

The fact that colloquial English does not itself allow (26), and therefore does not have (25), is presumably to be related to the fact that it does allow preposition stranding. For example:

(27) We want to know what you're thinking about.

One aspect of this relation can be expressed straightforwardly by claiming that the derivation of (27) involves movement of *what* through the specifier of the PP headed by *about*, essentially as proposed by Van Riemsdijk (1978, chap. 6).[20] Thus, both (27) and (25) will involve movement of the *wh*-phrase to Spec,PP, overtly in (27) (with subsequent movement to Spec,CP)[21] and covertly in (25), in those languages that permit (25).

Related to this discussion is the partial acceptability of (the noncolloquial) (28) and (29).

(28) ?The father of every eight-year-old girl$_i$ thinks she$_i$'s a genius.

(29) ?The author of no linguistics article ever wants it to go unread.

These suggest that the object of a postnominal *of* can to some extent be moved to Spec,DP in LF.

If I am correct in arguing that a specifier c-commands out of the phrase that it is the specifier of, then an obvious question arises concerning reflexives. Consider (30).

(30) *Every girl's father admires herself.

One possibility would be to say that in this sentence *every girl* is prevented from being the antecedent of the reflexive since there is a closer potential antecedent, *every girl's father*.[22] However, this would leave unexplained the fact that the sort of pronoun binding found in (15) is not sensitive to a parallel "closest antecedent" requirement, as can be seen in (31).

(31) Every girl's father thinks he knows what's best for her.

Here, *her* can be bound by *every girl* even though *every girl's father* is a potential pronoun binder, as shown in particular by the fact that it can actually bind *he*.

I will instead pursue an approach to (30) based on Szabolcsi's (1981; 1983; 1992) analysis of Hungarian possessives, as transposed to English in Kayne 1993, sect. 2.2. In a number of clear cases the possessor in Hungarian is preceded by an independent D^0, much as in the Italian example (32), with the difference that in Hungarian the prenominal possessor is not limited to being a pronoun.

(32) il mio libro
the my book

The crucial step in the transposition to English is to take the English prenominal possessor to likewise be preceded by D^0, which in English must be empty.

(33) [$_{DP}$ . . . D^0 [John ['s book]]]

There are now two relevant specifier positions, that associated with the head *'s*, in which *John* is found, and that associated with D^0, which is indicated in (33) by the three dots.

Szabolcsi argues that Spec,D^0 is an operator position. Let us assume that, insofar as the notion "antecedent" is concerned, operator positions, although essential to operator binding of a pronoun qua variable, are invisible to Conditions A, B, and C of the binding theory. Let us further assume that in English the possessor phrase, when an operator phrase, moves up in LF to Spec,D^0. Then the operator binding of the pronoun in (31) and (15) will be legitimate, since Spec,D^0 c-commands out of DP. (Similarly, in (17) *nobody* will c-command the polarity item at LF, and in (21) *who* will c-command C^0 at LF.)

On the other hand, (30) will be excluded, as desired, as follows: From its visible position in Spec,*'s*, the phrase *every girl* does not c-command *herself* (because DP dominates the former without dominating the latter). When *every girl* moves at LF to Spec,D^0, it does come to c-command *herself* (because in its LF position *every girl* is not dominated by DP). But

this c-command relation holds between an operator position and a reflexive, and therefore does not suffice to license the latter, by assumption.

The primary distinction here is thus whether a phrase reaches the highest specifier within DP. If it does, then it can c-command out of DP, by virtue of the definition (16) of c-command and the fact that specifier positions are instances of adjunction. If it does not, then it cannot. The clear contrast between the following two sentences is accounted for straightforwardly:

(34) *John considers John's father highly intelligent.

(35) ?John's father considers John highly intelligent.

(34) is a standard Condition C violation. (35) is substantially better because the first instance of *John*, in Spec,*'s* (not the highest specifier within DP), does not c-command the second.[23]

(36) does not violate Condition C for a similar reason: the pronoun does not c-command *John*.

(36) His father considers John highly intelligent.

Nor does *John* c-command *him* in (37), so Condition B is not violated.

(37) John's father considers him highly intelligent.

On the other hand, Loubana Mouchaweh has brought to my attention the fact that the counterpart of (36) is not possible in (her Damascus) Arabic.[24] From the standard c-command perspective, this apparent Condition C violation in Arabic is unexpected. From the present perspective, it might be accounted for if pronouns in (the relevant varieties of) Arabic did have to move to the highest specifier within DP.

3.6 Verb-Second Effects

In the previous section I discussed some issues related to the claim that specifiers c-command out of their containing category. This property is due to the union of two factors. The first is that the LCA forces specifiers to be analyzed as instances of adjunction (otherwise, a specifier and its sister phrase would be too "symmetric"). The second is the definition (16) of c-command in terms of category dominance (rather than segment dominance).

As noted at the end of section 3.4, these two factors have the desirable effect of limiting the number of specifiers of a given category to one. In addition, they have the more general effect of limiting a given phrase to having at most one adjoined phrase (including a specifier). If a phrase XP had a specifier YP, and if another phrase ZP were further adjoined to XP, then YP and ZP would c-command each other, leading to a violation of the LCA, as discussed earlier for (13).

This severe limitation on adjunction that the LCA derives surely appears to be too restrictive. I will nonetheless take it to be correct. That it is correct is suggested by the following considerations. First, it may provide a deep account of (at least one aspect of) the well-known obligatory verb-second effect found in the Germanic languages other than English. A German example would be (38).

(38) *Gestern Peter tanzte.
yesterday Peter danced

Taking *Peter* to be the specifier of IP, adjunction of *gestern* to IP is immediately prohibited.[25]

For the English sentence parallel to (38), which is grammatical, I am led to propose that a covert functional head above a root IP is available and that *yesterday* can adjoin to its projection.

(39) Yesterday Peter danced.

English actually does display the restriction on adjunction seen in (38).

(40) *Never Peter has danced so well.

If the auxiliary raises to a position above the subject, then the negative phrase can be initial.

(41) Never has Peter danced so well.

I take *never* here to be adjoined to the projection of the functional head that contains *has*. This strategy is not available with *yesterday*.

(42) *Yesterday did Peter dance.

This paradigm suggests that (39) is the covert equivalent of (42) and that the difference between *yesterday* and *never* is that only negative phrases (and phrases with *only*) require that the functional head above IP be overtly filled.[26]

That (39) and (41) are indeed parallel is suggested further by their similar behavior in embedded contexts.

(43) I didn't know *(that) yesterday Peter danced.

(44) I didn't know *(that) never had Peter danced so well.

Without the extra functional head above IP, (43) is prohibited as an instance of double adjunction to the same projection. With that extra head, (43), exactly like (44), is embeddable as a complement introduced by *that*, but not as a complement with no overt complementizer.[27]

Another aspect of obligatory verb-second is seen in (45).

(45) *Peter immer tanzt.
Peter always dances

In the English equivalent the adverb *always* can presumably be adjoined to a projection below that of the highest I. If so, then the proposal I made for (38) does not carry over to (45), which could instead be thought of as parallel to the French (46), for which the standard account is rather to require the verb to raise to the highest I node (see Emonds 1978; Pollock 1989).

(46) *Pierre toujours danse.
Pierre always dances

The idea that (38) and (45) are not entirely the same phenomenon is reinforced by the fact that Icelandic allows the equivalent of (45) with some adverbs (see Thráinsson 1985), but apparently not the equivalent of (38).

My proposal for (39) is best understood in the context of the following more fundamental question: why are there so many functional heads? This question is not particularly bothersome in the case of the interpretively contentful functional heads such as Tense. Tense is present in the overt syntax and at LF, and there are well-known advantages to taking it to be an independent head (see Chomsky 1957). Less obvious is the status of agreement. It is evident from languages like Chinese that overt agreement can be completely dispensed with. At the same time there is evidence for abstract Agr heads (even in Chinese; see Chiu 1991). Chomsky (1993) makes significant use of his generalized Agr_O hypothesis (and, of course, of Agr_S). Belletti (1990) and Cardinaletti and Roberts (1990) have given interesting arguments for a second Agr_S, above and beyond the familiar one. I have found evidence, reported in Kayne 1993, for an abstract Agr_S in participial clauses in Romance. Finally, Sportiche (1992) has argued that an extremely wide variety of phrases must be licensed via a spec-head relation with an appropriate head.

The theory developed here, based on the LCA and the characterization of c-command in terms of categories, provides at least a partial answer to this question. Assuming that phrases of various kinds must move out of their base position at some point in the derivation, the answer is that functional heads make landing sites available. Spec-head configurations are used for licensing for a principled and simple reason: there is no other possibility. Given that double adjunction to the same projection is prohibited,[28] there must, for every moved phrase, be a distinct head to whose projection it can adjoin as specifier.[29]

From this perspective, spec-head licensing can be broken into two parts. One is simply that every maximal projection[30] that is not the complement of some head must be the specifier of some head (since no other phrasal adjunction sites are available) and in that sense can be said to be licensed by the head it is specifier of.

The second aspect of spec-head licensing involves the question of matching/agreement—that is, the question, which I will not systematically pursue in this monograph, of what phrase can be the specifier of what head.[31] I would, however, like to call attention to one aspect of this question. Certain heads are intrinsically contentful, such as lexical heads and functional heads like Tense and Aspect. In some cases a moved phrase will become the specifier of a contentful head. In cases where movement is called for, but where no contentful head is available, the moved phrase must become the specifier of a head lacking intrinsic content. It may be that this is what is meant by Agr^0—namely, that Agr^0 is properly thought of as a label for head positions imposed upon phrase markers by the paucity of available adjunction sites, with this paucity following from the present theory.[32]

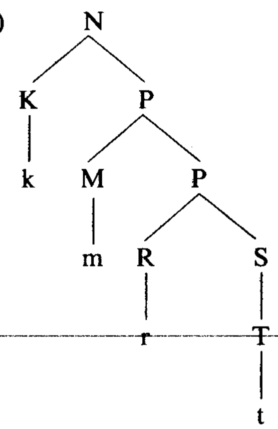
3.7 Adjunction of a Head to a Nonhead

Can a head be adjoined to a nonhead? Chomsky (1986a, p. 73) shows that such adjunction followed by further movement back to a head position leads to an undesirable result. I will now show that the desired prohibition follows directly from the theory developed here. In its essentials, the phrase marker that corresponds to Chomsky's case is (47).

P here is the nonhead to which the head M has adjoined. K is the next higher head to which M is to move. (Note that the argument that follows holds independently of where M originates.)

To see that (47) violates the LCA, consider k and m. K c-commands M, but M also c-commands K (since M is not dominated by P). Furthermore, although K c-commands P, P also c-commands K. Therefore, A for (47)

(47)

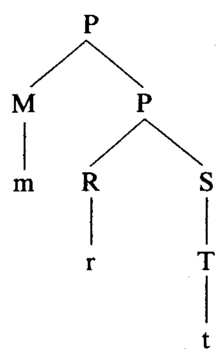


contains no pair that could lead, via the mapping d, to either $\langle k, m \rangle$ or $\langle m, k \rangle$ being in $d(A)$. Thus, $d(A)$ will fail the requirement of totality and hence fail to constitute the necessary linear ordering of the terminals of (47), as desired.[33]

Observe that although k and m yield a violation in (47) (essentially because they are too symmetric to one another), there is no parallel violation based on m and r. M c-commands R in (47), but R does not c-command M (since P dominates R without dominating M). Hence, M asymmetrically c-commands R, leading to $\langle m, r \rangle$ being in $d(A)$, so that m and r pose no totality requirement problem.

This means that (47) without K (and N), as shown in (48), is well formed.

(48)



Assume, however, that the highest element of a chain of heads must have a specifier, in the sense of having a phrase that asymmetrically c-commands it within its maximal projection (or within the maximal projection of the head it is adjoined to). Then (48) is not legitimate. Furthermore, adjoining a nonhead to P in (48) would not yield a specifier of M, since M would c-command that adjoined phrase (given that M is not dominated by P). Adjunction of a head to a nonhead is thus systematically unavailable.

Since specifiers are instances of adjunction, it follows that specifiers cannot be heads.[34]

Chapter 4
Word Order

4.1 The Specifier-Complement Asymmetry

I would now like to explore the relation between the LCA, repeated in (1), and the ordering, in terms of precedence/subsequence, of the terminals of a given phrase marker.

(1) *Linear Correspondence Axiom*
d(*A*) is a linear ordering of *T*.

Implicit in the earlier discussion was the assumption (not used until now) that the linear ordering of terminals constituted by d(*A*) must directly and uniformly provide the precedence/subsequence relation for the set of terminals.

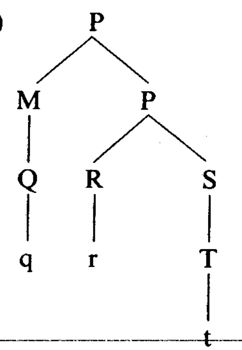
However, nothing said so far tells us whether it is precedence itself or rather subsequence that is provided. Put another way, the question is whether asymmetric c-command is mapped (by *d*) to precedence or to subsequence. If it is to precedence, then the following holds:

(2) Let X, Y be nonterminals and x, y terminals such that X dominates x and Y dominates y. Then if X asymmetrically c-commands Y, x precedes y.

Were asymmetric c-command to map to subsequence, then *precedes* in (2) would have to be replaced by *follows*. I will proceed to argue that (2) is true as stated, namely, that asymmetric c-command does map to precedence.

Let us temporarily hold the choice between precedence and subsequence in abeyance, however, and consider again a phrase marker representing a head with complement and specifier, as in (3).

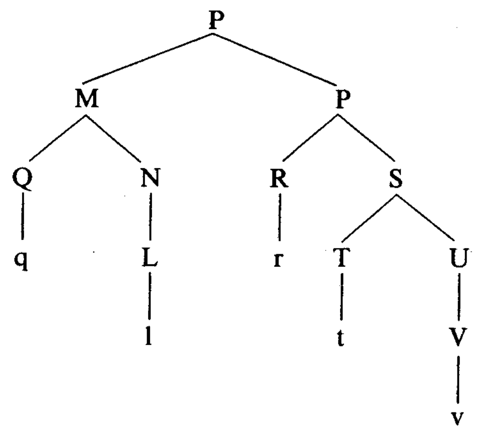
(3)



Since M asymmetrically c-commands R (the head)—that is, A contains ⟨M, R⟩—it follows that d(A) contains ⟨q, r⟩. Similarly, since R asymmetrically c-commands T, d(A) contains ⟨r, t⟩. It therefore follows that with respect to the ordering of terminals, q and t are necessarily on opposite sides of the head r.

A similar conclusion would hold if M dominated a more complex phrase than just Q and if S dominated a more complex phrase than just T, as in (4).

(4)



In (4) M asymmetrically c-commands R and R asymmetrically c-commands T, U, and V. Therefore, A for (4) contains ⟨M, R⟩, ⟨R, T⟩, ⟨R, U⟩, ⟨R, V⟩. From the fact that ⟨M, R⟩ is in A, it follows that ⟨q, r⟩ and ⟨l, r⟩ are in d(A). From the fact that ⟨R, T⟩ and ⟨R, U⟩ are in A, it follows that ⟨r, t⟩ and ⟨r, v⟩ are in d(A). In other words, all the terminals of the specifier M are on the opposite side of the head R (terminal r) from all the terminals of the complement S.

More generally put, no matter how complex the specifier or complement, it will always be the case, in any phrase marker, that specifier and complement are on opposite sides of the head. In other words, if we represent head, specifier, and complement as H, S, and C, then the conclusion so far is that of the six permutations of H, S and C, only two are permitted by the theory, namely, S-H-C and C-H-S. The other four (S-C-H, C-S-H, H-S-C, H-C-S) are all excluded by the requirement that specifier and complement be on opposite sides of the head.

The exclusion of S-C-H (e.g., of SOV) requires us to distinguish a complement position from the contents of that position. What I claim, and will return to in more detail below, is that SOV (and more generally S-C-H) is strictly impossible, in any language, if taken to indicate a phrase marker in which the sister phrase to the head (i.e., the complement position) precedes that head. On the other hand, SOV (and S-C-H) is perfectly allowable if taken to indicate a phrase marker in which the complement has raised up to some specifier position to the left of the head.

4.2 Specifier-Head-Complement as a Universal Order

We are now left with two constituent order possibilities, specifier-head-complement and complement-head-specifier. A rapid look at (a small subset of) the world's (presently existing) languages reveals that of the two orders, the former is a significantly more plausible universal than is the latter. Consideration of the relative order of head and complement alone is not sufficient to yield any firm conclusion, since both head-complement and complement-head orders are widely attested. On the other hand, the relative order of specifier and head is much more visibly asymmetric, in the following sense: although there may be some categories for which both orders are widespread, there are other categories where specifier-head order strongly predominates. (I know of no categories for which head-specifier is the cross-linguistically predominant order.)

In fact, CP is a category whose specifier, the typical landing site for moved *wh*-phrases, is visibly initial to an overwhelming degree.[1] Spec,IP (i.e., subject position) is clearly predominantly initial in its phrase.[2] That is straightforwardly true for SVO and SOV languages, and almost as obviously true for VSO languages, assuming the by now usual analysis of VSO order as deriving from SVO order by leftward V-movement.[3] According to Greenberg (1966, p. 76), the other types, OVS, OSV, and VOS, are "excessively rare."

From the present perspective, OSV would involve movement of the O past S to the specifier position of a higher head. OVS and VOS must not have S in a final specifier position, but must instead either have OV or VO moving as a unit leftward past S, or else V and O moving separately leftward past S,[4] with the expectation, then, that such languages should show OVSX and VOSX orders.

I conclude that specifier-head-complement, and not the reverse, is the only order available to the subcomponents of a phrase. Consider again, in this light, (4). The conclusion just stated that S-H-C order is the only one available means that the linear ordering d(*A*) containing, for example, ⟨q, r⟩ and ⟨l, r⟩ and ⟨r, t⟩ and ⟨r, v⟩ should be interpreted so that ⟨x, y⟩ means that the terminal symbol x *precedes* the terminal symbol y.

4.3 Time and the Universal Specifier-Head-Complement Order

This is not logically necessary. We can imagine a UG that would differ from the one I claim to be characterizing. This other UG would be identical to the actual one but would interpret ⟨x, y⟩ as meaning that x follows y. That would yield a perfectly valid linear ordering, one that would be the mirror image of the actual one. Languages compatible with this other UG would have C-H-S order instead of the actually valid S-H-C and would look like a set of strict mirror images of the languages we are familiar with.

I would now like to suggest a possible explanation for the fact that UG imposes S-H-C order on phrases, rather than the reverse. (The proposal will also account for the fact that UG does not allow languages a choice between S-H-C and C-H-S.) It is therefore necessary for me to explain why ⟨x, y⟩ is interpreted as 'x precedes y' rather than as 'x follows y'.

Recall from chapter 1 that the asymmetric c-command relation is significantly similar to the dominance relation (both are locally linear). Associated with the dominance relation on phrase markers is a "root node" that has the property of dominating every node in the phrase marker (except itself). In the usual phrase marker, no node has the property of asymmetrically c-commanding every node except itself. I would like to propose bringing asymmetric c-command and dominance more into parallel by postulating an abstract node A for every phrase marker, with the property that A asymmetrically c-commands every other node. This abstract node should be thought of as being adjoined to the root node.[5]

Since every other node dominates at least one (perhaps empty) terminal element, A should be taken to dominate a terminal element. There are two plausible candidates, either an abstract terminal a that has the property of preceding all the other terminals (i.e., an abstract beginning terminal) or an abstract terminal z that follows all the other terminals. I propose that the abstract root node for asymmetric c-command should be mapped by *d* into the abstract beginning terminal a.

The intuitive motivation for taking d(A) = a rather than d(A) = z is that a and z are not quite as symmetric as they might seem, in a way that favors a. Let us think of the string of terminals as being associated with a string of time slots. That by itself is not sufficient to induce an asymmetry between a and z. Let me then make the further claim that what is paired with each time slot is not simply the corresponding terminal, but the substring of terminals ending with that terminal (i.e., the substring produced up to that time).

In other words, a string of terminals abcdz (with a and z abstract) is mapped to a set of substrings.

(5) a, ab, abc, abcd, abcdz

An asymmetry between a and z has now appeared: a precedes every terminal in every substring, but z does not follow every terminal in every substring (since z figures in only one substring). If the abstract root node for asymmetric c-command needs to be mapped by *d* to a corresponding abstract "root node" for terminals, and if that root node for terminals must be in some fixed relation to every terminal in every substring, then that abstract terminal must be a and the fixed relation must be 'precedes'.

Let us consider, then, that d(A) = a. The question we are trying to answer is how to interpret ⟨x, y⟩, where ⟨x, y⟩ is in d(*A*). More specifically, the question is whether ⟨x, y⟩ is 'x precedes y' or 'x follows y'.

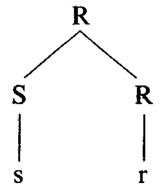
Assume the latter. Now by hypothesis the abstract node A asymmetrically c-commands Y, for all Y, so that ⟨A, Y⟩ is in *A*, for any phrase marker containing Y. Since d(A) = a, it follows that ⟨a, y⟩ is in d(*A*) (for all y dominated by Y). So that if ⟨x, y⟩ is 'x follows y', we conclude that 'a follows y', for all y. But a is the abstract beginning terminal. Thus, we have a contradiction. Therefore ⟨x, y⟩ cannot be 'x follows y' but must rather be 'x precedes y'.

From the fact that ⟨x, y⟩ is to be interpreted as 'x precedes y', it follows, as discussed in the preceding section, that the unique order of constituents provided by UG is S-H-C, as desired. In effect, the fact that UG provides S-H-C order (rather than the reverse) derives from the hypothesis that (5) (rather than a sequence of substrings working backward from the final terminal) is the correct way of representing the relation between terminals and time slots. This S-H-C property of UG, as well as the fact that UG does not make both orders available, is thus seen to be ultimately related to the asymmetry of time.

4.4 Linear Order and Adjunction to Heads

We have just seen how the fact that specifiers are adjoined to the head + complement constituent results in specifiers necessarily preceding the head (and complement). Now consider the adjunction of one head to another, as in (6).

(6)



As shown in section 3.2, S asymmetrically c-commands R. Therefore, ⟨s, r⟩ is in the d(*A*) of (6). By the results of the preceding section, it follows that s precedes r. In other words, the present theory has as a necessary consequence that an adjoining head (S) will invariably precede the head that it adjoins to (R).

Taking S to be a clitic, this derives the generalization that I proposed in earlier work to the effect that a clitic invariably precedes the head that it adjoins to.[6]

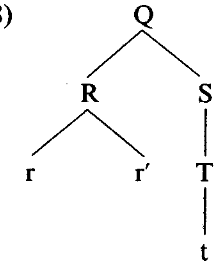
4.5 Linear Order and Structure below the Word Level

What does the present theory say about structure below the word level? Consider a head with internal structure, as in (7).

(7) Cats overturn chairs.

The verb *overturn* consists of *over* and *turn*. The relevant VP configuration is (8).

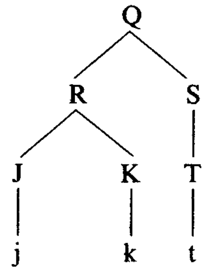
(8)



R asymmetrically c-commands T. Therefore, both r and r′ will precede t, apparently correctly. However, it is not at all clear that in a case like (7) we really want to say that there is a V node (corresponding to R) that dominates two terminals with no intervening structure whatsoever.

Let us assume, then, that rather than (8) the structure of (7) is (9).

(9)



Here j and k are the two morphemes that comprise the head R, and they are dominated respectively by the sub-word-level nonterminals J and K. Now R asymmetrically c-commands T as before, so that j and k will precede t, as expected. However, S now asymmetrically c-commands J and K, so that ⟨t, j⟩ and ⟨t, k⟩ will be in d(*A*). Consequently, t must precede both j and k. But that is a contradiction (violation of antisymmetry). The conclusion, then, is that (9) is not a possible representation for (7).

This conclusion could be evaded if we were to decide that J and K are really different from the usual nonterminals. Put another way, we might decide that J and K here do not belong to the set of nonterminals on which is defined the relation of asymmetric c-command that in this theory maps into linear precedence. Such a decision would have the effect of divorcing sub-word-level structure from phrase structure; that is, it would have the effect of making structure below the word level invisible to the LCA.