SYNTAX and SEMANTICS

VOLUME 25

Perspectives on Phrase Structure: Heads and Licensing

Edited by

Susan D. Rothstein

Department of English
Bar-Ilan University
Ramat Gan, Israel

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
San Diego New York Boston
London Sydney Tokyo Toronto

This book is printed on acid-free paper. ♾

Academic Press, Inc.
San Diego, California 92101

United Kingdom Edition published by
Academic Press Limited
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Syntax and semantics / edited by Susan D. Rothstein.
p. cm. -- (Syntax and semantics : v. 25)
Papers presented at a conference organized by the 1989 Summer Institute at the University of Arizona in Tucson.
Includes index.
ISBN 0-12-613525-8 (Hardcover). -- ISBN 0-12-606106-8 (Paperback)
1. Phrase structure grammar--Congresses. I. Rothstein, Susan Deborah. II. LSA Summer Institute (1989 : University of Arizona) III. Series.
P1.S9 vol. 25
[P158.3]
415--dc20 91-34189
CIP

PRINTED IN THE UNITED STATES OF AMERICA
91 92 93 94 9 8 7 6 5 4 3 2 1

To the memory of Harold and Hannah Levy
1909–1987
1920–1987

. . . כָּאֵלָה וְכָאַלּוֹן אֲשֶׁר בְּשֶׁלֶכֶת מַצֵּבֶת בָּם זֶרַע קֹדֶשׁ מַצֵּבְתָּהּ׃
Isaiah 6:13

Ernst, T. (1990a). *Levels and Locality in Phrase Structure*. Unpublished manuscript, University of Delaware, Newark.
Ernst, T. (1990b). *The Phrase Structure of English Negation*. Unpublished manuscript, University of Delaware, Newark.
Fukui, N. (1986). *A Theory of Category Projection and Its Implications*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
Greenberg, J. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language* (J. H. Greenberg, ed.), 2nd ed. MIT Press, Cambridge, Mass.
Greenberg, J. (1973). Numeral classifiers and substantival number. *Stanford Working Papers in Language Universals* **1**, 1–39.
Greenberg, J. (1975). Dynamic aspects of word order in the numeral classifier. In *Word Order and Word Order Change* (C. N. Li, ed.). University of Texas Press, Austin.
Hawkins, J. (1983). *Word Order Universals*. Academic Press, New York.
Hutchinson, J. (1976). *Aspects of Kanuri Syntax*. Doctoral dissertation, Indiana University, Bloomington.
Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Mass.
Jesperson, O. (1924). *The Philosophy of Grammar*. Norton, New York.
Lewis, M. B. (1969). *Sentence Analysis in Modern Malay*. Cambridge University Press, London.
Nagarajan, S. (1990). *Tamil Serial Verbs*. Unpublished manuscript, University of Delaware, Newark.
O'Siadhail, M. (1988). *Learning Irish*. Yale, New Haven, Conn.
Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* **20** (3), 365–424.
Rizzi, L. (1990). *Relativized Minimality*. MIT Press, Cambridge, Mass.
Roberts, C. (1986). *Modal Subordination, Anaphora, and Distributivity*. Doctoral dissertation, University of Massachusetts, Amherst.
Salleh, M. (1987). *Fronted Constituents in Malay*. Doctoral dissertation, University of Washington, Seattle.
Saul, J., and Wilson, N. (1980). *Nung Grammar*. Summer Institute of Linguistics, Dallas, Texas.
Speas, M. (1990). *Phrase Structure in Natural Language*. Kluwer, Dordrecht.
Sproat, R. (1985). Welsh syntax and VSO structure. *Natural Language and Linguistic Theory*, **3**(2), 173–216.
Stowell, T. (1981). *Origins of Phrase Structure*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.
Uriagereka, J. (1988). *On Government*. Doctoral dissertation, University of Connecticut, Storrs.

RELATIVE CLAUSES, LICENSING, AND THE NATURE OF THE DERIVATION

DAVID LEBEAUX

Department of Linguistics
University of Maryland
College Park, Maryland 20742

1. 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 of smaller units, in the course of the derivation. Thus, distinct licensing conditions are modeled derivationally. Move-α may be interspersed with these phrase structure composition operations freely. The primitives which pick out these separately generated substructures (in the base) are precisely the primitives of GB theory: Case theory, θ-theory, the Projection Principle, and so on. Moreover (I suggest) each primitive (ur-) structure is a 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 θ-theory do not simply mutually describe an existent phrase marker, but pick out distinct subrepresentations which are merged in the course of the derivation. This latter operation I call Merger (or Project-α).

This article is a continuation of work in Lebeaux (1988). That work attempted to develop a syntactic and language acquisition theory side-by-side. I 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

Syntax and Semantics, Volume 25
Perspectives on Phrase Structure

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 modeling 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 she or he hears. These individual phrase structure analyses come at some computational cost: that is, 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, that is, 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 coordinate, 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, 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 article, however, I concentrate on syntactic reasons for such an assumption, particularly with respect to adjunction of adjuncts.

2. ANTIRECONSTRUCTION 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 premovement site. For example, in (1a) 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 (1b) 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 (2c,d) should be bad, like their correspondents, (2a,b). If Condition C were placed at SS, then both (2c) and (2d) should be good. But they divide: (2c) is ungrammatical, while (2d) is not. (By "Condition C" here, I mean the disallowance of a name being c-commanded by a coindexed pronoun; cf. Lasnik, 1986.)[2]

Van Riemsdijk and Williams (1981) suggest that the criterial difference between (2c) and (2d) 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 [(2d)]; 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 [(2c)]. That it is the argument/adjunct distinction that matters, rather than the degree of embedding, can be seen by holding embedding constant and varying the argument/adjunct 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?*[3]

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 versus 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. *Whose examination (paper) near John$_i$ did he$_i$ peek at t?*

Finally, we note that the posthead genitive patterns with the adjunct, not the argument. This is to be expected, as (1) the posthead 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 posthead 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. 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 return below to a modulation of (9) necessary to explain A-chains.] It has been shown that the generalization in (9) covers five subcases:

(10) a. RC versus ***claim***-type complement
b. locative adjunct versus ***picture-of*** complement
c. locative adjunct versus 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 not to be acting as an equivalence class of positions. (I owe this representational insight to Jean-Roger Vergnaud.[4]) Finally, the anti-Reconstruction effects appear to be robust over a wide variety of sentence types. How might this be accounted for?

In the Introduction, I suggested 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 language 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 difference 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 itself 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, 1990). In particular, and simplifying, X′ theory builds the verb up to VP; the Projection Principle and the θ-criterion force 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 reapplies 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) obtained in this way, that is, by ap-

plying 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) $[_S\ [[_{NP}\ [_{N'}\ [_N$ *man*$]]]\ [_{V'}$ *saw* $[_{NP}\ [_{N'}\ [_N$ *woman*$]]]]]$

[The locative in (11) is intended to modify the woman.]

Evidently, 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 obtained 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 θ-criterion—that is, applying the modules which *must* apply in the base—the substructure (14b) would be picked out from the entire structure in (14a).

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

The retrieval of the substructure in (14b) from the full structure in (14a) 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 (14a) 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*

Similarly, 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. Both substructures, of course, obey X′ theory.

(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 the examples in (17)–(19) in which an object nominal head (e.g., *claim, examination*) 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, for example, *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 θ-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. This means that the Projection Principle + θ-criterion, X′ principles, and the Projection Principle + θ-criterion applies over again, and forces 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, and for which the phrase marker breaks up into substructures (along the lines of licensing conditions).

(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 represen-

tation 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) versus (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 θ-criterion.

This will break each tree into the relevant subtrees. Each is, roughly, a pure representation of the "argument-of" relation. As we may conceive of X′ principles, the Projection Principle, 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.[5]

4. ADJOIN-α, CONDITION C, AND THE FORMATION OF RELATIVE CLAUSES

Suppose that we do not take the conceptual filtering operation above to be merely conceptual, but to be actual, 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 modeled differently, quite differently, than the argument-of relation. The latter is modeled by each individual substructure. The former is modeled by the relation composing them in the course of the derivation. Let us call this operation *Adjoin-α*. The derivation as a whole then looks like this (see, also Kroch and Joshi, 1985; Chomsky, 1955):

(27) DS (substructures respecting ARG-of relation)

Move-α, Adjoin-α
SS

We may now explain the anti-Reconstruction effects. Suppose that the operation Adjoin-α exists, composing substructures, and applies freely within the derivation, along with Move-α, the minimal assumption. Of course, each individual substructure must obey the Projection Principle, so Adjoin-α applies only between substructures that are in an adjunct-of relation; those that 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-α would take place), whereas the latter would hold for *claim* complements. Suppose now that we take a derivation in which Move-α 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 DSs 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-α has applied. In the other derivation, it will be added before Move-α has applied. As above, we will call the operation composing the substructures Adjoin-α. 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.* $\xrightarrow{\text{Move-}\alpha}$
SS: **Whose claim that John$_i$ liked Mary did he$_i$ deny t?*
b. DS: s_1: *He$_i$ later denied which claim* $\xrightarrow{\text{Move-}\alpha}$
s_2: *That John$_i$ made*
s_1: *Which claim did he$_i$ deny t*] $\xrightarrow{\text{Adjoin-}\alpha}$
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*] $\xrightarrow{\text{Adjoin-}\alpha}$
s_2: *that John$_i$ made*]
**He$_i$ later denied which claim that John$_i$ made* $\xrightarrow{\text{Move-}\alpha}$
SS: **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 (29a) and the first derivation in (29b). 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 (29a), 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 as 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 (29b). 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 (29b). By the Projection Principle, the adjunct need not be part of the original s_1. Move-α may apply to s_1, moving the WH-phrase to the front. At that point, Adjoin-α 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 structurally 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 added is crucial, as only then will it escape the c-command domain at S-structure, assuming that Condition C applies everywhere. This latter shows that the mere statement of adjunctual relations being somehow "different" will not by itself be sufficient.

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

The other four types of antireconstruction effects given in (10) pattern identically. For example, the locative versus 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-α→
SS: *Which picture of $John_i$ did he_i like t?*

(31) DS: s_1: *He_i liked which picture* —Move-α→
s_2: *near $John_i$*
s_1: *Which picture did he_i like t?*] —Adjoin-α→
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 that may be added later. This is impossible in the derivation in (30), where through 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

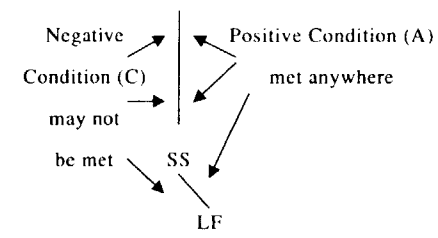
In the preceding section, I assumed that (1) the Projection Principle holds, and (2) the binding conditions, in particular Condition C, applied throughout the derivation. Note that if one takes the Projection Principle, and only the Projection Principle, to characterize the obligatory licensing relation 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).[^] 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 (Belletti and Rizzi, 1986; Burzio, 1989; Kayne, 1990; Lebeaux, 1988). (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)



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).[6] 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 (1978, 1986), in which a bound pronoun has been moved outside of the scope domain of a quantifier which binds it.

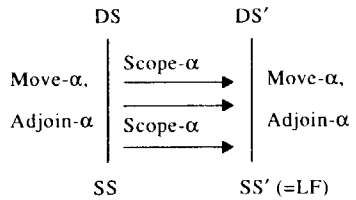
(34) a. *Which of his parents do you believe that every man likes t best?*
b. *How many of his papers do you think that every linguist wishes to rewrite t?*

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 configuration, that 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; Chomsky, 1987; Hornstein, 1984; Williams, 1987). 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), 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*movement 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 precursor, 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.

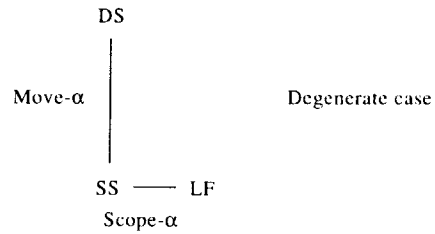
(35)



Scope-α here is simply Move-α, with the additional properties that scope taking operations seem to have. Here, Move-α and Adjoin-α 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-α). 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 ones 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, as 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 is not visible, without introducing another entire 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 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 that are adjoined in the course of the derivation, by Adjoin-α.

(38) s_1: *you saw the picture*] $\xrightarrow{\text{Adjoin-}\alpha}$ *you saw the picture*
s_2: *near him* *near him.*

What 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 place early in the derivation, prior to the point at which the WH-phrase to which it attaches is fronted over the binding quantificational phrase. [This would be the second derivation in (29b).]

(39) s_1: *you think every man saw which picture*] —Adjoin-α→
s_2: *near him*]
You think every man saw which picture near him. —Move-α→
Which picture near him do you think that every man saw t?

Thus 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-Reconstruction effects, the adjunct must be adjoined later, 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 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 that 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. $[N\ [\ldots BP_i \ldots name_k \ldots]]_j\ [_s\ QP_i \ldots [_s \ldots PN_k \ldots e_j \ldots]]$
b.*$[N\ [\ldots BP_i \ldots name_k \ldots]]_j\ [_s\ PN_k \ldots [_s \ldots QP_i \ldots e_j \ldots]]$

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), 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, as 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?*

(43) a. *Which part that he_i played with $Madonna_j$ did every aspiring $actor_i$ wish that she_j would support t?*
b.**Which part that he_i played with $Madonna_j$ did she_j think that every aspiring $actor_i$ had failed at t?*

In each case, the relative clause has a bound pronoun and a target conjoint name. In (41), these are *he*, bound by *every student*, and *Bresnan*, targeted as conjoint with *she*. Yet the resultant is grammatical in (41a), (42a), and (43a), but not in (41b), (42b), and (43b). The reason is that in (41a), 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 (41a), namely, the intermediate landing site. No such solution is available in (41b), because to be bound by the quantifier in the lowest clause—and assuming interpretation throughout—the adjunct must be adjoined prior to any WH-movement. 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) and (43).

Note that nothing in particular needs to be said about the grammar to explain the data set in (41), (42), and (43); 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).

4.2. A Conceptual Argument

In the discussion above, the lack of Condition C effects in structures like (44a) (the anti-Reconstruction structures), together with the intricate patterning of data in structures like (44b), was used to argue for an analysis of the derivation in which adjuncts were added-in in the course of the derivation, and Binding theory and, in general, indexing functions and their interpretation were stated throughout.

(44) a. *Which pictures that John took did he like t?*
b. *Which paper that he gave to Bresnan did every student/*she think that she/*every student would like t?* (where the two choices are linked)

The same effect might be obtained not in this fashion, but by assuming that Condition C applies at LF (or some post-Reconstruction level) and that it applies obligatorily for N's, but not full NPs. In such a case, the abrogation of Condition C effects for fronted adjuncts in (44a) could be obtained by assuming that N′ obligatorily reconstructs, and then triggers a Condition C violation at LF, while the outer shell of the NP need not reconstruct obligatorily, but only optionally. This would derive the contrast between (45a) and (45b).

(45) a. *Which pictures that John took does he like t?*
b.?**Which pictures of John does he like t?*

It is clear that, empirically, the above solution would work. The problem with it is conceptual. Namely, a particular part of the Reconstruction, that of the N′, is obligatory. This is questionable on two grounds: the obligatoriness itself (which is not generally found in the grammar), and in the particular picking out of the N′ as special. Note that there is no theoretical reason for supposing that the reconstruction statement would be the one in (46a), rather than that in (46b) or (46c).

(46) a. Reconstruct N′ obligatorily.
b. Reconstruct the full NP obligatorily.
c. Reconstruct the outer shells of NP (outside of N′) obligatorily.

Obviously, the statements in (46) would give rise to different—in the case of (46c), opposite—patterns of results. Yet if the grammar were stated in terms of Reconstruction, there would be no principled reason that the grammar would not contain the reconstruction statement in (46c) rather than (46a). This is not the case with the Adjoin-α formulation, because the obligatoriness of the N′ in the base follows directly from the Projection Principle from the fact that both X′ theory and the Projection Principle apply obligatorily in the base, and the base must therefore contain their transitive closure. See Lebeaux (1988) for further discussion.

5. 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 θ-theory, by SPEC–head agreement (Fukui and Speas, 1985; Lebeaux, 1988), by X′ principles (Chomsky, 1970; Jakckendoff, 1977; Speas, 1990), 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, namely, rewrite rules, that 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. Let us call the licensing in *Aspects* homogeneous, and the licensing in *LGB* heterogeneous.

(47) 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 modeled in a derivational way, with adjunct-embedding modeling 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 to generate 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.

(48) 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 counterintuitive 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 (Farmer, 1980; 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 (see also Partee, 1979).

(49) 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 (49) 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 marker as preexistent, 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 (49), as in (32), 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, for example, 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 (50), 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).

(50) 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, although 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 θ-theory and with respect to agreement (and the specifier system). Here, in this conference article, there is not enough room to fully explore the consequences of the adoption of (49) (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.

(51) *The man saw the woman.*

This is because the Case filter applies throughout; as 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, because, 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 (52) cannot be (52a), but must instead be something like (52b), where the insertion of lexical NPs has not taken place.

(52) *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 (53) must be (53b), and not (53a) (where the O here is meant as a real null determiner; the e is pure emptiness).

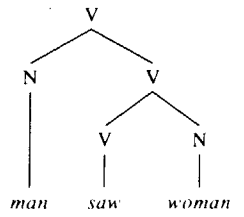
(53) *John saw Mary.*
a. [$_{DP}$ [$_{D}$ O] [$_{NP}$ *John*]] [*saw* [$_{DP}$ [$_{D}$ O] [$_{NP}$ *Mary*]]]
b. [$_{DP}$ [$_{D}$ O] [$_{NP}$ *e*]] [*saw* [$_{DP}$ [$_{D}$ O] [$_{NP}$ *e*]]]

Let us call the structures in (52b) and (53b) INDEXED STRUCTURES; these correspond, in some respects, to the structures prior to those in which names have been lowered, as in Montague (1974).

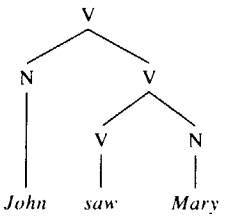
The structures in (52) and (53) suggest that aspects of the open-class (OC) and functional closed-class (CC) vocabulary should not only be distinguished theoretically, but are distinguished in terms of their primitive structures as well. Namely, there is a sort of CC functional grid into which the open class categories

are mapped. The residue of the above representations would be the following (the labeling here is provisional) (also see Garrett, 1975; Lapointe, 1985).

(54)



(55)



I have adopted here the labeling of Lebeaux (1988), where the representations in (54) and (55) were called θ subtrees, 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 (54) and (55) into the frames of (52b) and (53b) Project-α (or Merger).

We may think of the null categories in the indexed structures of (52) and (53b) 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 (56) (ignoring aspects of INFL, and so on).

(56) *The man saw the woman.*
a. substructure 1: the __ [*saw the* __]
b. substructure 2: [*man* [*saw woman*]]

The derivation now looks as follows:

(57) DS
|
Move-α, Project-α,
Adjoin-α
SS

The operations Project-α (or Merger) and Adjoin-α correspond to the satisfaction of particular licensing relations (adjunct predication, Case/θ-theory). A further question now arises: Where does Project-α take place? In the proposal of Lebeaux (1988), the substructures composing the sentence in (56) existed prior to DS, and they then made up the DS representation. However, if the above reasoning is correct, namely, it is for Case reasons that the NP portion of the phrase marker is kept uninserted, this cannot be, because 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 that triggers the Merger operation. As 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,

(58) $[e_i$ *seems* $[\text{pro}_j$ *to like* $\text{pro}_k]]$ __ Move-$\alpha \rightarrow$
$[\text{pro}_j$ *seems* $[e_j$ *to like* $\text{pro}_k]]$ __ Project-$\alpha \rightarrow$
John$_j$ *seems* $[e_j$ *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: (1) nothing may appear in the tree before it is licensed (the Principle of Licensing Well-formedness), and (2) 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). 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 versus 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.

(59) 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.

(60) 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 DSs of (60a,b) should be the following:

(61) 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 (61) 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).

(62) 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.*

[I have changed the reflexive to a pronoun in (62b) to avoid the extraneous Condition A violation.] But this creates difficulties, or seems to. For if Condition C applies throughout the derivation, then the structures in (61) should be out already in D-structure.

The problem, then, is that A-movement (NP movement), unlike A′-movement (WH-movement), does not seem to trigger a Condition C violation, even if the moved name is not in an adjunct.

But we now have a solution as well. A-movement, unlike A′-movement, involves the movement of indexed structures, that is, 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 (61b) is given roughly in (63):

(63) *e seems to himself* $[\text{pro}_i$ *to be a nice guy*$]$ __ Move-$\alpha \rightarrow$
pro_i *seems to himself* $[e_i$ *to be a nice guy*$]$ __ Binding $\rightarrow$
pro_i *seems to himself*$_i$ $[e_i$ *to be a nice guy*$]$ __ Project-$\alpha \rightarrow$
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.[7]

6. CONCLUSION

In this article I have argued for two very general restrictions in UG: a Condition on Licensing Well-Formedness, which disallows a subtree 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 derivation in which (1) the grammar was decomposed into pure and less complex substructures in the base, and (2) 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 acquisition.

7. APPENDIX: RECONSTRUCTION AND NESTING EFFECTS IN A-CHAINS

In this Appendix, I would like to point to some additional nesting effects with A-chains, that is, additional cases where the application of a negative condition

(Condition C) must be such that the subtree over which it applies is nested inside the subtree over which a positive condition (Binding Condition A or quantificational binding in general) applies, under conditions of movement over both of them. The basic cases of this type were given in (41) and (42) in the text for A′-chains.

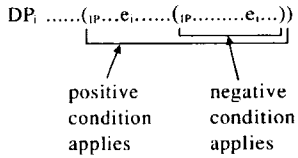
(64) a. *Which paper that he$_i$ gave to Bresnan$_j$ did every student$_i$ think that she$_j$ would like t?*
b.**Which paper that he$_i$ gave to Bresnan$_j$ did she$_i$ think that every student$_i$ would like t?*

(65) a. *Which stories that he$_i$ told to the President$_j$ did every Congressman$_i$ think that he$_j$ would believe t?*
b.**Which stories that he$_i$ told to the President$_j$ did he$_j$ think that every Congressman$_i$ lied about t?*

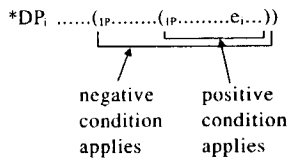
(66) a. *Which part that he$_i$ played with Madonna$_j$ did every aspiring actor$_i$ wish that she$_j$ would support t?*
b.**Which part that he$_i$ played with Madonna$_j$ did she$_j$ think that every aspiring actor$_i$ had failed at t?*

The extraction of the WH-phrase is taking place over the entire clause. In the (a) sentences, the negative condition, Condition C, would apply in the lowest IP (if one looked at the entire moved NP), while the positive condition, simply binding, would apply over the entire phrase marker (i.e., at the highest IP node). The negative condition is thus nested inside the positive condition. In the (b) sentences the reverse occurs: the positive condition, binding, would apply in the lowest clause (if one looked at the premovement structure), while the negative condition, Condition C, would apply over the entire IP. Thus, possible alignment of the conditions with respect to the structure is as is given in (67). This is for adjuncts in A′-chains.

(67) a. Negative Condition Nested Inside Positive Condition (Grammatical):



b. Positive Condition Nested Inside Negative Condition (Ungrammatical):



This need not be stated in the primitive basis, but rather follows from the interaction of true movement, the fact that indexing conditions apply over the entire derivation, and the possibility of the derivation being modeled by radically information-increasing operations; here, the possibility of adjunction of adjuncts in the course of the derivation. Simply put, because the target information that is contained in the relative clause is not present at DS, the structure does not fulfill the structural description of the negative condition (Condition C) at DS. At an intermediate stage in the derivation (DS, SS), the relative clause may be added and satisfy the structural description of the positive condition (quantificational binding). As the derivation is increasing in information, such an adjunction operation is possible; and the result is that negative conditions must always be nested inside of positive conditions, as the information in the moved element may escape the negative condition in the local tree.

Remarkably, a second instance of such nesting is found in A-chains as well. Before noting these, a partial correction in the generalizations found in the literature or reconstruction must be noted. It is sometimes said that the following holds:

(68) Elements in A′ chains reconstruct; elements in A chains do not (or act as if reconstructed, etc.).

This is to account for the fact that A-moved elements may bind reflexives in their surface position, for example, whereas A′-moved elements may not:

(69) *John seems to himself t to be a nice guy.*

I believe though that reconstruction, however implemented, treats binders and bindees differently:

(70) a. WH-moved elements may be bound from the surface position or any of the trace sites (both quantificational and anaphoric binding).
Example: *Which pictures of himself does John believe that Bill likes t?* (ambiguous)
b. WH-moved elements may not bind from any surface position except the DS position.

(71) a. NP-moved elements may be bound from the surface position or any of the trace sites (both quantificational and anaphoric binding).
Examples:
(i) *His mother tends t to seem to every man t to be the best woman there is.*
(ii) *His first girlfriend is known t to seem to every man t to be a sort of revelation.*
(iii) *His knowledge of history tends t to seem to every history professor t to be irrelevant in political discourse.*
(iv) *Pictures of himself are expected t to be provided t by every candidate who applies.*

b. NP-moved elements may bind from the surface position or any of the trace sites (both quantificational and anaphoric binding).

The empirical contention in (71a) is the crucial one here, for the point of view of this article [it may be that (71b) is simply indeterminate, or possibly wrong; this depends on the actual structure of chains, and their interpretation].

If the data in (71a) are correct, then A-chain sites may act as sites from which the element is bound, here the bound pronoun. It can be shown that in these cases it is not the case that the quantifier is directly scoped over the pronoun (one clause or two clauses up), apart from the implausibility of that claim in general, because if the pronoun is not in a chain one or two clauses up, but simply one or two clauses up in general (as, say, the PP off of a raising predicate), no such binding is possible. That is, it is possible to get the bound reading for the pronoun in (72a) but not (72b).

(72) a. *His mother seems t to have pleased t every man.*
(bound reading possible)
b. *John seems to him t to have pleased t every man.*
(bound reading impossible)

I have adopted the Belletti–Rizzi (1986) analysis of *please* here, but the main point is simply that *his mother* can be bound by *every man*, and not *him* by *every man*, precisely because *his mother* is in a chain whose terminal member is in the scope of *every man*, whereas *him* is not. (I have used the representational description here for convenience, but I do not believe that the binding takes place representationally, contra, e.g., Barss, 1986, for reasons that are given below.)

Let us therefore take (71a), repeated below, as empirically correct.

(73) A bindable element in an A-chain may be bound in its surface position or at any trace site.

Is the reconstruction down the A-chain obligatory? Clearly not, for two reasons. First, an A-moved element may bind an anaphor from its surface position; if reconstruction were obligatory, the reflexive would be expected to be unbound at LF, and the sentence bad.

(74) *John seems to himself t to be a nice guy.*

Second, if the A-moved element contains a name, this name does not trigger a Condition C violation in its DS position (as noted in the text):

(75) a. *John seems to himself t to be a nice guy.* (i.e., the same example)
b. *Those pictures of John seem to him t to be nice.*

It appears, then, that the empirical formulation in (73) is the correct one, and may be extended as follows:

(76) In $DP_i \ldots e_i \ldots e_i \ldots e_i$, an A-chain, the DP may be read as if in any of the trace sites (optionally).

Then in the case where the DP contains a pronoun that must be bound below, we may view it as "reconstructing" (or the derivational equivalent)—these are the examples in (71a). On the other hand, when the DP contains a name, we need not view it as reconstructing; these are the examples in (75b) and (62) in the main text. That is, the reconstruction is optional.

The representational description in (76) may appear at first to be the appropriate one. If this were true, then the chain-binding view of Barss (1986) (and obvious extensions of it) would be supported, where the chain is viewed simply as an equivalence class of positions, and binding is treated as a relation which holds between this class and some exterior element. However, the representational description in (76) is not accurate, and the representational description, while it *can* encode the necessary information, does not do so naturally.

The reason is that the binding of elements in A-chains exhibits exactly the same nesting effects which have been shown in (67) above for A′-chains, where negative conditions must be nested inside positive conditions. These nesting properties make sense in an essentially derivational and information-increasing derivation that is suggested here. Within a purely representational view, there is no reason that they take the form that they do, rather than, say, the reverse (where the positive condition would be nested inside the negative condition).

Let us consider this in more detail. I have stated the binding possibilities so far in the Appendix in terms of reconstruction. In the main text, however, I adopted the view that movement took place over indexed structure and that lexical insertion took place into that. The avoidance of Condition C violations in examples like those in (75b) was then dealt with by assuming that lexical insertion (of the DP, or possibly the NP part of the DP) was held back until the moved element in indexed structure was past the point where it would trigger a Condition C violation. This is shown in (77).

(77) a. e *seems to him* pro *to be very nice.*
DP: Those pictures of John
b. *after* Move-α (on indexed structure):
pro_i *seems to him* e_i *to be very nice*
DP: Those pictures of John
c. Insert-a (into indexed structure):
Those pictures of John$_i$ seem to him$_i$ e_i *to be very nice.*

Similarly, to catch the quantificational binding, we must allow that a DP or NP may be inserted at an intermediate point in the derivation, including the DS position, to be in the appropriate structural configuration for quantificational binding or anaphoric binding to be read off (i.e., a configuration in which c-command

holds). (Note that this differs from the position in the main text, where it is assumed that lexical insertion applies at the point of Case assignment. Here, I suggest that it may take place anywhere in the point of A-movement.) The relevant example is below. I simplify the representation to focus on the main point.

(78) a. *e tends e to seem to every man e to be the best woman there is.*
DP: his mother
b. after insertion (on indexed structure):
e tends e to seem to every man his mother to be the best woman there is.
c. Move-α applies twice:
His mother$_i$ tends e_i to seem to every man e_i to be the best woman there is.

This predicts that exactly the same nesting effects with respect to positive and negative conditions should apply in A-chains that appeared before in A′-chains, though here with respect to the entire content of the DP. That is, when the DP is inserted, it must be inserted at a point to (1) avoid a Condition C violation and (2) catch any relevant quantificational binding. If the quantifier to be caught is higher in the tree than the pronoun to be conjoint with, then such a point should be available—an intermediate A-site—whereas if the pronoun to be conjoint with is higher than the quantifier, no such intermediate A-site should exist, and the derivation should block. Precisely this configuration of data seems to occur.

(79) a. *(His$_i$ mother's)$_j$ bread seems to every man$_i$ t to be known t by her$_j$ to be the best there is.*
b.*(*His$_i$ mother's)$_j$ bread seems to her$_i$ t to be known t by every man$_j$ to be the best there is.*

(80) a. *His$_i$ picture of the President$_j$ seemed to every man$_i$ t to seen by him$_j$ to be a real intrusion.*
b.**His$_i$ picture of the President$_j$ seemed to him$_j$ t to seen by every man$_i$ t to be a real accomplishment.*

(81) a. *The President's$_i$ discussion with him$_j$ seemed to every man$_j$ t to be seen by him$_i$ t to be very beneficial.*
b.**The President's$_i$ discussion with him$_j$ seemed to him$_i$ t to be seen t by every man$_j$ to be quite an achievement.*

I have permuted the name and pronoun in the moved DP, (80) and (81), to show that this is irrelevant. Although the sentences are complex, the judgments seem to be relatively durable.

The same description that was given in (67) of the negative condition (Condition C) having to be nested inside the positive condition (binding) holds here as well. The only difference is that in this case we are talking about A-chains, not

A′-chains, and the direct content of the DP, rather than the content of the adjunct. It is the nesting property, however, that shows the content of the derivation. Note that this goes strongly against a purely representational view of the chain, in which all positions are held as interpretively equivalent and the chain itself is interpreted as if it were a single interpretive entity: the chain-binding view (e.g., of Barss, 1986). In such a view, the chain is interpreted as a single element, and the content of the chain is read as if it were available at any trace site.[8] I give below one simplified formulation of the chain-binding view.

(82) CHAIN-BINDING (REPRESENTATIONAL VIEW):
In the chain $DP_i \ldots e_i \ldots e_i \ldots e_i$, any element in the tree **chain-binds** an element in DP_i iff it c-commands a trace bound by DP_i (optional).

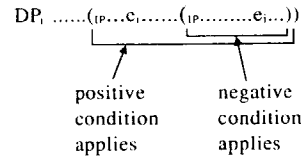
I have intentionally left (82) indeterminate between A- and A′-chains. The formulation in (82) is essentially equivalent, with respect to the positive condition of binding, to allowing optional multiple domination of the dislocated phrase by all of the mothers of the trace sites (that is, it is equivalent to the phrase-linking view of Peters-Ritchie, at least with respect to this positive condition).

The chain-binding view in (82) adequately predicts the data set in A-chains in (67) repeated below as (83a), because the quantifier may chain-bind the pronoun in the dislocated DP. Similarly, it adequately predicts the lack of a Condition C violation [given in (83b)], as the chain-binding is optional.

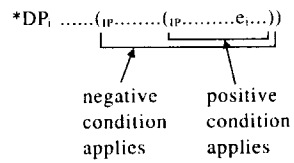
(83) a. *His mother tends t to seem to every man t to be the best woman there is.*
b. *Those pictures of John seem to him t to be quite wonderful.*

However, it does not predict the interaction between positive and negative conditions. Namely, it does not predict the statement in (67) repeated below.

(84) a. Negative Condition Nested Inside Positive Condition (Grammatical):



b. Positive Condition Nested Inside Negative Condition (Ungrammatical):



The content of (84) follows from the fact of a derivation (together with an information-increasing account within the derivation). The nesting condition must be

that in (84a) (as possible) rather than the reverse. In the purely representational view, an intermediate element in the tree may be viewed as binding the content of the DP or not; however, there is no *nesting* property in that account. The intermediate element either binds the content of the DP or not. This nesting property is rather a consequence of a derivational view, in which the subsequences that make up a chain are viewed in a particular order, that in (85).

(85) chain: $DP_i....e_i....e_i....e_i$

subsequences: 1. DP

2. $DP_i....e_i$

3. $DP_i....e_i....e_i$

4. $DP_i....e_i....e_i....e_i$

Although I have argued here directly from A-chains, the same holds for A′-chains, which have the same property.

I make two final points. I have argued above that a chain is not interpretively simplex, but rather that reference is required to the subsequences making it up [(85)] and given in the derivation. This does not mean that it is not grammatically simplex (as in Chomsky, 1981). Rather, I take it that grammatical statements like those in (86) are (basically) correct, and simply show that an interpretively nonsimplex element may act grammatically simplex.

(86) a. *A chain with lexical content must have Case assigned at (only) one point.*

b. *A chain must have a single θ-role.*

Second, although the nesting effects noted here are fairly strong evidence against a representational view of the interpretation of chains (the chain-binding view), they do not decide conclusively between either of two derivational approaches, namely, one in which an element is literally iteratively lowered down an A- or A′-chain (as in Chomsky, 1981, one formulation), and one in which interpretation takes place throughout the derivation, side-by-side with the derivation, as in the formulation of Jackendoff (1972) or the rule-by-rule approach of Montague (1974) and Bach (1977). But that is the subject for another article.

ACKNOWLEDGMENTS

This article is part of a larger project, outlined in Lebeaux (1988, 1989, in preparation). I thank Juan Uriagereka, Jean-Roger Vergnaud, Peggy Speas, Tom Roeper, Alan Munn, and Pierre Pica for helpful discussion, Kyle Johnson and Noam Chomsky for discussion and early and more recent encouragement, respectively and Susan Rothstein for editorial assistance. All errors are mine.

NOTES

[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 coindexed pronoun, which is stronger, and one which disallows the c-commanding of a name by a coindexed name, which is weaker. I follow him in this and call only the former Condition C. See Lasnik (1986) for further discussion.

[3] A reviewer notes that the possibility of disjointness (or not) also depends on the content of the genitive phrase: thus, (3c, d) vary both the adjunct argument distinction and the content of which versus whose. I believe that this is correct, and that part of the reason of the ungrammaticality of (3d) is due to the presence of the genitive *whose*. However, examples like (4), (5), and (6), (a) and (b) show that this cannot be all that is at stake; more interestingly, if we keep the genitive constant in (5), we note that the contrast still holds. The covariance with the true genitive is actually part of a larger phenomenon, where the penetrability of the fronted DP to Condition C effects in part covaries with opacity, though in a reverse direction from what might be expected, for example:

(i) a. *Pictures of John, he really likes t.*
b. *?A picture of John, he would really like t.*
c. *??Mary's pictures of John, he really likes.*

I am unclear as to judgments here, and the covariance must await further explication.

[4] Thus, Vergnaud suggests, contra Barss (1986), for example, that the binding conditions cannot be treated directly in terms of chain binding. See also the Appendix and Vergnaud's forthcoming work.

[5] Thus, each substructure will obey the projection principle.

[6] This 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.

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

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

REFERENCES

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Bach, E. (1977). The position of the embedding transformation in the grammar revisited. In *Linguistics Structures Processing* (A. Zamboli, ed.). North-Holland Publ., Amsterdam.

Barss, A. (1985). *Chain-Binding*. Paper presented at the West Coast Conference on Formal Linguistics.

Barss, A. (1986). *Chains and Anaphoric Dependencies*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Belletti, A., and Rizzi, L. (1986). Psych-verbs and theta-theory. *Lexicon Project Working Papers* **13.**

Brennan, V. (1989). *A Formal Semantics of Telegraphic Speech.* Unpublished manuscript, University of Massachusetts, Amherst.

Burzio, L. (1989). *Anaphoric Binding.* Unpublished manuscript.

Chomsky, N. (1955). *On the Logical Structure of Linguistic Theory* (printed in 1975). Plenum, New York.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* MIT Press, Cambridge, Mass..

Chomsky, N. (1970). Remarks on nominalization. In *Studies in Semantics in Generative Grammar.* Mouton.

Chomsky, N. (1981). *Lectures on Government and Binding.* Foris, Dordrecht.

Chomsky, N., and Lasnik, H. (1977). Filters and control. *Linguistic Inquiry* **8.**

Engdahl, E. (1978). Doctoral dissertation, University of Massachusetts, Amherst.

Engdahl, E. (1986). *Constituent Questions.* Reidel, Dordrecht.

Farmer, A. (1980). *Modularity in Syntax,* MIT Press, Cambridge, Mass.

Fukui, N., and Speas, M. (1985). Specifiers and projections. *MIT Working Papers* **8.**

Garrett, M. (1975). The analysis of sentence production. In *The Psychology of Learning and Motivation* (G. H. Bower, ed.), Vol. 9. Academic Press, New York.

Hamburger, H., and Crain, S. (1982). Relative acquisition. In *Language Development: Volume 1* (S. Kuczaj II, ed.). Erlbaum, Hillsdale, N.J.

Hornstein, N. (1984). *Logic as Grammar.* MIT Press, Cambridge, Mass.

Huang, J. (1982). *Logical Relations in Chinese and the Theory of Grammar.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar.* MIT Press, Cambridge, Mass.

Jackendoff, R. (1977). *X'-Syntax.* MIT Press, Cambridge, Mass.

Katz, J. J., and Fodor, J. (1963). The structure of a semantic theory. *Language,* 170–210.

Kayne, R. (1990). Romance clitics and PRO. *Proceedings of NELS* **20.**

Kroch, A., and Joshi, A. (1985) *The Linguistic Relevance of Tree-Adjoining Grammars.* University of Pennsylvania, Philadelphia.

Lapointe, S. (1985). A model of syntactic phrase combination in speech production. *Proceedings of NELS* **15.**

Lasnik, H. (1986). "Binding Theory," talk at the Princeton Conference on Formal Linguistics.

Lebeaux, D. (1988). *Language Acquisition and the Form of the Grammar.* Doctoral dissertation, University of Massachusetts, Amherst.

Lebeaux, D. (1989). Adjoin-a and the formation of relative clauses. In *Language Processing and Acquisition* (deVilliers and Frazier, eds.). Kluwer/Reidal, Dordrecht.

Lebeaux, D. (in preparation). Licensing and the nature of the derivation.

Marantz, A. (1984). *On the Nature of Grammatical Relations.* MIT Press, Cambridge, Mass.

Montague, R. (1974). *Formal Philosophy* (Richard Thomason, ed.). Yale University Press, New Haven, Conn.

Partee, B. H. (1979). Montague grammar and the well-formedness constraint. In *Syntax and Semantics* (F. W. Heny and H. S. Schnelle, eds.), Vol. 10. Academic Press, New York.

Rouveret, A., and Vergnaud, J. R. (1980). Specifying reference to subject. *Linguistic Inquiry.*

Speas, M. (1990). *A Theory of Phrase Structure.* Reidel, Dordrecht.

Stowell, T. (1982). *The Origins of Phrase Structure.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Tavakolian, S. (1978). *Structural Principles in the Acquisition of Complex Sentences.* Doctoral dissertation, University of Massachusetts, Amherst.

Uriagereka, J. (1989). Doctoral dissertation, University of Connecticut.

van Riemsdijk, H., and Williams, E. (1981). NP structure. *Linguistic Review* **1** (2).

Williams, E. (1987). Reassignment of functions at LF. *Linguistic Inquiry.*