Relative Clauses, Licensing, and the Nature of the Derivation

David Lebeaux
University of Maryland

RELATIVE CLAUSES, LICENSING, AND THE NATURE OF THE DERIVATION

DAVID LEBEAUX

UNIVERSITY OF MARYLAND

1.0 INTRODUCTION.

In this paper, I attempt to give empirical justification for the view that the phrase marker is not generated as is in the base, but rather is composed out of smaller units, in the course of the derivation. Thus distinct licensing conditions are modelled derivationally. Move-a may be interspersed with these phrase structure composition operations freely. The primitives which pick out these seperately generated substructures (in the base) are precisely the primitives of GB-theory: Case theory, theta theory, the Projection Principle, and so on. Moreover (I suggest) each primitive (ur-) structure is pure instantiation of a single licensing primitive. Two proposals in particular are assayed here. First, since by the Projection Principle, arguments but not adjuncts are required in the base, adjuncts may be added later, adjoined-in in the course of the derivation. Second, and more radically (Lebeaux, 1988), Case and theta theory do not simply mutually describe an existant phrase marker, but pick out distinct subrepresentations which are merged in the course of the derivation. This latter operation I call Merger (or Project-a).

The paper here is a continuation of work in Lebeaux (1988). That work attempted to develop a syntactic and language acquisition theory side-by- side. I will concentrate here on the syntactic side of that theory, developing it further, but would like to highlight a general reason from language acquisition for assuming that the phrase structure component is organized in such a way: so that the "start" structures are simple. This is the following. The child may be assumed to start from a universal linguistic specification, Universal Grammar (Chomsky, 1965), which has at least as much information in it as the final, in terms of universal specifications. This was formalized in Lebeaux (1988) by modelling the acquisition sequence as the removal of a certain sort of information, parentheses representing parameter-settings in a bleeding relation, in a representation. However, along with this (decreasing) universal specification, the child has individual analyses of each sentence that s/he hears. These individual phrase structure analyses come at some computational cost: i.e., a more complex phrase structure must be more complex to compute than a simpler one. Moreover, unlike the universal specification, it seems fairly clear that this specification of structures should get more complex over time, as the child's computational capacity increases. This can be seen most graphically, for example, in the change from telegraphic speech to nontelegraphic--i.e. from a speech in which CC functional elements are absent to one in which they are not. The simplification here seems principled. More subtly, different licensing conditions appear to come in at later times than others: for example, conjunction appears later than thematic complementation in the child's grammar; prepositional adjuncts

may come in later, and with different properties, than prepositional complements (Brennan, 1989); and relative clauses seem to be interpreted initially as co-ordinate, rather than NP-adjoined (Tavakolian, 1978) in initial stages[1]. These sorts of simplifications or differences between the child's grammar and the adult's are in part parametric, but seem to implicate some notion of simplicity as well: the problem, for the acquisition theorist, is to account for the simpler grammar in a way which does not seem to make it more complex: which would be the case, for example, if one assumed a deletion transformation. The hypothesis of substructures suggests a general solution to this problem: if composed substructures are what make up a phrase marker, and if the operations making up the substructure have some (fairly transparent) computational reflex (as in the theory of Miller and Chomsky, 1963), then the initial grammar will be computationally less complex, as it should be.

The above considerations are computational-language acquisition for the assumption of substructures. For the remainder of this paper, however, I will concentrate on syntactic reasons for such an assumption, particularly with respect to adjunction of adjuncts.

2.0 ANTI-RECONSTRUCTION EFFECTS.

For a number of processes in the grammar, it appears that interpretation of a moved element is most readily captured if it is "placed back" (Reconstructed) into its original pre-movement site. For example, in 1 a) the reflexive himself is outside of the c-command domain of name John at s-structure; it would be properly c-commanded if construed at the trace site. Similarly, the bound pronoun in 1 b) is outside the domain of the quantifier: by interpreting the phrase containing the pronoun as if it were in its trace site, the appropriate c-command relations are observed.

1 a) Which pictures of himself$_i$ does John$_i$ like t?
b) Which of his parents do you think that every man likes t best?

Along with these reconstruction type effects, implicating the DS (or pre-wh movement) position as the relevant one, a rather remarkable set of data is observed in van Riemsdijk and Williams (1981), with respect to Condition C.

2 a) *He$_i$ likes those pictures of John$_i$.
b) *He$_i$ likes those pictures that John$_i$ took.
c) ?*Which pictures of John$_i$ does he$_i$ like t?
d) Which pictures that John$_i$ took does he$_i$ like t?

The effect in 2), the so-called "anti-Reconstruction" effect creates a paradox that the placement of Condition C at some particular level (e.g. DS, or NP-Structure, or S-structure) cannot solve. If Condition C were placed at DS (or some pre-wh-movement level), then both 2 c) and d) should be bad, like their correspondants, a) and b). If Condition C were placed at SS, then both 2 c) and d) should be good. But they divide: 2 c) is ungrammatical, while 2 d) is not. (By "Condition C" here, I mean the disallowance of a name being c-commanded by a co-indexed pronoun, cf. Lasnik, 1986.[2])

Van Riemsdijk and Williams (1981) suggest that the criterial difference between 2 c) and d) has to do with degree of embedding of the name (when it is fronted). But a larger examination of the data shows that the contrast is more interesting yet. In fact, it implicates the argument/adjunct distinction, in an unusual way. If the fronted name is contained in an adjunct, then the resultant is good, even if this adjunct is embedded in an NP which is itself a complement of the main verb (2 d)); if the fronted name is contained in a complement (within the fronted NP), then the resultant is bad: it correctly triggers a Condition C violation (2 c)). That it is the argument/adjunct distinction that matters, rather than degree of embedding, can be seen by holding embedding constant, and varying the arg/adj relation. In 3), there is a clear-cut difference depending on whether the name is in a claim-complement or a relative clause, though both involve CP embedding. This contrast of course does not hold for the base form.

3 a) *He$_i$ denied the claim that John$_i$ made.
b) *He$_i$ denied the claim that John$_i$ likes Mary.
c) Which claim that John$_i$ made did he$_i$ later deny t?
d) *Whose claim that John$_i$ likes Mary did he$_i$ deny t?

Second, holding the PP embedding constant, the structures again differ, according to whether the PP is selected by the N head, or is a N'' adjunct.

4 a)?*Which pictures of John$_i$ did he$_i$ like t?
b) Which pictures near John$_i$ did he$_i$ look at t?

The same contrast appears in process vs. result nominals, even more strongly, where in the former case the complement of the N truly does have an argument-taking role (Lebeaux, 1988).

5 a)*Whose examination of John$_i$ did he$_i$ fear t?
b) Which examination near John$_i$ did he$_i$ peak at t?

Finally, we note that the post-head genitive patterns with the adjunct, not the argument. This is to be expected, since 1) the post-head genitive has a loose, "relation R" relation to the head, unlike that of the complement, 2) it is attached after subcategorized complements, and 3) it is an island for extraction.

6 a)?*Which pictures of John$_i$ does he$_i$ like t?
b) Which pictures of John's$_i$ does he$_i$ like t?

The lack of Condition C effects in partitive-type constructions suggests that they pattern with the post-head genitive: they appear to be acting as adjunctual type elements in this construction. Note the loose relation to the (null) head, and the lack of extractibility.

7) Which (ones) of John's$_i$ pictures does he$_i$ like t?
8)*Whose pictures does he like which of t?

3.0 ADJUNCT-FILTERING AND THE PROJECTION PRINCIPLE.

The foregoing suggested that the following generalization holds for the disappearance of Condition C effects in fronted constituents (Lebeaux, 1988).

9) Anti-Reconstruction Effects
Condition C effects are abrogated, when the fronted name is contained in an adjunct.

(I will return below to a modulation of 9) necessary to explain A-chains.)
It has been shown that the generalization in 9) covers 5 subcases.

10 a) RC vs. claim-type complement
b) locative adjunct vs. picture-of complement
c) locative adjunct vs. process complement
d) postposed genitive
e) partitive

Yet the generalization in 9) is an amazing one. While the vast bulk of the syntactic literature has been confined to discussing the lack of extractability out of adjunct phrases, the CED type effects of Huang (1982),however reconstructed, the effects in 10) do not have to do with extractability, but with Condition C, a totally different type of long distance relation. Second, the abrogation of Condition C effects does not hold for all adjuncts, but only those

that have been fronted. Thus any explanation which, for example, had adjuncts as not part of the representation at all, would not explain the fact that Condition C effects are not abrogated when the adjunct is in its DS-position. Note as well that the chain here appears to not be acting as an equivalence class of positions. Finally, the anti-Reconstruction effects appear to be robust, over a wide variety of sentence types. How might this be accounted for?

I suggested above that from the point of view of initial grammatical competence, it would make sense for the grammar to specify substructures, simpler structures which could be used in initial stages of languge acquisition. This would reduce the computational load on the child. The computational load might then be viewed as a function of the complexity of the individual units, and the complexity of the operation of composition. Let us suppose that such substructures exist, and, further, that they are picked out by the grammar, rather than being (say) late simplifications by the parsing or production device. The "degradation" of the system in initial stages, in the sense of the differnce between that system and that of the adult, would not be a true degradation, but rather a result of the child having recourse to a simpler system:simpler in the sense that fewer units would be composed, or the composition operation would be the default. But the grammar itself would be well-formed, just partial.

Such a view entails that the grammar (not the parser) pick out well-formed subparts. What would these be? The proposal here is that they are along the lines of distinct licensing relations, and the derivation itself models the distinct licensing relations (in distinct ways).

Consider the sentence in 11). Taking the verbal head (or verb + Infl) for free, we may view the phrase marker as the result of the various grammatical well-formedness rules building the structure out from it, by virtue of their licensing character (Speas, 1989). In particular, and simplifying, X' theory builds the verb up to VP; the Projection Principle and the theta criterion forces a subject and object into the representation; Case theory forces the object to be an NP (in the case where it is); and X' theory re-applies to expand the NP subject and NP object down to the N node. (I ignore for now the DP hypothesis, returning to it, crucially, later.) The subrepresentation of 11) gotten in this way--i.e. gotten by applying the obligatory modules of DS, the Projection Principle and X' theory--is shown in 12).

11) The man saw the woman near the bridge.

12) (($_{np}$ ($_{n'}$ ($_{n}$ man))) ($_{v'}$ saw ($_{np}$ ($_{n'}$ ($_{n}$ woman)))

Evidentally, a substructure is created in which 1) the determiner has dropped out, and 2) the adjunct has dropped out. Let us temporarily just stipulate that the determiner is part of the basic representation (this might be gotten by assuming DPs). Then the licensing operations of the Projection Principle and X' theory, applying outward from the verbal head to the structure in 11) would generate the following substructure, filtering out the adjunct.

13) The man saw the woman.

Similarly, in 14), starting with the verbal head + INFL complex for free, and freely applying 1) X' Principles, and 2) the Projection Principle and the theta criterion--that is, applying the modules which must apply in the base-- the substructure 14 b) would be picked out from the entire structure in 14 a).

14 a) John saw the woman who he knew.
b) John saw the woman.

The retrieval of the substructure in 14 b) from the full structure in 14 a) may be considered to be the consequence of a conceptual operation by which an "argument skeleton" (Lebeaux, 1988) is retrieved: that is, the full structure in 14 a) breaks up into two argument skeletons, subparts which are pure representations of the "argument-of" relation, and between which an "adjunct-of" relation holds. This is shown in 15).

15 a) John saw the woman who he knew.
b) s_1: John saw the woman
s_2: who he knew

The full sentence in 16) breaks up into two argument skeletons, each of which is a pure instantiation of the "argument-of" relation, and between which an "adjunct-of" relation holds.

16 a) John left, because he wanted to.
b) s_1: John left
s_2: because he wanted to

Conceptually, then, in 14)-16) I have filtered the full phrase marker through the Projection Principle, allowing it to break the phrase marker into subparts (each subpart is that over which the Projection Principle obligatorily holds).

Consider now those examples in which an object nominal head does itself take a complement.

17) John denied the claim that Bill was a werewolf.
18) I feared the examination of the students.
19) Steve liked the picture of Fred.

Starting with the main verb + INFL FN, e.g. deny, the verb will require the object NP and the subject NP nodes by the Projection Principle. These nodes will expand to the N head by X' theory. However, in this case, the nominal itself (claim, examination, picture) selects an object, and theta marks it. That is, an "argument-of" relation holds between the nominal head and its complement. By the Projection Principle, this argument of relation must be present whenever the head is. But the nominal head itself is required to be present by the main verb. So the Projection Principle + theta criterion, X' principles, and the Proj Pr + theta criterion applies over again, forcing the complement of the nominal to be in the single main structure as well. That is, starting with the main verb, the entire structure is "chained" or "netted" into the representation in the case where the nominal head takes a complement. These are in contrast to those phrase markers in which there is a true adjunct in the representation.

20) John saw Mary.
s_1: John saw Mary

21) John denied the claim that Bill was a werewolf.
s_1: John denied the claim that Bill was a werewolf.

22) I saw the picture near Mary.
s_1: I saw the picture
s_2: near Mary

23) John left because he wanted to.
s_1: John left
s_2: because he wanted to

24) I feared the examination of the students.
s_1: I feared the examination of the students.

25) I denied the claim that Bill made.
s_1: I denied the claim
s_2: that Bill made

In 22), 23), and 25) there is an element bearing an adjunct relation. As such, it starts off in a separate structure, each individual s_i being a pure representation of the argument-of relation. In 20), 21), and 24), all elements are in an argument-of relation, either directly or recursively. Thus the RC of claim is not part of the main structure, while the complement clause of claim is (21) vs. 25)). The substructures are generated out from some starting head element (here a verb + INFL) along the following lines:

26) Apply the transitive closure of X' principles, the Projection Principle, and the theta criterion.

This will break each tree into the relevant subtrees. Each is, roughly, a pure representation of the "argument-of" relation. Since we may conceive of X' Principles, the Proj Pr, and so on applying immediately in the base (as principles), there need be no "derivational lag" in forming the structure over the verbal head: rather, each of the substructures may be thought of as instantaneously there. [3]

4.0 ADJOIN-A, CONDITION C, AND THE FORMATION OF RELATIVE CLAUSES.

Suppose that we do not take the conceptual filtering operation above to be actual not conceptual, so that the units involved are actual grammatical (sub-)structures. Second, suppose that the direction is the reverse of that given above, so that the units are not filtered out of a complete structure, but rather are the initial primitive units that go into the making of the full structure. We arrive at the position that the initial structures making up the full structures in 20)-25) are those immediately below them. These structures must be composed in the course of the derivation. Note that the adjunct-of relation is modelled differently--quite differently--than the argument-of relation. The latter is modelled by each individual substructure. The latter is modelled by the relation composing them in the course of the derivation. Let us call this operation Adjoin-a. The derivation as a whole then looks like this. (See, also, Joshi and Kroch, 1985.)

27) DS (substructures respecting arg-of relation)

Move-a, Adjoin-a

SS

We may now explain the anti-Reconstruction effects. Suppose that the operation Adjoin-a exists, composing substructures, and applies freely within the derivation, along with Move-a: the minimal assumption. Of course, each individual substructure must obey the Projection Principle, so Adjoin-a only applies between substructures which are in an adjunct-of relation: those which are in an argument-of relation are forced to be in the same structure at all levels of representation, by the Projection Principle. The former case would hold, then, for relative clauses (i.e. Adjoin-a would take place), while the latter would hold for claim-complements. Suppose now that we take a derivation in which Move-a has applied, to some relevant NP. There is only one derivation for the single tree structure when it is solely a reflection of the argument-of relation (e.g. claim-complements): that in which movement has applied. On the other hand, two potential derivations underlie the DS's which contain an adjunct, say an adjunct relative clause (and which are thus a collection of two trees). In one derivation, the relative will be added after Move-a has applied. In the other derivation, it will be added before Move-a has applied. As above, we will call the operation composing the substructures, Adjoin-a. Some sample derivations are below.

28) Base structures:
a) He denied the claim that John liked Mary. (complement case)
s_1: He denied the claim that John liked Mary
b) He later denied the claim that John made. (RC case)

s^1: He later denied the claim
s_2: that John made

29) Derivations:
a) DS: *He_i denied whose claim that $John_i$ liked Mary. --Move-a-->
SS: *Whose claim that $John_i$ liked Mary did he_i deny t?

b) DS: s_1: He_i later denied which claim —Move-a→
s_2: that $John_i$ made

s_1: Which claim did he_i deny t] —Adjoin-a→
s_2: that $John_i$ made

SS: Which claim that $John_i$ made did he_i later deny t?

DS: s_1: He_i later denied which claim] —Adjoin-a→
s_2: that $John_i$ made

—Move-a→
*He_i later denied which claim that $John_i$ made

*Which claim that $John_i$ made did he_i later deny t?

Recall that the anti-Reconstruction effects appear only when the name is embedded in an adjunct and this adjunct is fronted (perhaps as in a relative clause within an NP). A maximally simple account for this is now available. The relevant contrast is between the derivation in 29 a) and the first derivation in 29 b). Assume that Condition C applies throughout the derivation, over purely structurally defined c-command. Assume further that distinct substructures have no c-command relation holding between them (at the point that they are distinct). Then in 29 a), by the Projection Principle, the argument CP of claim must be present at all levels of representation. In particular, it must be present in the base, and prior to wh-movement. But then the name John within that CP is present. But since Condition C applies throughout the derivation, this means that Condition C will apply, and there is no reading under which John is coreferent with the pronoun he, which c-commands it at a pre-wh-movement level of representation. This is the expected enforcement of Condition C effects. The truly interesting case is that in 29 b). Here the construction is unexpectedly good: coreference is possible in "Which claim that John made did he later deny t?". This is precisely what is expected under the first derivation in 29 b). By the Projection Principle, the adjunct need not be part of the original s_1. Move-a may apply to s_1, moving the wh-phrase to the front. At that point, Adjoin-a may apply, adjoining the adjunct containing the coreferent name to the already moved wh-phrase. However, at no time will the coreferent name John be within the c-command domain of the pronoun he. That is, even though Condition C applies throughout the derivation (over pure structually defined c-command), "Which claim that $John_i$ made did he_i later deny t?" will escape a Condition C violation, because at the point that the wh-phrase is within the c-command domain of the pronoun, the relative clause containing the name is not part of the representation. It is adjoined-in later, after the movement has applied removing the wh-phrase from the c-command domain of the pronoun. Thus the structural description triggering Condition C effects is never satisfied, and the structure itself is predicted to be good: as it is.

Note that both of the conditions necessary for the abrogation of Condition C effects, that the name is contained in an adjunct, and that the adjunct is fronted over the coreferent pronoun are expected, given this account: the adjunctual character is necessary, because only then will the name not be within the c-command domain of the pronoun prior to wh-movement. The fact that the adjunct has been is crucial, since only then will it escape the c-command domain at s-structure, assuming that Condition C applies everywhere. This latter show that the mere statement of adjunctual relations being somehow "different" will not by itself be sufficient.

Finally, note that the presence of a derivation which blocks--the second in 29 b)--is irrelevant, since it is sufficient to have one derivation which goes through.

The other four types of anti-Reconstruction effects given in 10) pattern identically. For example, the locative vs. the complement PP (in NPs) would have different derivations, with only the latter obligatorily present in the base.

30) DS: *He$_i$ liked which picture of John$_i$ ----------Move-a-->
SS: *Which picture of John$_i$ did he$_i$ like t?

31) DS: s_1: He$_i$ liked which picture ——Move-a——>
s_2: near John$_i$

s_1: Which picture did he$_i$ like t?]—Adjoin-a—
s_2: near John$_i$

SS: Which picture near John$_i$ did he$_i$ like t?

In the derivation in 31), the name escapes the c-command domain of the pronoun, because it is part of an adjunct which may be added later. This is impossible in the derivation in 30), where by the transitive closure of the Projection Principle and X' theory, the complement of picture is forced to be in the initial representation. The same holds for all the other cases.

4.1 WHERE DO THE BINDING CONDITIONS APPLY?: THE INTERACTION OF CONDITION -C AND QUANTIFICATIONAL BINDING.

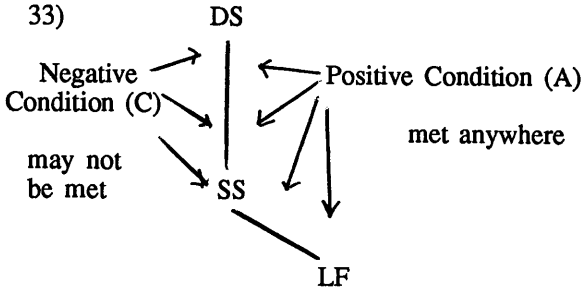
In the preceding section, I assumed that: i) the Projection Principle holds, and ii) the Binding Conditions, and in particular Condition C, applied throughout the derivation. Note that if one takes the Projection Principle, and only the Projection Principle, to characterize the obligatory licensing relation in the base, then it would actually take an additional stipulation to force adjuncts such as relative clauses to be present there. From these two maximally simple assumptions, the anti-Reconstruction effects followed.

By the contention that the Binding Theory applies throughout the derivation, I mean the following: that the positive indexing condition, Condition A, applies throughout the derivation, attempting to index an (unindexed) anaphor with an antecedent in an appropriate structural configuration (Belletti and Rizzi, 1986, Lebeaux, 1988).[4] If the anaphor is appropriately indexed at LF, the derivation holds: otherwise, the structure is ungrammatical. On the other hand, a negative indexing condition such as Condition C again applies throughout the derivation, and if its structural description is met at any point in the derivation, then the sentence is assigned a *, and the derivation blocks (the * may not be removed). This proposed universal, called a Positive Conditions Somewhere/Negative Condition Nowhere view, has the following form (Lebeaux, 1988; Kayne, 1989; Belletti and Rizzi, 1986 ; Burzio, 1989) (These views are similar: the one below, and in Lebeaux, 1988, is, I think, the strongest and most explicit.)

32) Conditions on Indexing (UG):
a) Positive conditions on indexing must be met somewhere in the derivation.
b) Negative conditions on indexing may not be met anywhere in the derivation.

In Lebeaux (1988), I took 32) to apply to Binding Theory and Control. This is shown in 33).

33)



LICENSING AND THE NATURE OF THE DERIVATION

Let us now consider an empirical extension. I have suggested above that the binding principles apply throughout the derivation (over pure structurally defined c-command)[5]. Assume that the quantificational binding does likewise, registering the appropriate binding relations throughout the derivation, over pure structurally defined c-command, rather than at some post-LF or post-SS reconstruction level. Consider now, the set of constructions first studied extensively by Elisabet Engdahl (Engdahl, 1980, 1986), in which a bound pronoun has been moved outside of the scope domain of a quantifier which binds it.

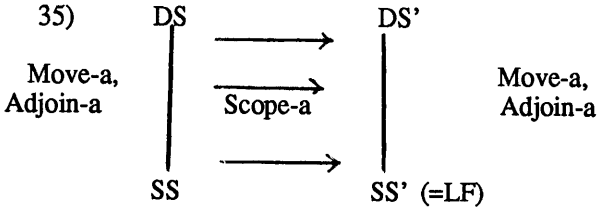
34 a) Which of his parents do you believe that every man likes t best?
b) How many of his papers do you think that every linguist wishes to rewrite t?

In each case, the pronoun is bound, yet it is one clause above the quantifier that it is bound to, and thus outside its scope domain. In general, such data suggest that it is the wh-trace site, rather than the s-structure configuation, which determines the appropriate representation feeding the semantics, for moved bound pronouns.

Such data have been treated in the literature in one of two ways. The first, the predominant one in the literature, has allowed the pronoun (and any other relevant material) to "reconstruct" back into the trace site, at some post-LF or post-SS level (e.g., Barss, 1985, Hornstein, 198 , Williams, 1987, Chomsky, 1985). Such reconstruction may involve actual movement back into the trace site, or, in another version, may be kept in place and simply read as if it were back in the trace site (e.g., Barss, 1985). Let us call the latter proposal, quasi-Reconstruction. In either case, this "reflects back" some aspect of the previous derivation, into the post-SS or post-LF part of the derivation, undoing the relevant operation (or reading it as if undone), in order to recover the original relations.

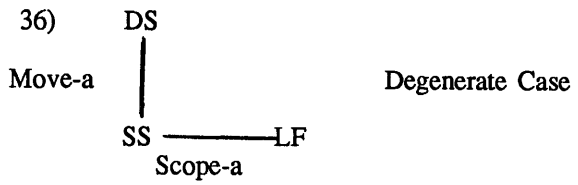
The alternate possibility, which I would like to explore and ultimately adopt, is that scope relations are read off of the pre-movemet structure. This view is more in line with an architectural position, that semantic interpretation applies throughout the derivation, rather than being read exclusively off of s-structure. A precusor, in this regard, is Jackendoff (1972): Bach (1977), Marantz (1984), and earlier versions of the Projection Rules of Katz and Fodor (1963) would fall into the same camp. The position of the interpretive component in the grammar would then be the following.

35)



Scope-a here is simply Move-a, with the additional properties that scope taking operations seem to have. Here, Move-a and Adjoin-a apply to the derivation from DS to SS, composing substructures and initiating movement. The linked interpretive structure has identical operations, with the additional proviso that scope relations may be determined at any point in the derivation (by Scope-a). In the degenerate case, namely where there is no movement involving elements taking mutual scope, the grammar will reduce to that of Chomsky and Lasnik (1977).

36)



In other cases, in particular in those in which a quantifier or wh-phrase has been moved up the tree and yet takes scope or gets bound below, the architecture in 35) has advantages over the classical Chomsky-Lasnik one, I believe. Reconstruction, in this regard, can be viewed as a measure of the degree to which the classic Chomsky-Lasnik architecture, with a single level SS feeding LF, fails to account for intermediate structures being the appropriate one determining scope relations. Reconstruction itself is a way of "leaking" information from such levels back into the post-SS derivation. In an architecture like that of Jackendoff (1972), such leaking would presumably not be necessary, since the intermediate levels would directly feed the semantics, and register the scope relations directly. Of course, this leaves a very large technical question of precisely how this would be done (in current theory). See Lebeaux (in preparation) for further discussion. (A less radical version of the dual derivation idea, brought to my attention by Juan Uriagereka, would be to assume that it is end-of-cycle structures which feed interpretation. This would get around the fact that LF movement isn't visible, without introducing an entire other derivation.)

Putting aside the exact mechanism for now, let us assume that scope relations are read directly off of the pre-wh-movement structure, registering the appropriate scope relations. This will work appropriately for the Engdahl-type sentences, repeated below.

37) You think that every man likes which of his parents best.
(scope relations of every and his read off prior to wh-movement)

Consider now the case of adjuncts. I have suggested earlier, that adjuncts do not form a unitary representation with the main clause in the base, but are part of a collection of substructures which are adjoined in the course of the derivation, by Adjoin-a.

38) s1: you saw the picture] —Adjoin-a→ You saw the picture
s_2: near him near him

What about if the adjunct contains a bound pronoun, and is part of a wh-noun phrase which will turn out fronted? In this case, the adjoining of the adjunct must take early in the derivation, prior to the point at which the wh-phrase which it attaches to is fronted over the binding quantificational phrase. (cf. This would be the second derivation in 29 b) above.)

39) s_1: you think every man saw which picture] —Adjoin-a→
s_2: near him

—Move-a→
You think every man saw which picture near him

Which picture near him do you think that every man saw t?

Where in 39), the correct c-command relations are obtained by adjoining the adjunct early, prior to movement.

But now consider the following situation. For certain adjunct cases, those given for anti-Reconstructin effects, the adjunct must be adjoined lated, after it has moved, to escape a Condition C violation. For other cases, those given immediately above, the adjunct must be adjoined early in the derivation in order for a pronoun to "catch" the appropriate c-command relation with a quantifier. But this then makes a subtle prediction: that if an adjunct contains both a pronoun to be quantificationally bound, and a name which is to be conjoint to a c-commanding pronoun, then the sentence will be grammatical or not, depending on the ordering of the quantifier and the pronoun in the major tree. The relevant abstract structural configurations are shown in 40).

40 a) $(N\ (...BP_i..\ name_k..))_j\ (_s\ QP_i.....(_s....PN_k....e_j....))$
b) $*(N\ (...BP_i..\ name_k..))_j\ (_s\ PN_k.....(_s....QP_i....e_j....))$

In 40 a), where the quantifier is in the higher clause and the pronoun is in the lower (both part of the main argument skeleton), then the adjunct may be adjoined in the

intermediate landing site to catch the quantificational binding and at the same time escape the Condition C violation. On the other hand, if the pronoun is in the higher clause and the quantifier in the lower (both part of the main argument skeleton), no such adjunction will be possible: since the adjunct must be adjoined prior to any wh-movement to be within the scope domain of the quantifier--at this point, however, a Condition C violation will be triggered.

In fact, precisely this configuration of facts occurs.

41 a) Which paper that he gave to Bresnan did every student think that she would like t?

b)*Which paper that he gave to Bresnan did she think that every student would like t?

42 a) Which stories that he told to the President did every Congressman think that he would believe t?

b)*Which stories that he told to the President did he think that every Congressman lied about t?

In each case, the relative clause has a bound pronoun and an target conjoint name. In 41), these are he, bound by every student, and Bresnan, targetted as conjoint with she. Yet the resultant is grammatical in 41-42) a), but not in 41-42) b). The reason is that in 41 a), for example, the adjunct must be adjoined past the start of wh-movement to escape the Condition C violation with she. The point at which it is adjoined must be within the scope of the quantified NP, every student. Such a position exists in the representation in 41 a): namely, the intermediate landing site. No such solution is available in 41 b), since in order to be bound by the quantifier in the lowest clause--and assuming interpretation throughout--the adjunct must be adjoined prior to any wh-movment. But at this point, the sentence triggers a Condition C violation (under the coreferent interpretation of she and Bresnan). Hence it would be expected to be ungrammatical--and it is. The same holds for 42).

Note that nothing very particularistic needs to be said about the grammar to explain the data set in 41)-42): all that needs to be assumed is that adjuncts may adjoin in the course of the derivation (the minimal assumption, given the Projection Principle), and that both binding principles and scope interpretation apply throughout (on pure structurally defined c-command). For additional argument, see the Appendix and Lebeaux (1988, 1989, in preparation).

5.0 CONCLUSION.

In this paper I have argued for two very general restrictions in UG: a Condition on Licensing Well-Formedness, which disallows a sub-tree from appearing in a containing tree before it is licensed in it, and a version of indexing relations (Binding Theory and Control theory) in which indexing functions of this kind apply throughout the derivation. These two broad assumptions were themselves chosen to be compatible with a view of the derivation in which i) the grammar was decomposed into pure and less complex substructures in the base, ii) these were composed (and generated) along the lines of distinct licensing relations, in the course of the derivation. This latter view was justified both empirically (the various sorts of anti-Reconstruction effects, the interaction of Condition C and quantificational binding, the lack of Condition C for NP movement), and in terms of the need to isolate smaller substructures for language acquistion.

FOOTNOTES

This paper is part of a larger project, outlined in Lebeaux (1988, 1989, in preparation). I would like to thank Juan Uriagereka, Peggy Speas, Tom Roeper, Alan Munn, and Pierre Pica for helpful discussion, and Kyle Johnson and Noam Chomsky for discussion and early and more recent encouragement, respectively. All errors are mine.

[1]But see Hamburger and Crain (1982) for a different point of view.

[2]Lasnik (1986) differentiates Condition C into two conditions, one which disallows the c-commanding of a name by a co-indexed pronoun, which is stronger, and one which disallows the c-commanding of a name by a co-indexed name, which is weaker. I follow him in this, and will call only the former "Condition C". See Lasnik for further discussion.

[3]Thus each substructure will obey the Projection Principle.

[4]Tis indexing may be different than the initial NP indexing. Uriagereka (1988) notes that it involves the assignment of a feature like a gamma feature, rather than a regular index.

[5]I leave aside the important question of how copular constructions and the like are treated (cf. Barss, 1985, 1986).

[6]See Lebeaux (in preparation) for further discussion.

REFERENCES

Abney, S. (1987), The English Noun Phrase in Its Sentential Aspect, MIT PhD. dissertation.
Bach, E. (1977), "The Position of the Embedding Transformation in the Grammar Revisited," in A. Zamboli, ed., Linguistics Structures Processing, North Holland.
Barss, A. (1985), "Chain-Binding," West Coast Conference on Formal Linguistics.
Belletti, A. and L. Rizzi (1986), "Psych-verbs and Theta-theory," Lexicon Project Working Papers 13, MIT.
Bowerman, M. (1973), Early Syntactic Development, Cambridge Univ. Press, Cambridge.
Brennan, V. (1989), "A Formal Semantics of Telegraphic Speech," unpublished paper, UMass.
Chomsky, N. (1955), On the Logical Structure of Linguistic Theory, printed in 1975, Plenum Press.
Chomsky, N. (1965), Aspects of the Theory of Syntax, MIT Press.
Chomsky, N. (1981), Lectures on Government and Binding, Foris.
Chomsky, N. (1989), "Notes on the Economy of Derivations," MIT Working Papers.
Chomsky, N. and H. Lasnik (1977), "Filters and Control," Linguistic Inquiry 8.
Engdahl, E. (1978), UMass PhD dissertation.
Fukui, N. and M. Speas (1985), "Specifiers and Projections," MIT Working Papers 8.
Garrett, M. (1975), "The Analysis of Sentence Production," in G.H. Bower, ed. , The Psychology of Learning and Motivation, vol. 9, Academic Press, New York.
Jackendoff, R. (1972), Semantic Interpretation in Generative Grammar, MIT Press.
Kroch, A. and A. Joshi (1985), The Linguistic Relevance of Tree-Adjoining Grammars, Univ. of Pennsylvania.
Lebeaux, D. (1988), Language Acquisition and the Form of the Grammar, UMass PhD. dissertation.
Lebeaux, D. (1989), "Adjoin-a and the Formation of Relative Clauses," in deVilliers and Frazier, Language Processing and Acquisition, Kluwer/Reidel.
Lebeaux, D. (in preparation), "Licensing and the Nature of the Derivation".
Marantz, A. (1984), On the Nature of Grammatical Relations, MIT Press.
Partee, B.H. (1979), "Montague Grammar and the Well-Formedness Constraint," in F. Heny and Schnelle, eds., Syntax and Semantics 10.
Rothstein, S. (1981), The Syntactic Forms of Predication, MIT PhD. dissertation.
Speas, M. (1989), A Theory of Phrase Structure, to appear, Reidel.
Stowell, T. (1982), The Origins of Phrase Structure, MIT dissertation.
Tavakolian, S. (1978), Structural Principles in the Acquisition of Complex Sentences, UMass dissertation.
Vainikka, A. (1986), "Case in Finnish and Acquisition," unpublished manuscript.
Van Riemsdijk, H. and E. Williams (1981), "NP Structure," The Linguistic Review, 1.2.
Williams, E. (1987), "Reassignment of Functions at LF," Linguistic Inquiry.

APPENDIX: LICENSING AND THE DERIVATION.

Let us consider the nature of licensing in the grammar. Elements in the phrase marker may be licensed in the derivation in a heterogeneous number of ways: in particular, by Case theory, by theta theory, by SPEC-Head agreement (Fukui and Speas, 1985, Lebeaux, 1988), by X'-principles (Chomsky, 1970, Jakckendoff, 1977, Speas, 1989), and so on. Let us distinguish classical Aspects-type theories (Chomsky, 1965) from GB theory (Chomsky, 1981) by noting that in Aspects, there was a single type of licensing relation, rewrite rules, which licensed elements phrase-structurally in the base, via the mother-daughter relation (I intentionally ignore movement, and the special role of lexical insertion in that theory.) In GB-theory, there are a number of distinct licensing relations, even within the phrase structure component itself. Call the licensing in Aspects, homogeneous; the licensing in LGB, heterogeneous.

LICENSING AND THE NATURE OF THE DERIVATION

1 a) homogeneous licensing: Aspects
b) heterogeneous licensing: LGB

In the theory outlined above, I have suggested that the heterogeneous licensing of LGB has a derivational reflex: namely, that distinct licensing relations are modelled in a derivational way, with adjunct-embedding modelling its (distinct) licensing relation in a way different than direct argument structure.

In addition, another distinction holds between the two theories, corresponding to the full anchoring of the base in Aspects. In Aspects, no element appeared in the phrase marker prior to the point that it was licensed there. The phrase structure rules applied recursively the base toenerate the relevant structures. This is not true in current views of GB. In particular, assuming that relative clause licensing takes place (in general, or at least optionally) at some post-DS point, the relative clause would be licensed after the point that it is assumed to be present in the structure--that is, if we adopted the standard assumption that it is generated in the base.

2) RC formation (current versions of GB)
The man who I saw left early. (DS)
The man_i (who I saw)$_i$ left early. (RC predication and licensing)

It seems to me, however, that there is something very counter-intuitive about the conventional view. Namely, it requires that an element be "hanging around" in the base, prior to the point that it was licensed at all in the derivation. But what, we might ask, allows it to be in the phrase structure tree at that point? Phrase structure rules have been dispensed with (Stowell, 1982). It seems clear that the only thing that allows an element to be present is the licensing relation itself: that is, it is the licensing relation itself which composes or "nets" the phrase structure element into the representation, or allows it to be visible. But if this is true, then the element cannot be present in the base (or at least cannot be visible there): the relevant licensing relation allowing it into the phrase marker has not yet applied.

The comments in the previous paragraph, namely that the licensing relation itself composes the phrase marker and nets elements into it, suggest the following condition on derivations.

3) Principle of Licensing Well-Formedness (UG)
A subtree T_s may not appear in a major tree T_m prior to the point in the derivation that T_s is licensed in T_m
(T_s, T_m relative).

Note that if 3) holds, then a derivation is a "real thing": it cannot be collapsed into the representational mode. Further, a principal property (perhaps, the principal property) of the derivation is to compose the phrase marker: movement also occurs, perhaps as a side effect of the composition that goes on. Rather than viewing the phrase-maker as pre-existent, and having various modules and principles apply throughout the derivation, we view the modules and principles as applying to compose the phrase marker in its derivation. Finally, we note that the thrust of 3), as in 32) above, is to specify well-formedness conditions throughout the derivation, rather than earmarking them for any particular level. To the extent to which some module of the theory may be thought of as applying at a level, e.g. Case theory applying at some post-DS point, we would view the module as applying throughout, and vocabulary over which it was defined being inserted only at the point at which it "really" applied. That is, if the Case filter applies throughout, and has the following form (4), then lexical NPs can only be inserted when Case is assigned, and not before: otherwise, the lexical NP without Case would violate the Case Filter (at a pre-Case assignment point of the derivation).

4) Case Filter (Rouveret and Vergnaud, 1980)
*NP if lexical and no Case.

In the case of adjuncts, the effect of the Principle of Licensing Well-Formedness is fairly easily calculable, and, while not minor, does not radically alter the form of the grammar: it requires that adjuncts be embedded from a separate structure. The situation is clearly more interesting with respect to Case and theta theory, and with respect to agreement (and the specifier system). Here, in this conference paper, there is not enough room to fully explore the consequences of the adoption of 4) (in the interpretation above), but I would like to indicate briefly where these consequences would be. First, note that the following D-structure would not be well formed.

5) The man saw the woman.

This is because the Case filter applies throughout: since Case assignment applies at some post-DS point, the Case filter would be violated at DS, and the sentence marked as ungrammatical. This difficulty cannot be repaired by assigning Case at DS, since, if we assume that passive morphology prevents the assignment of accusative Case, it is necessary to have Case assignment apply after affix-hopping, to prevent accusative Case from accidentally being assigned to the passive object.

Let us provisionally assume the DP hypothesis (Abney, 1987). Then the DS of 6) cannot be 6 a), but must instead be something like 6 b), where the insertion of lexical NPs has not taken place.

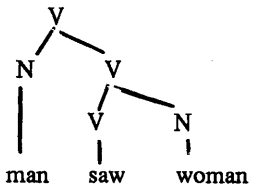
6) The man saw the woman.
a) ($_{dp}$ ($_{d}$ the) ($_{np}$ man)) (saw ($_{dp}$ ($_{d}$ the) ($_{np}$ woman)))
b) ($_{dp}$ ($_{d}$ the) ($_{np}$ e)) (saw ($_{dp}$ ($_{d}$ the) ($_{np}$ e)))

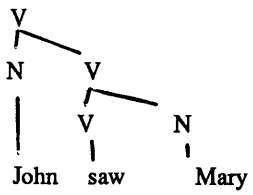
Similarly, the DS of 7) must be 7 b), and not 7 a) (where the O here is meant as a real null determiner; the e is pure emptiness).

7) John saw Mary.
a) ($_{dp}$ ($_{d}$ O) ($_{np}$ John)) (saw ($_{dp}$ ($_{d}$ O) ($_{np}$ Mary)))
b) ($_{dp}$ ($_{d}$ O) ($_{np}$ e)) (saw ($_{dp}$ ($_{d}$ O) ($_{np}$ e)))

Let us call the structures in 6 b) and 7 b), indexed structures.

The structures in 6) and 7) suggest that aspects of the open class (OC) and functional closed class (CC) vocabulary should not only theoretically be distinguished, but are distinguished in terms of their primitive structures as well. Namely, that there is a sort of CC functional grid, that the open class categories are mapped into. The residue of the above representations would be the following (the labelling here is provisional). See, also, Garrett (1975), Lapointe (1985).

8)


9)


I have adopted here the labelling of Lebeaux (1988), where the representations in 8) and 9) were called theta sub-trees, and identified as a sort of "heavy" lexical entry (a lexical entry into which OC insertion had applied), and identified as well with telegraphic speech. For current purposes, we might wish to posit NP phrasal nodes instead, and a small V' projection: what is important is that a coherent subrepresentation makes up the residue.

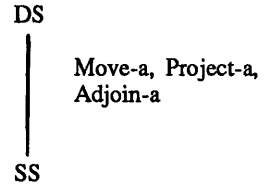
Let us call the operation which projects the OC representations in 8) and 9) into the frames in 7) and 8), Project-a (or Merger).

We may think of the null categories in the indexed structures of 5)-6) as particular types of null categories: little pro's, perhaps.

Note that the general goal of decomposition of the phrase marker, already taken one step with the splitting off of adjuncts, is taken one step further: two substructures underlie the sentence in 7) (ignoring aspects of INFL, and so on).

7)The man saw the woman.
a) substructure 1: the __ (saw the __)
b) substructure 2: (man (saw woman))

The derivation now looks as follows:

8)


The operations Project-a (or Merger) and Adjoin-a correspond to the satisfaction of particular licensing relations (adjunct prediation, Case/theta theory). A further question now arises: where does

LICENSING AND THE NATURE OF THE DERIVATION

Project-a take place? In the proposal of Lebeaux (1988), the substructures composing the sentence in 7) existed prior to DS, and they then made up the DS representation. However, if the above reasoning is correct--namely, that it is for Case reasons that the NP portion of the phrase marker is kept uninserted-- this cannot be, since if the merger operation took place in the base, the resultant would immediately be marked ungrammatical. The requirement, then, is that the substructure 2, the OC structure, be projected into the CC frame at the same point, or after, Case assignment applies. Let us provisionally assume that it is Case assignment itself which triggers the Merger operation. Since Case assignment applies after NP movement (i.e. A-movement) in general (to prevent accusative Case from being assigned accidentally), Merger, which involves lexical insertion of the NP, must take place after NP movement. What this means is that NP movement will be defined on the closed class frame--the indexed structure itself--and lexical insertion applies after that. Simplifying:

9) (e_i seems (pro$_j$ to like pro$_k$)) --Move-a ---------->
(pro$_i$ seems (e_i to like pro$_k$)) --- Project-a ----->
John$_i$ seems (e_i to like Mary)

It is on the CC frame with the indices that the (NP-) indexing relations have been determined.

Note that while we have adopted a fairly "radical" syntactic representation, the ideas forcing this adoption are not particularly radical at all: i) that nothing may appear in the tree before it is licensed (the Principle of Licensing Well-Formedness), and ii) that conditions on indexing apply throughout (Condition on Indexing).

We are finally in a position to solve an apparent problem with the anti-Reconstruction effect alluded to in 9) in the main text above. Recall that we stated Condition C throughout the derivation, applying whenever its structural description was met. The distinction between a name in a fronted adjunct vs. a fronted complement (within an NP), where the fronted argument triggered a Condition C violation and the fronted adjunct did not, was captured by supposing a late adjunction of the adjunct.

10 a) Which claim that John$_i$ made did he$_i$ later prove t?
b)*Whose claim that John$_i$ likes Mary did he$_i$ deny t?

However, there is a serious problem with this generalization, for NP movement (A-movement) cases.

11 a) John's$_i$ mother seems to him$_i$ t to be wonderful.
b) John$_i$ seems to himself$_i$ t to be a nice guy.
c) The pictures of John$_i$ seem to him$_i$ t to be gorgeous.

The DS's of 11 a) and b) should be the following:

12 a) e seems to him$_i$ (John$_i$'s mother to be wonderful)
b) e seems to himself$_i$ (John$_i$ to be a nice guy)

But note that the corresponding tensed DS to the sentences in 12) are bad (with coreference) suggesting that the pronoun in the to-phrase does indeed c-command the downstairs subject (and that Condition C is violated).

13 a)*It seems to him$_i$ that John's$_i$ mother is wonderful.
b)*It seems to him$_i$(self) that John$_i$ is a nice guy.

But this creates difficulties, or seems to. For if Condition C applies throughout the derivation, then the structures in 12) should be out already in D-structure.

The problem, then, is: A-movement (NP movement), unlike A'-movement (Wh-movement), doesn't seem to trigger a Condition C violation, even if the moved name isn't in an adjunct.

But we now have a solution as well. A-movement, unlike A'-movement, involves the movement of indexed structures, i.e. structures in which NPs have not yet been inserted (though determiners have been). This was itself forced by the application of Case theory. But then the derivation of 62 b) is given roughly in 14):

14) e seems to himself (pro$_i$ to be a nice guy) ---Move-a --->
pro$_i$ seems to himself (e_i to be a nice guy) ---Binding->
pro$_i$ seems to himself$_i$ (e_i to be a nice guy) ---Project-a->
John$_i$ seems to himself$_i$ (e_i to be a nice guy)

At no time is Condition C violated. The other derivations follow the same pattern.[6]