Norbert Hornstein, Jairo Nunes and Kleanthes K. Grohmann
Understanding Minimalism: An Introduction to Minimalist Syntax

February 1, 2004

CHAPTER 6

Phrase Structure

6.1. Introduction

Recall from section 1.3 that one of the "big facts" regarding human languages is that sentences are composed of phrases, units larger than words organized in a specific hierarchical fashion. This chapter is devoted to phrase structure. The starting point for our discussion will be X'-Theory, the module of GB responsible for determining the precise format of licit phrases and syntactic constituents in general.

One of the main motivations for the introduction of X'-Theory into generative grammar was the elimination of a perceived redundancy in the earlier *Aspects*-model. The *Aspects*-theory of the base included two kinds of operations. First, there was a phrase structure component based on a variety of context free phrase structure rules (PS rules) such as those in (1) below. (1a), for instance, states that a sentence S expands as (is formed by) *NP Aux VP* and (1b) says that a VP expands as a V with optional NP, PP, and S complements. The application of these sorts of rules generates phrase markers (trees) with no lexical items at the terminals, as illustrated in (2).

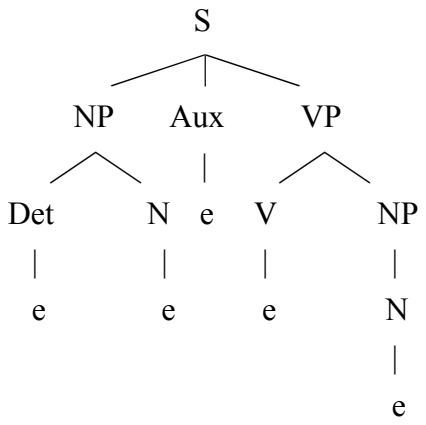
(1) *Basic Phrase Structure Rules*

a. S → NP Aux VP

b. VP → V (NP) (PP) (S)

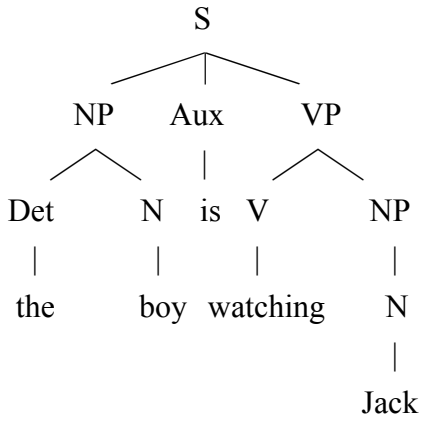
c. NP → (Det) N (PP) (S)

(2)



Lexical elements were then introduced into the empty terminal positions (designated by *e* in (2)) by a process of lexical insertion, yielding phrase markers like (3).

(3)



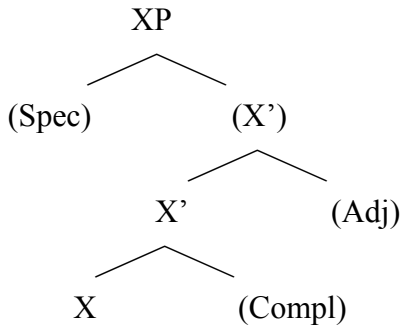
So dividing the task of building initial phrase markers contains an unfortunate redundancy.[1] To see this, consider what sorts of verbs can be inserted into the VP of (2), for instance. Only transitive verbs like *watch* and *kiss* yield an acceptable sentence if inserted. Intransitive verbs like *sleep* or *cough* don't take objects and so don't "license" enough of the available portions that the phrase structure affords, and ditransitive verbs like *give* or *put* are not provided with enough empty positions for all their arguments. In effect, the rules for lexical insertion must code the argument structure of the relevant lexical heads and match them to the possible phrase structure that the PS rules make available. In other words, the information about possible phrase structures is coded twice, once in the PS rules and a second time in the lexical entries.

X'-Theory was intended to eliminate this redundancy by dispensing with PS rules and construing phrase structure as the syntactic "projection" of the argument structure of a lexical head. It incorporates several distinctive claims, providing a recipe for how such "projection" from argument structure takes place. Under one of its more common formulations, the recipe

[1] See e.g. Chomsky (1965, 1970), Lyons (1968), and Jackendoff (1977).

has the general format along the lines of (4), where a head X projects a maximal constituent XP by being optionally combined with a complement, a number of modifiers (adjuncts), and a specifier that "closes off" the projection of X.

(4)



In the sections that follow we'll review the main properties encompassed by the general schema in (4), as well as the motivation for their postulation, and discuss if and how such properties can be derived or incorporated in a minimalist system. The chapter is organized as follows. In section 6.2 we review the main properties of phrase structure that X'-Theory intends to capture. In section 6.3, we discuss a "bare" version of phrase structure, according to which the key features of phrase structure follow from the internal procedures of the structure building operation Merge, coupled with general minimalist conditions. Section 6.4 shows how structures formed by movement also fall under the bare phrase structure approach and introduce the copy theory, according to which traces are copies of moved elements. Finally, section 6.5 concludes the chapter.

6.2. X'-Theory and Properties of Phrase Structure

6.2.1. Endocentricity

One of the key ingredients of the recipe for projecting phrases provided by X'-Theory is endocentricity. The general X'-schema in (4) embodies the claims that every head projects a phrase and that all phrases have heads. Support for this endocentric property of phrases comes from distributional facts. A single verb like *smile*, for instance, can be an adequate surrogate for the VP in (5) below, but the sequence adjective plus PP can't, as illustrated in (6). In other words, endocentricity imposes hierarchy of a specific kind onto linguistic structures, allowing for phrases structured as in (7a), but not as in (7b), for instance.

(5) [John will [$_{VP}$ drink caipirinha]]

(6) a. [John will [smile]]
b. *[John will [fond of caipirinha]]

(7) a. VP → V
b. *VP → A PP

The endocentricity property coded by X'-Theory thus says that whenever we find phrases, we find morphemes that serve as heads of those phrases and that these heads are relatively prominent in not being further embedded within other phrases of a distinct type. It's not merely the case that verb phrases must contain verbs; they must *prominently* contain them. The phrase in (8a), for instance, contains the verb *like*, but it's a noun phrase rather than a verb phrase because the verb is too deeply buried within another phrase to serve as the head of the whole.

(8) a. books that I like
b. [[books [that I like]]

Endocentricity also affords a local way of coding another interesting fact about natural languages: that words "go" with some words and not others. An example or two should make what we mean here clear. Consider a sentence like (9).

(9) Rhinos were/*was playing hockey.

(9) displays subject-predicate agreement. The plural subject *rhinos* requires that the form of the past tense of *be* come out as *were*. In an example like (9), we can state the required relation very locally: the predicate immediately following or next to the subject must agree with it in number properties.

Consider now a slightly more complex case.

(10) Rhinos playing on the same team were/*was staying in the same hotel.

Observe that the very same restriction witnessed in (9) holds in (10); that is, the verb agrees in number with *rhinos* and must be plural. However, in this instance, there is no apparent local linear relation mediating the interaction of *rhinos* and *were* as they are no longer linearly contiguous, at least not evidently. In fact, matters are much worse than this. Once we consider (9) and (10) together, it's easy to see that any number of words can intervene between the subject element coding number and the predicate, without altering the observed agreement requirement. How then can this restriction between subject and predicate be

locally stated?

Endocentricity comes to the rescue. If we assume that phrases are projections of their heads as endocentricity mandates, then the number specification of an NP can be seen as a simple function of the number specification of its head. In the case of (10), for instance, the subject NP triggers plural agreement in virtue of the plural specification of its head *rhino*s, as illustrated in (11).

(11) [[$_{NP}$ [$_{N'}$ *rhinos*] [playing on the same team]] *were* staying in the same hotel]

Observe that the NP projected from *rhinos* does abut *were* and hence the same locality requirement that holds between *rhinos* and *were* in (9) can be seen to be present in (10), as well, once some phrase structure is made explicit and we assume that there is a tight relationship between a phrase and its head, i.e. if we assume that phrases obey an endocentricity requirement.

Notice further that if agreement could peruse all the constituents of the subject, the verb *be* in (10) could in principle agree with *team*, which is actually linearly closer to it, and surface as *was*. The fact that this doesn't happen illustrates what may be called the *periscope property* induced by endocentricity: subject-predicate agreement is allowed to look into the subject NP and see its head, but nothing else.

Let's now consider the sentences in (12).

(12) a. John ate bagels.
b. *John ate principles.
c. *John ate principles of bagel making.

(12b) is a funny sentence. Why? Presumably because principles are not things that one eats. This contrasts with (12a), since bagels are quite edible. Observe that the oddity of (12b) doesn't diminish if we add more elements to the phrase. Arguably, (12c) is odd for the same reason that (12b) is (principles are not edible). This in turn constitutes another example of the periscope property. Consider why. The object of a verb like *eat* should be something edible. To determine if an object denotes something edible, one need only look and examine its head. If the head is a food product like *bagels*, then all will go swimmingly. If the head is something like *principles*, then no matter what else edible we put in the phrase, the sentence will retain its oddity. Thus, the contrast between (12a) and (12c) is due to the fact that the head of the object NP is *bagels* in the former, and *principles* in the latter; crucially, *bagel* in (12c) is too buried to be seen by *ate*.

Accordingly, there are also no known cases where a syntactic relation cares about anything, but the head. For example, there are no verbs that select NPs with certain

determiners, say *three* but not others, say *every*, or verbs that like some kinds of nominal modifiers for their complements, say PPs, but not others, say APs. Thus, although the verb *eat* imposes restrictions on the head of its complement, it seems to have no effect on what sorts of specifiers or modifiers this head may take, as illustrated in (13).

(13) a. John *ate* [$_{NP}$ Bill's/no/every *bagel*].
b. I *ate* [$_{NP}$ a big fat greasy luscious chocolate square *bagel* with no hole].

To sum up, endocentricity is a well motivated property of the phrase structure of natural languages and is captured under the general X'-schema in (14).

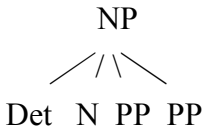
(14) XP → … X …

Before we move on, it's important to point out that endocentricity is not an intrinsic property of any phrase-structure system. The PS rule in (1a), repeated below in (15), for instance, is not endocentric. However, if endocentricity is an inherent property of all structures in natural languages, they should have no rules like (15). Research in the 1980s about functional heads both in the clausal and in the nominal domain indeed led to this conclusion and to the complete abandonment of PS rules. We return to this issue in section 6.2.5 below, where we discuss the structure of functional projections.

(15) S → NP Aux VP

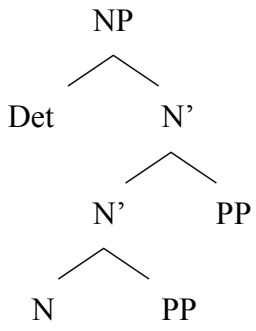
6.2.2. Binary branching

One further property of phrase structure incorporated into standard versions of the X'-schema is binary branching.[2] Within these versions of X'-Theory, multiple branching structures such as (16), for instance, came to be replaced by binary branching structures like (17).

(16)


[2] See especially Kayne (1984) on binary branching in phrase structure.

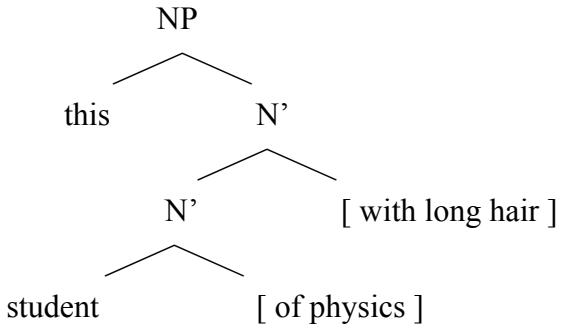
(17)



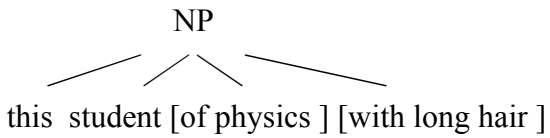
Binary branching was motivated for a mix of aesthetic and empirical reasons.[3] Let's consider one empirical argument. It's a standard assumption that syntactic processes and operations deal with syntactic constituents. Pronominalization is one such process. Consider the sentences in (18) below, for instance. In English, the pronoun *one* may replace *student of physics* in (18a) and *student of physics with long hair* in (18b).[4] Thus, each fragment that is pronominalized should be a syntactic constituent (a node in a syntactic tree) in the relevant NP structure. In other words, in order to capture the pronominalization facts in (18), there should be a node dominating only *student of physics* and excluding everything else and another node dominating *student of physics with long hair* and excluding everything else. These requirements are met in the binary branching structure in (17), as shown in (19a), but not in the multiple branching structure in (16), as shown in (19b).

(18) a. John met this student of linguistics with long hair, and Bill met that one with short hair.
b. John met this student of linguistics with long hair, and Bill met that one.

(19) a.



b.



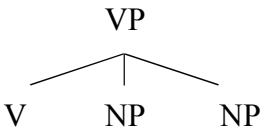
[3] See Kayne (1984) for relevant discussion.

[4] This test goes back to Baker (1978); see also Hornstein and Lightfoot (1981) and Radford (1981), among others, for early discussion.

Research in the 1980s generalized binary branching to all lexical and functional projections, with very interesting empirical consequences.[5] Take double object constructions such as (20) below, for example. If their VP were to be assigned a ternary branching along the lines of (21), neither complement should be more prominent than the other, for they c-command each other. However, binding and negative polarity licensing, which both require c-command, show that this can't be the case. Under the structure in (21), the anaphor in (22b), for instance, should be bound by *the boys* and the negative polarity item *anyone* in (23b) should be licensed by the negative quantifier *nothing*.

(20) John gave Bill a book.

(21)

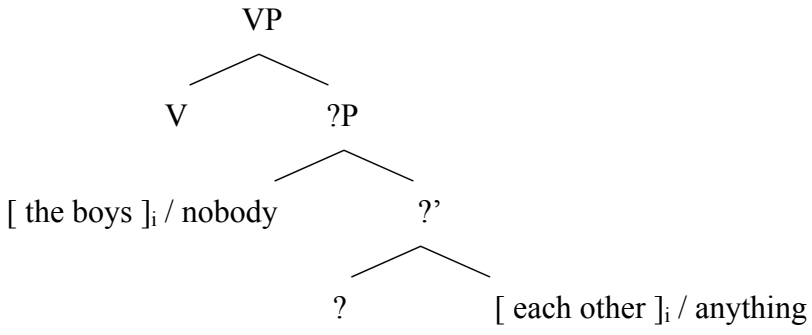


(22) a. Mary showed [the boys]$_i$ [each other]$_i$
b. *Mary showed [each other] [the boys]$_i$

(23) a. John gave nobody anything.
b. *John gave anyone nothing.

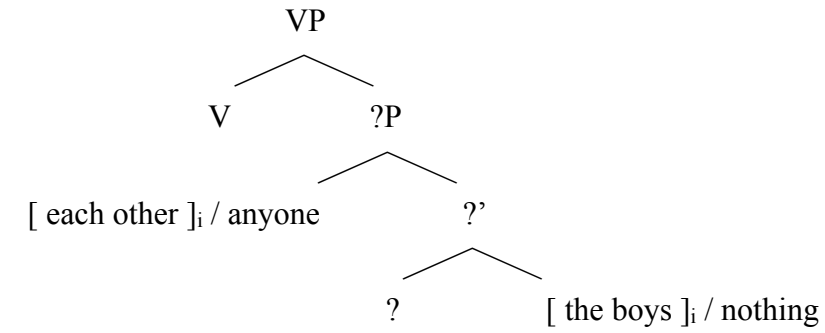
By contrast, if only binary branching is permitted, the contrasts in (22) and (23) can be accounted for if the phrase structure of double object constructions is actually more complex, with an extra layer of structure, as illustrated in (24).

(24) a.



[5] See e.g. Kayne (1984), Chomsky (1986a), and Larson (1988).

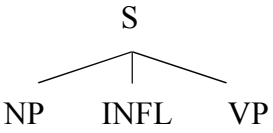
b.



Given that in (24) the dative c-commands the theme, but not the opposite, the anaphor and the negative polarity item are licensed in (24a), but not in (24b); hence the contrasts in (22) and (23).

The assumption that all phrases are organized in terms of binary branching also led to the reevaluation of the clausal skeleton given in (25) below. We'll get back to this issue in section 6.2.5 below.

(25)

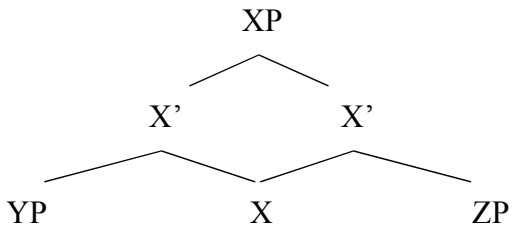


EXERCISE 6.1:
What could the extra projection *?P* in (24) be? Given our discussion of ditransitive predicates in section 3.3, discuss if and why the structure you proposed in your answer to exercise 3.7 is more adequate than the one in (24).

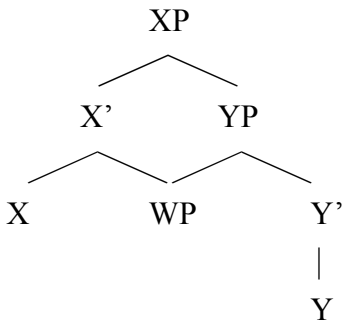
6.2.3. Singlemotherhood

Another property of phrase structure in natural languages is that syntactic constituents are not immediately dominated by more than one constituent. That is, syntactic constituents don't have multiple mothers. There seem so be no syntactic process that requires structures such as the ones below, for instance, where X in (26a) is the head of more than one phrase, and the complement of X in (26b) is also the specifier of Y.

(26) a.



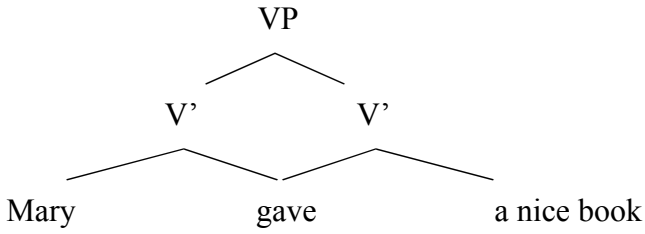
b.



It's important to stress that there is nothing crazy about the structures in (26) by themselves.[6] Notice that they are endocentric and binary branching, like all the licit structures we have been examining thus far. One could even hypothesize that the structure in (26a), where X has two complements, would serve well to represent double object constructions, as shown in (27), or that the structure in (26b) would provide a nice account for the fact that in constructions involving headless relative clauses, the moved *wh*-phrase may function as the complement of the matrix verb (see section 2.3.2.3), as illustrated in (28).

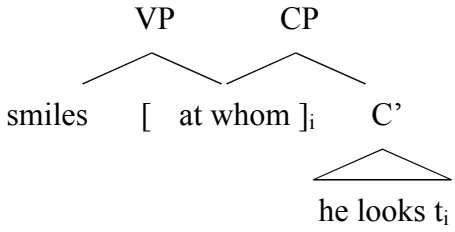
(27) a. John gave Mary a nice book.

b.



(28) a. John always smiles at whom he looks.

b.



However, as discussed in section 6.2.2, facts regarding binding and negative polarity licensing show that in double object constructions, the dative must c-command the theme, which is not the case in (27b), where neither c-commands the other. In turn, if the structure in (28b) were allowed, VP-preposing should in principle target only the main verb and the moved PP, leaving CP stranded, contrary to fact, as illustrated in (29).

---

[6] See McCawley (1981), Cann (1999), and Starke (2001) for relevant discussion.

(29) John said that he would smile at whom he would look, and
   a. smile at whom he looked, he did.
   b. *smile at whom, he did, he looked.

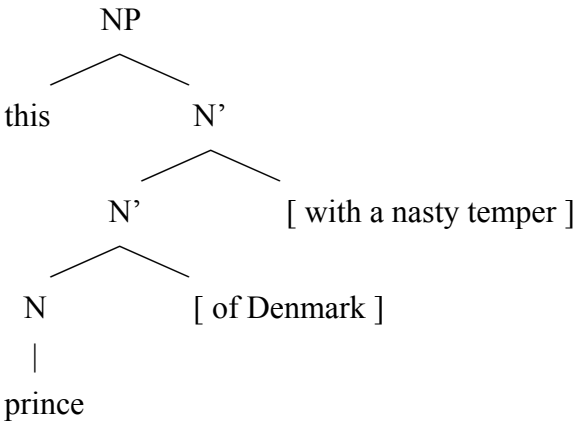
To sum up, despite the plausibility of multiple immediate dominance, it seems to be a fact that human languages simply don't work this way, and singlemotherhood is also a property of natural language phrases.

6.2.4. Bar-Levels and Constituent Parts

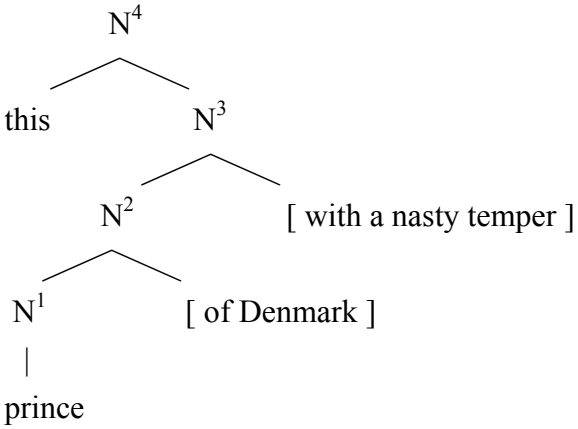
Consider now the two possible representations for the phrase in (30) given in (31).

(30) this prince of Denmark with a nasty temper

(31) a.



b.

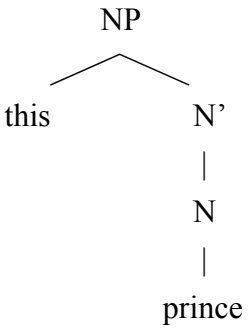


(31a) illustrates our familiar sandwich-like organization of X'-Theory: the bottom (the head), the top (the maximal projection), and the filling (the intermediate projections); in other words, three levels are encoded. (31b), on the other hand, differs in that it registers the total number of nominal projections (four in this case). At first sight, these appear to be just

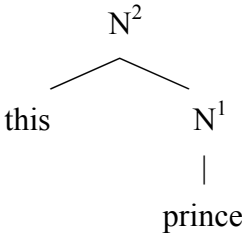
notational variants recording the same information. However, they actually make distinct empirical predictions when we also consider the two representations in the case of the simpler phrase in (32).

(32) this prince

(33) a.



b.



According to the counting approach, the constituent *prince* will always be of the same type (N^1), regardless of whether or not it occurs in more complex structures. By contrast, under the X'-approach, *prince* doesn't have the same status in (30) and (32); in (32), in addition to counting as an N, it's also an N' as well (cf. (31a) and (33a)). In other words, the counting approach makes the prediction that if some syntactic process affects *prince* in (32), it may do the same in (30); the X'-approach, on the other hand, doesn't make such a prediction because *prince* doesn't necessarily have the same status in these phrases. Let's then see how the two approaches fare with respect to the *one*-substitution facts in (34).

(34) a. John likes this prince and I like that one.
b. *John likes this prince of Denmark and I like that one of France.

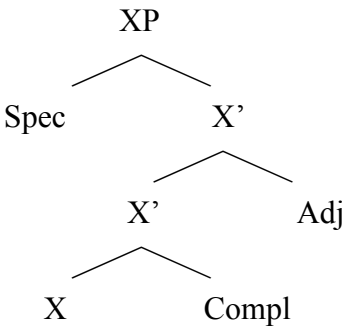
In (34a), *one* is a surrogate for *prince* and we have a well-formed sentence. Thus, under the counting approach, we should get a similar result in (34b), contrary to fact. Under the X'-approach, on the other hand, the contrast in (34) can be accounted for if *one* targets N'-projections; hence, it may replace the N'-projection of *prince* in (34a) (cf. (33a)), but there is no such projection in (34b) (cf. (31a)).[7] Facts like these require that an adequate theory of

[7] These data get reanalyzed in section 6.2.6 below without the use of N'.

phrase structure in natural languages resort to the three-way bar-level system distinguishing heads, intermediate projections, and maximal projections.

In addition to encoding this three-way distinction, the general X'-schema in (35) also functionally identifies three constituent parts — complements, modifiers (adjuncts), and specifiers — which are mapped into their hierarchical positions according to the principles in (36).

(35)



(36) *Principles of Phrase Structure Relations*

a. Complements are sisters to the head X.
b. Modifiers are adjuncts to X'.
c. Specifiers are daughters to XP.

That complements and modifiers are semantically distinct is easy to see. In the verbal domain, for instance, complements are generally obligatory, whereas adjuncts are optional, as illustrated in (37).

(37) John fixed *(the car) (yesterday).

Furthermore, whereas the head and the complement form a single predicate, a modifier adds a further specification to an existing predicate. Compare the adjunct structure in (38a) with the complement structure in (38b) below, for example. (38a) says two things about Hamlet: that he is a prince and that he is from Denmark. (38b), on the other hand, says just one thing about him: that he has the property of being a prince of Denmark; in fact, it's quite meaningless to paraphrase (38b) by saying that Hamlet is a prince and is of Denmark.

(38) a. Hamlet is a prince from Denmark.
b. Hamlet is a prince of Denmark.

What X'-Theory does with the mapping principles is (36) is state that in addition to lexical information (the difference between *from* and *of* in (38), for instance), the hierarchical

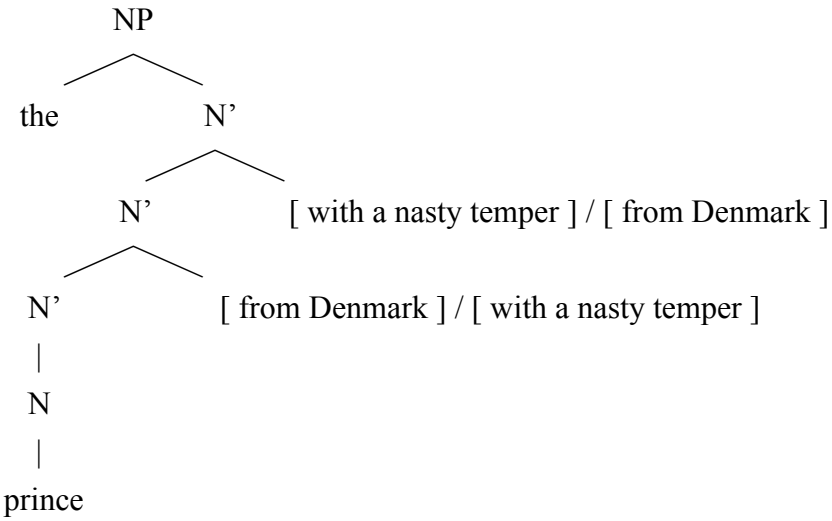
configuration is crucially relevant for the interpretation of complements and modifiers. This can be clearly seen by the contrast between (39) and (40).

(39) a. the prince from Denmark with a nasty temper
b. the prince with a nasty temper from Denmark

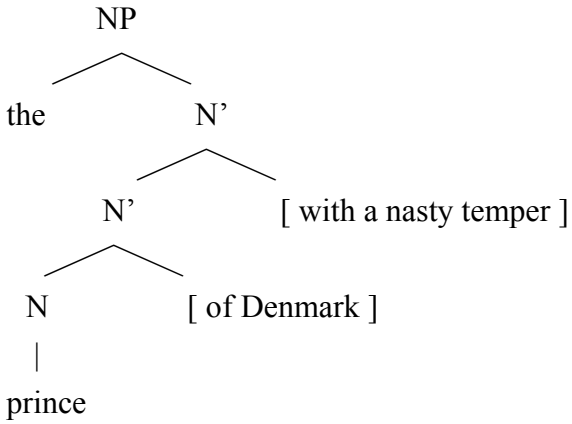
(40) a. the prince of Denmark with a nasty temper
b. *the prince with a nasty temper of Denmark

Whereas the adjuncts can freely interchange in (39), that is not the case of the complement and the adjunct in (40). This contrast in word order is accounted for by the mapping principles in (36). In (39), word order doesn't matter as long as (36b) is satisfied and each of the adjuncts is mapped as a sister of N', as shown in (41) below. In (40), on the other hand, only the order in (40a) can comply with both (36a) and (36b), as shown in (42a); the order in (40b) requires that *of Denmark* appears as a sister of N', as shown in (42b), yielding a conflict with the lexical specification of *of* and violating (36a).

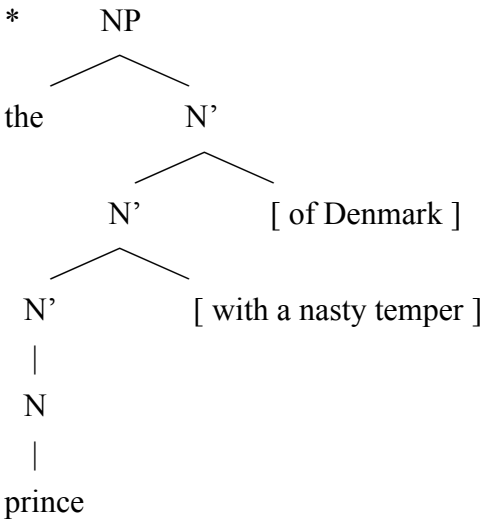
(41)



(42) a.



b. *



As for the functional identification of specifiers in (36c), the guiding intuition was that any head could project as many intermediate projections as there were adjuncts, but some specific projections would close off projections of that head. For instance, whereas one could keep indefinitely adding adjunct PPs to N'-projections and getting another N', once a determiner was added, we would obtain an NP and no further projection from the relevant N head would further take place. Distributionally, this would account why adjuncts can iterate, but determiners can't, as shown in (43).

(43) a. the prince from Denmark with a nasty temper
b. *this the prince from Denmark

To sum up, the key properties embodied in the X'-schema in (35) and the mapping principles in (36) are reasonably motivated and invite closer scrutiny from a minimalist perspective. We have already seen in section 5.4.2.1, for instance, that if *v*Ps allow more than one Spec, the system may get simpler. But before getting into a detailed discussion of phrase structure from a minimalist point of view, let's first briefly examine the consequences of assuming X'-Theory for the structure of functional heads.

EXERCISE 6.2:
Try to build an argument based on syntactic constituency that VPs should also involve three bar-levels. Consider how VP ellipsis, VP fronting, and *do so* might be employed for collecting evidence.

EXERCISE 6.3:
Some prepositions may be used to introduce both complements and adjuncts, as illustrated in (i). Based on this ambiguity, explain why (ii) has just one of the two potential readings it could have. (Assume the rough bracketing provided here.)

(i) a. books on linguistics
b. books on the floor

(ii) books [on chairs] [on tables]

6.2.5. Functional Heads and X'-Theory

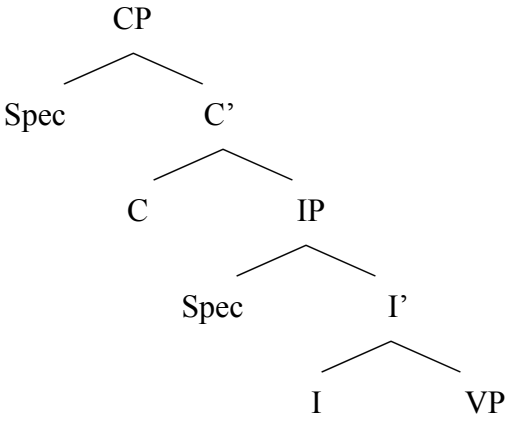
As mentioned in section 6.1, one of the main motivations behind X'-Theory was the elimination of PS rules. Two such rules, however, still made their way into GB, namely, the rules for clausal structure in (44).

(44) a. S' → Comp S
b. S → NP Infl VP

(44a) was in fact more congenial to X'-Theory, in that it was endocentric (Comp was taken to be the head of S'[8]) and binary branching; its difference from the standard X'-schema was that it had just two levels: the head and the maximal projection. (44b), by contrast, was far from meeting X'-postulates: it was not endocentric, it had ternary branching and the issue of bar-levels was even worse, for S was not taken to be a maximal projection.

Research in the mid 1980s led to the conclusion that PS rules could be completely eliminated from the grammar and that the clausal structure could be roughly organized along the lines of (45).[9]

(45)



[8] See Bresnan (1972).

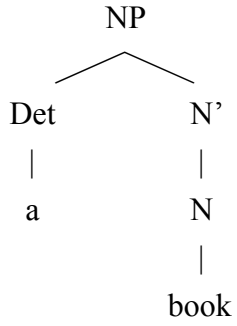
[9] See Fassi Fehri (1980), Stowell (1981), and Chomsky (1986a) for relevant discussion.

In (45), the complementizer C takes a projection of Infl (= I) as its complement and Infl, in turn, takes VP as its complement; [Spec,CP] is the position generally filled by moved *wh*-elements (or their traces) and [Spec,IP] is the position traditionally reserved for syntactic subjects.

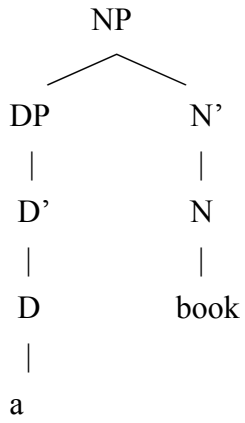
Later research within GB has reexamined the structure in (45), suggesting that Infl (see section 4.3.1) and C should be split into several heads — such as T(ense), Agr(eement), Asp(ect), Top(ic), Foc(us), etc. — each of which projecting a distinct phrase.[10] Although there is disagreement with respect to the number of such phrases and the dominance relationship among them, researchers generally agree on one point: all of these phrases are in compliance with the postulates of X'-Theory.

A similar reevaluation took place with respect to nominal domains. At first sight, the traditional structure in (46a) below required just a minor readjustment: in order for a well formed X'-structure to obtain, the determiner would have to project. (46b) should in principle fix this problem. However, by inspecting the projected structure of DP in (46b), one could not help but wonder what kind of complement a D head (= Det) could take or whether it could it take a specifier.

(46) a.



b.

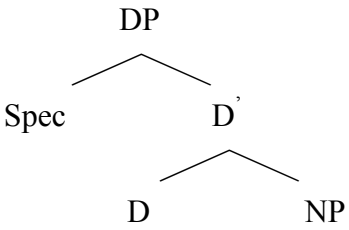


Addressing similar questions, research in the 1980s pointed to the conclusion that a

[10] See Pollock (1989), Belletti (1990), Chomsky (1991), Rizzi (1997), Cinque (1999), and the more recent collections of papers in Cinque (2002), Belletti (2004), and Rizzi (2004).

better representation for a phrase such as *a book*, rather than (46b), should actually be along the lines of (47), where the determiner takes NP as its complement.[11]

(47)



The structure in (47) receives support from very different sources. First, it still captures the old intuition that, in general, once a determiner is added to a structure, no further projections of N are possible. But it also has room to accommodate interesting cases such as (48) below, where a *wh*-element precedes the determiner and we are still in the "nominal" domain. (48) receives a straightforward analysis if we assume the structure in (47), with the *wh*-phrase in [Spec,DP].

(48) [[how good] a story] is it?

The structure in (47) also captures the fact that in many languages determiners and clitic pronouns are morphologically similar or identical, as illustrated in (49) below with Portuguese.[12] Pronouns, under this view, should be D-heads without a complement.

(49) *Portuguese*

a. João viu o menino.
João saw the boy
'João saw the boy.'

b. João viu-o.
João saw-CL
'João saw him.'

Further examination of the structure of DP, like what happened in the clausal domain, opened the possibility that there should be additional layers of functional projections between DP and NP.[13] Again, these analyses generally agreed that the extra layers of functional structure were organized in compliance with X'-Theory.

---

[11] See Brame (1982), Szabolcsi (1983), Abney (1987), and Kuroda (1988) for relevant discussion.

[12] See Postal (1969) and Raposo (1973) for early discussion.

[13] Bernstein (2001) provides a recent overview of the "Clausal DP-Hypothesis" and plenty of references on the finer structure of DP developed in the wake of Brame (1982), Szabolcsi (1983), and Abney (1987).

Since a detailed discussion of the competing alternatives for clausal and nominal domains would derail us from our discussion of the general properties of phrase structure, from now on we'll assume the structures in (45) and (47) for concreteness.

EXERCISE 6.4:
Try to build additional arguments for the structure in (45) and (47) in your language by using traditional tests for syntactic constituents.

EXERCISE 6.5:
In section 6.2.1, we saw that the periscope property induced by endocentricity ensures that, for selectional purposes, a given head only sees the head of its complement and nothing else. Assuming the clausal structure in (45), that would imply that a verb that selects a CP for a complement should see only the head C, and that should be it. However, the data in (i) and (ii) seem to show that the matrix verb is seeing more than the head of its complement. In (i) it seems to select the tense of the embedded clause, whereas in (ii) it appears to impose restrictions on the specifier of the embedded CP. How can these facts be reconciled with the periscope property?

(i) a. John wants Bill to win.
b. *John wants that Bill will win.

(ii) a. John believes that Bill won.
b. *John believes how Bill won.
c. *John wonders that Bill won.
d. John wonders how Bill won.

EXERCISE 6.6:
In exercise 6.5, we saw that verbs appear to select the tense of their clausal complement. Things may seem more complicated in face of the following generalization: in English, if a verb requires that the [Spec,CP] of its complement be a *wh*-phrase, it imposes no restriction on the tense of the embedded clause. This is illustrated in (i) and (ii) below. Show how your answer to exercise 6.5 can also account for this generalization.

(i) a. *John wondered/asked that Bill won.
b. John wondered/asked how Bill won.

(ii) a. John wondered/asked how Bill will win.
b. John wondered/asked how to win.

EXERCISE 6.7:
In section 6.2.1, we saw the effects of the periscope property induced by endocentricity in two different processes involving nominal domains: subject-verb agreement and selectional restrictions on complements. Reexamine these two processes assuming the DP structure in (47), showing what assumptions must be made in order for the DP-approach to capture the periscope property.

6.2.6. Success and Clouds

X'-Theory became one of the central modules of GB as it made it possible to completely dispense with PS rules. This was particularly noticeable in its successful utilization in the analysis of functional projections. Interestingly, however, progress in the description of specific syntactic constituents under X'-Theory ended up somewhat clouding this bright and blue sky.

Consider, for example, the assumption that XPs don't have multiple specifiers. The main motivation behind it was distributional in nature. Determiners were analyzed as [Spec,NP] and negation as [Spec,VP], for instance, because once they were added in the structure, no further nominal or verbal projection would obtain. Notice, however, that this continues to be true even in the structures in (50) below, where D and Neg are heads that respectively take NPs and VPs as complements. In other words, what was seen as a requirement on the number of specifiers turned out to be a reflex of the fact that D and Neg, like any other head, project when they take a complement.[14]

(50) a. $[_{DP}$ D NP]
b. $[_{NegP}$ Neg VP]

Intermediate vacuous projections illustrate a similar case. It's reasonable to say that a given head, say the verb *smiled*, projects a VP, given that it may occupy VP slots, as exemplified in (51) below. However, why should it also project an intermediate V'-projection?

(51) John $[_{VP}$ won the lottery] / $[_{VP}$ smiled].

Vacuous V'-projections were taken to be useful in the characterization of mono-argumental verbs as unaccusative or unergative (see section 3.4.2), as shown in (52) below. However, with the introduction of light verbs in the theory (see section 3.3.3), the distinction

[14] In fact, as Chomsky (1999: 39, n. 66) puts it, "[i]t is sometimes supposed that [the possibility of multiple specifiers] is a stipulation, but that is to mistake history for logic."

can be made with no resort to vacuous projections, as shown in (53) (see section 3.4.2). The automatic projection in three bar-levels therefore has lost much of its appeal in the verbal domain.

(52) a. unaccusative verbs: $[_{VP}$ V DP]
b. unergative verbs: $[_{VP}$ DP $[_{V'}$ V]]

(53) a. unaccusative verbs: $[_{VP}$ V DP]
b. unergative verbs: $[_{vP}$ DP $[_{v'}$ *v* $[_{VP}$ V]]]

The same can be said with respect to the nominal domain. Recall from our discussion in section 6.2.4 that the pronoun *one* appears to be a surrogate for N'-projections, explaining the adjunct-complement contrast between (54a) and (54b), for instance, which in turn requires that there be a vacuous N'-projection of *prince* in (54a).

(54) a. John likes this prince from Denmark and I like that one from France.
b. *John likes this prince of Denmark and I like that one of France.

Upon closer inspection, we can however see that this analysis crucially relies on two assumptions that now may not look as well grounded as before: first, that the determiner is the specifier of NP and second, that adjuncts are sisters of X' (the mapping principle in (36b)). As mentioned in section 6.2.5, it has now become a consensus that determiners take NPs as their complements. Besides, as discussed in chapter 3, there are strong reasons to believe that external arguments are generated within their theta domains (the Predicate-Internal Subject Hypothesis), more precisely, as sisters of an intermediate projection. Under this picture, a phrase such as (55a), for instance, should be represented along the lines of (55b), where *John* is generated in [Spec,NP] and moves to [Spec,DP].

(55) a. John's discussion of the paper
b. $[_{DP}$ John$_i$ $[_{D'}$'s $[_{NP}$ t$_i$ $[_{N'}$ discussion of the paper]]]]

The question now is how the interpretive component distinguishes adjuncts from external arguments if they may be both sisters of N'. One can't simply say that specifiers are different in that they close off projections, for the distributional facts that motivated this assumption have received alternative explanations on more reasonable grounds. As mentioned above, determiners establish the upper boundary of a nominal projection, for instance, not because they are specifiers but because the merger of D and NP yields DP. Furthermore, we may need more than one specifier at least for *v*Ps, if the computation of

locality is to be simplified, as discussed in section 5.4.2.1.

One possibility for accommodating these worries is to give up the mapping principle in (36b) (viz. that modifiers are adjuncts to X') and assume that modifiers are actually adjoined to XP. This in effect provides a much more transparent mapping from structure to interpretation: arguments are dominated by XP and adjuncts are adjoined to XP. Under this scenario, the contrast in (54) may be accounted for without resorting to vacuous N'-projections, if *one* is a phrasal pronoun and can't replace simple lexical items. That is, it can't target *prince* in (56b), but it can in (56a), because in the latter *prince* is also an NP.

(56) a. [$_{DP}$ this [$_{NP}$ [$_{NP}$ prince] [from Denmark]]
b. [$_{DP}$ this [$_{NP}$ prince of Denmark]]

The points above serve to show that much of the motivation for the initial postulates of standard X'-Theory got bleached as a deeper understanding of the structure of specific constituents was achieved. X'-Theory is therefore ripe for a minimalist evaluation. We should distinguish which of its properties reflect true properties of phrase structure in natural languages and investigate if such properties may follow from deeper features of the language faculty. This is the aim of next section.

EXERCISE 6.8:
Check if the analysis of (54) along the lines of (56) can also be extended to (i) without resorting to vacuous N′-projections or making any other amendments.

(i) John likes this prince from Denmark with the nasty temper, but I like that one with the sweet disposition.

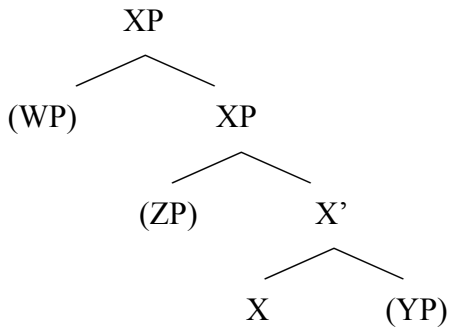
6.3. Bare Phrase Structure[15]

6.3.1. Functional Determination of Bar-Levels

Let's start our discussion with the qualm concerning bar-levels mentioned above. Take the X'-schema in (57), which incorporates the assumption made in section 6.2.6 that modifiers are adjoined to maximal projections.

[15] This section is primarily based on Chomsky (1995: sec. 4.3).

(57)



YP, ZP, and WP in (57) are, respectively, the complement, the specifier and an adjunct of the head X. Given that the actual realization of the projections of YP, ZP, and WP is regulated by other modules of the grammar (the Theta Criterion, for instance), they are in principle all optional. If none of them is realized, as illustrated by *John* in (58) below, then the three-bar level distinction seems to be motivated just on theory-internal grounds, for independent empirical motivation for it has considerably dimmed, as discussed in section 6.2.6. The schema in (57) also invites a related question: *why* is it that only maximal projections can function as complements, specifiers or modifiers?

(58) Mary saw [$_{NP}$ [$_{N'}$ [$_{N}$ John]]].

These sorts of worries may be seen as different facets of the fundamental question of how to interpret the claim that a phrase consists of parts with various bar-levels. Abstractly speaking, one can conceptualize the difference between X, X', and XP in two rather different ways. First, they may differ roughly in the way that a verb differs from a noun, that is, they have different intrinsic features. Alternatively, they can differ in the way that a subject differs from an object, namely, they differ in virtue of their relations with elements in their local environment, rather than inherently. On the first interpretation bar-levels are categorial features, on the second relational properties.

The three-bar level analysis of *John* in (58) is clearly based on a featural conception of phrase structure. To compare it with a relational way of conceptualizing projections, let's assume the definitions in (59)-(61) and examine the structure in (62), for instance.[16]

(59) *Minimal Projection: X^0*

A minimal projection is a lexical item selected from the numeration.

[16] These definitions are taken from Chomsky (1995: 242-243), who builds on work by Fukui (1986), Speas (1986), Oishi (1990), and Freidin (1992); the relational understanding of projection levels goes back to Muysken (1982). See also Chomsky (1998, 1999, 2000, 2001), Grohmann (2003b, 2004), Oishi (2003), and Rubin (2003, 2005) for further discussion.

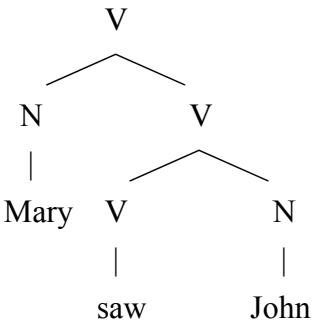
(60) *Maximal Projection: XP*

A maximal projection is a syntactic object that doesn't project.

(61) *Intermediate Projection: X'*

An intermediate projection is a syntactic object that is neither an X^0 nor an XP.

(62)



According to (59)-(61), *Mary*, *saw*, and *John* in (62) are each an X^0 (they are lexical items). The N-projection dominating *Mary* and the one dominating *John* are also interpreted as maximal projections since they don't project any further. The same can be said of the topmost V-projection; it's also a maximal projection. The V-projection exclusively dominating *saw* and *John*, on the other hand, is neither a minimal projection (it's not a lexical item), nor a maximal projection (it projects into another V-projection); hence, it's an intermediate projection. In other words, the definitions in (59)-(61) are also able to capture the fact that phrase structure may involve three levels of projection.

But it has additional advantages, as well. First, observe that there is simply no room for suspicious vacuous intermediate projections under this relational approach. In (62), for instance, the N-projection dominating *John* is both a minimal and maximal projection; hence, it can't be an intermediate projection, according to (61).

The relational approach also derives the claim that complements, modifiers, and specifiers are maximal from a more basic assumption: an expression E will establish a local grammatical relation (either Spec-head, modification, or complementation relation) with a given head H only if E is immediately contained within projections of H. Let's call this assumption the *Strong Endocentricity Thesis*. According to this thesis, heads actually project structure via the complement, modifier, and specifier relations.[17] Thus, by being immediately contained by a projection of X, a complement, a specifier, or an adjunct of X are necessarily maximal according to (60), because they don't project further. To put this in different words,

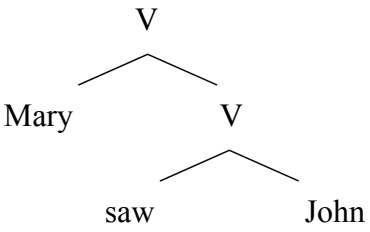
[17] This would make a lot of sense if these relations were ultimately discharged in a neo-Davidsonian manner with specifiers, complements, and modifiers anchored to the semantic values of heads (see Parsons 1990, Schein 1993, and Pietroski 2004). Thus, verbs denote events, complements and specifiers are thematic relations to events, and modifiers are properties of events.

the phrasal status of complements, specifiers, and adjuncts follows from the fact they enter into a local grammatical relation with a given head, and need not be independently postulated.

Bar-levels under the conception of phrase structure embodied in (59)-(61) are, therefore, not an inherent property of nodes in the tree, but rather the reflex of the position of a given node with respect to others. From a minimalist point of view, this is an interesting result. Recall that one of the features that ensure internal coherence to the minimalist project is the Inclusiveness Condition, which requires that LF objects be built from features of the lexical items in the numeration (see section 2.4). In order to encode maximal and intermediate projections, the featural approach to phrase structure in (57) tacitly relies on the theoretical primes expressed by the symbols " ^{0} ", " ' ", and " P " (as in N^0, N', and NP, for instance), which can't be construed as lexical features. By contrast, under the relational approach, the double role played by *John* as a head and as a phrase in (62), for instance, is captured without the postulation of non-lexical features.

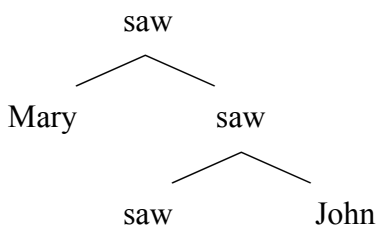
In fact, this observation may call into question the very distinction between terminal nodes and lexical items. In some sense, this distinction still keeps the same kind of redundancy perceived between PS rules and argument structure in the lexicon (see section 6.1). The lexical entry of *John*, for instance, arguably includes the information that *John* is a noun. That being so, what information does the categorical label N in (62) convey that *John* doesn't already convey? In other words, what piece of information would be lost if (62) were replaced by the structure in (63)?

(63)



One could say that this redundancy between terminal nodes and lexical items could be tolerated, for categorial nodes appear to be independently required to specify the properties of projections other than heads. In (63), for instance, we need to register that *[saw John]* is a verbal rather than a nominal constituent. It should be observed that what is actually required is a labeling mechanism to encode the relevant properties of nonminimal projections; however, this doesn't imply that this mechanism should necessarily involve categorial features. The structure in (64), for instance, works pretty well in the sense that it encodes the fact that the constituents *[saw John]* and *[Mary saw John]* are of the same relevant type as *saw*.

(64)



In the discussion that follows, we'll be assuming the projection-notation as in (64) instead of (62), guided by the intuition that we independently need lexical items, though we may not require categorial nodes.[18] But it's important to stress that the notation in (64) is just one way to encode the "projection" of the head. There are others conceivable that may as well do the job. We return to this issue below.

To summarize, the relational conception of bar-levels presents several advantages over a featural approach from a minimalist perspective: (i) it distinguishes different levels of projections in compliance with the Inclusiveness Condition; (ii) it doesn't have vacuous projections; (iii) it derives the fact that complements, specifiers, and adjuncts are maximal projections; and (iv) it allows the elimination of the distinction between terminal nodes and lexical items.

Assuming such a relational approach, we now turn to the mechanics of how phrase structure is built.

6.3.2. The Operation Merge

As discussed in section 2.3.2.1, one of the "big facts" about human languages is that sentences can be of arbitrary length and within GB, this recursion property was encoded at D-Structure. It was shown, however, that grammatical recursion is not inherently associated with DS. One can ensure recursion in a system that lacks DS by resorting to an operation that puts lexical items together in compliance with X'-Theory. We referred to this operation as *Merge*. Given that DS was abandoned for conceptual and empirical reasons (see section 2.3.2) and that much of the motivation for standard X'-Theory lost weight with later developments on phrase structure within GB (see section 6.2.6), it's now time to examine the details of the operation Merge.

Building a phrase involves at least three tasks: combining diverse elements, labeling the resulting combination, and imposing a linear order on the elements so combined. We'll leave the issue of linearization for chapter 7 and concentrate on how we combine elements and how we label the resultant combinations. For concreteness, take the derivation of the VP

[18] Some recent research in the framework of Distributed Morphology (see Halle and Marantz 1993, among others) pursues the idea that categorial information is defined relationally (see Marantz 1997 and subsequent work).

in (65) below. We know that *at John*, for instance, is a PP. But how can this be obtained from the independent lexical items *John* and *at*?

(65) [$_{VP}$ Mary [$_{V'}$ looked [$_{PP}$ at John]]]

Let's start by bringing the Strong Endocentricity Thesis into the picture. According to this thesis, local grammatical relations to a head X such Spec-head, complementation, and modification can only be established under projections of X (see section 6.3.1). Furthermore, the Extension Condition requires that such relations be established by targeting root syntactic objects (see section 2.3.2.3). That is, if the computational system establishes a head-complement relation between the lexical items *looked* and *at* by combining them, the lexical item *John* will not be able to later establish a head-complement relation with *at* by being combined with it. Finally, let's invoke the general (substantive) economy guidelines of Last Resort, according to which there are no superfluous steps in a derivation; in other words, every operation must have a purpose (see section 1.3). Thanks to this Last Resort property of syntactic computations, the combination of *Mary* and *John* as a syntactic object, for instance, is not an option because no local grammatical relation can be established between them.

With these considerations in the background, suppose that what the operation Merge does is combine elements to form a set out of them, as illustrated in (66).

(66) {at, John}
⇧
at $\Leftrightarrow_{Merge}$ John

The set in (66) should be a new syntactic object with subparts that are themselves syntactic objects. But this definitely can't be the whole story. *At* and *John* in {at, John} are in too symmetrical a relation with respect to each other (they are just members of a set) and such symmetry arguably can't ground the asymmetric relations of Spec-head, complementation, and modification. Once no local grammatical relation can be established, economy should prevent the formation of the set in (66) from taking place. Notice that this reasoning also explains why *at* and *John* in (66) can't both project: again, if that happens, there will be no asymmetry between these elements to anchor the Spec-head, complementation and modification relations. In other words, a local relation can be established only if there is some asymmetry between the members of the set and such asymmetry may be reached if one of them labels the resulting structure. This is what is meant by projection of a head.

The question then is which of the constituents projects. Of course, we know the result: the head projects. But the question is *why* this is so, that is, why can *John* not project in (66), for instance? Although at this point we can't go much beyond speculation, this seems to be

due to the fact that it's the head that has the information that it requires a Spec or a complement or is compatible with specific kinds modifiers — and not the opposite. Thus, it's a property of *at* in (66) that it requires a complement, but it not a property of *John* that it requires a head to be the complement of. If something along these lines is correct, a head may project as many times as it has specifications to be met.

To put this in general terms, in addition to providing information regarding the immediate constituents of the syntactic object resulting from merger, the system must also signal the relevant properties of the new object, whether it's a VP or a PP, for instance. In other words, we need to label the resulting object. If the potential relation between *at* and *John* is such that the former may take the latter as complement (and not the opposite), *at* projects by labeling the structure as in (67) below. According to the functional determination of bar-levels discussed in section 6.3.1, the resulting syntactic object in (67) is a maximal nonminimal projection, *John* is both minimal and maximal, and *at* is a minimal nonmaximal projection.

(67) {at, {at, John}}
⇧
at $\Leftrightarrow_{\text{Merge}}$ John

It's worth emphasizing that what is important here is that the constituent is labeled as having the relevant properties of its head and not how such labeling is annotated. We'll use the additional set notation in (67) because it's the one more commonly found in the literature, but it should be borne in mind that it would have been just as good for our purposes if *at* in (66) were underlined or received a star. This doesn't mean that the issue has no importance, but rather that at the moment it's not clear how exactly labeling should be technically implemented.

In fact, depending on its exact formulation, labeling may indeed be at odds with the Inclusiveness Condition in the sense that it may be adding features in the structure that may not be present in the numeration. In addition, given the Strong Endocentricity Hypothesis, the headness information encoded by a label is largely a function of the local grammatical relation being established (Spec-head, complementation, or modification). All of this brings the question of whether labels are really necessary.[19]

Even if the content of a label can be independently determined, it still arguable that labels are required in the system as optimal design features. Let's consider why by examining the derivational steps in (68) and (69) below. In (68), the PP of (67) merges with *looked*,

[19] The whole set of issues that surround labeling (whether labels can be derived, if they are even necessary, whether they violate the Inclusiveness Condition, etc.) is currently a major focus of research. For relevant discussion, see Uriagereka (2000a), Boeckx (2002), and Collins (2002).

which projects under a complementation relation, yielding a verbal projection. In (69), such verbal projection merges with *Mary* and another verbal projection is obtained, this time in virtue of a Spec-head relation.

(68) {looked, {looked, {at, {at, John}}}}
⇧
{at, {at, John}} ⇔$_{Merge}$ looked

(69) (looked, {Mary, {looked, {looked, {at, {at, John}}}}}
⇧
{looked, {looked, {at, {at, John}}}} ⇔$_{Merge}$ Mary

Notice that in both (68) and (69), the system doesn't need to compute the relations previously established in order to determine whether another local relation can be obtained. That is, by looking at the label of (67), the system has the information that this complex object is of a type that can enter into a complement relation with *looked*. Likewise, the label of the resulting object in (68) also allows the system to determine that such an object may enter into a local relation with *Mary*.

Now suppose we don't have labels. How does the system know that *Mary* may enter into a local relation with the relevant complex syntactic object in (68)? Or, putting this another way, if operations are carried out for some grammatical end, how does the system know that *Mary* can be merged with the label-less set {looked, {at, John}}? Apparently, by backtracking and determining first the kind of licensing/projection resulting from merging *at* and *John*, and then the kind of licensing/projection resulting from merging *looked* and the previously identified projection. This is obviously not very efficient, for the relation between *at* and *John*, for instance, in a sense gets repeatedly reestablished as more complex objects are formed. Besides, although such backtracking is manageable for simple objects such as the VP under discussion, recall that sentences in natural languages can have an unbounded number of recursions. Thus, the determination of the type of complex syntactic objects may be intractable if the constituent type is not encoded locally. Labels are in this regard a way of reducing the complexity of the task to a minimum: the system may simply check the label of a complex syntactic object to determine whether or not it can enter into a local relation with another syntactic object.

Consider in this regard an expression with a specifier, as well as a complement. What sort of locality could exist between *Mary* and *looked* in (65)/(69), for instance, without labeling? Note that the specifier and the head alone don't form a constituent and, assuming binary branching (see section 6.3.3 below), they are not immediate constituents of a larger object either. So, assuming that natural languages exploit at least two head-X relations (head-

specifier and head-modifier) in addition to head-complement, such relations can be locally coded only if we allow the head to label all of its projections. In effect, labeling not only allows head-to-head relations to be locally stated, but also makes it possible to locally state *several* grammatical relations to the head, and this perhaps explains *why* natural languages have labeled constituents where the label codes information of the head.

Assuming that this suggestion is on the right track, we can also appreciate the role of the Inclusiveness Condition in the reasoning. The Inclusiveness Condition is more of a meta-theoretical condition in that it sets up boundaries for minimalist analyses; in particular, a minimalist analysis should refrain from adding theoretical entities that can't be construed as features of the lexical items that feed the derivation. An unavoidable violation of Inclusiveness can, however, serve to illustrate deeper properties of the system as it strives for optimal design. In the case at hand, despite the fact that labels may be at odds with Inclusiveness, they may also be the optimal way of allowing multiple relations with a head and determining the properties of a complex syntactic object, *all in a local manner*.

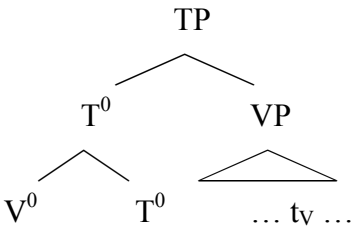
Let's recap. Minimalist commitments induce us to ask *why* each of the features found in phrase structure should hold true. What is it about language that gives it these features and not others? Why are constituents labeled? Why do heads project? These are tough questions and the suggestions above may well be on the wrong track. However, whatever the degree of our success in addressing these questions, it should not obscure the value and interest of the questions themselves. We noted in chapter 1 that one of the "big facts" about natural languages is that they have both words and phrases made up of words. Once this is noted, an operation like Merge, a grammatical operation that combines words into bigger and bigger units, is a natural feature of the system. What is less clear, however, is that labeling is also conceptually required given the "big facts" surveyed at the outset. Why do derived units need to have heads? We have suggested here that labeling is the optimal solution to a fact about words (they impose conditions on one another) and the basic relations among words (they enter into relations of specification, modification and complementation to heads). The Strong Endocentricity Thesis amounts to saying that there are local grammatical bounds on the influence words can lexically exert on one another. We have conjectured that this, in turn, is possibly related to issues of computational efficiency as it puts a very local bound on word-to-word interactions. This looks like a good design feature. If this is indeed the case, then labeling can be seen as a solution to the following problem: allow words to interact but in a tractable manner.

So far, we have discussed complex syntactic objects involving complements and specifiers. What about adjuncts? How can they be distinguished from specifiers once the system allows as many specifiers as Spec-head relations licensed by a given head?

How to deal with adjunctions is a vexed problem within generative grammar, one that has never been adequately resolved. The properties of adjuncts are quite different from those

of complements or specifiers. They don't enter into agreement relations, they appear to have different Case requirements from arguments, they are interpreted as conjuncts semantically, and they come in a very wide variety of category types. Thus, it's not clear what features, if any, are checked under merger by adjunction. Even more unclear is how exactly adjuncts syntactically relate to the elements that they modify. Recall that although forming a constituent with the modified projection, an adjunct is not dominated by the resulting syntactic object. This can be illustrated by head adjunction. Take V-to-I movement, for instance, now understood as V-to-T (section 5.4.2.3), which generates the structure in (70).

(70)



The verb and T (formerly Infl) in (70) clearly form a constituent, for T-to-C (formerly I-to-C) movement pied-pipes the verb adjoined to T. On the other hand, the moved verb can't be dominated by the structure resulting from adjunction; otherwise, it will fail to c-command its trace. That is why adjuncts are taken to be *contained* — not dominated — by the adjunction structure (see the discussion in section 5.4). Furthermore, we also want to say that adjunction of V to T doesn't disrupt the head-complement relation between T and VP. To borrow Haegeman's (1994) metaphor, being an adjunct is like being on a balcony: in some sense you are both inside and outside the apartment.

Translated in formal terms, being on a balcony amounts to saying that an adjunct doesn't change the label and bar-level of its target, though forming a constituent with it. To take a concrete example, if *hit John* in (71) is a non-minimal maximal projection labeled *hit*, the adjunction structure *hit John hard* in (72) should be characterized in the same way *and* — here comes the tricky part — preserve the previous bar-level specification about *hit John*; that is, *hit John* in (72) should remain a nonminimal maximal projection.

(71) {hit, {hit, John}}

(72) {?, {{hit, {hit, John}}, hard}}

If the label of (72) were just *hit*, the constituent in (71) would have projected, becoming an intermediate projection (a non-minimal non-maximal projection) with *hard* as its Spec. In other words, if the labels of adjunction structures were like the labels of projection structures, there would be no way to distinguish specifiers from adjuncts. We thus need another kind of

label to make the appropriate distinctions. (73) below, which revives the old notation of Chomsky-adjunction, may well serve these purposes.[20]

(73) {<hit, hit>, {{hit, {hit, John}}, hard}}

The pair <*hit, hit*> is taken to mean that the structure in (71), whose label is *hit*, determines the label of the structure in (73), but doesn't project. If (71) doesn't project in (73), it remains a non-minimal maximal projection, as desired.

Again, the notation above is nothing more than that: a notation. If it's not clear what the appropriate technical implementation of labeling under regular projection should be, labeling under adjunction gets even murkier.[21] However, the relevant questions about adjunction concern not the technology to get the empirical job done, but why it has the properties it has, rather than others. To date, no good answer has been forthcoming and we provide none here. For concreteness, we'll assume that the distinction between merger by projection and merger by adjunction in terms of their different labels reflects the different nature of the grammatical relations each operation establishes. In the sections that follow, we'll keep using the traditional bracket or tree notation, which are much easier to process visually, unless a substantial issue may be at stake.

To summarize, this section has reviewed the mechanics of phrase construction under the operation Merge. Merge is conceptually necessary given the obvious fact that sentences are composed of words and phrases. We have tried to provide some conceptual motivation for labeling as well. Whatever the insight gained by going down the road sketched above, many questions remain. For example, say we grant that labeling is in service of locality, *why* is it that we distinguish modifiers from specifiers from complements? Is this a semantic distinction projected into the syntax or is it an irreducibly syntactic categorization? Moreover, *why* are complements sisters of heads, while specifiers are sisters of intermediate projections, and not the opposite? What in the end distinguishes specifiers from modifiers? These are questions we have left to one side not because they are unimportant, but because we currently have no compelling suggestions, let alone answers. Many questions remain open that we are confident that readers of this book will one day successfully address.

---

[20] Whenever an expression is Chomsky-adjoined to an XP, the resultant structure bears the same label as the target of the adjunction. In (i), the adjunct *at six* is Chomsky-adjoined to the VP. Note that the constituent without *at six* is a VP as is the VP plus *at six*.

(i) John [$_{VP}$ [$_{VP}$ ate a bagel] [at six]]

[21] For technical definitions of dominance, containment, and c-command using the set notations such as (71) and (73), see Nunes and Thompson (1998).

EXERCISE 6.9:
Under traditional X'-Theory, the representation of multiple specifiers is indistinguishable from the notation of adjuncts to intermediate projections, as illustrated by the *v*P structure in (i), which is formed after the object moves to the outer [Spec,*v*P]. Provide the bare phrase structure representation of (i) and explain why it can't be confused with an adjunction structure.

(i) $[_{vP}$ OB $[_{v'}$ SU $[_{v}$ *v* $[_{VP}$ V t_{OB}]]]]

EXERCISE 6.10:
Chomsky (1995) has suggested that what prevents the projection of two merged elements in a range of cases is that their features are such that they can't form a composite label, if we understand a label as being composite in the sense of the union or intersection of the features of merged elements. For example, under the assumption that a verb has the set of features {+V, -N} and a noun has the set of features {-V,+N}, if a verb and a noun merge and both project, the intersection of their features would be the null set and the union would be the set {+V, -N, -V, +N}, with incompatible properties. Notice however that this suggestion opens the possibility that if features don't conflict, double projection should in principle be possible.

Having these observations as background, discuss if they could provide a viable way to explain periscope effects where a verb selects a noun buried within a DP-structure (see exercise 6.5). What would be the advantages and disadvantages of such an alternative analysis?

6.3.3. Revisiting the Properties of Phrase Structure

Leaving aside the issue of bar-levels, which was addressed in section 6.3.1, let's now reconsider the other properties of phrase structure discussed in section 6.2 from the point of view of the "bare" phrase structure approach reviewed in section 6.3.2. Let's start with binary branching.

As discussed in section 6.2.2, the fact that phrase structure in natural languages displays binary branching is reasonably well motivated on empirical grounds. That being so, we should now face the question of *why* the language faculty should restrict syntactic objects this way. Minimalism may offer a possible answer. We noted that in building a sentence, we begin with lexical atoms and combine them via Merge to form larger and larger units. What is the nature of Merge? *If* it's an operation that combines at most two elements per operational step, then the fact that there is binary branching reflects the basics of this operation. Is there some reason for why it should be that Merge involves at most two elements per step? Perhaps. Minimalism puts a premium on simple assumptions and asks that they be accorded methodological privilege in the sense of being shown to be inadequate

before replaced. This has a potential impact on the specifics of the Merge operation as follows: What is the simplest instance of merger? What are the minimal specifications for a Merge operation that respect the "big facts" we know about natural language?

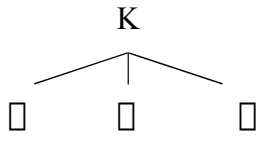
One thing we know is that Merge must be recursive. It can apply both to basic lexical expressions and to items that have themselves been formed via applications of Merge. This simply reflects the fact that there is no upper bound on sentence size. Second, it must be the case that Merge can combine *at least two* lexical items and form them into a constituent. We know this on two grounds. First, because this is the minimum required to get recursivity off the ground. We can't get larger and larger units unless we can repeatedly combine at least two units together again and again. Second, we have plenty of evidence that we need a two-place Merge operation to code some of the most basic facts, like the formation of unaccusative or transitive predicates, for instance. In other words, we need Merge to be able to form simple structures such as (74).

(74) [$_{VP}$ arrived he]

Now for a minimalist maneuver. It's clearly necessary that Merge be able to take at least two arguments; all things being equal, it would be nice (on methodological grounds) if we could strengthen this, so that it's also true that Merge take at most two arguments. In other words, seeing that two is the minimum required to meet the "big fact" of recursion in natural languages, it would be nice if it were the maximum as well. Note that this argument is very similar in form to the one that restricted levels to LF and PF (see chapter 2). We need at least these two to deal with sound/sign-meaning pairs; so, methodologically, we should try and make do with only these two. So too here: we need at least a two-place Merge operation; we should thus try and make do with at most a two-place Merge operation. That being so, binary branching follows straightforwardly. Consider the details.

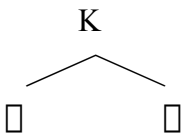
Suppose we take three lexical items, □, □, and □ out of a numeration and try to form a ternary branching structure K as illustrated in (75), by simultaneously merging them.

(75)

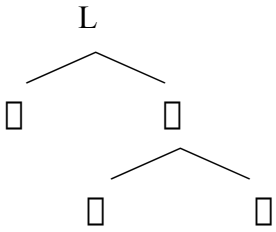


If Merge is a two-place operation, however, it can only manipulate two elements at a time, and a structure such as (75) can't be generated. Merge should first target two of the lexical items, say □ and □, forming K, and then combine K with the remaining lexical item, as shown in (76). But notice that only binary branching structures are yielded.

(76) a.



b.

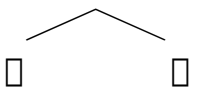


So it's perhaps plausible that binary branching is a reflection of the simplicity of language design: a two-place Merge operation is the minimum required to allow recursion (a "big fact"). Methodologically, it would be best if that were all that was required. Binary branching suggests that, at least in this respect, we live in the best of possible worlds. Pangloss be praised!

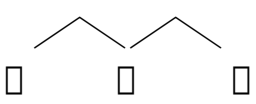
As for endocentricity (see section 6.2.1), it arguably follows from the interaction between Last Resort and the asymmetric nature of the local grammatical relations of head-complement, Spec-head, and modification. The Last Resort condition demands that every operation must serve a grammatical purpose. In the case at hand, if two elements are combined by Merge, either a head-complement, Spec-head, or modification relation must obtain in order for it to be licensed. Having one of these elements label the resulting structure creates an asymmetry between them that may ground these asymmetric relations. In fact, given the suggestion in section 6.2.2 regarding the inherent features of the head and their role in projection, the constituent containing the head will always project. Thus, any complex syntactic object will have its properties determined by one of its immediate constituents; that is, syntactic objects are always endocentric.

Finally, let's consider the singlemotherhood property, according to which a syntactic constituent can't have multiple mothers. Suppose, for instance, that after having merged and , forming K, we try to merge with , forming L, as illustrated in (77).

(77) a.



b.



The step illustrated in (77b) is however precluded by the Extension Condition, which requires that Merge target root syntactic objects. That is, once K is formed in (77a) its constituents are no longer available for further merger. Addition of □ in the structure will have to be through merger with K, as seen in (76b).

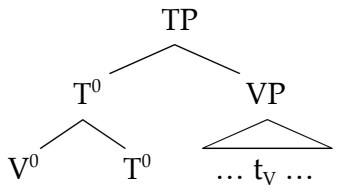
Notice that it also possible to conceive of the Extension Condition as a reflex of simplicity in the system. If only root syntactic objects are merged, as in (76), there is no change in constituency of the syntactic objects already built; only further layers of structures are added. Thus □, for example, is the sister of □ and is immediately dominated by K in both (76a) and (76b). Non-cyclic merger as in (77), by contrast, not only adds new structures, but also alters the constituency relations previously established; the sisterhood and immediate dominate relations involving □ are not the same in (77a) and (77b).

To summarize, the discussion above suggests that many of the properties of phrase structure in natural languages captured by X'-Theory can receive a more principled account if we assume a two-place structure building operation such as Merge, coupled with general minimalist principles of economy and methodological simplicity.

EXERCISE 6.11:
Discuss whether vacuous intermediate projections can be generated if structures are built by applications of Merge as described in section 6.2.2. In particular, what prevents an element from merging with itself?

EXERCISE 6.12:
Consider the structure in (i), where the verb has adjoined to T in violation of the Extension Condition. Lay out the problem and discuss possible scenarios under which such movement could comply with the Extension Condition.

(i)



6.4. The Operation Move and the Copy Theory

To this point, we have mainly discussed what we might term the "base configurations" of phrases, those formed by a series of Merge operations. Let's now address the question of how structures formed by movement are generated. Recall that within GB, movement proceeds by

filling empty positions projected at DS or adjoining to structures projected at DS, in accordance with the Structure Preservation Condition. In section 2.3.2, however, we saw not only that there is no need for all the structure building operations to precede movement, but also, and more importantly, that there is empirical evidence showing that structure building and movement operations should actually be interspersed (see sections 2.3.2.3 and 2.3.2.4). Having these considerations in mind, how should we understand the operation Move under the context of the bare phrase structure discussed in the previous sections?

Take the movement illustrated in (78) below, for instance. Part of the description of the movement in (78) is identical to the Merge operation depicted in (79). In both cases, the syntactic object labeled TP in (78a) and (79a) merges with another syntactic object, *a man* in (78b) and *there* in (79b), establishing a Spec-head relation and further projecting, thus becoming an intermediate projection.

(78) a. $[_{TP}$ T $[_{VP}$ arrived $[_{DP}$ a man]]]
b. $[_{TP}$ $[_{DP}$ a man $]_i$ $[_{T'}$ T $[_{VP}$ arrived t_i]]

(79) a. $[_{TP}$ T $[_{VP}$ arrived $[_{DP}$ a man]]]
b. $[_{TP}$ there $[_{T'}$ T $[_{VP}$ arrived $[_{DP}$ a man]]]

In other words, a movement operation appears to take Merge as one of its components.[22] Under this view, then it's not at all that surprising that merge and movement can alternate.

What are then the other components? Well, we have to say that somehow a trace is inserted in the object position of *arrived* in (78b) and this seems to put us in a corner. On the one hand, the empirical motivation for traces is overwhelming, as any cursory look in the GB literature can show. On the other hand, traces are by definition theoretical primes inserted in the course of the computation and are not present in the numeration, which is at odds with the Inclusiveness Condition.

Upon closer inspection, it may be that the size of the problem is actually related to the way in which it was presented. In fact, we don't have overwhelming evidence for *traces* and, for that matter, not even for *movement*. After all, nobody would bother to check if the speed of the DP in (78b) was within legal limits… In order words, what we actually have is an amazing set of facts that show that elements that appear in one position may get interpreted in a different position, the so-called *displacement property* of human languages (one of the "big facts"). The question that we have to address then is: can we account for this property within the bounds of minimalist desiderata?

The structure building part of movement, as we have seen, can be naturally captured by Merge. What we have to come up with is a solution for the "residue" of movement that is

[22] We address this Merge-over-Move preference in terms of economy in section 10.2.2.

congenial to Inclusiveness. A conceivable way to meet this requirement is to assume that a trace is actually a copy of the moved element.[23] As a copy, it's not a new theoretical primitive; rather, it is whatever the moved element is, namely, a syntactic object built based on features of the numeration. In other words, if traces are copies, Inclusiveness is pleased. Under this view, the movement depicted in (78) should actually proceed along the lines of (80), where the system makes a copy of *a man* and merges it with TP in (80a).

(80) a. [$_{TP}$ T [$_{VP}$ arrived [$_{DP}$ a man]]]
b. *Copy DP*: [$_{DP}$ a man]
c. *Merge DP and TP*: [$_{TP}$ [$_{DP}$ a man] [$_{T'}$ T [$_{VP}$ arrived [$_{DP}$ a man]]]]

Note that treating movement as simply the sequence of operations Copy and Merge leads us to expect that whatever principles apply when Merge alone (i.e. without Copy) obtains should also hold when movement (Copy and Merge) takes place. Consider, for example, the fact that Merge alone is subject to Last Resort, that is, it must serve some purpose. The same is observed with respect to movement. The merger in (80c), for instance, is licensed by Last Resort in that it allows the strong feature of T and the Case feature of both T and *a man* to be checked.

Now consider the issue of how the label of the constituent resulting from movement is determined. In particular, one wonders *why* the whole expression in (80c), for instance, is labeled TP, or put more generally, *why* the target of movement projects. Well, what else could it be? Recall that the Strong Endocentricity Thesis requires that in order for a local grammatical relation (Spec-head, head-complement, or head modifier) to be established, the head of the constituent must project. In the case of (80c), the checking relations mentioned above should take place under a Spec-head relation with T; hence, the head T projects and the resulting projection is a TP. According to a suggestion made in section 6.3.2, this is arguably related to the fact that it makes sense to say that T in (80a) needs a specifier, but it doesn't make any sense at all to say that *a man* in (80a) needs a head to be the specifier of. The important thing is that this is not different in essence from the (simple) merger in (79a): the Strong Endocentricity Thesis requires that T projects, as shown in (79b), in order for the Spec-head relation afforded by Merge to be established, and this is again arguably due to the fact that it's an inherent property of T that it requires a specifier, but it's not an inherent property of *there* that it requires a head to be the specifier of.

If we assume that the grammar only looks at what it has in deciding what to do next and doesn't "remember" earlier operations (in other words, if tree building is Markovian), then the fact that what is merged in movement is a copy is irrelevant to the merge operation applied. As far as the grammar is concerned, both applications of Merge are identical and so

[23] See Chomsky (1993) and Nunes (1995, 1999, 2001, 2004), among others.

should be subject to identical principles. Recall the suggestion in section 6.3.2 that labeling could be understood as a feature of optimal design of the system in that it allows structure building to work with the current information available, with no need to backtrack to earlier stages of phrase-structure building. That this line of reasoning also yields the desired empirical outcomes in the context of movement is quite pleasing and buttresses the assumption that movement is not a primitive operation, but the combination of the operations Copy and Merge.

At this point, the reader might however ask if this way of satisfying Inclusiveness is not too extravagant: the cost being the introduction of a new operation, Copy, and a new problem: why is the structure in (80c) not pronounced as (81), with the two links of the DP-chain phonetically realized or, to put in general terms, why can a trace not be phonetically realized?

(81) *A man arrived a man.

As it turns out, the alternative sketched above seems to be neither theoretically costly, nor empirically problematic. First, it seems that we independently need an operation like Copy.[24] To see this, let's examine what we mean when we say that we "take" an item from the lexicon. Clearly, this is not like taking a marble from a bag containing marbles. In the latter case, after taking the marble, the bag contains one less marble. In contrast, consider the (simplified) numeration that feeds (80) given in (82) below, for instance. When we say that we took those four items from the lexicon to form N in (82), we definitely don't mean that the lexicon has now shrunk and lost four items. Rather, what we are tacitly assuming is that numerations are formed by *copying* items from the lexicon. Thus, once the system independently needs such copying procedure, it could as well use it in the syntactic computation, as illustrated in (80).

(82) $N = \{arrived_1, a_1, man_1, T_1\}$

Second, we do indeed find instances where traces are pronounced, as illustrated in (83), where the intermediate traces of *met wie* 'with who' are realized.[25]

[24] See Hornstein (2001).

[25] The Afrikaans datum is taken from du Plessis (1977).

(83) *Afrikaans*

Met wie het jy nou weer gesê *met wie* het Sarie
with who have you now again said with who did Sarie
gedog *met wie* gaan Jan trou?
thought with who go Jan marry
'Who(m) did you say again that Sarie thought Jan is going to marry?'

Cases such as (83) suggest that the realization of copies is more a matter of the phonological component, rather than syntax *per se*. We'll return to this issue in chapter 7 and discuss a plausible explanation for why in general a chain doesn't surface will all of its links phonetically realized, as shown by (81).

Finally, by assuming that traces are actually copies, we may be able to account for binding facts within minimalist boundaries. Consider the sentence in (84), for instance, which should be represented as in (85), under the trace theory of movement.

(84) Which picture of himself did John see?

(85) [[which picture of himself]$_i$ did [John see t_i]]

In (85), the anaphor is not bound by *John*, but the sentence in (84) is nevertheless acceptable. In order to account for cases like this, GB requires additional provisos. For instance, Binding Theory should be checked at DS, prior to movement of *which picture of himself*, or at LF, after the moved element is "reconstructed," that is, put back in its original position; alternatively, the notion of binding should be modified in such a way that *John* in (85) gets to bind *himself* in virtue of its c-commanding the trace of the element containing *himself*.[26]

Leaving a more detailed discussion of Binding Theory to chapter 8 below, what is relevant for our purposes is that the copy theory accounts for (84), without extra machinery. As seen in (86), the copy of *himself* in the object position is appropriately bound by *John*, as desired.

(86) [[which picture of himself] did [John see [which picture of himself]]]

To summarize, the copy theory of movement seems to be a worth pursuing approach to the displacement property of human languages, in that it's tuned to minimalist worries and has some empirical bite both on the PF and LF sides. In the chapters that follow, we'll examine several other issues that also point to the conclusion that movement is just the result of applications of Copy and Merge.

[26] See Barss (1986), for instance, for a proposal along these lines.

EXERCISE 6.13:
In section 2.3.2.2, it was proposed that the TRAP, as defined in (i), would prevent a derivation of (ii) along the lines of (iii), with raising to a thematic position. As seen in this section, the copy theory takes movement to be the combination of the operations Copy and Merge. If this is so, how is the derivation in (ii) to be blocked? Or, to put it in more general terms, given the theoretical framework developed thus far, should it be blocked? If so, why?

(i) *Theta-Role Assignment Principle (TRAP)*
θ-roles can only be assigned under a Merge operation.

(ii) Mary hoped to kiss John.

(iii) [$Mary_i$ hoped [t_i to kiss John]]

6.5. Conclusion

Generative grammar has had many illuminating things to say about phrase structure. Minimalism has adopted the main results of these earlier approaches, largely encompassed by X'-Theory, and has tried to rationalize and explain the various properties of phrase structure on grounds of economy, simplicity, and optimal design. This, in turn, has led to very interesting questions and minimalism has raised them to prominence even if it has not yet offered fully compelling answers

This chapter has argued in particular that the key properties of phrase structure follow from the inner workings of the structure building operation Merge, coupled with general minimalist conditions, yielding what was referred to as a bare phrase structure. In addition, it was proposed that Move is not a primitive operation of the system, but the result of the interaction between the operations Copy and Merge (the copy theory of movement). Recent developments in the theory of movement strengthen the theoretical appeal of such an approach with very interesting empirical evidence, as we'll see in the chapters that follow.

Recap of Terminology

The terminology introduced in this chapter is summarized in (□), while (A)-(C) lists new principles we are going to adopt in the remainder of the book.

(□) *Bare Phrase Structure:* extension from X′-Theory which replaced PS rules
endocentricity — binary branching — singlemotherhood — bar-levels
[functional determination of bar-levels vs. three-bar level system]
specifiers vs. adjuncts — functional heads/projections
Copy Theory: Merge — Copy — Move

(A) *Minimal Projection:* X^0 (= (59))
A minimal projection is a lexical item selected from the numeration.

(B) *Maximal Projection: XP* (= (60))
A maximal projection is a syntactic object that doesn't project.

(C) *Intermediate Projection: X′* (= (61))
An intermediate projection is a syntactic object that is neither an X^0 nor an XP.