

Introduction to Transformational Grammar

Kyle Johnson
University of Massachusetts at Amherst

Fall 2002

Contents

Preface	iii
1 Theories	1
2 Phrase structure and arguments	13
3 Order of Complements	33
4 Argument Movement	49
5 Verb Movement	71
6 Determiner Phrases and Noun Movement	105
7 Complement Structure	131
8 Subjects and Complex Predicates	163
9 Rethinking Things	207
9.1 Cleaning up	207
9.2 Towards Deriving \bar{X} Theory	214
9.2.1 Kayne's "Antisymmetry" hypothesis	214
9.2.2 The "Bare Phrase Structure" proposals	231
10 Quantifier Raising	239
11 Introduction to the Binding Theory	271

Preface

These are the, always evolving, notes from an introductory course on syntactic theory taught at the University of Massachusetts at Amherst. Its target audience is first-year graduate students, but no background exposure to syntax is presupposed. These notes augment a set of readings, which, in Fall 2002, were:

- Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: M.I.T. Press, 1965. Chapter 1.
- Stowell, T. “Origins of Phrase Structure.” Doctoral Dissertation, Massachusetts Institute of Technology, 1981. Chapters 1 & 2.
- Sportiche, D. “A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure.” *Linguistic Inquiry* 19 (1988): 425-49.
- Pollock, J.-Y. “Verb Movement, UG and the Structure of IP.” *Linguistic Inquiry* 20 (1989): 365-424.
- Vikner, S. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press, 1995. Chapter 3.
- Chomsky, N. “Some Notes on Economy of Derivation and Representation.” In *Principles and Parameters in Comparative Grammar*, ed. Freidin, R. 417-54. Cambridge, Massachusetts: MIT Press, 1991. pp. 417-454.
- Kayne, R. “Unambiguous Paths.” In *Connectedness and Binary Branching*, 129-64. Dordrecht: Foris Publications, 1984. pp. 129-164.
- Abney, S. “The English Noun Phrase in Its Sentential Aspect.” Doctoral Dissertation, Massachusetts Institute of Technology, 1987. Chapters 1 & 2.
- Bernstein, J. B. “Topics in the Syntax of Nominal Structure Across Romance.” Doctoral Dissertation, City University of New York, 1993.
- Larson, R. “On the Double Object Construction.” *Linguistic Inquiry* 19 (1988): 335-92.
- Johnson, K. “Object Positions.” *Natural Language and Linguistic Theory* 9 (1991): 577-636.

- Kratzer, A. “Severing the external argument from its verb,” In *Phrase Structure and the Lexicon*, eds. Rooryck, Johan and Zaring, Laurie. 109-137. Kluwer Academic Publishers, 1996.
- von Stechow, A. “The Different Readings of *wieder* “again”: A Structural Account.” *Journal of Semantics* 13 (1996): 87-138.
- Kayne, R. S. *The Antisymmetry of Syntax*. Cambridge, Massachusetts: M.I.T. Press, 1994. Chapter 1.
- Chomsky, N. “Bare Phrase Structure.” In *Government binding theory and the minimalist program*, ed. Webelhuth, G. 383-439. Oxford: Oxford University Press, 1995.
- Fiengo, R., and R. May. *Indices and Identity*. Cambridge, Massachusetts: MIT Press, 1994. Sections 6.1-6.2.
- Hornstein, N. “An Argument for Minimalism: The Case of Antecedent-Contained Deletion.” *Linguistic Inquiry* 25 (1994): 455-80.
- Chomsky, N. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris Publications, 1981. Chapter 3.
- Huang, C.-T. J. “Logical Relations in Chinese and the Theory of Grammar.” Doctoral Dissertation, Massachusetts Institute of Technology, 1982. Sections 5.0-5.5.
- Chomsky, N. *Knowledge of Language*. New York, New York: Praeger Publishers, 1986. Chapter 3, pp. 164-186.
- Rizzi, L. “Null Objects in Italian and the Theory of *pro*.” *Linguistic Inquiry* 17 (1986): 501-58.
- Huang, C.-T. J. “On the Distribution and Reference of Empty Pronouns.” *Linguistic Inquiry* 15 (1984): 531-74.
- Chomsky, N. “On Wh-Movement.” In *Formal Syntax*, eds. Culicover, P., T. Wasow, and A. Akmajian. New York, New York: Academic Press, 1977.
- Kayne, R. “Connectedness,” in *Connectedness and Binary Branching*. Dordrecht, Holland: Foris Publications, 1984.
- Chomsky, N. “Beyond Explanatory Adequacy.” (2001): manuscript.

Chapter 1

Theories

Syntactic Theory, and to a slightly weaker extent, Linguistic Theory, in the United States has been very heavily influenced by learnability considerations in the latter half of this century, thanks largely to the writings of Noam Chomsky. If we decide that syntactic theory is charged with the duty of modeling part of our knowledge of language, that is that it concerns cognition, psychology or whatever, then one of the questions that arises is how this knowledge is acquired by the child. A number of considerations combine to make this task look very difficult indeed: the complexity of the acquired grammar, the nature of the data available to the child, etc. In addition, it is made even more puzzling by the fact that children appear to learn some particular language with relative ease in a very short period of time and that the course of acquisition goes through a set schedule of stages. It is clear that a very large problem arises: If our grammars of particular languages are correct, then how can these possibly be learned?

It is Chomsky's proposal that Syntactic Theory itself should contribute to solving this dilemma. The classical formulation of this idea (see Aspects and SPE) characterizes the situation as follows. Think of a grammar of L (G_L) (this is what Chomsky (1986b) calls "I-Language") as a set of rules that generates structural descriptions of the strings of the language L (Chomsky's E-language). Our model of this grammar is descriptively adequate if it assigns the same structural descriptions to the strings of L that G_L does. We can think of the learning process as involving a selection from the Universe of G_L s the very one that generates the strings of the L to be acquired.

The learning problem can now be stated in the following terms: how is it that the learning procedure is able to find G_L when the universe of G s is so huge and the evidence steering the device so meager.

One step towards solving this problem would be to hypothesize that the universe of G s has structure (i.e, is not so large). This is Chomsky's proposal. It amounts to the claim that there are features of G s which are built-in: certain properties which distinguish the natural class of G s from the rest. There is a kind of meta-grammar of the G s, then, which is sometimes referred to with the label Universal Grammar (though it should be noted that this term has found a multitude of vague, but probably other, uses). Chomsky further hypothesizes that these properties are biologically given: that it is a fact of the human brain/mind construction that the class of G s are the way they

are. This argument, the one that leads from the observation that G_L s have features that are too complex to be learned to the conclusion that the universe of Gs is constrained is often called “The Poverty of the Stimulus” argument. It is a classic from Epistemology, imported with specific force by Chomsky into linguistics.

This way of setting up the problem, note, allows for the Universe of Gs to be larger than the learnable Gs. There could be, for instance, constraints imposed by the parsing and production procedures which limit the set of Gs that can be attained. And it’s conceivable that there are properties of the learning procedure itself — properties that are independent of the structure of Gs that Universal Grammar imposes — that could place a limit on the learnable Gs. Universal Grammar places an outside bound on the learnable grammars, but it needn’t be solely responsible for fitting the actual outlines of that boundary. It’s therefore a little misleading to say that the set of “learnable Gs” are those characterized by Universal Grammar, since there may be these other factors involved in determining whether a grammar is learnable or not. I should probably say that Universal Grammar carves out the “available Gs,” or something similar. But I will instead be misleading, and describe Universal Grammar as fixing the set of learnable Gs, always leaving tacit that this is just grammar’s contribution to the learnability question.

Chomsky proposes, then, that the goal of syntactic theory should be to reveal this structure on the universe of Gs. In particular, he proposes in *Aspects of The Theory of Syntax* that syntactic theory should include an evaluation metric which “ranks” Gs. A syntactic theory that has this feature he calls explanatory. Thus “Explanatory theory” has a specific, technical, sense in linguistic theory. A theory is explanatory if and only if it encapsulates the features that ranks Gs in such a way that it contributes to the learnability problem, distinguish the learnable Gs from the unlearnable ones. This criterion can help the syntactician decide whether the model of G_L he or she has proposed corresponds exactly to G_L . In particular, the many descriptively adequate models of G_L can be distinguished on this basis: we should select only those that are ranked highly by the evaluation metric. These grammars meet the criterion of explanatory adequacy.

A very important role, therefore, is played by the evaluation metric. At the time of *Aspects*, the learning procedure was conceived of as a process very much like that which the linguist goes through. The child builds a battery of rules which generate the strings of L. The evaluation metric steering this process was thought to have essentially two parts: a simplicity metric, which guides the procedure in its search through the space of grammars, and inviolable constraints, which partitions the set of Gs into the learnable ones and the unlearnable ones. Thus, for example, rules that used fewer symbols were defined as “simpler” than ones that used a greater number of symbols one, mnemonic, way of defining a simplicity metric. Inviolable constraints might be those, for example, expressed as part of Theory which places constraints on phrase structure grammar, and therefore simply removes from the universe of Gs a great many possible Gs. Let’s call these models of Gs “rule based,” because the simplicity metric is defined as rule construction procedure, and the companion picture of the acquisition process the “Little Linguist” model.

To take a concrete example, if \bar{X} Theory — the theory that expresses constraints on phrase structure in UG — imposes the following constraint on all phrase structure rules, then the evaluation metric leaves to the learner only the matter of filling

in the variables W, X, Y and Z, discovering their linear order, and determining what coöccurrence restrictions there are on the phrases.

$$\begin{aligned} XP &\rightarrow \{ (ZP), \bar{X} \} \\ X' &\rightarrow \{ \bar{X}, (YP) \} \\ X' &\rightarrow \{ X^0, (WP) \} \end{aligned}$$

(Understand “ $\{\alpha, \beta\}$ ” to signify that α and β are sisters, and “ (α) ” to indicate that α is optional.) As the child goes from step to step in matching the grammar he or she is constructing with the information coming in, these are the only decisions that have to be made. If we imagine that this set of options were to be operationalized into a concrete decision tree, then we could see this as constituting a kind of “simplicity metric.” It would constitute a procedure for searching through the space of learnable grammars that ranks the grammars. Additionally, Theory provides information which places an absolute cap on the possible phrase markers. In this respect it also illustrates an inviolable constraint.

Let’s consider another example involving transformational rules one that Chomsky often points to. Transformational rules map one level of representation, D-structure, to another, S-structure, typically by way of moving constituents. Interestingly, it appears that all such rules are “structure dependent.” That is, they make reference to the relative structural positions of moved thing and target of movement. They don’t, for example, make reference to points in a string on the basis of their position relative to some numerical count of formatives. Thus Wh-Movement moves maximal projections that meet certain criteria to particular positions in a phrase marker. And this operation is governed by a set of constraints that make reference to the relation between these points solely in terms of structure. There is no rule, for example, like Wh-Movement which affects terms based on how far apart they are numerically. Thus, the learning procedure will never have to entertain the hypothesis that G_L should contain such rules.

In both cases, the classic argument for distinguishing the inviolable constraint from the simplicity metric follows very closely the logic of the poverty of stimulus argument. Because it is difficult to see (maybe even provably impossible) how such things as \bar{X} Theory or structure dependence could be learned, they must belong to the features that define the universe of Gs. And because they are overarching properties of the rules in some G_L a guide or template they also have the right form to be inviolable constraints.

There is another argument towards the same end which has gained increasing influence in the last couple decades; and this one comes to us through the narrowly linguistic study of language typology, and only tangentially from learnability considerations. I will call it “Humboldt’s argument,” though it no doubt has an earlier champion. Humboldt’s argument is based on the observation that there are certain properties that appear to hold true of all G_L s. This can be explained, Humboldt argues, only if the universe of Gs is constrained to just those which have the relevant, universal, properties. Like Chomsky, Humboldt relates this to the construction of the mind, and uses the language of learnability in his account. He puts it this way:

Since the natural inclination to language is universal to man, and since all men must carry the key to the understanding of all languages in their

minds, it follows automatically that the form of all languages must be fundamentally identical and must always achieve a common objective. The variety among languages can lie only in the media and the limits permitted the attainment of the objective.

(von Humboldt (1836))

(One might read the last sentence of this passage as making the distinction, touched on above, between aspects of Universal Grammar (“the media”) and the limits our cognition places on exploiting UG (“the limits permitted the attainment of the objective”).) So, like Chomsky, he supposes that there is a Universal Grammar, a feature of the mind, which constrains the form that languages may have. But his perspective is different from Chomsky’s. He expresses the notion of Universal Grammar not in terms of learning theory, or through the glass of the Poverty of the Stimulus argument, but from the perspective of language variability. He links limits on language variability to a universal ability he sees in human psychology to acquire a language. There is a weakness to Humboldt’s argument. There is another possible explanation for typological similarities: the thesis of monogenesis. If all languages descend from a common one, then features that are shared among them could all simply be vestiges of the ancestral language that historical change has left untouched. I think it’s possible to read Sapir as advancing this alternative. Sapir is commonly associated with the position exactly opposite to Humboldt’s; in Sapir’s words:

Speech is a human activity that varies without assignable limit as we pass from social group to social group, because it is a purely historical heritage of the group, the product of long-continued social usage.

(Sapir 1921, p. 4)

But, perhaps because of his vagueness, it’s possible to credit Sapir with a more sophisticated view. One that assigns the universal properties of languages to the detritus of historical change:

For it must be obvious to any one who has thought about the question at all or who has felt something of the spirit of a foreign language that there is such a thing as a basic plan, a certain cut, to each language. . . . Moreover, the historical study of language has proven to us beyond all doubt that a language changes not only gradually but consistently, that it moves unconsciously from one type towards another, and that analogous trends are observable in remote quarters of the globe.

(Sapir (1921, p. 120-121))

Perhaps the common properties of extant (and known) languages are a function of two facts: all languages descend from a common language, and the forces that cause languages to change are not fully random they preserve certain features and change others only according to some “basic plan.” Could it be, then, that the similarities in languages are all due to the laws of diachrony?

Note, however, that even if we grant monogenesis, this thesis entails that language variation is solely the product of historical change, as Sapir’s quote(s) makes clear. So we expect that languages vary in features which historical change can affect,

but will remain similar in those ways that are immutable. Which of the features appear as language universals, then, is determined by the internal mechanisms of historical change, and the limits thereon. What are the internal mechanisms of historical change?

I don't know that anyone knows. But the only proposal I know of is that historical change is a by-product of language acquisition. And language acquisition, the poverty of the stimulus argument tells us, is capped by Universal Grammar. So even granting the diachronic argument for language universals, we see that as historical change weeds out the mutable properties from the immutable ones, it will leave the properties that characterize Universal Grammar. The antidote for the argument I have blamed on Sapir, then, involves bringing the poverty of the stimulus argument into play. I don't know if Humboldt's argument can stand against this alternative unaided.

But even if it can't, it provides us with another way of viewing how to factor out the components of the evaluation metric. Following the logic of Humboldt's argument, what we expect is that language comparison should give us a means of separating inviolable constraints from the simplicity metric. The inviolable constraints will be (among) those things found in all languages; the differences in languages are to be credited to the simplicity metric. Put somewhat differently, an explanatory theory is to give us both how languages cannot be constructed, and how their construction can vary. The data it must fit, then, emerges only once languages are compared: for not only does this allow the universals to be clearly discerned, but it is only through this means that the particulars of language variation are known.

When this method of factoring out the universals in G_s is followed in earnest, a rather different picture of various G_L s emerges; and a very different conception of the language acquisition procedure becomes available. This course is meant to illustrate these emerging pictures in detail.

We'll begin, really, somewhere near the beginning. We set out to give a characterization of the syntax of one particular language: English. By "syntax," I mean:

- The process by which knowers of the language recognize legitimate arrangements of words, or, to put it in jargon, the principles which enable individuals competent in the language to form "grammaticality judgments."
- The processes by which knowers of the language are enabled to extract the meaning of sentences by virtue of their word arrangements.

It is a model of these processes which we are interested in creating: a grammar of the syntax of English. At heart, then, syntax is aimed at characterizing how "arrangements of words" are formed and contribute to forming the meanings of sentences.

Consider the first of these tasks: modeling the processes by which grammatical arrangements are recognized. Our first observation is that, to a very large extent, this process or these processes make use of very little information in the words whose arrangement is being determined. In fact, all that seems to matter is the morpho-syntactic "category" they belong to. That this is what's relevant can be demonstrated by virtue of the "Novel Form" argument, which goes as follows:

- (1) a. A "novel" word can be introduced into an individual's vocabulary.
- b. If enough information is introduced with the novel word to enable the individual to recognize its category, then

- c. The individual knows which arrangements it can grammatically combine in.
- d. Hence, it must be category membership to which these processes refer.

Now, it turns out that certain strings of categories have a characteristic distribution. For example, the string Det+Adj+N can appear immediately before a verb, immediately after a verb, immediately after a preposition, and immediately preceding the 's of the genitive Case:

- (2) a. the happy woman left.
- b. I met the happy woman.
- c. I talked to the happy woman.
- d. the happy woman's friend

Next, note that this string of categories can be "coordinated" with a like string of categories. Coordination involves the use of words called "conjuncts," words such as *and*, *or*, *nor*, etc. Thus, we find:

- (3) the happy woman and an unhappy man
- (4) a. * the angry and an unhappy man
- b. * the and an unhappy man

Finally, with respect to all these distributional tests, the strings Det+N+P+N, N+P+N, Adj+N, N, and (infinitely) many others also pass. We need some way of describing the fact that these strings are "the same," and different from, say, P+N which has a distinct distributional pattern.

We call this family of strings a phrase, and write a Phrase Structure Rule to describe which strings belong to this family. In the case at hand, this rule might look like:

- (5) $\alpha P \rightarrow (\text{Det}) (\text{Adj}) N$

This leaves out the strings Det+N+P+N and N+P+N. These strings, however, involve another phrase, made up of the string P+N. This string, along with any string that conforms to the template P or P+NP has a characteristic distribution: this family of strings can appear immediately after a noun, a verb, and can be conjoined with another like string. Unlike α Ps, however, this family of strings cannot occur immediately preceding the verb (except in one special case, that I use right below), nor can it immediately precede the genitive "s".

Like α Ps, β Ps may be coordinated with other β Ps, but not with other α Ps, as the following examples illustrate.

- (6) a. Under the bed and behind the sofa are usually good places to find money in my house.
- b. * Under the bed and the dining room table are usually good places to find money in house.

Hence, just as with the family of strings that our incomplete rule describing NPs, this family of strings is in some significant sense "the same thing."

Putting this together, we come up with the pair of Phrase Structure rules below.

- (7) a. $\alpha P \rightarrow (\text{Det}) (\text{Adj}) N (\beta P)$
 b. $\beta P \rightarrow P (\alpha P)$

Now note that these two rules together have the property that they are recursive. This is an important aspect of Phrase Structure rules, for it is the primary means by which we describe the infinity of grammaticality judgments. These two PS-rules are able to characterize infinitely many and infinitely long strings of words. This is a correct result, for that is in fact what we are able to do.

Still another Phrase Structure rule is required to account for the fact that the family of strings that include V , $V+\alpha P$, $V+\beta P$, $V+\alpha P+\beta P$, and an infinite set of other such strings act the same. These strings have a characteristic distribution: they can appear immediately following a verb, they may appear immediately following the initial αP (or βP in the one special case of (6)) which starts a sentence, but they may never appear immediately following a PP (unlike an NP) nor may they appear in the sentence initial position that αP s and βP s can. And like αP s and βP s, coördination treats members of this family as alike and distinct from NPs and βP s.

- (8) a. Mary walked and talked.
 b. Mary visited Paul and kissed Barry.
 c. Mary talked to Paul and met with Barry.

These facts call for a Phrase Structure rule like the following:

- (9) $\gamma P \rightarrow V (\alpha P) (\beta P)$

Now note that there is a common property to all these Phrase Structure rules. In each case, all of the constituents are optional, except one. Thus, a verb is the only necessary member of a γP , a noun the only requisite member of an αP and a preposition is all that's required to make a βP . Further, the converse also turns out to be true: whenever there is a preposition, there is a βP , wherever a noun is found, there is an NP, as so on. Thus, nouns and αP , prepositions and βP , verbs and γP are in one-to-one correspondence. This is a very pervasive property of Phrase Structure rules. Whereas Phrase Structure rules vary to a considerable degree across languages, this property of them seems almost uniformly to be true. We'll confront two apparent counterexamples from English shortly, but these are probably only apparent counterexamples. (We'll see the solution to one of them in a few classes.) So far as I am aware, there is no clear counterexample to this generalization. We call this property of Phrase Structure rules *endocentricity*; and we call the word that must be a member of the phrase its *head*. Finally, it is common practice to name the phrases after their heads, so we'll rename αP , NP, βP PP and γP VP. Thus, we now have:

- (10) a. $NP \rightarrow (\text{Det}) (\text{Adj}) N (\text{PP})$
 b. $PP \rightarrow P (NP)$
 c. $VP \rightarrow V (NP) (\text{PP})$

In addition to these three Phrase Structure Rules, we'll need quite a few others. Indeed, the principle of endocentricity leads us to expect that for every category, there will be a Phrase Structure rule that builds a phrase headed by that category. For

example, corresponding to the category Adjectives, there is a rule that builds adjective phrases:

$$(11) \quad AP \rightarrow A \text{ (PP)}$$

The presence of PPs within Adjective phrases is supported by the existence of strings like:

- (12) a. She is interested in syntax.
She is interested.
b. He seems happy with linguistics.
He seems happy.

We'll put off, for a moment, writing a phrase structure rule that allows APs to appear in this position immediately after a verb but as can be seen from these examples, this is a position in which Adjectives are defined, as are A+PP strings. The coördination test also treats A and A+PP strings as being the same:

- (13) a. She is happy and interested in syntax.
b. He seems bored but happy with linguistics.

Finally, we'll need a Phrase Structure rule that tells us how these various phrases are put together to form a sentence. This brings us to our first apparent counterexample to endocentricity:

$$(14) \quad S \rightarrow NP \text{ VP}$$

But if we look harder, we find that sentences are, in fact, in a one-to-one correlation with some category; that is, they are endocentric. The reason we have to look hard is because that category includes not just free morphemes, but bound ones as well.

- (15) a. Jerry leaves.
b. Sally left.
c. Sam has left.
d. Sarah **had** left.
e. Martha **should** leave.
f. George **might** have left.
g. Laura desires [Sal **to** leave].
h. Larry remembers [Jim leav**ing**].

These boldfaced terms have similar distributions: they are found either immediately preceding the verb (if they are free) or affixed onto the verb (if they are bound). For this reason, and because they are never found together (i.e., they are in complementary distribution) we believe they belong to the same category. Every S has one of these; or, putting it differently: Ss have the same distribution as this word class. The category these morphemes belong to is called "Inf" or "tense," after the affixes found in finite clauses. So we should rebuild and rename Ss along the following lines:

$$(16) \quad IP \rightarrow NP \text{ I VP}$$

There's another sort of "sentence" that we encounter in examples like the following (17).

- (17) a. Mary said **that John likes chocolate**.
 b. Mary recalled the rumor **that John likes chocolate**.
 c. **That John likes chocolate** bothers Mary.
 d. Jerry is angry **that John likes chocolate**.

Note that the string following the word *that* meets the conditions imposed by the rule that builds sentences. The word *that* is called a complementizer and it is the head of the phrase (actually clause) found in these sentences. This clause is called sometimes an "S-bar" (designated with \bar{S}), but more modernly a "Complementizer Phrase" (CP). CPs conform to the requirements of the following Phrase Structure rule.

- (18) $CP \rightarrow C IP$

Other complementizers are *if* and *whether*, as found in the following examples.

- (19) a. I wonder if Mary likes chocolate.
 b. I asked whether Mary likes chocolate.

Having introduced this constituent, we will now need to revise our previous Phrase Structure rules to include the positions where they may lie. In addition, we'll need to include a position for adverb phrases. This yields the following battery of rules.

- (20) a. $IP \rightarrow \left\{ \begin{array}{c} NP \\ CP \end{array} \right\} I VP$.
 b. $NP \rightarrow (Det) (Adj) N (PP) (CP)$
 c. $VP \rightarrow V (NP) (PP) (CP)$
 d. $AP \rightarrow A (PP) (CP)$

Note the option of having a CP in place of an NP at the beginning of a sentence. These two options are disjunctively available; a relationship which is encoded by use of the curly brackets.

There's a couple of refinements/additions to the Phrase Structure rules we've so far developed that we'll now consider. First, notice that an AP may include a certain sort of "adverb" preceding the head adjective, as in the following examples.

- (21) a. Mary is loudly angry.
 b. John is marvelously happy.

Note that not just any sort of adverb can be found here:

- (22) a. * Mary is loudly angry.
 b. * John is slowly happy.

Here, perhaps, the meaning of the words involved does matter: maybe there is some semantic constraint responsible for this division. Or perhaps adverbs divide into sub-categories, or perhaps there is some syntactic constraint(s) on the distribution of adverbs that accounts for this. Let's assume that adverbs do not divide into different

sub-categories and that the restrictions illustrated here are essentially semantic, and therefore don't need to be reflected in our phrase structure rules. This leads, then, to the following modification of the rule that characterizes adjective phrases.

$$(23) \quad AP \rightarrow (Adv) A (PP) (CP)$$

Next, observe that Adverb Phrases may also precede the verb heading a VP, as in:

- (24) a. Mary quickly walked home.
b. Mary often talks to Jane.

And so, the VP rule should be redrafted to include this possibility.

$$(25) \quad VP \rightarrow (Adv) V (NP) (PP) (CP)$$

In addition, note that VPs may occur immediately following a verb, as in:

- (26) a. Mary has walked.
b. Mary has talked to John.
c. Mary has visited Gary.

Interestingly, if the verb heading a VP is followed by another VP, nothing else may follow the head verb. Thus: *Mary has on the platform walked* is ungrammatical. We need, therefore, to modify the VP Phrase Structure rule in such a way that the head verb is followed by a VP, or by the expansion previously arrived at, but no combination thereof. This can be done with the aid of curly brackets in the following way:

$$(27) \quad VP \rightarrow (Adv) V \left\{ \begin{array}{c} (NP) \quad (PP) \quad (CP) \\ \quad \quad VP \end{array} \right\}$$

Further, it is possible to find APs embedded within VPs, as we've seen above; some additional examples are:

- (28) a. Sally remains angry at Jim.
b. Frank is happy with himself.

When APs follow verbs, then may be preceded by, at most, a PP, as in (29).

$$(29) \quad \text{Jerry seem } [_{PP} \text{to Bill}] [_{AP} \text{happy with his kumquats}].$$

So we change the rule that characterizes VPs to:

$$(30) \quad VP \rightarrow (Adv) V \left\{ \begin{array}{c} (NP) \quad (PP) \quad (CP) \\ \quad \quad VP \\ \quad \quad (PP) AP \end{array} \right\}$$

Next, note that the NP Phrase Structure rule will need to be modified to account for strings such as:

$$(31) \quad \text{the very angry woman}$$

This indicates that not just an adjective precedes the head noun, but in fact an Adjective Phrase; thus:

(32) $AP \rightarrow (Det) (AP) N (PP) (CP)$

Next, notice that adverbs may be preceded by a very tiny class of adverbs, as in:

- (33) a. Mary walks very quickly.
b. Mary talks extremely loudly.

Let's call the class of adverbs that can be found in this position Degree words; indeed, let's suppose that they form a different category from adverbs. This leads us to the following new Phrase Structure rules.

(34) $AdvP \rightarrow (Deg) Adv$

Finally, consider that part of the NP rule that introduces determiners. Determiners include words like *the*, *a*, *that* (not to be confused with the complementizer *that*), *every*, *some*, *all*, etc. Interestingly, it's very rare that we find determiners combining with other words to form a phrase that combines with a noun that follows. A couple of these rare examples are given in:

- (35) a. all but three dogs
b. more than most people

I don't know precisely what the Phrase Structure rule is that determines which strings may stand in this position. Nonetheless, one common approach to these cases is to imagine that determiners head their own anemic phrases, which are then positioned withing NPs. We will revisit this idea; for now let's imagine that determiner phrases are made up of nothing but determiners.

(36) $DP \rightarrow Det$

We'll also need to update the Phrase Structure rule forming NPs, then. But before we do this, let's consider strings like the following.

- (37) a. Mary's book
b. the man's toy
c. the man on the table's nose

These examples involve a possessive or genitive phrase. Note that this phrase is an NP with the morpheme *'s* appended to the end. Further, note that this genitive phrase never co-occurs with a DP:

- (38) a. * the Mary's book
b. * the the man's toy
c. * a the man on the moon's nose

They are in complementary distribution, a fact that calls for an account. One very typical explanation for complementary distribution is to understand the two things as different instances of the same thing. In this situation, that can be done by rigging the NP Phrase Structure rule in such a way that it either produces a DP or a genitive phrase in the same position. This is done with curly brackets, as in the following example.

$$(39) \quad NP \rightarrow \left\{ \begin{array}{c} (DP) \\ (NP'_s) \end{array} \right\} (AP) N (PP) (CP)$$

One final Phrase Structure rule is required by the sorts of examples we've so far reviewed. This is the Phrase Structure rule that generates coordinated phrases. This can be done with the following.

$$(40) \quad \alpha \rightarrow \alpha \text{ Conj } \alpha$$

This rule says that a phrase of any category can be made up of two other such phrases with a conjunct stuck between them. Conjuncts, recall, are *and* and *or*, principally.

Summarizing, we've now introduced the following battery of Phrase Structure rules:

$$(41) \quad \begin{array}{ll} \text{a. } IP \rightarrow \left\{ \begin{array}{c} NP \\ CP \end{array} \right\} I VP \\ \text{b. } NP \rightarrow \left\{ \begin{array}{c} (DP) \\ (NP'_s) \end{array} \right\} (AP) N (PP) (CP) \\ \text{c. } VP \rightarrow (Adv) V \left\{ \begin{array}{ccc} (NP) & (PP) & (CP) \\ & VP & \\ & (PP) AP & \end{array} \right\} \\ \text{d. } DP \rightarrow Det \\ \text{e. } AdvP \rightarrow (Deg) Adv \\ \text{f. } AP \rightarrow (AdvP) A (PP) (CP) \\ \text{g. } CP \rightarrow C IP \\ \text{h. } PP \rightarrow P (NP) \\ \text{i. } \alpha \rightarrow \alpha \text{ Conj } \alpha \end{array}$$

Chapter 2

Phrase structure and arguments

The technique we have employed up to now to uncover the constituent structures of strings of words (when they are grammatical) has relied on a procedure that identifies “substitution classes.” We get this method from Zellig Harris,¹ who in turn inherits it from the procedure the Structuralists, mainly Leonard Bloomfield, used to discover the constituents of words. Indeed, it is precisely this technique which yields the structure of vocabularies into the categories: Noun, Verb, Adjective, etc.

An important property of this procedure, which Harris noted, is that it should not equate the class itself with its head. He points out, for example, that while singular nouns are in the same substitution class as are plural ones, a plural noun cannot substitute for a singular one when it combines with the plural morpheme. (That is, we cannot form from this procedure a doubly pluralized noun. Note that this relies on Harris’s presupposition that the rules of syntactic composition range over morphemes, and not just words.) This is the reason, then, why our phrase structure rules look like (1) and not (2).

(1) $NP \rightarrow (Det) (Adj) N (PP)$

(2) $N \rightarrow (Det) (Adj) N (PP)$

Or, to put it somewhat differently, we do not want these rules to be recursive with respect to their head.

When this procedure of discovering substitution classes is pursued further, we find that there are little phrases inside those we’ve identified so far. For example, in the position marked by “___” in (3), we find the family of strings in (4). Some examples are in (5).

(3) Det ___ V

(4) $\{N, A N, N PP, A N PP, A A N, N PP PP, A A N PP, A A N PP PP, \dots\}$

¹See Harris (1946).

- (5) the woman left.
 the happy woman left.
 the woman with a hat left.
 the happy woman with a hat left.
 \vdots

And coördination also reveals that this set of strings form a family:

- (6) The woman and happy man left.
 The happy woman and man with a hat left.
 \vdots

Now this family of strings does not appear to be the family we have called NP. There are two, related, reasons for this. First: there are grammatical strings from the second family which cannot be substituted for instances of the first family:

- (7) The woman left.
 *Woman left.

Second: a close inspection of the set that the second family is made up of indicates that it does not share Harris's property: indeed, this family is recursive with respect to a substring that includes its head. So, we set up something like this:

- (8) a. $NP \rightarrow \left\{ \begin{array}{c} NP's \\ DetP \end{array} \right\} \bar{N}$
 b. $\bar{N} \rightarrow Adj \bar{N}$
 c. $\bar{N} \rightarrow \bar{N} PP$
 d. $\bar{N} \rightarrow N$

Note how these rules encode the "optionality" of Adj and PP differently than the optionality of Det. And note, further, that they are all endocentric on N. (They also leave out the position of CP; this is because fitting CPs into this structure poses a problem that we will return to in just a moment.)

We find the existence of very similar subphrases within VPs as well. Consider, for instance, the environment in (9), which permits the family of strings in (10), as (11) exemplifies.

- (9) NP ____ CP
- (10) {V, Adv V, V PP, Adv V PP, Adv Adv V, Adv Adv V PP, V PP PP, Adv V PP PP, ...}
- (11) Sally said that Jerry left.
 Sally quickly said that Jerry left.
 Sally said to Peter that Jerry left.
 Sally quickly said to Peter that Jerry left.
 Sally carefully said to Peter on Tuesday that Jerry left.
 \vdots

And, as before, coördination recognizes this family.

- (12) Sally shouted and whispered that Jerry left.
 Sally loudly shouted and whispered that Jerry left.
 Sally shouted to Peter and quietly whispered that Jerry left.
 ⋮

Again, this subphrase is recursive and headed. So we have something like:

- (13) a. $VP \rightarrow \bar{V}$
 b. $\bar{V} \rightarrow Adv \bar{V}$
 c. $\bar{V} \rightarrow \bar{V} PP$
 d. $\bar{V} \rightarrow V$

These batteries of rules leave out the expansions of VP which introduce NPs, CPs, APs, and VPs. And, I've made the strange decision to characterize VP as dominating a \bar{V} and nothing else. (The first of these rules says that VPs s and nothing more.) We'll return to these niceties.

A similar situation arises in Adjective Phrases too:

- (14) Det N ____ V
 (15) {A, Adv A, Adv Adv A, A PP, Adv A PP, Adv A PP PP, ...}
 (16) A child unaccompanied is likely to cause trouble.
 A child often unaccompanied is likely to cause trouble.
 A child unaccompanied by a guardian is likely to cause trouble.
 A child often unaccompanied by a guardian to the zoo is likely to cause trouble.
 ⋮

As before, this family is recursive and headed. And, as before, it is visible to coördination as well.

- (17) A child unaccompanied and unloved is likely to cause trouble.
 A child unaccompanied by a guardian and completely unloved is likely to cause trouble.
 A child frequently unaccompanied and completely unloved by her parents is likely to cause trouble.
 ⋮

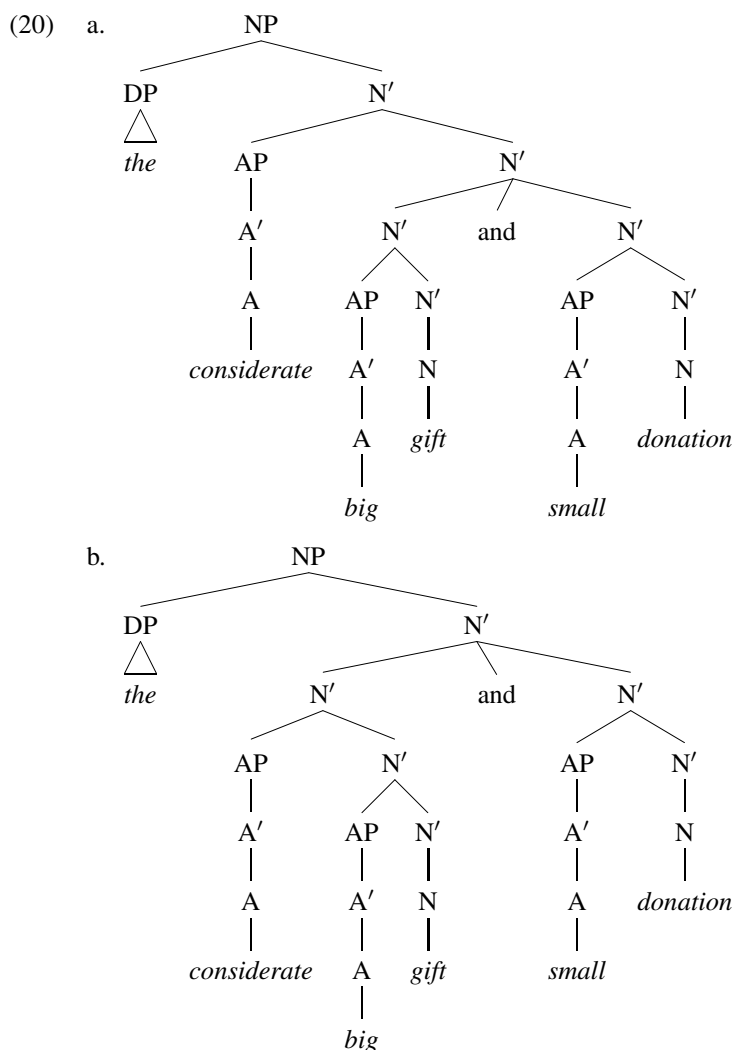
We need to revise the AP rule to something like:

- (18) a. $AP \rightarrow \bar{A}$
 b. $\bar{A} \rightarrow Adv \bar{A}$
 c. $\bar{A} \rightarrow A$

Note that I have left out CP, as in the other rules; and, as with the VP rule, characterized AP as consisting of just an \bar{A} and nothing else. Both matters we'll take up shortly.

Now, one feature of this method of representing these subfamilies in particular, a feature of our decision to do so with a group of recursive rules expanding, and is that it allows for two separate parses of examples such as (19).²

(19) the considerate big gift and small donation



(There is a third parse here too.) We might note that there are two meanings attached to this string as well, having to do with how the meaning of *considerate* is combined with the meanings of the rest of the parts. There is some reason for thinking that these two syntactic representations map onto those two interpretations. For one thing, the number of meanings and the number of parses matches.

²“X’” is a notational variant of “X.”

The strings belonging to Adverb Phrases are so simple that it is difficult to know whether they contain the substructure we've found in these other families. Nonetheless, they do have a recursive part and this might be construed, on analogy with these other cases, as evidence for substructure:

- (21) Sally carefully spoke.
 Sally very carefully spoke.
 Sally very, very carefully spoke.
 Sally very, very, very, carefully spoke.
 ⋮

The coördination phenomenon also seems to suggest subphrases, at least if our decision about the meaning-form mapping made above is correct.

- (22) Sally spoke [almost [very rapidly] and [quite softly]].

So, let's convert the AdvP rule we have built to something like:

- (23) a. $\text{AdvP} \rightarrow \overline{\text{Adv}}$
 b. $\overline{\text{Adv}} \rightarrow \text{Deg } \overline{\text{Adv}}$
 c. $\overline{\text{Adv}} \rightarrow \text{Adv}$

The rule building sentences, IPs, is similarly meager. But it too shows some signs of the subfamilies which we have discovered in NPs, APs and VPs. This is indicated by coördination in examples such as (24).

- (24) Jerry [can speak loudly] but [can't speak clearly].

And, when we add to our observations that adverbs can fall to the left of Infl, we discover the recursive flag of these intermediate families:

- (25) Jerry evidently can't speak.
 Jerry evidently deliberately can't speak.
 Jerry evidently occasionally deliberately can't speak.

(These are all somewhat strained, I grant you, but I think still grammatical.) This calls for a change along the following lines:

- (26) a. $\text{IP} \rightarrow \left\{ \begin{array}{c} \text{NP} \\ \text{CP} \end{array} \right\} \bar{\text{I}}$
 b. $\bar{\text{I}} \rightarrow \text{AdvP } \bar{\text{I}}$
 c. $\bar{\text{I}} \rightarrow \bar{\text{I}} \text{ AdvP}$
 d. $\bar{\text{I}} \rightarrow \text{I VP}$

Note how in this rule, unlike the others we've formulated, the rule that terminates the recursion has more than just the "head" of the phrase in it, in this case it also introduces the VP. This is required because VPs are not recursively introduced, and the method we have adopted of representing recursion in these phrases is built into the structure of the substitution classes.

Actually something similar is true for the rules that build APs, NPs and VPs as well. In the case of VPs, the NP and CP parts of their family are not recursively introduced. So we should change the terminal expansion to:

$$(27) \quad \overline{V} \rightarrow V \text{ (NP) (CP)}$$

And similarly, the CP parts of the AP and NP families are not recursively introduced, so the terminal expansions of these families should be changed to:

$$(28) \quad \begin{array}{l} \overline{A} \rightarrow A \text{ (CP)} \\ \overline{N} \rightarrow N \text{ (CP)} \end{array}$$

So this corrects the omission of CP and NP in our original formulation of these rules, though, as foreshadowed above, this will produce a difficulty.

To see this difficulty, consider how our structural method of stopping the recursion relates the terms that are within some phrase. We expect that those terms which are introduced in the terminal expansion “ $\overline{X} \rightarrow X \dots$ ” (that is, the non-recursively introduced terms) will form the most inclusive substitution class of the phrase involved. There are some kinds of phenomena which suggest that this expectation is fulfilled. There are processes, for example, in which a rather surprisingly short string can substitute for one or another of the families we have discovered. This happens under conditions of anaphora.³ For example, the term *one* which prosodically looks like a word can semantically substitute for any of the expansions under anaphoric conditions.

- (29) a. I will examine the blue book about language if you will examine the brown one.
 one = “book about language”
 b. I will examine the big blue book about language if you will examine the small one.
 one = “blue book about language”
 c. I will examine the long book about language if you will examine the one about Quarks.
 one = “long book”

The reason we think that *one* not only semantically is an \overline{N} , but is also syntactically an \overline{N} , is because of the contrast in (30), a contrast which also supports our treatment of the non-recursive parts of NP.

- (30) a. I will examine the long proof that language exists if you will examine the short one.
 one = “proof that language exists”
 b. *I will examine the long proof that language exists if you will examine the one that it doesn’t.
 one = “long proof”

³“Anaphora” refers to processes in which a phrase in one position refers, in some fashion or other, to the same things that another phrase, in a different position, refers to. For instance, in the sentence: *Mary knows that she is smart*, it is possible for *she* to refer to the same individual that *Mary* refers to. In such a case, we say that *she* is **anaphoric** to, or with, *Mary*.

What this contrast indicates is that *one* must “stand in” for the noun and the CP that follows, and cannot stand in for the noun by itself. This is explained if *one* can stand in for an \bar{N} , because there is no \bar{N} , under the current rule set, that fails to contain these both. It isn’t, incidentally, that there is some semantic constraint on *one* that prevents it from standing in for something that has the meaning of a single noun, because that is possible in cases like (31).

(31) I will examine the book on the shelf if you will examine the one on the table.

The difference between (31) and (30b) is just whether the material that combines with *one* is allowed to be a sister to an \bar{N} or not: PPs are (look back at (8)), and CPs aren’t.

Similarly, the \bar{V} family can be anaphorically connected to other \bar{V} s by way of silence.

- (32) a. Although Sally shouldn’t Δ , Jerry must leave town.
 Δ = “leave”
 b. Although Sally can carelessly Δ , Jerry must carefully read *Aspects*.
 Δ = “read *Aspects*”
 c. Because Jerry frantically read *Aspects* after dinner, Sally did Δ just before class.
 Δ = “frantically read *Aspects*”

This process of anaphora — called “VP Ellipsis,” though it might be more accurate to call it “ \bar{V} Ellipsis” — reveals that the non-recursive parts of the VP family are trapped within the smallest subfamily.

- (33) a. * Although Sally shouldn’t Δ Chicago, Jerry must leave New York.
 Δ = “leave”
 b. * Although Sally didn’t Δ that she was tired, Jerry said that he would sleep.
 Δ = “say”

These processes also indicate that there are at least some PPs that must be part of the terminal expansions of \bar{N} and \bar{V} .

- (34) a. ?? I will listen to this long examination of quarks, if you will listen to the *one* of syntax.
one = “long examination”
 b. * Although Sally didn’t Δ about George, Jerry will carelessly talk about Sal.
 Δ = “carelessly talk”

So we should change these rules to:

- (35) $\bar{N} \rightarrow N$ (PP) (CP)
 $\bar{V} \rightarrow V$ (NP) (PP) (CP)

This way of distinguishing the recursive and non-recursive parts also predicts that the non-recursive parts will always come between the head of their phrase and the recursive parts. This is not entirely true, however. It does seem true sometimes, as in (36) and (37).

- (36) a. Jill ate it at noon.
 b. *Jill ate noon it.
- (37) a. Jill ate spätzle at noon.
 b. *Jill ate at noon spätzle.

But for other cases it seems uncertain, or downright wrong:

- (38) a. Jill ate the rotting kumquats.
 b. Jill ate at noon the rotting kumquats.
- (39) a. ??Jill said [that you shouldn't eat kumquats] at noon.
 b. Jill said at noon [that you shouldn't eat kumquats].

This then, is the difficulty in trying to place CPs, and certain NPs, within VP (and other phrases, as we'll see). Let's set this problem aside, momentarily it will be the subject of a lot of our work in the next couple days.

Now there is a similarity to the organization of the family of substitution classes that make up NP, VP, AP, AdvP, and IP. The other phrases: PP, CP, DegP and DetP are too anemic for us to see that structure, so we don't know, empirically, whether or not they have it. (Except for DP and PP, where we will eventually see that they do.) But, following Chomsky's injunction to make our job solving the "Poverty of the Stimulus" problem, we would do well to accept as the null hypothesis that they are in fact organized along the same guidelines. This is because doing so is a step towards shrinking the space of grammars through which the learning device has to search. Here, then, is an illustration of how explanatory adequacy can help make a decision between two descriptively adequate grammars.

So giving all these phrases the shape that NPs, VPs, etc. do, we end up with a family structure of substitution classes like that below.

$$\begin{array}{ll}
 (40) \quad \overline{CP} \rightarrow \overline{C} & IP \rightarrow \left\{ \begin{array}{c} (NP) \\ (CP) \end{array} \right\} \overline{I} \\
 \quad \overline{C} \rightarrow C \, IP & \overline{I} \rightarrow I \, VP \\
 \\
 \quad NP \rightarrow \left\{ \begin{array}{c} (DetP) \\ (NP's) \end{array} \right\} \overline{N} & VP \rightarrow \overline{V} \\
 \quad \overline{N} \rightarrow AP \, \overline{N} & \overline{V} \rightarrow AdvP \, \overline{V} \\
 \quad \overline{N} \rightarrow \overline{N} \left\{ \begin{array}{c} PP \\ AP \\ CP \end{array} \right\} & \overline{V} \rightarrow \overline{V} \left\{ \begin{array}{c} PP \\ AdvP \\ CP \end{array} \right\} \\
 \quad \overline{N} \rightarrow N \, (PP) \, (CP) & \overline{V} \rightarrow V \left\{ \begin{array}{c} PP \\ AP \\ CP \end{array} \right\} \\
 \\
 \quad AP \rightarrow \overline{A} & AdvP \rightarrow (DegP) \, \overline{Adv} \\
 \quad \overline{A} \rightarrow \overline{A} \, PP & \overline{Adv} \rightarrow Adv \\
 \quad \overline{A} \rightarrow AdvP \, \overline{A} & \\
 \quad \overline{A} \rightarrow A \, (PP) \, (CP) \, (IP) &
 \end{array}$$

$$\begin{array}{l} \text{PP} \rightarrow \overline{\text{P}} \\ \overline{\text{P}} \rightarrow \text{P NP} \end{array}$$

$$\begin{array}{l} \text{DetP} \rightarrow \overline{\text{Det}} \\ \overline{\text{Det}} \rightarrow \text{Det} \end{array}$$

These all conform to the following shape.

$$\begin{array}{l} (41) \quad \overline{\text{X}} \text{ Skeleton: } \text{XP} \rightarrow (\text{ZP}) \overline{\text{X}} \\ \quad \quad \quad \overline{\text{X}} \rightarrow \text{QP } \overline{\text{X}} \\ \quad \quad \quad \overline{\text{X}} \rightarrow \overline{\text{X}} \text{ WP} \\ \quad \quad \quad \overline{\text{X}} \rightarrow \text{X (YP) (UP)} \end{array}$$

Where ZP is called the *Specifier* of XP, WP, QP are called *Adjunct(s)*, and YP/UP the *Complements* of X.

Now it should be said that these rules leave out considerable detail. In particular, there are a wide range of things that can stand in adjunct position which are not indicated in these rules. For example, $\overline{\text{V}}$ can have an AP adjoined to it, as in:

$$(42) \quad \text{Sandy saw a man today [angry at her].}$$

And, as noted earlier, an $\overline{\text{N}}$ can have certain kinds of CPs adjoined to them:

$$(43) \quad \text{the [book [that Mary read] [which no one will admit writing]]}$$

I will continue to leave out this detail, invoking it where necessary as we go along.

This procedure, as we've noted, characterizes grammatical arrangements of words in terms of the words' categorial status. It throws sets of words together into nouns, verbs, adjectives and the like irrespective of the particular verbs, nouns, adjectives, etc. that they are. Thus it forms some pretty odd sentences:

- $$(44) \quad \begin{array}{ll} \text{a.} & \text{Jerry danced with pickles.} \\ \text{b.} & \text{Jerry danced at noon at midnight.} \\ \text{c.} & \text{Jerry slowly stood still.} \\ \text{d.} & \text{a green idea} \end{array}$$

Odd, but still recognizable as grammatical strings of words.

But some combinations which these rules characterize seem to go bad in a very different way; consider:

- (45) a. Jerry laughed Mary.
 b. Sam gave it at Jill.
 c. Sally died that you should eat better.
 d. Jim claimed to Kris.
 e. Jerry slapped.

These don't go together into weird meanings; instead, they just don't go together.

What's wrong, here, is that we've matched up verbs with material following them incorrectly.

- (46) a. Jerry laughed.
 b. Sam gave it to Kris.
 c. Sally died.
 d. Jim claimed that you should eat better.
 e. Jerry slapped Mary.

Here, then, is something more particularly about the words themselves that seems to be relevant to the procedure that recognizes grammatical strings.

There's another respect in which the particular choices of words — in this example, verbs — seems to play a role in the syntax. Consider the different contributions the NP *Tuesday* makes in these sentences:

- (47) a. I danced Tuesday.
 b. I remember Tuesday.

In the first case, we say that *Tuesday* is a modifier. In this case, it modifies the sentence's meaning by restricting the times the events denoted by "I left" to just those that transpire on Tuesday. But this is not the role it has in the second case. Here it refers to the thing remembered. We say in this case that it is an argument of the relation that remember denotes.

A similar contrast can be seen in the pair in (48).

- (48) a. I kissed her on the bus.
 b. I put her on the bus.

Again, *on the bus* is a modifier in the first case. It locates the events denoted by "I kissed her" to just those that took place on board the bus. In the second case, by contrast, it names one of the locations related by put. It is an argument.

The semantic role an argument has in some sentence is determined by the word for which it is an argument. The meaning that modifiers contribute to the sentence they're part of are considerably more constant.

There's a way of talking about argumenthood that is commonplace, and which we inherit from Gruber (1965). Gruber, like Fillmore in a paper published at about the same time, was concerned with the problem of verb meaning, and in particular with finding a theory that restricted the kinds of argument types that verbs permit. He speculated that there was a finite, in fact quite small, set of argument types, or 'roles',

that could be put together by verbal meanings. In particular, he argued that all roles which verbs combined were ones that had to do with the logic of motion.

So, for example, a verb like *send* involves three terms, one that can be seen as indicating the Source of the motion, another that denotes the moved term and a third that names the Goal of that motion. Gruber called the role borne by the term undergoing motion ‘Theme.’

- (49) Jeff sent his book to Ray.
 Source Theme Goal

This requires that the logic of motion admit of extensions, as in cases like the following.

- (50) Jeff showed his book to Ray.
 Jeff pleases Ray.

It should be noted, however, that there are cases which even metaphorical extensions of the logic of motion look unlikely to characterize:

- (51) Jeff made Ray happy.
 Jeff kissed Ray.
 Jeff wrote his book for Ray.

This way of constraining the roles that verbs may manipulate has been pretty firmly abandoned, I believe, as a consequence. And in its place, a method has been pursued that predates Gruber’s hypothesis, one that tries to see the roles which verbs use as a product of a small number of elemental predicates, like CAUSE, MOVE, etc., which make up their meaning. Still, the language we inherit comes from Gruber, who named the roles that he conjectured verbs combined after his chief one: Theme. He called them “Thematic Roles,” usually abbreviated to Theta-Role or θ -role.

The relation between verbs and their arguments expressed by θ -roles can be seen as a special instance of a more general relationship which goes under the name selection, or sometimes s-selection (for “semantic” selection). This refers to the connection between a verb’s (or other similar term’s) meaning and the semantic value that its arguments deliver. θ -roles express a similar function: they name the meaning that an argument’s semantic value must be compatible with. But the relation holds for other cases too, where the language of θ -roles doesn’t extend so easily. One of those places is where verbs connect with clauses of various types. So, a verb’s meaning determines somehow whether the clause it combines with must have the meaning of an interrogative or a declarative, for example.

- (52) a. Martha knows that John has left.
 b. Martha said that John has left.
 c. * Martha wonders that John has left.
 (53) a. * Martha knows whether John has left.
 b. Martha said whether John has left.
 c. Martha wonders whether John has left.

We say of these cases that verbs select or s-select a question or declarative. Note that some verbs are compatible with either, as is *say*.

Though it is hard to see these differences as fitting the roles that θ -roles typically name, I will use the language of θ -roles to describe this relation too. I am keeping here with the sloppy usage often employed in the literature.

Now, what the contrast between (45) and (46) indicates is that verbs also specify what “category” their argument must be. So, as the pair below shows, this is another property that distinguishes verbs.

- (54) a. Jerry likes Mary.
 b. * Jerry likes to Mary.
- (55) a. * Jerry talks Mary.
 b. Jerry talks to Mary.

We say that verbs are *subcategorized* by the category of their argument. Or — this term has been relexicalized — that verbs subcategorize their arguments. Sometimes this is also described as a verb *c-selecting* its argument.⁴

Jackendoff (1977) argues that arguments (when they follow the head they are an argument of) are necessarily in complement position. This is supported by contrasts like:

- (56) a. Although Sally didn’t Δ Tuesday, she will dance Monday.
 b. * Although Sally didn’t Δ Tuesday, she will remember Monday.
- (57) a. Although Sally won’t Δ on the bus, she will kiss her in the car.
 b. * Although Sally won’t Δ on the bus, she will put her in the car.

Because the phrase following the verb is an argument in (56b) and (57b), it must be within the \bar{V} which elides, whereas in (56a) and (57c), the phrase following the verb is a modifier and can therefore remain outside the ellipsis.

Jackendoff’s thesis is also supported by similar contrasts involving *do so* anaphora, which, like \bar{V} Ellipsis, finds \bar{V} s.

- (58) a. Sam talked to Mary on Tuesday, and Sally *did so* on Thursday.
 b. Gerry eats chocolate after dinner, and Sandy *does so* before lunch.
- (59) a. ?* Sam talked to Mary and Sally *did so* to George
 b. * Gerry eats chocolate, and Sandy *does so* marzipan.
 c. * Mag proved that she loved chocolate, and Holly *did so* that she loved marzipan.

If we examine the position that *do so* may stand, we will find that it has the same distribution as \bar{V} s: it may appear between a subject NP and a sentence final PP, as in (58). In this respect, then, *do so* is like *one* in that it is a lexical expression of a certain phrase. If this is granted, then the ungrammaticality of the examples in (59) indicates that the material following the verb in these cases must be within the smallest \bar{V} . This would explain why this material cannot be positioned outside of *do so*. And what distinguishes the cases in (58) and (59) is that the phrases following *do so* in (58) are arguments, whereas those in (59) aren’t.

⁴The terms *s-select* and *c-select* come from Pesetsky (1982).

It's difficult to be certain of the argument-status of terms which combine with nouns, but to the extent that the system of θ -roles we just reviewed can be identified in NPs, Jackendoff's claim seems correct here as well.

- (60) a. I will examine your careful placement of it on the canvas, if you will examine my clumsy *one*.
one = "placement of it on the canvas"
- b. *I will examine your careful placement of it on the canvas, if you will examine my clumsy *one* on the collage.
one = "placement of it"
- c. I will examine your careful discussion of it in the manifesto, if you will examine my clumsy *one* in the summary.
one = "discussion of it"

The ungrammaticality of (60b) follows from what we have discovered about *one* pronominalization if the PP *on the canvas* must be within the smallest \bar{N} . The difference in grammaticality between (60a) and (60b), then, reflects the fact that *on the canvas* must be within the \bar{N} which has been lexicalized as *one*. The difference between (60b) and (60c), flows from the fact that *on the canvas* is an argument of *placement* in (60b), whereas the PP in (60c), *in the manifesto*, isn't. It follows, that is, if Jackendoff's conjecture that arguments, but not modifiers, must be positioned within the smallest of the phrase that contains them.

The judgments I report for (60) are often not sharply perceived, perhaps because of the difficulty in determining whether the PPs involved are arguments or not. So, let me offer another example of the same kind, but for which judgments tend to be a bit more pronounced.

- (61) a. I'll examine your long, careful discussion of it, if you'll examine my short *one*.
one = "careful discussion of it"
- b. *I'll examine your long, careful discussion of it, if you'll examine my short *one* of it.
one = "careful discussion"
- c. I'll examine your long, careful discussion in that book, if you'll examine my short *one* in this book.
one = "careful discussion"

As in (60), the contrast between (61a) and (61b) will follow if *of it* must be positioned within the smallest \bar{N} . The contrast between (61b) and (61c) corresponds to the differing argument-status of the PPs involved: *of it* is more strongly perceived as an argument of *discussion* than is *in that book*. As with (60b,c), then, this contrast supports the hypothesis that arguments and modifiers are fit into phrases in different positions.

Okay, to summarize: we're looking for a way to factor into our procedure for recognizing grammatical sentences enough of the meanings of the words involved to guarantee that Verbs and Nouns (at least) combine with the arguments they select and subcategorize. Moreover, when these arguments follow them, we must find a way of guaranteeing that they are in the non-recursive \bar{X} .

Let's concentrate, to begin with, on "complements," the arguments that show up after the verb or noun. We can ensure that these arguments are in the non-recursive part of the \bar{X} if we force them to bear a θ -role, and allow θ -roles to be assigned only to these complement positions. We need also to describe the fact that when a verb has a θ -role, there must be an argument present in the syntax which bears that θ -role. It is customary to divide this task into two parts, which can be expressed as follows:

- (62) THE THETA CRITERION (preliminary)
- a. For every θ -role there is a position to which that θ -role is assigned.
 - b. For every θ -position, there is something with an appropriate semantic value that occupies that position (i.e., the argument).

It is usual to strengthen the Theta Criterion to a bijection, because of cases like (63).

- (63) Sally showed John *doesn't mean* Sally showed John himself.

So we change this to:⁵

- (64) THE THETA CRITERION
- a. For every θ -role there is exactly one position to which that θ -role is assigned.
 - b. For every θ -position, there is exactly one thing with an appropriate semantic value that occupies that position (i.e., the argument).

When we add to this the fact that verbs also specify the categories of their arguments, we get something like:

- (65) for every θ -role assigned by some X^0 , X^0 c-selects the phrase that bears that θ -role.

As we'll see momentarily, this statement of the relation is too strong.

We need to worry about cases like the following, of course, in which there appears to be an optional argument.

- (66) a. Martha ate (pie).
b. It seems (to me) that Marty left.

Here we might imagine either that there actually is an object in these cases that bears the θ -role, or, alternatively, that something relaxes the condition guarantees that every θ -role projects a position. The common wisdom is that both possibilities exist — we will return to this issue in some detail later. For now, let us imagine that there is a lexically determined process which allows θ -roles for certain predicates to not project a position.

Some have suggested that (62b) and (65) should be collapsed, and in particular, that there is a means by which the categorial type of some argument can be

⁵The Theta Criterion is also often formulated in terms of a bijective relation between θ -roles, or θ -positions, and arguments. That is, it is sometimes written to say: "For every θ -role (or positions) there is exactly one argument and for every argument there is exactly one θ -role (or position). (In *Lectures on Government and Binding* it is formulated in various ways, including these two.) The difference between this alternative formulation and the one I've given here is that mine does not force every argument to receive a θ -role, whereas the alternative does. I've decided to place this requirement in another principle, which we'll come to shortly.

determined from its θ -role. (The reason c-selection is usually thought to be derivable from s-selection, rather than the other way round is tied to Chomsky's "epistemological priority" argument, see Pesetsky (1982).) Grimshaw (1979) provides a way of viewing this hypothesis which has gained some popularity. Her idea is that one of the functions that makes up the learning device assigns a categorial status to arguments on the basis of their θ -role. She calls this function "Canonical Structural Realization" (CSR). She sketches how this function might work by way of examples that compare CPs with NPs.

So let's look at some of the facts she considers. Note first that CPs may distinguish themselves as according to whether they denote Propositions, Exclamatives or Questions. Let's suppose that these can be assimilated to the language of θ -roles. These θ -roles can sometimes be borne by NPs too:

- (67) a. John asked me $\left\{ \begin{array}{l} \text{what the time is} \\ \text{the time} \end{array} \right.$ (Question)
- b. I'll assume $\left\{ \begin{array}{l} \text{that he's intelligent} \\ \text{his intelligence} \end{array} \right.$ (Proposition)
- c. Bill couldn't believe $\left\{ \begin{array}{l} \text{how hot it is} \\ \text{the heat} \end{array} \right.$ (Question)

In these cases, then, the verbs s-select either Q, P or E and c-select either an NP or CP.

There are other verbs, however, which s-select these very same θ -roles, but c-select only CPs.

- (68) a. John wondered $\left\{ \begin{array}{l} \text{what the time was} \\ \text{*the time} \end{array} \right.$ (Question)
- b. I'll pretend $\left\{ \begin{array}{l} \text{that he's intelligent} \\ \text{*his intelligence} \end{array} \right.$ (Proposition)
- c. Bill complained $\left\{ \begin{array}{l} \text{how hot it was} \\ \text{*the heat} \end{array} \right.$ (Exclamative)

Here then, we have a special instance of the difference in s-selection and c-selection that needs to be overcome if one is to be derived from the other.

Grimshaw's suggestion is that the CSR of Questions, Propositions and Exclamatives is CP and that those verbs which allow these θ -roles to be borne by NPs are learned on a piece-by-piece basis. Thus, this is a partial collapse of c-selection to s-selection. And it predicts that every verb that s-selects a Q, P or E will c-select a CP; that is, there should be no verbs that express these θ -roles with an NP only. This seems to be correct.

Whether or not this project can be maintained for the situation involving the relation between CPs and NP and the θ -roles they bear, I don't think a parallel story holds for the complements of other categorial type. Moreover, the scheme Grimshaw proposes won't help determine which verbs select non-finitees as opposed to finite clauses, which also seems to be a rather language particular fact. So, from now on let us assume that c-selection is at least in part independent of s-selection, and determined on a verb-by-verb basis.

Interestingly, however, it looks like the program which would derive the category type from θ -role might fare better when external arguments are concerned. Here, first, it looks like we get only NPs and CPs phrases of other types are banned from Specifier position (at least in English).⁶ And second, when a θ -role is consistent with either NP and CP, any kind of CP is possible as is an NP:

$$(69) \left\{ \begin{array}{l} \text{That John left} \\ \text{To have to leave} \\ \text{Leaving} \\ \text{The fact} \end{array} \right\} \begin{array}{l} \text{bothers Mary.} \\ \text{makes Mary happy.} \end{array}$$

And only when the θ -role is incompatible with the meanings that CPs yield are they banned from Specifier position:

$$(70) \left\{ \begin{array}{l} * \text{That John left} \\ * \text{To have to leave} \\ * \text{Leaving} \\ \text{John} \end{array} \right\} \begin{array}{l} \text{kisses Mary.} \\ \text{likes Mary.} \end{array}$$

So, let us conclude that only complements are c-selected. This will require weakening (65) to something like (71).

- (71) a. An X^0 c-selects its complements.
b. If an X^0 c-selects Y, then it θ -marks Y.

We can summarize what we've discovered so far as follows.

- (72) a. If a verb has a θ -role, then there is exactly one syntactic position to which that θ -role is assigned.
b. A θ -marked position must be occupied by something with the appropriate semantic value.
c. A verb c-selects its complements.
d. c-selection entails s-selection (aka θ -role assignment).

The statements in (72a) and (72b) are the Theta Criterion, whereas those in (72c) and (72d) concern the relation between c-selection and s-selection which we've just reviewed. The Theta Criterion insists that for every θ -role that some verb has, there will be a unique position occupied by an argument in the sentence holding that verb. (72c) and (72d) determine whether that argument will be c-selected or not.

To force arguments to be within the smallest \bar{X} , it will now be sufficient to force the θ -position for that argument to be within the smallest \bar{X} . We want this effect for complement arguments only — we don't want to force "subject" arguments into \bar{X} — so one way of doing this would be to restrict those positions that are c-selected to just those within the smallest \bar{X} . This would mean that we'd have two principles: one that determines the c-selected position for verbs, and another, yet to be determined, which locates the s-selected position for subjects. We're going to see, however, that

⁶With the exception of cases like "Under the bed is a slipper," plausibly instances of impersonal constructions with inversion; see Stowell (1981) and Rochement and Culicover (1990).

the procedure for locating the θ -positions for both subject and object arguments is the same, and so we won't take precisely this course.

Instead, we will follow a popular view of these principles that is first found in Chomsky's *Lectures on Government and Binding*. There he formulates what he calls "The Projection Principle," which is responsible for mapping the argument structure of a verb — or head more generally — into a syntactic representation. I will formulate his principle as follows:⁷

(73) THE PROJECTION PRINCIPLE

- i. For α , a position, if α is a sister to X^0 , then X^0 c-selects α 's contents.
- ii. If α s-selects β , then α and β are sisters.

The second part of the Projection Principle does what we are in search of. It forces arguments of a verb to be in the lowest \bar{X} , for only in that position will it be a sister to the verb.⁸ Note that this principle is not restricted to verbs and their projections, it spreads what we've discovered about VPs to all other categories. This, so far as I know, is correct.

As presently formulated, the second part of the Projection Principle wrongly forces subjects into the smallest \bar{V} of the verb that assigns it a θ -role. We will see, however, that this problem is only apparent. Once we discover what is truly responsible for assigning the subject its θ -role, this problem is resolved.

The first part of the Projection Principle is just (72c). It has the interesting consequence of preventing non-arguments from standing in the smallest \bar{X} . Thus, the Projection Principle not only has the consequence of forcing arguments into the smallest \bar{X} , but also of forcing non-arguments out of this position. Whether this stronger result is correct is rather difficult to determine. We will eventually examine some phenomena that bear on it.

An interesting consequence of the Projection Principle is that it factors into the lexical specification of the verbs everything needed to know what sort of phrases will be found in the non-recursive part of \bar{X} s. Stowell argues, in fact, that it should only be found here. This sort of information doesn't properly reside in the phrase structure rules, since it is information that is tied to the particular choice of words, and not the pure form that sentences may take. In fact, the information phrase structure rules give about the contents of the smallest \bar{X} can now be seen as merely a summation of what is possible across particular lexical items filling the head slot. Thus, we should factor out of the phrase structure rules information which concerns the categorial nature of the complements involved. We can do the same for the subject arguments as well, since their categorial nature is derived from their θ -role.

What we have seen, then, is that the phrase structure rules can be stripped of a great deal of their information; indeed, what is left is largely what the skeleton expresses and the categorial specification of non-arguments. There is some hope, I think, for the view that the categorial specifications for non-arguments will follow entirely from the meanings that categorial types may have. So it might be that the fact that \bar{V} can combine with non-argument PPs, AdvPs, CPs and NPs may follow entirely from

⁷Chomsky's own formulation builds in various other properties that we will encounter later on; see in particular the discussion on pp. 34–48, especially p. 38, in *Lectures on Government and Binding*.

⁸X and Y are *sisters* if every phrase including one includes the other.

the meanings that categories of these types may carry. Thus, CPs may denote “reasons” (say, as in *because* clauses) and AdvPs can denote manners, and PPs can denote locations or times, as can NPs, and these are just the sorts of things that allow for combination with \bar{V} s to form other \bar{V} s. Similarly, it might be that \bar{N} s may only combine with the types that PPs, AdjPs and CPs belong to because these are the only types that, once joined with \bar{N} s, produce another \bar{N} . Let us suppose that this is the case. (Note that the range of categories possible in these positions is relatively free, suggesting that there are few, if any, constraints on category type.) Since the inventory of categories varies from language to language we might, just to be safe, factor this information out of the phrase structure rules into a language particular set of statements of the form in (74).

- (74) a. If α modifies \bar{N} , then α must be ...
 b. If α modifies \bar{V} , then α must be ...
 :
 :

in which the “...” will carry lists of category types.

If this project is successful, then the PS-rules of English collapse in full to the \bar{X} Skeleton. Some have suggested, in fact (see Travis (1984), for example), a picture of language variation that makes the hierarchical arrangements of constituents that the \bar{X} Skeleton, together with the PROJECTION PRINCIPLE and THETA CRITERION and whatever yields (74), completely immutable. All that varies is the linear order in which the terms that follow the arrows in the \bar{X} Skeleton may have. So, the meta-grammar might look something like:

- (75) a. $XP \rightarrow \{\alpha, \bar{X}\}$
 b. $\bar{X} \rightarrow \{\bar{X}, \beta\}$
 c. $\bar{X} \rightarrow \{X^0, \gamma\}$

The linear arrangements of these constituents must then be determined by the language particular part of the grammar. So far as I know, it is an informal agreement among syntacticians that whatever it is that determines the order of Specifier and \bar{X} is independent of what determines the order of heads and their complements. There is no widely agreed upon account of what is responsible for this factor, so let’s leave this for the future. It is also sometimes thought that the linear order of adjunct and \bar{X} is fixed independently of the order of head and complement. In English, for example, complements always follow their heads, whereas adjuncts can often precede them. Indeed, there is considerable freedom in the order of adjuncts and \bar{X} in English, a freedom which is not mimicked by the head-complement relation.

On the other hand, Greenberg’s typological work suggests that the ordering of these two phrases are related (see Dryer (1992)). What fixes the order of adjunct and \bar{X} is not well understood, so let’s leave that too to the future. As for the relation between head and complement, it is sometimes held that there is a “headedness parameter” that specifies whether the head of a phrase may come initially or finally in its (immediate) projection. This predicts that complements will either all precede or follow their heads, and not come among them. While this is not superficially true (German/Dutch, for

example), it does look like this could be true of the underlying arrangements of these constituents.

So, you might imagine that there is an evaluation metric that allows for choices of the following sort:

- (76) a. Specifier: [first, last]
 b. Projection of X: [first, last]

This would connect the linear order of head to complement with the linear order of head to adjunct, which Dryer's work suggests might be there.⁹

We might see "languages," then, as compilations of the \bar{X} Skeleton plus value assignments to the options that determine the ordering of constituents plus a lexicon, whose members determine the categorial specifications for arguments. Thus, there's a sense, then, in which "languages" do not actually have PS-rules — these are just the epiphenomena that emerge when the various factors of UG and language particular information are combined.

However, note that this brings with it the conclusion that what fixes the order of complements will not be part of PS-rules. Because this information was conveyed in PS-rules by way of referencing category type, it is no longer possible for the PS-rules to convey this when they have been purged of categorial information. There must be another component of the grammar which expresses this information.

⁹See also Saito (1985) and Saito and Fukui (1998).

Chapter 3

Order of Complements

We've made a whole-scale revision to the information that phrase structure rules provide, placing much of what they formerly said into the lexicon and general statements about the structural relationship that arguments and non-arguments have to the phrases they are within. Concretely, in place of the elaborate phrase structure rules we began with, we've adopted the set of statements below.

- (1) \bar{X} Skeleton: $XP \rightarrow (ZP) \bar{X}$
 $\bar{X} \rightarrow QP \bar{X}$
 $\bar{X} \rightarrow \bar{X} WP$
 $\bar{X} \rightarrow X (YP) (UP)$
- (2) THE THETA CRITERION
 - a. For every θ -role there is exactly one position to which that θ -role is assigned.
 - b. For every θ -position, there is exactly one thing with an appropriate semantic value that occupies that position (i.e., the argument).
- (3) THE PROJECTION PRINCIPLE
 - i. For α , a position, if α is a sister to X^0 , then X^0 c-selects α 's contents.
 - ii. If α s-selects β , then α and β are sisters.
- (4)
 - a. An X^0 c-selects its complements.
 - b. If an X^0 c-selects Y , then it θ -marks Y .
- (5)
 - a. If α modifies \bar{N} , then α must be AP, PP or CP
 - b. If α modifies \bar{V} , then α must be AdvP, PP, CP or NP

⋮

The Skeleton forces a certain shape on all arrangements of word classes, setting up the head, intermediary projection, maximal projection arrangement. The linear ordering it imposes on the terms in each of its expansion reflect , perhaps, a particular setting of the universal options of Specifier first or last, and \bar{X} first or last. The linear ordering of $\{\bar{X}, WP\}$ and $\{\bar{X}, QP\}$ terms will have to be determined on the basis of the particular phrases that fill these positions — what controls this remains unknown. The Projection Principle in conjunction with (4) ensures that θ -positions are sisters to the terms that

assign the θ -roles, and that when the term assigning a θ -role is a head, that this position is also c-selected. This leaves the problem of correctly determining the subject's θ -position — a problem whose solution we are working towards. The Theta Criterion ensures that for every θ -role associated with some predicate, there will be exactly one θ -position in the syntax, and that this position will be occupied by an argument. This, together with the Projection Principle will correctly place arguments deeper than non-arguments within the phrase that contains them. Finally (5) lists the categorial status that modifiers may have, depending on the term that is being modified.

We ended last chapter with the observation that this system will require some additions to capture all of the information that our former phrase structure rules expressed. In particular, we noted that the linear arrangement that the complements have will need to be derived in some fashion. Before turning to that particular need, however, let's make a few other additions, addressing some other information loss that has occurred in our transition from phrase structure rules to this system.

Consider first the optionality of phrases. The Skeleton makes all of the phrases within some maximal projection optional. The presence of modifying phrases is, in fact, completely optional. The presence of complements is determined by the existence of θ -roles: if the head of the phrase has θ -roles, then phrases in these positions will be obligatory, forced by the Theta Criterion and the Projection Principle. What about the phrase in Specifier position? In general, the phrases in these positions are optional.

Occasionally, phrases in specifier positions are forced by processes that are not well understood. For instance, the presence of something in Specifier of NP seems to be determined by whether the head noun is singular or plural:

- (6) a. I like horses.
- b. I like the horse.
- c. *I like horse.

This might be due to a semantic effect, though there is considerable language variation here whose source is not known. Or another possibility is that there is a determiner present even in (6a), but that it is silent. This would allow for the possibility that Specifier of NP is obligatorily filled, accounting, then, for the ungrammaticality of (6c). Let's leave the status of the specifier of NP open, for the moment. We'll have a chance to revisit this issue when we examine more closely the structure of noun phrases.

Phrases in specifier of IP, however, buck the trend and are always obligatory. This was one of the consequences of our original phrase structure rules that has been lost. In cases where the IP contains a subject argument, the obligatoriness of this subject is plausibly derived in the same way that the obligatoriness of complements is: by the Theta Criterion and the Projection Principle. Somehow or other, the Projection Principle is going to have to be fixed so that it guarantees that there is a θ -position for the subject argument, and the Theta Criterion will force an argument into this position. If the θ -position happens to be specifier of IP, then this will guarantee the presence of something in specifier of IP. But, interestingly, even in IPs whose verbs do not have a θ -role associated with a subject argument the presence of something in specifier of IP is obligatory. The verb *seem*, for instance, has only one θ -role, and that is assigned to its clausal complement. And yet, as (7) shows, an IP containing this verb must surface with something in its specifier position.

- (7) a. It seems that we are behind.
 b. *Seems that we are behind.

The *it* in (7a) appears to have no meaning whatsoever, and is merely present in order to occupy specifier of IP. This it is called an “expletive” or “pleonastic” term, to indicate its lack of semantic content.

To recapture this bit of information, Chomsky proposes in *Lectures on Government and Binding* adding another statement to the Projection Principle which simply requires that specifier of IP be filled. This is known as the extension to the Projection Principle:

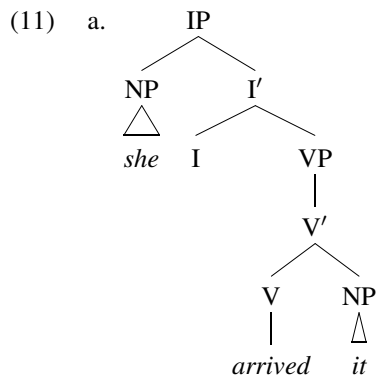
- (8) EXTENSION OF THE PROJECTION PRINCIPLE
 The Specifier of IP must have a phrase in it.

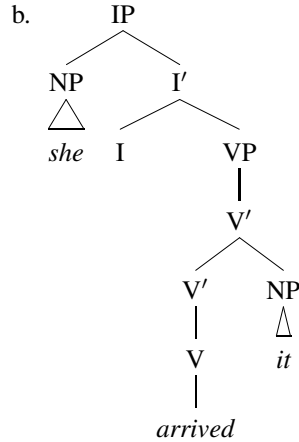
In the normal case, when the verb of a sentence has a “subject” θ -role, the Extended Projection Principle will be satisfied by the presence of an argument phrase, whose presence will also satisfy the Theta Criterion. But in the relatively rare case when the verb does not have a subject θ -role, it will still demand the presence of something, and the expletive is invoked as a consequence. Notice that this system restricts the use of the expletive to just those circumstances where there is no subject θ -role. When the verb of some sentence has a subject θ -role, the Theta Criterion will require that the Specifier of IP be occupied by an argument and expletives, by virtue of their semantic vacuity, cannot function as arguments. This, then, is why (9) does not allow the *it* in specifier of IP to be interpreted as an expletive.

- (9) It discovered the problem

In fact, the distribution of expletives is extremely limited. They are found only in Specifier positions to which no θ -role is associated. As we’ve seen, they’re not found in Specifier positions associated with θ -roles, but they are also not found in complement or adjunct positions. Thus an example like (10) is ungrammatical on either of the parses indicated in (11).

- (10) *She arrived it.





The parse in (11a) is already blocked by the Projection Principle, the first clause of which requires that things in complement position be c-selected by the neighboring head. This isn't the case in (11a).

I don't know of anything in the literature that is explicitly designed to exclude (11b), so I suggest that something along the lines of (12) is responsible.

- (12) $\{\overline{X}, \alpha\}$ iff α modifies \overline{X} , where “ $\{\gamma \beta, \alpha\}$ ” indicates that α and β are daughters of γ .

(Note that the “ $\{ \}$ ” notation expresses hierarchical relations without expressing linear ones.) This is a strengthening of one of the ways of defining the syntactic relation in which modification is expressed that was the object of our first homework assignment. It not only allows one thing to modify another only if they are sisters, but further requires the modification requirement if one of the sisters is an \overline{X} and both are dominated by an \overline{X} . Because expletives, by virtue of being semantically vacuous, cannot modify, they will be banned from this position.

There are other ways of blocking (11b), of course, and there is no particular reason to believe that this is the correct method. But let's adopt this principle until something better comes along.

There is one last fact in this domain that requires addressing. This is that when there is no subject θ -role, only an expletive can satisfy the Extended Projection Principle. Placing an argument in the Specifier of IP in such a case is ungrammatical, as a comparison between (7) and (13) indicates.

- (13) *Jerry seems that we are behind.

Many formulations of the Theta Criterion target this fact, requiring that there be a θ -marked position for each argument. We have a different option. Because modification is restricted to just adjunct positions, semantically contentful phrases in specifier positions are not going to be able to modify. If the only other way a meaningful phrase can be put together with the rest of a sentence is by way of a θ -role, then the ungrammaticality of (13) will follow.

Let's return now to our original problem: when there are two or more complements, what determines their order? This information, encoded in our former phrase

structure rules, is now impossible to encode in this fashion, since we have decided to follow Stowell and put this information in the subcategorization frames of the lexical items.

One thing our present system does do, is impose a linear order on complements and non-complements. In particular, because complements are trapped inside the lowest and non-complements are forced out of that, we should expect to find that complements precede non-complements when they both follow the head.

- (14) The Projection Principle entails: If an argument, X and a non-argument, Y, both fall linearly on the same side of the head, then X will come closer to the head than Y.

In fact, as we've already seen, this is only very partially true in English. It does seem true sometimes, as in (15).

- (15) a. Jill ate it at noon.
b. *Jill ate at noon it.
- (16) a. Jill ate squash at noon.
b. *Jill ate at noon squash.

But for other cases it seems uncertain, or downright wrong:

- (17) a. Jill ate the rotting kumquats at noon.
b. Jill ate at noon the rotting kumquats.
- (18) a. ??Jill said [that you shouldn't eat kumquats] at noon.
b. Jill said at noon [that you shouldn't eat kumquats].

So, we can see already that there is some work to be done by principles that impose a linear order on complements. If we've got everything right up to now, they seem equipped with the power to pervert what appears to be the hierarchical relationship between argument and non-argument.

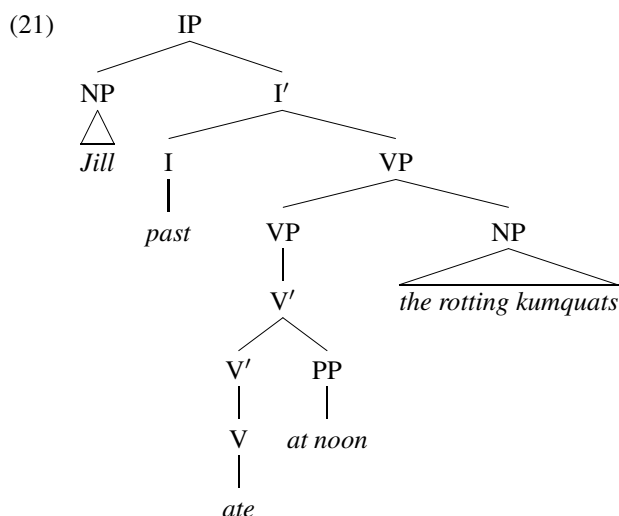
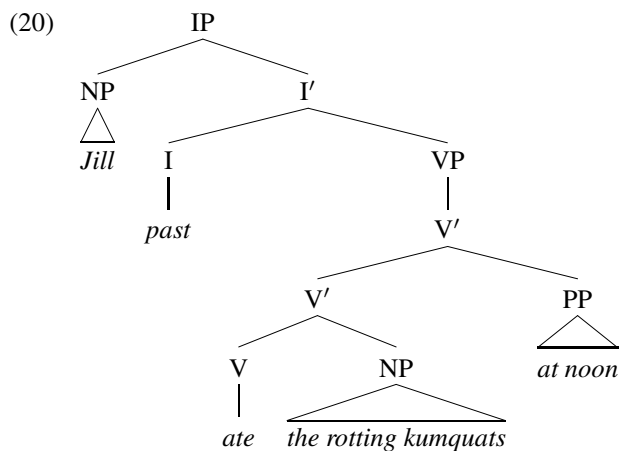
Ross (1967) proposed an influential way of thinking of the alternations we see here. He suggested that we see the cases where the argument follows the non-argument as arising by way of a Transformation: a rule that forms sentences not by combining word-classes as our phrase structure rules have but instead forms them from other sentences by moving things around. Transformations were introduced by Zellig Harris, and pushed in the direction that we currently understand them by Chomsky, to give a way of accounting for certain process that seem to require information about larger portions of strings than do phrase structure rules. As our discussion of phrase structure rules has indicated, it is possible for all of these rules to make reference to adjacent word classes, with the only exception (apparently) to this being these cases of multiple complements. But there are principles of sentence construction that need to look at quite distant points of strings — as we shall have occasion to see in detail — and these perhaps are better described by way of a separate process.

What Ross proposed for these cases is that there is a rule of NP Shift, which forms from sentences where the argument precedes the non-argument, sentences where that argument has been “shifted” to the right. His rule looks like this:

(19) NP SHIFT

Right adjoin a phrase to the first VP containing that phrase.

This rule will form the shifted sentence in (17b) from the unshifted (17a) as follows:



We can think of this sentence, then, as having two parses. One that satisfies the Projection Principle and Theta Criterion, and the other which doesn't, but is mapped from the first by NP Shift. This sentence is a set, then, made up of the two phrase markers shown above. We can see the members of this set as being generated in the following fashion: one member is produced by the Extended Projection Principle, \bar{X} Skeleton, the Theta Criterion, and the rule for modification in relation to an inventory of lexical items. The parse this produces is sometimes called a D-structure. The D-structure is the first element of the set. The other members of the set are produced by the action of transformational rules acting on elements already found in the set. We will eventually have to control the method by which transformations produce additional elements of

the set, but in the simple cases we will be concerned with right now, there are only two members of the set: the D-structure and another formed by a transformation acting on that D-structure. We only speak one element of these sets. In the case of NP Shift, it is the parse produced by the transformation. The parse that is spoken is called the “S-structure,” and these sets of parses are called “Derivations.”

Consider now how complements are ordered linearly. To start with, observe that unlike how a verb selects the category of its arguments, the way in which complements are ordered does not vary with the choice of head. We will not want to encode, as we did earlier, the order of complements on a verb by verb basis. This is something that emerges independently of the verbs. In general, as our earlier phrase structure rules encoded, verbal complements are ordered as follows.¹

(22) DP + PP + CP

- (23) a. I told Mary that she should join.
 b. * I told that she should join Mary.
 c. Mary explained to me that I should join too.
 d. ?* Mary explained that I should join too to me.
 e. Sally told the story to Bill.
 f. ?? Sally told to Bill the story.

Stowell suggests that systems of “Case assignment” are responsible for ordering these terms, and that they trigger a special instance of NP Shift. NPs differ from other kinds of phrases in IndoEuropean in being able to host Case morphology. In English this happens with pronouns only. The particular Case borne by a pronoun is determined by its syntactic position. In languages that are richer in Cases than English is we can see that the Case borne by a NP is determined by a term in proximity to the NP. In German, for instance, a certain class of prepositions and verbs determine Accusative Case for their complements, while others determine Dative Case. It is also often the case in other IndoEuropean languages that NPs other than just pronouns can bear Case morphology sometimes on the head of the NP, sometimes on the determiner, sometimes spread across modifiers of various sorts within the NPs. Let’s imagine then that, in general, NPs must be related to Case assigners. Or:

(24) CASE FILTER

An NP must be assigned Case if it is an argument.

I’ve restricted the Case Filter to argument NPs because, as we’ll see, adjunct NPs do not seem to be positioned in a way that suggests they are sensitive to Case assignment. Moreover, typically the Case morphology they bear is fixed, and not sensitive to Case assigners. The Case filter, then, requires that some parse in a sentence’s derivation puts every NP that sentence contains in a Case marked position.

As I noted before, Case marked positions are ones that are close to terms that are responsible for assigning (i.e, determining) the Case. So, all we have to do now is know what those terms are and what “close” means, and we’ll be able to use the Case filter to distinguish sentences in terms of grammaticality. The “object” Cases —

¹ As a simplifying measure, we consider only finite clauses.

so-called Accusative and Dative and Locative, and a host of others — are assigned by particular lexical items. In English we have only Accusative, and it is assigned by certain verbs and many prepositions. What assigns the “subject” Cases Nominative, in the IndoEuropean languages is less hard to identify. At present I will simply say that something assigns Nominative Case to the Specifier of finite IPs. What “close” means will be the subject of some scrutiny for us in the weeks to come. I will start out defining a particular version of “close,” whose particulars I will defend below.

- (25) a. Specifier of finite IP is assigned Nominative Case.
- b. X^0 assigns its Case to position α only if X^0 c-commands and is close to α .
- (26) α C-COMMANDS β iff:
 - i. every phrase that contains α contains β , and
 - ii. α does not contain β .
- (27) α is KLOSE to β iff there is no more than one phrase that contains β but not α .

In much of the literature, you will find that “ α c-commands and is close to β ” is rendered as “ α governs β .” So an alternative formulation is:

- (28) a. Specifier of finite IP is assigned Nominative Case.
- b. X^0 assigns its Case to position α only if X^0 governs α .
- (29) α GOVERNS β iff:
 - i. α c-commands β , and
 - ii. there is no more than one phrase that contains β but not α .

Stowell (1981) proposes to derive the ordering of complements by way of the Case Filter. His first suggestion is that Case is assigned not only under government but also under adjacency; thus:

- (30) α assigns Case to β only if α and β are linearly adjacent.

This will guarantee that an NP comes immediately adjacent to the verb (its Case assigner), and therefore before all other complements.

With regard to the relative order of PP and finite CP, Stowell suggests using Ross’s NP Shift operation. He argues that if we can make CPs obligatorily undergo this operation, then we can not only derive why they follow other complements, but also derive that they tend to follow other non-complements as well. All we need do, then, is find a way of making CPs especially partial to NP Shift. He speculates that finite CPs, like NPs, must receive Case but that unlike NPs they cannot sit in Case-marked positions at S-structure. Because of this final requirement, they must be moved by S-structure to some non-Case marked position. NP Shift is capable of doing this, and thus, in cases where a complement CP shows up string finally in a VP, it has satisfied Stowell’s injunction against surfacing in Case marked positions by undergoing NP Shift. There is, surprisingly, a certain amount of evidence for this picture.

First, as predicted, and briefly noted in (17b), it really does seem to be the case that finite CPs must follow all non-complements as well.

- (31) a. Mary believes sincerely that Joan should leave.
 b. * Mary believes that Joan should leave sincerely.
 c. Sally shouted loudly that Milly should read.
 d. * Sally shouted that Milly should read loudly.
 e. Sam remembered yesterday that Sally was president.
 f. * Sam remembered that Sally was president yesterday.

This is to be expected if finite CPs are necessarily “extraposed,” as NP Shift is sometimes called when it applies to clauses.

But there is an important confound here. There is a processing effect, known as “minimal attachment,” which leads us to favor placing phrases into the most recently processed clause. This effect would also push us towards judging sentences, like those in (31b,d,f), in a way that favors placing the modifier in the embedded clause. Maybe this blocks our ability to recognize the other reading as grammatical.

Consider, however, situations where the finite clause is the “subject” of the clause. Here too, following an argument from Koster (1978), we see that there is some reason for thinking that it isn’t actually in Nominative Case-marked, Specifier of IP position. Koster’s argument makes reference to a process that is found in certain question-types in English. Normally, in these question contexts, it is possible to move I^0 to the front of a sentence, as in (32).

- (32) Will Mary put the book on the table?

How precisely this is done is the subject of a later class. What’s relevant here is that this process is blocked if something has moved and adjoined to the left of IP, as in (33).

- (33) a. On the table, Mary will put the book.
 b. * Will on the table, Mary put the book?

Now, interestingly, for many speakers of English the presence of a finite CP as a subject of a sentence also blocks movement of I^0 .

- (34) a. That Mary has left should be obvious.
 b. * Should that Mary has left be obvious?
 c. That Sally sleeps late bothers Mittie.
 d. * Does that Sally sleeps late bother Mittie?

This would follow if finite CPs are driven from the nominative Case marked Specifier of IP, and adjoin to the left of IP in these cases. Stowell’s method of forcing NP Shift to apply to complement CPs would extend to this scenario as well. CPs start out in the nominative Case-marked position, but are driven from there in order to form an S-structure.

This evidence all points to the fact that finite CPs move. But is there evidence for the motivation for this movement that Stowell proposes? In particular, is there motivation for the fact that finite CPs, like NPs, require Case?

One piece of suggestive evidence comes from the class of verbs that permit both NPs and finite CPs. These are only just those verbs that already exceptionally allow two NPs: *promise*, *tell*, *show*, etc.

- (35) a. Mary promised me that she would sing.
 Mary promised me the ring
 b. Jerry told me that he can't stand Mary's singing.
 Jerry told me the story.
 c. Sheila showed me that she cares.
 Sheila showed me her concern.

This isn't completely the case, as (36) is a counterexample.

- (36) a. Mary persuaded Bill that he should go.
 b. * Mary persuaded Bill the fact

But, so far as I know, (36) is the only counterexample. To the extent that there is a match in the verbs which accept NP CP and those which accept NP NP complements, there are grounds for believing that their surface positions are governed by the same, or similar, principles. And to the extent that the dominant principle is the Case Filter, then there is reason to conclude from these data that CPs are subject to the Case Filter as well.

This match between the distribution of NPs and CPs should be completely general if Stowell is correct. Indeed, finite CPs are distributed in sentences much like NPs. We find them as complements to verbs, as we have seen, and in the subject position of other finite clauses, but not in the subject position of infinitives (as we shall see shortly). These are just the Case-marked positions. But there are several differences in their distribution. In English, finite CPs are probably never found as the complements to a preposition, though of course NPs are. The only potential counterexample comes from temporal prepositions, as in (37).

- (37) a. I left before Mary arrived.
 b. * I left before that Mary arrived.

Similarly, both adjectives and nouns can take CP complements, but not NP complements.

- (38) a. Sean is unhappy that he had to sing.
 b. * Sean is unhappy that.

- (39) a. the proof that lemons cure cancer
 b. * The proof this fact.

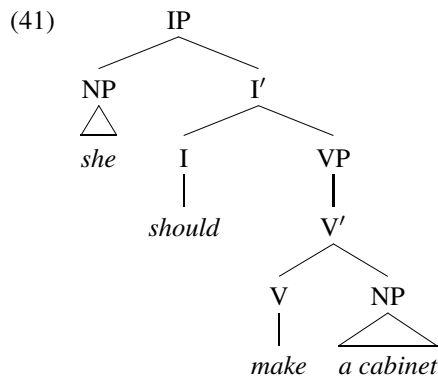
If this has a Case Theoretic explanation, then Stowell's theory is in trouble. But it could also be that this arises because of some property of c-selection.

Though the evidence suggests that Stowell's ideas meet with some success, there are problems for this system. One which threatens Stowell's Adjacency Condition on Case assignment, and its use in fixing the order of complements, concerns so-called "double object" constructions, as in:

- (40) a. Mary showed Bill the picture.
 b. Bill baked Sally a cake.

In these examples, how can the second NP receive Case? We need an account that not only explains how both these NPs can pass the Case Filter, but one that also explains the fact that the first is the argument that it is (Goal) and the second is the Theme. We will have to develop some of the rest of the system that is responsible for linearizing arguments before we can engage this difficulty. So, let me continue to ask for indulgence.

There is a use that we can put the Case filter to that is somewhat related to Stowell's. The Case filter plays a role in solving a problem we have postponed discussing for some time now: How is it that the second clause of the Projection Principle be correct, given the presence of subjects whose θ -role is determined by a verb. The second clause of the Projection Principle, recall, allows θ -roles to be assigned under sisterhood only. Subjects are obviously not sisters to the verbs they get their θ -role from:



I also remind you, now that this tree has re-revealed it, that there is something wrong in our phrase-structure rules characterizing VPs as well. Remember that one of the hedges I made was in calling the family of strings that describes something different than the family of strings that VPs describe. We haven't seen any reason for doing this; and yet it was this hedge that was one of the steps I took towards collapsing the battery of phrase structure rules into the Skeleton.

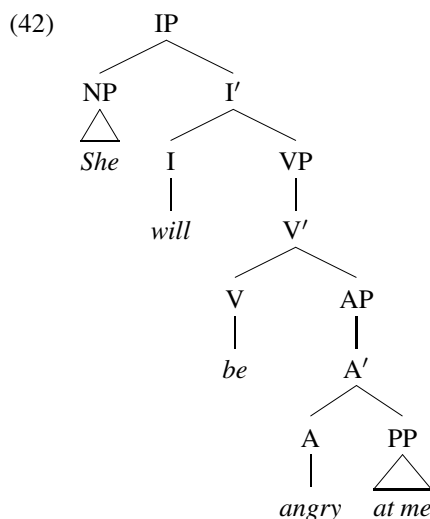
I had a similar hedge, incidentally, in connection with APs and PPs, which is reflected in (42) on the following page. It's now time to close this problem, at least for VPs and APs.

What is it that distinguishes VPs from \bar{V} s and APs from \bar{A} s? VPs and APs can have "subjects" in them, and \bar{V} s and \bar{A} s can't. This can be seen in certain special circumstances, of which (43) are examples.

- (43) a. I will let [her make a cabinet].
 b. This will make [her angry at me].

As indicated by the brackets, it is thought that the strings following *let* and *make* form a single phrase which is serving as the complement to these verbs.

Why don't we instead believe that these verbs are followed by two complements, the first of which is *her*? Because this would make this first phrase an argument



of the verb, thus place it in complement position, and we can see from other examples that this isn't the case. In (44), for example, this position can be occupied by an expletive.

- (44) a. I will let [it seem that there is a mistake on the handout].
 b. This will make [it obvious that there is a mistake on the handout].

We've already discovered that expletive *it* can only stand in Specifier positions. Thus, the positions following *let* and *make* in these cases are Specifier positions.

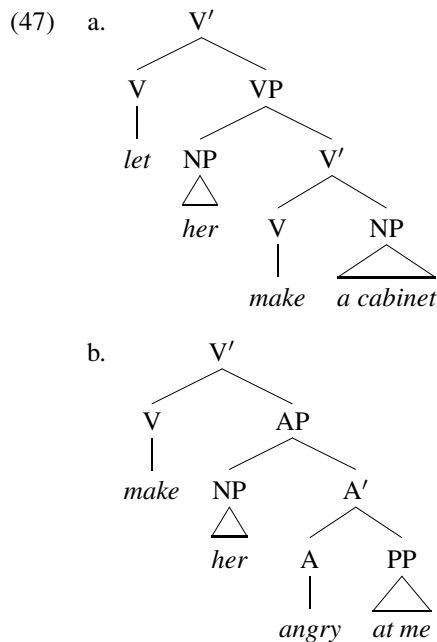
Next, note that these are Specifier of AP and VP, and not Specifier of IP. This, we know, because strings which follow these specifier positions are characterized by the rules which generate \bar{V} s and \bar{A} s, and not those which characterize \bar{I} s:

- (45) a. * I will let [her should make a cabinet].
 b. * I will let [her to make a cabinet].
 c. * I will let [her makes a cabinet].

Finally, why don't we think that these NPs are not inside the \bar{V} or \bar{A} ? I can't show you a reason for this in the case of \bar{A} ; but for \bar{V} we can tell this from the action of the *do so* phenomenon. Recall that *do so* replaces \bar{V} s; so the contrast in (46) indicates that the NP in (43a) is not in \bar{V} .

- (46) a. I will let [her make a cabinet] and I will let [him do so] as well.
 b. * I will let [her make a cabinet] and I will let [do so] as well.

From these observations we will want to give the examples in (43) the parses in (47).



These complements to *let* and *make* are sometimes called “small clauses,” and it is to Stowell, once again, that we owe this way of viewing them.

Here, finally, I am able to show you why we defined *klose* the way we did in connection with the principles governing Case assignment. In particular, these examples indicate why *klose* wasn’t defined as sisterhood, or something along those lines, as it could be if we restrict our attention just to cases where complements are involved. In these examples, the NPs get Accusative Case (witness the form of the pronouns), and this by virtue of their proximity to the verb on their left. But, as we’ve seen, these NPs are not complements to the Case assigning verb, and hence not sisters to the verbs. Nonetheless, the verbs govern the NPs because they c-command them and there is no more than one projection above the NPs that does not contain the verbs (there is just one).

Now, these examples were introduced in order to solve the problem for the Projection Principle which subjects pose. So, let’s see how these cases help.

Note, first, that the subject of the small clause in (43a) bears the same θ -role that the parallel subject in (41) bears. Maybe, then, despite appearances, the position the subject in (41) gets its θ -role from is Specifier of VP. That’s not where we see it, but maybe that’s because it moves. This isn’t quite yet a solution to the problem for the Projection Principle because even if the subject of (41) were in Specifier of VP to receive its θ -role, it still wouldn’t be sister to the verb it gets its θ -role from.

But wait: who said it gets its θ -role from the verb? True, the examples we’ve looked at indicate that the meaning the subject contributes to the sentence varies as a function of the verb. But it’s possible, I suppose, that we could describe this fact in terms of the subject getting its θ -role from. Indeed, this is just the position that Marantz advances. He notes that the specific content of the subject’s contribution to the sentence

varies as a function not just of the verb, but also of the verb in combination with the material that follows it; witness:

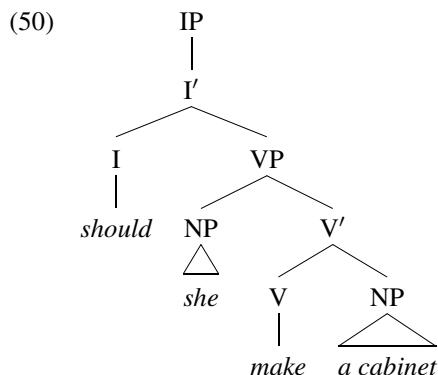
- (48) a. She should make a cabinet.
 b. She should make an effort.
 c. She should make like a bird.
 d. She should make nookie.

This doesn't seem to be true of the θ -role complements bear, however. It isn't possible to find examples where the specific content of the object's θ -role varies as a function of, say, the contribution the subject makes. (We will eventually see that other phrases, however, do seem to make a contribution that influences an object's θ -role.)

So maybe we should rescue the Projection Principle by allowing certain \bar{V} s to assign θ -roles.

- (49) If α assigns a θ -role, then α is a head or the maximal \bar{X} in some XP. The "external" θ -role of a predicate is assigned by the maximal \bar{X} .

This would make the D-structure representation for (41) something like (50).



This solution, then, claims that, in some weird way, subjects are not visibly in the position from which their θ -role derives. Maybe this seems too far to go for the Projection Principle's sake, but we have other cases where this type of discontinuity is required anyway.

Consider, for instance, what happens when (43b) involves the Passive:

- (51) She was made [angry at me].

The θ -role that *she* bears is the same one that it bore in (43b), and yet it shows up in a different Case and in a different position. Or consider the similar (52).

- (52) She seems [angry at me].

Again, *she* is the subject of *angry at me*, not *seems*. And yet, just as in (51), the subject of *angry at me* shows up in a higher Specifier of IP.

(52) is like another kind of example we shall have occasion to encounter frequently in this class:

(53) She seems [to be angry at me].

In this example too, *she* receives its θ -role from the lower clause: *she* is the external θ -role bearer of *angry at me*.

In all of these examples, we seem to need to let arguments appear in positions distant from their θ -role assigners. Whatever we come up with for these examples can be applied to the simpler case in (50).

How are we to account for these situations — ones where an argument shows up somewhere other than its θ -marked position? It is common to follow Ross's technique for solving cases where the linear relationships we'd expect from the Projection Principle fail to materialize. We speculate that the subject in these situations undergoes a Transformation which moves the argument into the Specifier of IP we see it in.

Returning now to (50), note that we would expect the subject to be unable to stand in the position this parse places it because of the Case filter. Specifier of VP in this example, unlike the one in (43), is not Case marked. In particular, *make* does not assign Accusative Case to *she* in (3.46) since *make* does not c-command the Specifier of VP. By contrast, *make* does c-command *her* in (47), and this is what allows it to assign Case to *her* and, consequently, for this subject to surface in Specifier of VP.

Interestingly, the ungrammaticality of (54) suggests that the Specifiers of the APs in (51) and (52) are not Case marked.

- (54) a. * It was made [her angry at me].
- b. * There was made [a child angry at me].
- c. * It seems [her angry at me].
- d. * There seems [a child angry at me].

At least from what we've seen up to now, these examples should be grammatical: the THETA CRITERION and the PROJECTION PRINCIPLE are satisfied. If *her* and *a child* do not stand in a Case marked position, however, the ungrammaticality of these examples would follow. But, at least in the case involving *make*, it is unlikely that this is because *made* fails to govern the Specifier of AP which holds *her*. We've seen in (43b) that when *make* is in the active mood, it is capable of governing and Case marking the subject of the lower small clause, and we have not detected any reason for believing that passivizing *make* causes a change to this feature of the parse. We do have evidence, on the other hand, that Passive prevents a verb which is normally capable of assigning accusative Case from assigning that Case. (We will come to this evidence only very slowly, I'm afraid.) We won't look at Passive in detail right away, but in general a description of its effects is:

- (55) PASSIVE
 - a. Add passive morphology and embed under *be*, and
 - b. Rob the passivized verb of its Accusative Case, and
 - c. Rob the passivized verb of its external θ -role.

This produces changes to a verb that causes it to be able to be placed into a sentence like (56b) in addition to one like (56a).

- (56) a. Jerry likes her.
b. She is liked.

Thus, the ungrammaticality of (54a) can be credited to the Case filter again, but now not because of a failure of government, but instead because the passive *made* does not assign Case.

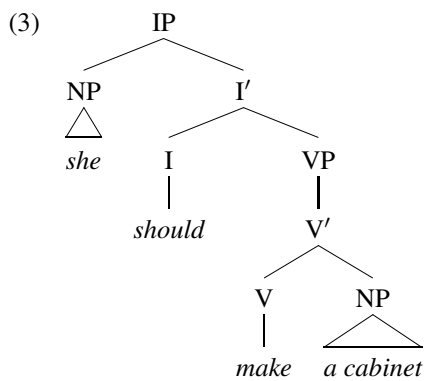
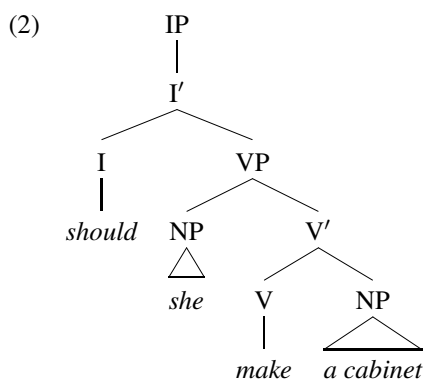
We suspect the same is true for (54b): in general, it appears that *seem* has no Accusative Case to assign. There are, then, at least some verbs which do not have Accusative Case to assign: those that have undergone the Passive operation are in this class too.

Chapter 4

Argument Movement

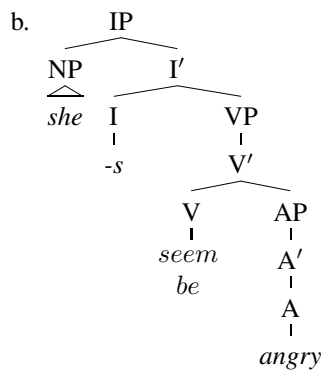
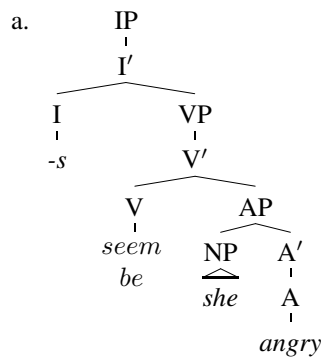
We left examining the possibility that the subjects of clauses get their θ -role from \bar{V} , in Specifier of VP, and from there move into the position from which their Case derives. So, a typical finite clause, like (1), might have the underlying representation in (2) and the surface representation in (3).

(1) She should make a cabinet.

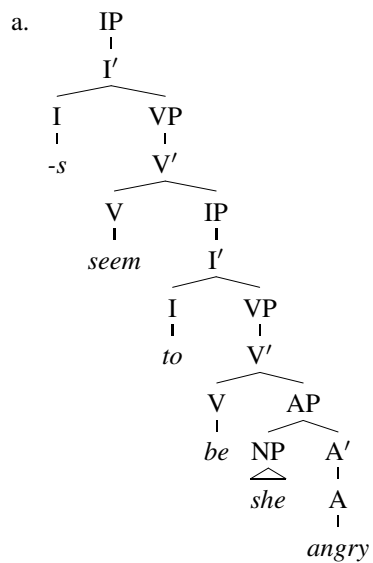


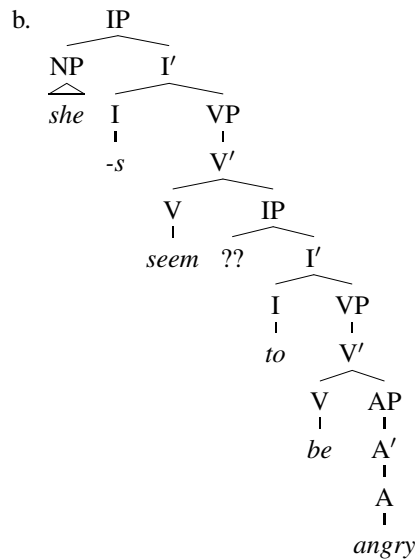
I tried to suggest that we find a similar situation in (4) and (5).

- (4) She seems angry.
She is angry.



- (5) She seems to be angry.





Notice that in this last example we have a puzzle concerning the Extension to the Projection Principle — what lies in the Specifier of the infinitival IP? One popular idea is that the derivation includes still another parse, one in which *she* stands in the Specifier position of the lower IP. On this view, the Extension to the Projection Principle is satisfied in this sentence because its derivation has an element in which it is satisfied. Let's adopt his solution for the moment; we'll return momentarily to some evidence in support of it.

We compared cases such as these with ones like (6) on the next page, where Case is delivered directly to the position in which the external θ -role is assigned.

The example in (5) has a superficial similarity to (7), but there is an important difference between them.

(7) She tried [to be angry]

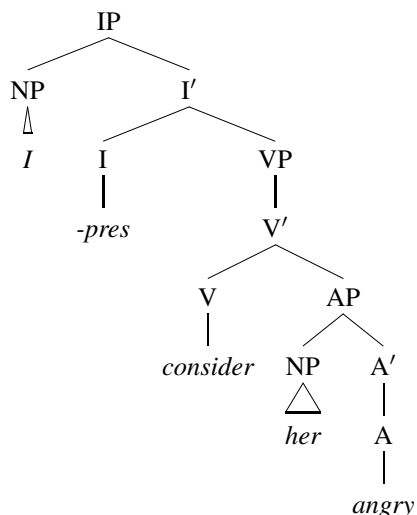
In (5), the count of θ -roles and arguments matched, as expected. But in (7), this isn't the case. Both *tried* and *angry* assign an external θ -role, and yet there is only one argument which seems to bear them both: *she*. This is precisely what the Theta Criterion prohibits. Recall, that the Theta Criterion is designed in this way in an attempt to capture the fact that (8a) doesn't have the meaning that (8b) does.

- (8) a. I showed John.
b. I showed John himself.

So, either we should find another explanation for the fact that (8) illustrates, or we should find a way of resolving (7) with the Theta Criterion.

Note also that (7) seems to be a violation of the Extension to the Projection Principle. Moreover, it's not just that (7) seems to counterexemplify these principles, it perversely can't satisfy them. If the Specifier of the embedded IP is given content, and in so doing hosts an argument for the θ -role that *angry* assigns, the result is ungrammatical.

- (6) I consider her angry.



- (9) *She tried [him to be angry].

Infinitives with this peculiar mix of properties are called control clauses, another example of a which is (10).

- (10) Sally prefers to eat chocolate.

Again, as with (7), there is expected to be a subject argument for *eat chocolate* here. Interestingly, however, in this case such an argument can be found:

- (11) a. Sally prefers for him to eat chocolate.
 b. Sally prefers him to eat chocolate.

We think that the *for* in (11a) is a Complementizer, for several reasons. One is that the *for*+infinitive string appears to be a single phrase, rather than two independent ones. This is indicated by, for instance, the behavior of these phrases in the context of cleft constructions.

- (12) a. It's for him to eat chocolate that Sally would prefer.
 b. *It's to him how to eat chocolate that Sally should explain.

The ungrammaticality of (12b) derives from the fact that there is only room for one phrase between the *it's* and the *that* of these clefts, and in (12b) two things, a PP and an infinitive, have been shoved into this spot. Thus the grammaticality of (12a) would seem to argue that we do not want to parse the underlined string as a PP followed by an infinitive, but instead as something that makes a single phrase. To the extent that an infinitive is an IP, and that the *him* in (11) is its subject, then one of the few ways of doing this is to let *for* be a complementizer.

Moreover, the optionality of *for* is reminiscent of an ability that the complementizer *that* has:

- (13) a. Sally said that he eats chocolate.
b. Sally said he eats chocolate.

Moreover, the constraints on C^0 Deletion, or whatever process is responsible for the optionality of the complementizer *that*, are mimicked by the optionality of *for*. In general, the complementizer *that* can go unspoken only in contexts where the CP it heads is a complement to a verb. That is why *that* cannot go missing in when it heads a CP that has been clefted, as in (14). A parallel constraint on *for* ellipsis is indicated by (15).

- (14) a. It's that he eats chocolate that Sally said.
b. *It's he eats chocolate that Sally said.
(15) a. It's for him to eat chocolate that Sally would prefer.
b. *It's him to eat chocolate that Sally would prefer.

In general, the constraints on where *that* can go missing are parallel to those on the optionality of *for*, suggesting that they are both subject to an ellipsis process that singles out complementizers.

So, what makes the subject of the control clause in (11) able to be expressed, whereas the one in (7) can't be? One thought is that this correlates with the presence of *for*. Indeed, the control clause following *try* is unable to host *for*.

- (16) *She tried [for him to eat chocolate].

This makes sense, actually, if we view it through the lens of the Case Filter. Granting that *for* is a Case assigner, we can say that the subject of infinitives is allowed to be expressed if Specifier of IP is Case-marked. (Here we have to assume that *for* assigns Case before it deletes.)

Now, this suggests that the infinitive following *try* isn't an IP. This is because *try* can assign Case, witness:

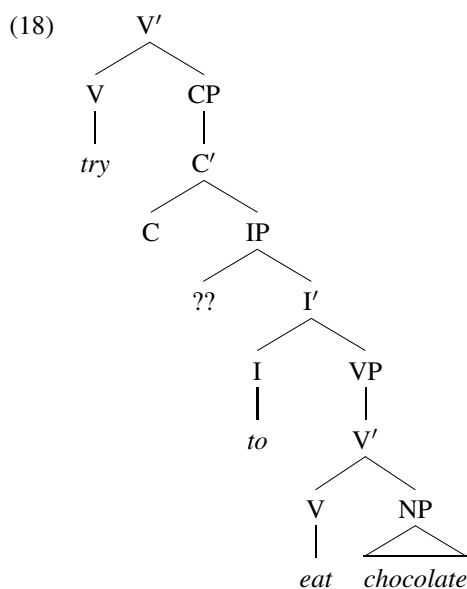
- (17) Sally tried the task.

So the reason that (9) is ungrammatical can't be because *try* is a verb like *seem* which doesn't support accusative Case assignment. But if the infinitive following *try* is a CP with no lexical item associated with its head, as in (18) on the following page, then the failure of Case assignment into the infinitive's Specifier position will follow. Of course, this doesn't yet help us with the problem this example poses for the Projection Principle and the Theta Criterion. But it shows us, anyway, that we are not looking for something that distinguishes infinitival clauses from finite ones.

Instead, it looks like we are searching for something that distinguishes Case marked subject positions from non-Case marked subject positions. In fact, as (19) shows, there is a perfect correlation with Case marking concerning whether or not the Theta Criterion is satisfied.

- (19) *Sally prefers [for to eat chocolate].

This correlation can be expressed this way with (20).



- (20) a. If the Specifier of IP is Case marked, then the external θ -role bearer must be expressed in it.
 b. If the Specifier of IP is not Case marked, then the external θ -role bearer cannot be expressed in it.

There's another puzzle that control infinitives pose, one that's closely tied to the puzzle of the missing subject. That's that when the \bar{V} in a control infinitive has no θ -role to assign, the presence of the missing subject creates ungrammaticality!

- (21) a. * Sally tried [to seem [that she likes chocolate].
 (compare: Sally said [that it seems [that she likes chocolate].)
 b. * She prefers [to seem [that she likes chocolate]].
 c. She prefers [(for) it to seem [that she likes chocolate]].

This suggests, somewhat surprisingly, that the Theta Criterion is in force. One way to capture the correlation in (20), and the facts above, is to imagine that there is an element that bears the external θ -role even in those cases where there apparently isn't any subject; imagine it's a kind of pronoun, call it PRO, and let it be licensed in Specifier of IPs that don't have Case assigned to them.

- (22) PRO is ungrammatical, unless in a non-Case marked Specifier of IP.

For the facts above to be captured, we must understand PRO to be an argument; in particular, this is what will force the clauses with PRO to have an external θ -role. We will have more to say about where PRO is found in a few paragraphs.

There is independent support for this solution which we'll come to shortly.

So what we've seen so far is that the Case Filter has applications beyond Stowell's. It can be used to characterize when the overt NP arguments are found and when the covert ones are found. And, now returning to the topic we began with, it can be used to characterize the positions in which arguments appear, even when these aren't the positions in which their θ -roles are determined.

These latter situations, ones where arguments are spread across two positions, can be characterized in a number of ways. Some, for instance, have attempted to formulate a rule that allows θ -roles to be "percolated" up the tree so that it gets assigned directly to the position that the argument lies in. But another method — the one that we assumed in the previous chapter — is to adopt the kind of solution that Ross did for the similar problem with complement positions. That is to rely on a Transformation, maybe like this one:

(23) ARGUMENT MOVEMENT

Move a term which is related to a θ -marked position.

I use this vague "related to a θ -marked position," rather than the more straightforward "is in a θ -marked position" in anticipation of later developments.

Sportiche's article presupposes this solution to the argument spread phenomenon, and argues that Argument Movement also plays a role in relating an external θ -role bearer in Specifier of IP with a θ -marked position in Specifier of VP. Thus, his article is a defense of the solution to the problem of external θ -role assignment that we have adopted. His argument is based on the phenomenon of Quantifier Float, which can be illustrated by the English and French examples in (24).

- (24) Les enfants (*tous) ont (tous) vu ce film (tous).
The kids (all) have (all) seen this film (*all).

Roughly speaking, "Q-float" names a process which relates a certain class of quantifiers with the terms that function as their restrictors. These examples have a meaning very like that found in:

- (25) All the kids have seen this film

So, again, there is a kind of discontinuous dependency in these examples. Note also that there are language particular differences in this process that are reflected in the differing grammaticality judgements for some of the French and English examples.

I will try to rest from this phenomenon an argument for using Argument Movement to characterize the argument spread phenomenon. We will see other, perhaps more compelling, arguments for Argument Movement as we go along. But the Q-float phenomenon does contain the ingredients for one such argument, and so we shall begin our catalogue of reasons to believe in Argument Movement here.

Sportiche's argument relies on the observation that the relation between the Floated Quantifier and the argument it is related to mimics the relation that is found in cases where an argument is spread across two positions. Thus, for example, this discontinuous relation is constrained in such a way that the position the argument is found in always c-commands the other, θ -marked, position. Examples like (26) are ungrammatical.

- (26) a. * It seems [to think [that she is possible that I like chocolate]].
 (compare: She seems [to think [that it is possible that I like chocolate]].)
 b. * It seems to [her mother] [– angry at me].
 (compare: She seems to my mother [angry at me].)

If we express the argument spreading phenomenon in terms of Argument Movement, these facts would require that we place the following constraint on it.

- (27) When an argument moves from α to β , β must c-command α .

A similar constraint is seen in Q-Float. It isn't possible for (28b) to have an interpretation like that which (28a) gets, presumably because it is not possible for Q-float to relate these two structures. And because (28b) can have no other interpretation — there are no other plural NPs for *tous* to be construed with — the sentence is ungrammatical.

- (28) a. [L'auteur de tous ces livres] a vu ce film.
 the author of all these books has seen this film
 b. * [L'auteur de ces livres] a tous vu ce film.
 the author of these books has all seen this film

The “Argument Spreading” phenomenon is also constrained in such a way that the position the argument appears in cannot be related to a position within certain kinds of clauses. For example, it cannot relate an NP outside a control infinitive to a position within that control infinitive, as in (29).

- (29) *Marty was preferred [to be angry at me].
 (≈ It was preferred that Marty be angry at me.)
 (compare: Marty was made [to be angry at me].)

If we express the argument spread phenomenon in terms of movement, and remembering that control infinitives are CPs, we can constrain it with something like (30).

- (30) An argument cannot move from α to β if α is within a CP and β is not within that CP.

The Q-Float relation observes a similar constraint:

- (31) a. * Les enfants l'ont persuadé [de tous acheter ce livre].
 the kids him-have persuaded to all buy this book
 b. * The kids have persuaded him [to all buy this book].

Sportiche suggests that the similarity of these constraints can be explained if “Q-Float” is a product of the same process that produces argument spread, and since he assumes that this is Argument Movement, he concludes that Q-float is produced by Argument Movement. In particular, he suggests that from a D-structure representation which puts the quantifier and NP together is derived an S-structure by moving the NP and stranding the quantifier.

The pre-Sportiche accounts of this fact are not very attractive. Jaeggli (1982), for example, suggests that *tous* is an anaphor, like the reflexive pronoun *herself*, *himself*, which, interestingly, is subject to the same constraints we've just identified for Q-float.

This seems stipulative, however, as there is no independent way in which *tous* is like an anaphor. (And as we will see, there is direct evidence against this.)

A more promising account comes from Belletti (1982), who suggests that *tous* undergoes the Movement operation and joins its argument, yielding a representation where these two are combined. (A level of representation that feeds the interpretive component). This would capture these properties of the relation, and in particular, their similarity with the Movement operation.

Sportiche, by contrast, argues that Movement moves the NP, separating it from its quantifier, again explaining the similar constraints on these two operations. He provides the following argument from French data against Belletti's alternative. First, he notes that in addition to Q-float, French allows for an operation that places the quantifier to the left of the argument it is construed with. This phenomenon, first described in detail in Kayne (1975), is called "Leftward *tous* Movement," or "*L-tous*," and is illustrated by:

- (32) a. Jean aurait aimé oser rencontrer tous les enfant.
 John would-have liked dare meet all the kids
 'John would have liked to dare meet all the kids.'
- b. Jean aurait tous aimé oser les rencontrer.
 John would-have all liked dare them meet
 'John would have liked to meet them all.'

Sportiche suggests that Belletti's proposal therefore requires positing two independent *tous* movement rules whereas his does not. Not an overwhelmingly strong argument, it should be noted. It's not terribly clear that Sportiche's proposal for Q-float would capture *L-tous*, after all, and this might mean that he'd need two rules here too. And in any case, what's wrong with two rules? They could well be different processes; after all, English has Q-float but not an equivalent of *L-tous*.

In any case, if the similarities in the Q-float phenomenon and the argument spreading phenomenon are not accidental, then Sportiche's conclusion that they should trace back to the same process is warranted. And if this conclusion is warranted, then examples like (33) indicate that argument spread is implicated in the positioning of subjects of simple monoclausal examples.

- (33) The kids should all eat.

More particularly, *all* in this example marks the position that the θ -role which *the kids* bears is assigned. Because *all* in (33) is linearly in a position that is consistent with putting it within Specifier of VP, this can be seen as support for the decision to let external θ -roles be assigned by \bar{V} s to their sister positions. Thus it supports our solution to the conditions under which external θ -roles are assigned. This is the point of Sportiche's argument — to furnish support for this conclusion.

We can also see in the Q-float phenomenon an argument on behalf of Argument Movement as a means of modeling the argument spread phenomenon. This argument is built on Sportiche's conclusion that the argument spread phenomenon and the Q-float phenomenon should have the same underlying source, and the observation that the class of quantifiers which can float are just those which appear to be in construction with a maximal projection:

- (34) a. Both the children have seen this film
 b. All the children have seen this film
 c. * Some/many/two/several the children have seen this film.
 d. * The children have some/many/two/several seen this film.

The exception to this is *each*.

- (35) a. * Each the children have seen the film.
 b. The children have each seen the film.

But *each* has slightly different behavior than the other floatable Qs. Its distribution is different, for example. Unlike the other quantifiers, it can be found in circumstances like (36), in which it relates NPs, rather than merely taking one as its restrictor.

- (36) John and Mary saw *The Fugitive* and *The Shining* each.

So let's set *each* aside. See Safir (1987) for some discussion.

This parallelism can be seen as a consequence of another condition on Argument Movement, at least under Sportiche's account Q-float. Argument Movement seems unable to affect \bar{X} projections. This is, presumably, what is behind contrasts like (37).

- (37) a. Mary's mother seems [to like chocolate].
 b. * Mother seems [Mary's – to like chocolate].

In fact, this constraint seems to cut across movement operations. It's also true of Ross's NP Shift rule, as the contrast in (38) indicates.

- (38) a. I gave to Mary my photos of the hike through Glacier national park.
 b. * I gave my to Mary photos of the hike through Glacier national park.

And in Topicalization contexts, which can be characterized as coming about through a rule which moves a term to "the front" of an IP, there are also contrasts of this sort:

- (39) a. Mary's books about quantifiers, I've always enjoyed.
 b. * Books about quantifiers, I've always enjoyed Mary's.

There's support, then, for something like:

- (40) \bar{X} s cannot be moved.

The match in (34), then, can be explained by way of (40), on Sportiche's account. Only quantifiers that have maximal projections as sisters should be strandable. And this looks roughly right.

There's a very popular, and potent, alternative to Sportiche's account which you should keep in mind. I'll introduce it now, and offer a problem for it, but keep it at hand because it may begin to look more appealing when we discover some of the problems for Sportiche's analysis. This alternative collapses the conditions on so-called "subject oriented" (see Jackendoff (1972)) adverbs with those of Q-float. Note that these adverbs have a distribution very much like floated quantifiers.

- (41) The kids (deliberately) have (deliberately) seen (*deliberately) this film (*?deliberately).

To get the interpretation right, we would need a semantics that lets a floated quantifier ‘distribute’ the predicate that follows it over the parts of the subject. This approach to Q-float has many advocates: Bowers (1993), for example, and Kayne (1975). For the proper kind of semantics, see Roberts (1987), especially her chapter 3. Sportiche’s (best) argument against this account is that the distribution of subject-oriented adverbs and floated Q are not always the same cross-linguistically.

There’s another kind of argument against this which is found in Giusti (1990). Her argument is typological, and based on where floated quantifiers can be found related to objects. Note that floated quantifiers related to objects are, by themselves, something of a puzzle for an approach that would make them predicate distributors. What Giusti observes is that the languages which allow quantifiers to be related to objects are those for which we have independent evidence that objects can move.

Thus, German/Dutch and Icelandic, but not English or Danish, can have a floated quantifier related to an object, as in (42) and (43).

- (42) a. Der Lehrer har die Schüler alle gelobt.
the teacher has the students all praised
‘The teacher has praised all the students.’
(German)
- b. De leraar heeft de kinderen allen geloofd
The teacher has the children all praised.
‘The teacher has praised all the students.’
(Dutch)
- c. Stúdentarnir lasu greinina allir.
students-the read article-the all
‘The students read all the articles.’
(Icelandic)
- (43) a. *The teacher has praised the children all.
b. *Laereren roste eleverne alle
teacher-the praised children-the all
‘The teacher praised all the students.’
(Danish)

She argues that this can be related to the fact that definite objects in Dutch, German and Icelandic can scramble, i.e., move leftwards, whereas this is not possible for objects in English and Danish. That leftwards movement is possible in these languages is, at least partly, indicated by the fact that they can precede adverbs which are normally found at the left edge of VPs. We have the contrasts, then, between examples such as (44) and those in (45).

- (44) a. Der Lehrer hat die Schüler ohne Zweifel gelobt.
the teacher has the students without a doubt praised.
‘The teacher has without a doubt praised the students.’
(German)

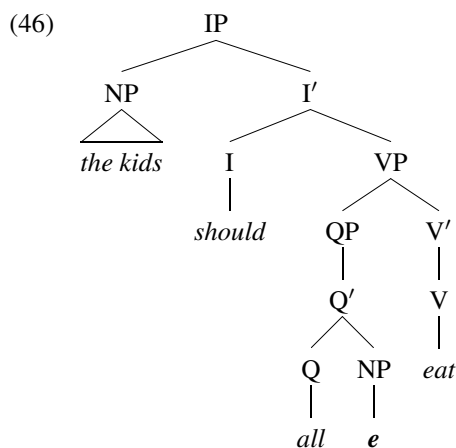
- b. De leraar heeft de kinderen gisteren geloofd.
the teacher has the children yesterday praised.
'The teacher has yesterday praised the students.'
(Dutch)
- c. Stúdentarnir lasu greinina ekki.
students-the read articles-the not
'The students didn't read the articles.'
(Icelandic)

- (45) a. *The teacher has praised the children not.
b. *Laereren roste eleverne uden tvivl.
teacher-the praised students-the without a doubt.
'The teacher without a doubt praised the students.'
(Danish)

It is difficult to imagine how any other account of Q Float could capture this correspondence (if it is one).

Let's adopt Sportiche's account of Q-float, then, and see what it teaches us about Argument Movement.

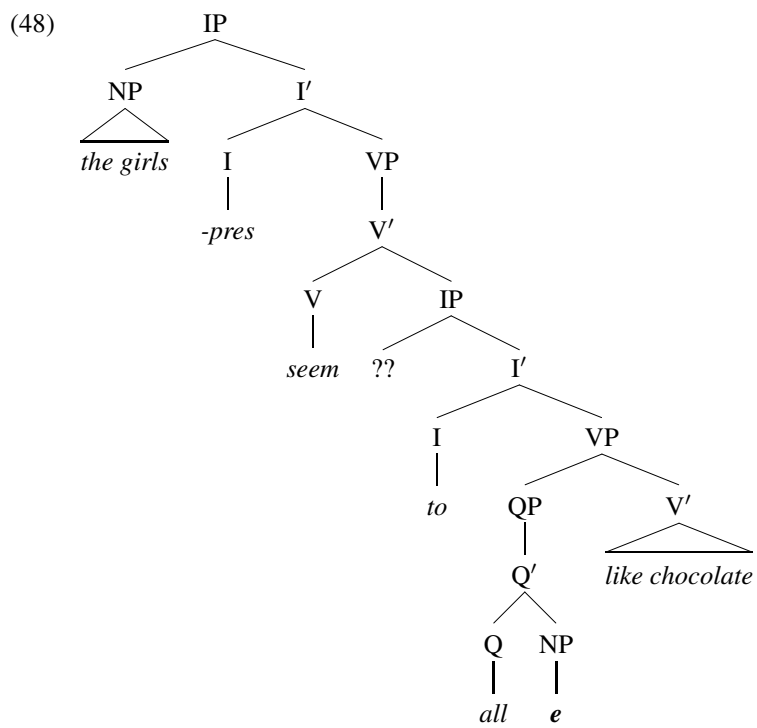
If we call the phrases that put *all* and *both* in construction with an NP, QPs, then a parse for a simple case like (33) would then look like (46), under Sportiche's analysis.



In (46), I've marked with "e" the position that the moved NP occupied at D-structure. I will sometimes resort to this device to represent more than one element of a derivation in a single parse.

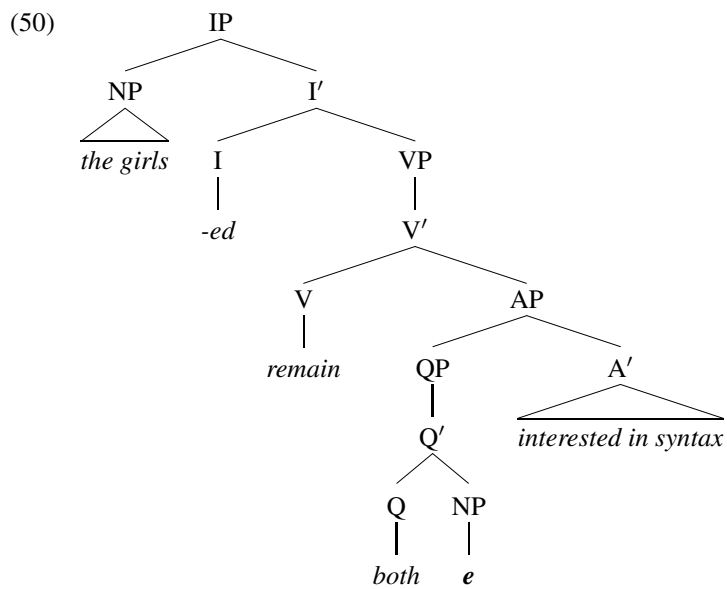
A slightly more complex case, one involving two clauses, like (47), gets the representation in (48).

- (47) The girls seem to all like chocolate.



And an example involving a “small clause,” such as (49) gets a representation like that in (50).

(49) The girls remained both interested in syntax.



In each of these cases, note, the floated quantifier stands in the position in which the relevant θ -role is assigned, and the NP part of this argument has moved into the relevantly Case marked position.

On Sportiche's account of Q-float, the floated quantifiers can be seen as indicators of the positions that arguments occupy not only in the D-structure element of the derivation, but in all the non-S-structure parses of the derivation. It gives us, in other words, a useful glimpse at these otherwise invisible parses. We learn from this that phrases can take smaller steps than necessary in their journey from a θ -marked position to a Case marked position. So in (51), for example, it appears that *the children* has made an intermediary stop in the Specifier of the VP headed by *have*.

- (51) The children might all have eaten chocolate.

And in (52) we see that *the girls* can make a stop in Specifier of the infinitival IP.

- (52) The girls seem all to like chocolate.

This, then, supports the method of satisfying the Extended Projection Principle in these cases that we adopted above. Let the derivation of this sentence involve a parse in which the subject argument occupies the embedded Specifier of IP, and allow the Extended Projection Principle to be satisfied if any parse in a derivation meets its requirements, and the grammaticality of these examples follows.

Incidentally, the grammaticality of Argument Movement out of the clausal complements to *seem* is one of the reasons we parse these complements as simple IPs, rather than as the CPs that we have parsed all other clausal complements. It's in this way that the ability of Argument Movement to escape the complements to *seem* but not the complements of *prefer*, as in (29), is derived. The condition on Argument Movement in (30) forbids movement out of CPs, recall, and this is what we credited the ungrammaticality of (29) with. If this constraint has been accurately rendered, then this will require that the complement to *seem* be something other than a CP, and given its shape, this points to making it a simple IP.

We see something similar in the contrast between the clausal complements to the verbs *try* and *believe*. We've already seen that the accusative Case that *try* assigns cannot be assigned into the "subject" position of a clause it embeds. This we credited to the presence of a CP node between *try* and the embedded Specifier of IP. By contrast, *believe* is capable of assigning its accusative Case to the "subject" position of the clause it embeds. There is a contrast, then, between (53a) and (53b).

- (53) a. * I tried [_{CP} [_{IP} him to be happy]].
b. I believe [_{IP} him to be happy].

If we have defined "klose" correctly — this, recall is the locality condition on Case assignment — then this means that the infinitival complements to *believe* are, like those to *seem*, simple IPs. And, indeed, Argument Movement is possible from the complements to *believe* but not *try*. This can be seen when these verbs are passivized which, as we've seen, sends the arguments that are dependent on these verbs for accusative Case in search of a Case-marked position. Passivizing (53b) is fine, but passivizing (53a) is ungrammatical.

- (54) a. *He was tried [_{CP} [_{IP} e to be happy]].
 b. He was believed [_{IP} e to be happy].

Another case like *try* is *prefer*, which, as we saw earlier, has the complementizer *for* heading the infinitival CP it selects as complement. Because this complementizer can delete, it is possible for *prefer* to appear in sentences that look very much like those that *believe* can appear in:

- (55) I preferred [him to be happy].

But the Case assigned to *him* in this circumstance is assigned by the deleted *for*, and not by *prefer*. In fact, these clauses are not transparent for Case assignment and, as expected, they are not transparent for Argument Movement either:

- (56) *He was preferred to be happy.

We will see this correlation between being “transparent” for Argument Movement and being “transparent” for Case assignment in several places. This correlation is accounted for here by giving the transparent infinitives a “non-CP” status, and then rigging the conditions on Argument Movement and Case assignment so that they are sensitive to the presence of this CP.

There is another part to this correlation that has to do with the distribution of the silent external θ -role bearer we have discovered in control infinitives. In describing where we have discovered this item, PRO, we came up with (22).

- (22) PRO is ungrammatical, unless in a non-Case marked Specifier of IP.

Restricting PRO to Specifier of IP prevents it from being in complement position — and this is what we want, since it isn’t the case that obligatory complement θ -roles can be assigned on silent objects. And requiring the Specifiers PRO appears in to be non-Case-marked ones prevents PRO from appearing as the subject of a finite clause, and is also, we speculated, responsible for the contrast in (57).

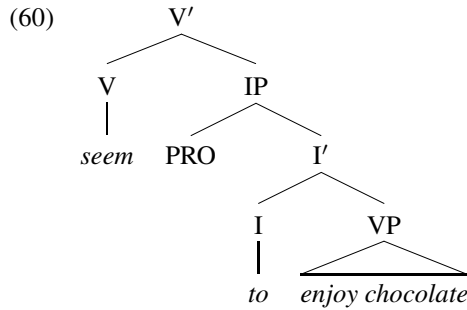
- (57) a. She prefers [_{CP} [_{IP} PRO to drink kava]].
 b. *She prefers [_{CP} for [_{IP} PRO to drink kava]].

If *for* is a Case assigner, then the position PRO sits in in (57b) has to be a Case marked position, whereas the one it sits in in (57a) might manage not to be. (For instance, we might imagine that *for* deletes in the derivation leading to (57a) before it assigns its Case.)

That there is something wrong with the description in (22) can be seen, now, from the behavior of the infinitive that follows *seem*. Recall that we’ve decided that *seem* doesn’t assign Accusative Case — this is why the subject argument in the infinitive beneath *seem* can’t appear with Accusative Case, in a way parallel to what would happen with a small clause beneath *let* or *make*. But, nonetheless, the infinitive following *seem* cannot have PRO in it. A sentence like (58) doesn’t have an interpretation like (59).

- (58) It seems [to enjoy chocolate].
 (59) It seem that someone enjoys chocolate.

The only way for (58) to be grammatical is for *it* to be understood as the argument bearing the external θ -role assigned by *enjoy chocolate*; (58) is ungrammatical if *it* is understood to be an expletive and the bearer of the external θ -role is PRO. Something, then, must prevent PRO from finding itself in the position indicated in (60).



(22) won't do this, since the Specifier of the IP in (60) is, as required, not Case marked. In fact, (60) is part of a larger generalization, which links the distribution of PRO to the correlations we just looked at between “transparent for Case assignment” and “transparent for Argument Movement.” PRO cannot appear in “transparent” clauses:

(61) PRO cannot be in the Specifier of a transparent infinitive.

There are a variety of ways of formulating a law that covers these various observations about where PRO can be found. One way, inspired by *Lectures on Government and Binding*, flows from the observation that what is similar about the two ungrammatical examples in (62) that is not true in the two grammatical examples in (63), is that PRO is close to a word.

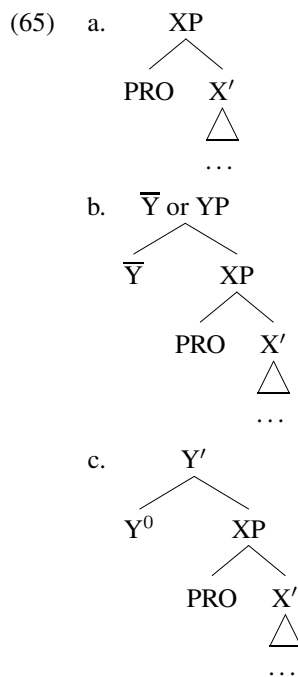
- (62) a. * It seems [_{IP} PRO to enjoy chocolate].
 b. * Sal prefers [_{CP} for [_{IP} PRO to drink kava]].
- (63) a. Jerzy tried [_{CP} [_{IP} PRO to enjoy chocolate]].
 b. Sal prefers [_{CP} [_{IP} PRO to drink kava]].

Because of the presence of the CP (and \bar{C}) in (63), there are two many phrases “between” PRO and the higher verb. In (62a), by contrast, there is only one phrase “between” PRO and *seem*; and similarly, there is only one phrase “between” PRO and *for* in (62b). We could describe this contrast, then, with (64).

(64) PRO may not be c-commanded and close to a lexical item.

On this view, modeling the “transparent” vs. “non-transparent” (or “opaque”) difference of clauses to Case assignment and Argument Movement by crediting “opaque” clauses with CP status, and “transparent” clauses with having no CP, allows the fact that PROs cannot be found in “transparent” clauses to fall into this correlation. That is, if (64) plays a role in determining where PRO can be, then it is no accident that PRO is not found in transparent clauses, if transparent clauses are IPs. IPs that are embedded as complements to some word will not put enough stuff between their Specifiers and that selecting word to satisfy (64).

For the time being, then, let's adopt (64) as part of our description of where PRO can be. In fact, (64), when combined with some of the other principles we've seen, will dramatically restrict the distribution of PRO, and comes very close to doing everything that (22) did. If PRO does not have a meaning that makes it an appropriate modifier, then it will not be capable of standing in adjunct position, since these positions we have reserved for terms that modify. Nor will it be able to stand in complement position, since this puts it in a position that is close to the word that selects it.¹ Thus, (64) will force PRO to occur only in Specifier positions. Further, (64) will allow PRO to exist in a Specifier position in just the configurations in (65), where Y^0 is a position that is not occupied by a lexical item.



When the phrase whose Specifier sits in is a complement, as in (65c), it must be embedded under a head position that is not occupied by a lexical item. The only two such head positions that we have seen, so far, are both C^0 s, and C^0 s always embed IPs. Therefore, when PRO is in a complement's Specifier, it must be in the Specifier of an IP. So, (64) lets PRO exist only in the Specifiers of unembedded phrases (i.e., (65a)), and the Specifiers of phrases in adjunct or Specifier position (i.e., (65b)), or in the Specifiers of embedded IPs. We've not seen PRO in the positions that (65b) describes, but we will. It turns out that PRO is not possible in the Specifier position of unembedded phrases, as in (65a). Because the only unembedded clauses we are likely to see are finite IPs, we'll need (66) in addition to (65), then.

(66) PRO cannot be in the Specifier of a finite IP.

¹(64) will allow PRO in complement position, but only if the head PRO is a complement does not surface as a lexical item. If verbs could delete, for instance, (64) would let PRO exist as the objects of deleted verbs.

This will also prevent PRO from being in the Specifiers of IP that are embedded under a C^0 from which the complementizer *that* deletes, as in (67).

- (67) *Jerzy said [_{CP} [_{IP} PRO likes chocolate]].

We'll have a chance to reexamine this way of describing where PRO exists; but let's adopt (65) and (66) in the meanwhile. They are accurate enough for our purposes.

Let's return now to the contribution that Argument Movement makes.

We've already seen that Argument Movement is constrained in particular ways. It can move the argument it affects only to c-commanding positions, and it cannot move that argument out of a CP. Roughly speaking, then, every grammatical sentence must have a D-structure representation that places arguments which need Case in spots from which they can reach their assigned positions. If they are not directly placed in Case marked positions, then they must be positioned so that they are c-commanded by an empty Case marked position which is within all the same CPs that they are. These constraints then are responsible for blocking a variety of sentences. They play, in other words, a small role in characterizing the grammatical from the ungrammatical sentences of English. For example, they prevent (68).

- (68) a. *It appears [to **e** prefer [for Mary to be possible that this is a sentence]].
 (compare: Mary appears [to prefer [for it to be possible that this is a sentence]].)
 b. *Mary appears [_{CP} that [_{IP} it seems [_{IP} to like chocolate]]].
 (compare: It appears that Mary seems to like chocolate.)

In (68a), the S-Structure shown is prevented because it would involve moving an argument NP, here *Mary*, to a position that does not c-command the position it moved from (indicated here with a "e"). And the S-Structure in (68b) is prevented by the constraint that blocks relocating an argument out of a CP.

There are additional constraints on Argument movement as well, ones that are responsible for preventing the formation of a variety of non-sentences that would currently be permitted. One of these constraints prevents Argument Movement from relating an argument to more than one Case Marked position. This, for instance, is what prevents (69).

- (69) *Mary seems to [that Jerry left].

This constraint has been formulated in a variety of ways; (70) will do for the moment.

- (70) Move α to a Case marked position only if not doing so would violate the Case filter.

There is one other salient constraint on Argument Movement (or A-Movement, as I will call it from now on). This is a locality constraint like our CP-based one, but because we'll need to develop some other material before we'll be able to see it, let's stick with these. Presently, then, we've got something like (71).

(71) ARGUMENT MOVEMENT

Move a term from α to β β an empty position licensed by \bar{X} Theory,²

if:

- i. not doing so would violate the Case filter, and
- ii. β c-commands α , and
- iii. there is no CP that contains α but not β .

While the normal case in which an argument NP finds itself satisfying the Case filter and Theta Criterion by virtue of two positions, involves moving that NP from one position to the other, there are situations where the relationship between Case Marked positions and θ -marked positions seems to be achieved without Argument Movement. These situations involve, instead, the expletives we have run into. For example, in certain circumstances we find that an NP can appear in its θ -marked, but non-Case Marked position, just when an expletive can stand in the Case Marked position we would have understood A-Movement to involve:

- (72) a. There might be a spider walking up my neck.
- b. There seems to be a spider walking up my neck.

Interestingly, the relationship between the expletive and the argument in these sorts of situations is subject to the same constraints that A Movement is. In particular, the expletive cannot be separated from the argument by a CP, and it must c-command the argument. This is indicated by the ungrammaticality of examples such as (73).

- (73) a. * It appears [to be **a woman** preferring [for **there** to be possible that this is a sentence]].
(compare: There appears [to be a woman preferring [for it to be possible that this is a sentence]].)
- b. * There appears [_{CP} that [_{IP}it seems [_{IP}to be a woman liking chocolate]]].
(compare: It appears that **there** seems to be **a woman** liking chocolate.)

We imagine, therefore, that there is some way in which these relationships are the same as those meted out by A Movement. There are various ways in which people have tried to collapse these two — we won't examine them now. One way of talking about this, however, which we might want to adopt in the meanwhile, is that there is some more abstract relationship which either A Movement or this expletive strategy can achieve. It is sometimes said that either of these strategies form an "A Chain," for instance, where you can think of an "A Chain" being simply a "spread" argument. With this terminology, we want to redraw the conditions we've placed on A Movement so that they hold of A Chains, which is what is done in (74) on the next page. This reformulation will require us to recraft the Case filter as well, so that it can be satisfied by an A-Chain; let's adopt (75).

(75) CASE FILTER

If NP is an argument, then it must occupy a Case marked position, or be part of an A-chain which does.

²I write it this way to distinguish Argument Movement from the NP Shift transformation, which, as you'll recall, creates structure that is not permitted by the \bar{X} skeleton.

- (74) Where α, β are positions licensed by \bar{X} Theory, (α, β) form an A-Chain iff:
- i. α holds an expletive and β an argument, or
 - ii. α holds a term moved from β .
- (α, β) is an A-Chain only if:
- i. not forming a chain would violate the Case filter, and
 - ii. α c-commands β , and
 - iii. there is no CP that contains β but not α .

We should note that the A Movement strategy and the expletive strategy are not entirely interchangeable. The expletive strategy is employable in a narrower range of cases; it is subject to its own additional constraints. It is not employable when the argument fails to be of a particular semantic type, as the contrasts with (76) indicate, and the sorts of predicates surrounding the argument seem to matter too, as the ungrammaticality of (77) shows.

- (76) a. * There might be Frank walking up my neck.
 b. * There seems to be Frank walking up my neck.
- (77) a. * There might have a woman walked up my neck.
 b. * There seems a woman pleased with the chocolate.

It's not just these various constraints that place a bound on the kinds of word-orders that are grammatical; they also put constraints on the kinds of the information that verbs contribute. To see this, let's consider the space of configurations that these constraints on A Chains permit.

One can think of Argument Movement, or the A Chain relation, as a function from a representation satisfying the Projection Principle and Theta Criterion to a representation that satisfies the Case Filter. (Moreover, the Extension of the Projection Principle can be thought of as a global constraint on derivations: there must be one representation in the derivation that satisfied its requirements for each Specifier of IP.) θ -roles are determined by verbs and the \bar{V} s they project, and Accusative Case is also controlled by the verb. So verbs will play a big role in setting up the environments in which Argument Movement is called to resolve. The conditions on Argument Movement place a cap, then, on the kinds of situations that verbs can create. They won't allow verbs to exist which create conflicts which Argument Movement can't fix.

Verbs set up situations by creating θ -marked and Case marked positions. Thus, for example, we have seen that there are verbs which assign one external θ -role, and verbs which assign an external and an internal θ -role and Accusative Case. These are sometimes called intransitive and transitive verbs, respectively; examples are in (78).

- (78) a. Sally slept. (intransitive)
 b. Sally likes kiwis. (transitive)

We have also seen verbs that have no external θ -role but do assign an internal θ -role. One of these we have seen in situations such as:

(79) Sally appears [to like kiwis].

And others we have seen formed by the process of passivization:

- (80) a. Sally was considered [e unhappy].
b. Sally was considered e.

It may well be that some intransitive verbs fall into this last class too; a candidate is *appear*.

- (81) a. A ghost appeared.
b. There appeared a ghost.

We shall see evidence for this characterization of some intransitives later.

Burzio (1986) argues on behalf of both these classifications of intransitives and uses the terms *unergative* and *ergative* to distinguish them.

- (82) Intransitives
a. A ghost should sleep. (unergative)
b. A ghost should appear. (ergative, aka unaccusative)

Finally, we might have verbs which assign neither Accusative Case nor any θ -roles:

- (83) It rained.

Although here, intuitions about the referential status of it seem to divide, which some claiming it is referential, hence an argument, and others claiming that it isn't.

So let's consider, then, the space of verb types that we might expect.

	θ -roles	Accusative Case	No Accusative Case
(84)	External, no internal	??	unergatives (<i>sleep</i>)
	No external, internal	not found	ergatives (<i>appear</i>)
	external and internal	transitives (<i>like</i>)	not found
	no external, no internal	??	<i>rain?</i>

Burzio discovered the two gaps in this paradigm, and stated them as a generalization:

- (85) Burzio's Generalization
a. If a verb assigns Accusative Case, then it determines an external θ -role.
b. If a verb determines an external θ -role (and an internal θ -role?), then it assigns Accusative Case.

The second of these generalizations might be derivable from the conditions we have seen on A Chains. To see this, consider the syntactic frames that our theory would let this verb be inserted into. One of these is (86).

- (86) [_{IP} should [_{VP} Smith V Jones]].

If V assigns these two θ -roles, but no Case to *Jones*, then there is no way both of these NPs are going to be able to satisfy the Case Filter. There are more NP arguments than there are Cases. So, if such a verb is to survive, the only environment it will be able to appear in are sentences which have two Case marked positions. These two positions will both have to be Specifiers, because these are the only positions reachable by A Movement. Thus, we're looking for contexts like:

- (87) [IP- should V₁ [XP - [VP Smith V₂ Jones]],
where both “-” positions are Case marked.

Now the first part of Burzio's Generalization tells us that V₁ cannot assign Accusative Case. If it did, then it would also assign an external θ -role, and that's going to bring the count of things that need Case to be one more than there are Case marked positions. As a consequence, the Case marked position inside XP is going to also have to get its Case from some place other than V₁. So far, the only ways we have seen for this to be done are if XP is in a CP:

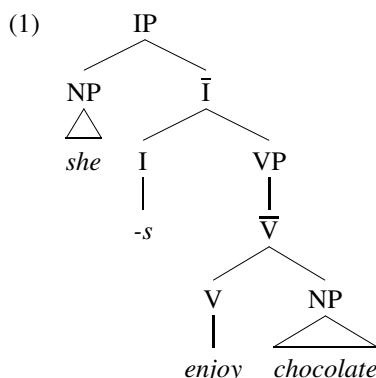
- (88) [IP- should V₁ [CP for [IP - to [VP Smith V₂ Jones]]]].
[IP- should V₁ [CP that [IP - I⁰ [VP Smith V₂ Jones]]]]

But now, there is no way for the arguments of V₂ to get to into the higher Case marked position, because to do so they would have to violate that constraint which prevents movement out of CP. Hence: If the only way for a Specifier to get Case is for it to be embedded inside a CP, then the second of Burzio's generalization will follow from the constraints on A Chains.

Chapter 5

Verb Movement

One problem with our characterization of the grammatical English sentences that we have left untouched up to now has to do with the fact that the heads of sentences, i.e. IPs, are often not free words. Instead, many finite clauses are headed by bound morphology, as in (1).



Recall that we reached this conclusion because it is morphology of this sort subject agreement/tense morphology — that is in the one-to-one relation with the family of strings that we call IP. So, what we would like to understand is how it is that this bound morpheme manages to find itself on the verb that follows. If the phrase markers are to be preserved by this process (that is, if the process involved is syntactic), then two possibilities are that the verb moves or the bound morpheme does.

There is some evidence that in certain cases of this sort, the verb moves. The rule responsible is, therefore, sometimes called Verb Movement. Its classic description is found in Chomsky (1957) and Emonds (1976), and elsewhere.

Let's call the type of argument that Emonds' is an instance of the "Correlation Argument," because it's based on the correlation that holds between a verb's syntactic position and its inflection. The fact that accounts like Emonds' hope to derive is the dependence that (appears) to hold of a verb's syntactic position on its inflectional category. This can be seen from the relative positions of auxiliary verbs in English finite

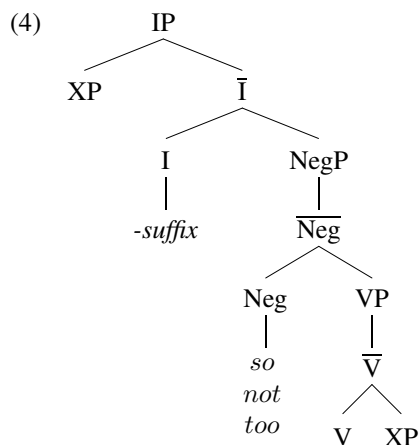
clauses. An auxiliary verb cannot follow *not*, *so*, *too* (“polarity terms”) if it is inflected with the tense/subject-agreement morpheme. That is, we find paradigms like those in (2).

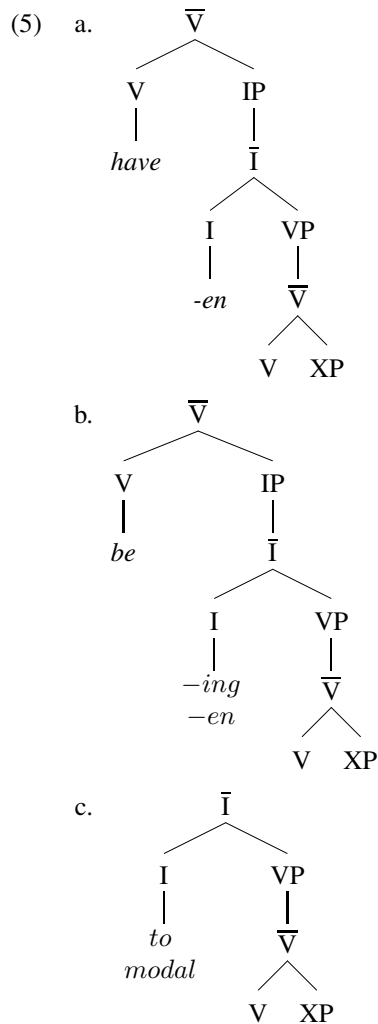
- (2) a. Gary has not been eating.
 b. Gary is not eating.
 c. * Gary not has been eating.
 d. * Gary not is eating.
 e. Gary has too/so been eating.
 f. Gary is too/so eating.
 g. * Gary too/so has been eating.
 h. * Gary too/so is eating.

A similar correlation is found across clause types. Auxiliary verbs can precede polarity terms in finite clauses (when they are appropriately inflected), but never in subjunctive clauses, where the tense/subject-agreement morpheme is absent:

- (3) a. I request that Gary not be eating when we arrive.
 b. I demand that Gary not have left by the time it starts.
 c. I require that Gary not leave.
 d. *? I request that Gary be not eating when we arrive.
 e. ?? I demand that Gary have not eaten when we arrive.
 f. * I require that Gary leave not.

An Emonds/Chomsky-type account of this phenomena lets phrase structure rules fix the location of the inflectional morphemes involved, and moves the appropriate verb to this location by way of a rule that moves heads. Thus, simple finite clauses are assigned a D-Structure representation like that in (4), and clauses with a complex tense (i.e., an auxiliary verb), have the D-Structure representations in (5).





(Note that this conforms to \bar{X} Theory, which is what motivates most of the rest of the detail of these phrase markers.) This gives a straightforward account of the correlation described above. Verbs stand in the positions determined by the inflections they bear because they are moved to the positions that the phrase structure rules assign to these affixes.

A well-known exception to this scheme is encountered when English “main” verbs are considered. In these situations, the verbs do not occupy the positions determined by the inflections they bear, but remain instead in the location given to main verbs by the phrase structure rules; thus (6) is ungrammatical.

(6) *Gary likes not/too/so apples.

The account of these situations that we inherit from Chomsky (1957) involves lowering the inflectional affix onto the main verb, a rule termed Affix Hopping. This

rule is blocked, for mysterious reasons, when negation is present (as the ungrammaticality of (7) indicates), and instead the dummy verb *do* is used to support the otherwise stranded affix, as in (8).

- (7) * Gary not likes apples.
- (8) Gary does not like apples.

Let's for the moment capture this part of the process with:

- (9) AFFIX HOPPING
Attach I^0 to the main verb immediately "next" to it.
(Condition: an item in Neg^0 may not intervene.)

That auxiliary verbs do have a different position when they are inflected is confirmed by the action of rules that affect VP. These rules seem to necessarily strand the verb inflected for agr/tense, indicated by the following examples of VP deletion and VP topicalization.

- (10) a. Sam is eating pickles, and Mike Δ is too.
- b. * Sam is eating pickles, and Mike Δ too.
- c. I claimed that Mary is eating pickles, and eating pickles she is.
- d. * I claimed that Mary is eating pickles, and is eating pickles, she.

This all points to the following rule:

- (11) VERB MOVEMENT
Adjoin an Auxiliary verb to I^0 .

Note, incidentally, that Verb Movement is forced in these contexts. That is, for the VP affecting rules to have the consequence that the finite verb is necessarily outside of them, it must be that the verb cannot remain unmoved. So, in addition to a rule which moves the verb, it is also necessary to explain (or describe) what forces this movement.

One suggestion, perhaps the most popular one, relies on the observation that the situations where movement to I^0 is forced are those where I^0 holds a bound morpheme. Therefore, if we assume that two morphemes can only be treated as a single word only if they share the same X^0 , as in Lasnik (1981). This idea is sometimes called the "Stray Affix Filter," and might be formulated as follows.

- (12) STRAY AFFIX FILTER
A bound morpheme must coexist with its stem under a common X^0 at S-structure.

Verb Movement, then, in conjunction with the Stray Affix Filter will guarantee that auxiliary verbs are moved out of the VP they head by S-structure in order they provide the affix in I^0 position with a stem.

But as noted above, Verb Movement does not affect "main" verbs. This corresponds to the observation that main verbs do not precede polarity terms. And we can see from VP constituency tests that finite main verbs also do not appear to be outside VP, as do finite auxiliary verbs.

- (13) a. * Sam slapped Mike, and Mary slapped Δ too.
 b. * I claimed that Mittie put the book on the table and the book on the table, she put.

So, there must be another way that main verbs and the inflection they bear come together; that is, the Stray Affix Filter must be satisfied in a way different than through the agency of Verb Movement in these cases. This is where Affix Hopping comes into play.

These rules can be embellished with another, which plays a role in forming sentences with a different communicative function: questions. One sort of question involves relocating the finite verb to the beginning of the sentence, as in the following.

- (14) a. Have you eaten pickles?
 b. Should you eat pickles?

These are called Yes/No questions; another sort of question which involves a relocation of the finite verb seeks more specific information and is called a Wh-Question.

- (15) a. Which pickles have you eaten?
 b. What should you eat?

In both of these types of questions, note, there is a rule that moves I^0 to some sentence-initial position. And as a consequence of this rule's targeting I^0 , only material which can stand in this position can be so affected. Therefore, main verbs in English cannot undergo this process, and instead the dummy stem *do*, appears.

- (16) a. * Eat you pickles?
 Which pickles eat you?
 b. Do you eat pickles?
 Which pickles do you eat?

What is the landing site for this I^0 movement operation? The common answer to this question, based on work of the Dutch linguists Jan Koster and Hans den Besten,¹ is that it is C^0 . The reason for this speculation is because the rule involved here seems to be in complementary distribution with complementizers. That is, its effects are blocked in those cases where we believe that complementizers are present in this position. Thus, there is a distinction between (17a) and (17b).

- (17) a. Have you eaten?
 b. * I remember (that) have you eaten.
 c. Which pickles have you eaten?
 d. * I remember that which pickles have you eaten.

Instead, Yes/No questions that are embedded are marked with a special complementizer: *whether*. And wh-questions in embedded contexts involve no evidence of a C^0 at all.

Under this view, the rule involved in these root questions, then, is:

¹See den Besten (1983) and Koster (1975).

- (18) I-TO-C MOVEMENT
Adjoin I^0 to C^0 .

We must also add to this account the convention that everything under the same X^0 be a single word: this will prevent two (or more) free morphemes from coexisting under the same X^0 , and therefore as a special case prevent I^0 from combining with a C^0 that holds a complementizer. This idea can be combined with the Stray Affix Filter into what I'll call "The Word Criterion."

- (19) THE WORD CRITERION
If α and β coexist under the same X^0 , then they must form a word.

Okay, so this is, roughly, a sketch of the rules that go into making up our knowledge of this fragment of English syntax. We've got three rules, one that moves auxiliary verbs to I^0 , another that moves I^0 to C^0 and a third that lowers I^0 onto main verbs. The first and last are driven by the Stray Affix Filter — forcing V^0 to I^0 or Affix Hopping whenever I^0 has a bound morpheme — and the second arises in questions. Both are subject to the Word Criterion, which prevents I^0 , C^0 or V^0 from combining whenever they cannot form a word.

There's one last feature of these rules that we should consider. In all the cases so far examined, the V^0 that moves to I^0 is always the one that is closest to it. It appears that this is not just an accident of the examples we have chosen, instead it looks like it is a necessary feature of the verb movement rule. Consider, for example, the following case.

- (20) a. Mary -s have -en be -ing eat.
b. * Mary is had eating.

We might expect to be able to get (20b) from (20a) if *be* could move past *have* and -en to adjoin to -s. Then the Stray Affix filter could be satisfied with respect to -en if it moved and adjoined to *have*. Because this doesn't seem possible, it looks like the rule that moves verbs is constrained in such a way that it is unable to move past other verbs.

From the example in (20) it looks like we need to say either that verbs cannot move past another V^0 or, perhaps, more generally, that they are unable to move past another X^0 . If we adopt this more general version of the constraint, then we can simplify our description of the syntax of verb position in the direction of reducing the two rules I^0 -to- C^0 and V^0 -to- I^0 to just one rule. We might try to do with just:

- (21) VERB MOVEMENT
Adjoin an X^0 dominating an auxiliary verb or modal to another X^0 .

Now the fact that only verbs which have first moved to I^0 can get into C^0 will follow from the constraint on verb movement, which we can give as (22).

- (22) A verb may not move past an X^0 .

There is some independent support for a general constraint of the sort that (22) describes. In the syntax of Noun Incorporation, which arises in some languages when a noun is adjoined to a verb there are some facts which suggest that something like

(22) is at play. In Mark Baker's influential study of this phenomena,² for example, he argues that (22) is responsible for blocking nouns from Incorporating into a verb across another verb, or across a preposition. This constraint is usually expressed as the "Head Movement Constraint," the name given to it by Lisa Travis.

(23) THE HEAD MOVEMENT CONSTRAINT

No X^0 may move past a Y^0 that c-commands it.

So here is the first feature of the grammar of English verb placement that we can assign to Universal Grammar. It obeys the Head Movement Constraint, and in this partakes of a feature found common to other instances of X^0 Movement cross-linguistically.

Let's now consider phenomena that look similar in some of the languages closely related to English. Let's start with examples from the Germanic family. In German we find that the position of the finite verb depends, roughly speaking, on whether the clause it is in is embedded or not. In embedded clauses, the verbs stack up at the end of the sentence in the mirror image of their arrangement in English: the finite verb comes last.

- (24) a. ... daß Hans das Buch kauft.
 ...that John the book buys
 b. ... daß Hans das Buch gekauft hat.
 ...that John the book bought has
 c. ... daß Hans das Buch gekauft haben muß.
 ...that John the book bought have must.

From this we can conclude, perhaps, that VPs are head-final. Further, if there are IPs and a method of combining verbs with the inflection so introduced, they too are head final in German. Then there might be a rule of verb movement in German that, as in English, moves the verb into the I^0 that inflects it. But this rule will move the relevant verbs to the right, rather than to the left.

Interestingly, we find a different word-order in root, or independent, clauses. Here, the inflected verb no longer comes finally in the series; but instead it appears immediately following the subject. So unlike (24), we find only the word order in (25).

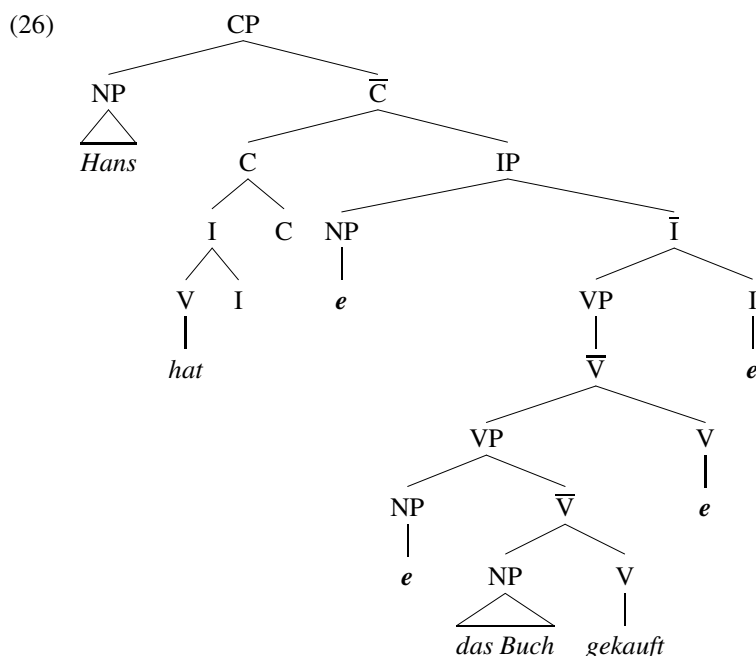
- (25) a. Hans kauft das Buch.
 John buys the book
 b. Hans hat das Buch gekauft.
 John has the book bought
 c. Hans muß das Buch gekauft haben
 John must the book bought have

This seems to suggest that there is a movement rule which relocates finite verbs into the position immediately following the subject. Using the logic of the correlation argument, we might imagine that the position where finite inflection is in German immediately follows the subject, and it's into this position that finite verbs are driven by the Stray Affix Filter, or the equivalent in German.

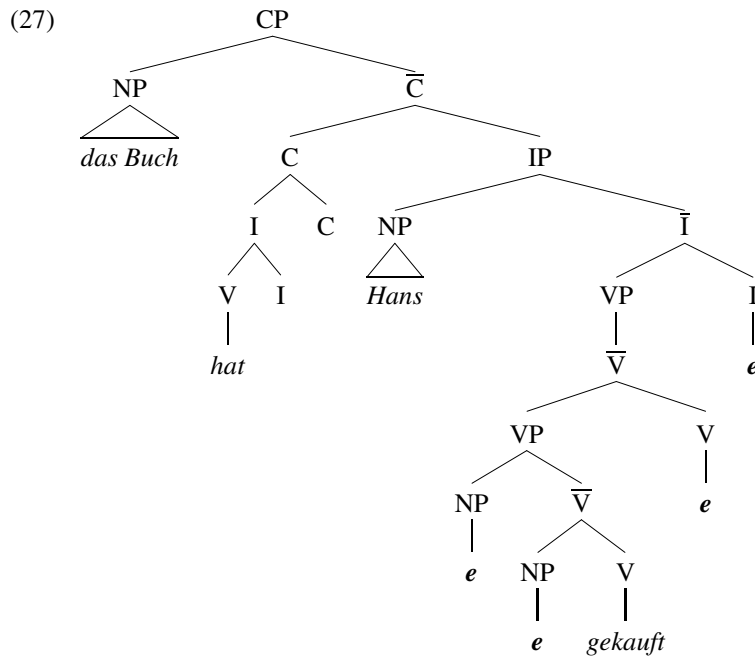
²See Baker (1988).

But this would miss the fact that the position of the finite verbs differs for embedded and independent clauses. What we want is some way of forcing verbs to move into the post-subject position in root clauses only. This suggests that it is not the finite distinction that is responsible for verbs' position in root clauses, but something else. Something that distinguishes root from embedded clauses.

We've already seen a similar difference in the grammar of English: recall that I^0 -to- C^0 movement is restricted to root clauses in English. Perhaps the verbs are moving through I^0 into C^0 in cases like (25), then. This would credit German and English with two differences. On the one hand there is the difference in headedness that we see most directly in embedded contexts. And then there is something that allows/forces subjects to move past C^0 in embedded clauses. We might imagine that this second force, whatever it is, is like the process that moves wh-phrases in English into Specifier of C^0 . Thus, the examples in (25) might get a parse like that in (26). (I've placed in (26) "e" in those positions from which something moved. So, for instance, in Specifier of IP is a *e* that corresponds to the NP, *Hans*, which has moved into Specifier of CP.)



Now note in this situation that the subject precedes the verb because of a movement operation which brings it into the Specifier of the X^0 that the finite verb has moved into. This leads to the expectation that we might find other phrases preceding the finite verb; and we do. It is not just the subject that may immediately precede the finite verb in root contexts, any phrase can. When some other phrase comes before the verb, the subject (typically) immediately follows the finite verb and the phrase that shows up at the front is understood to be "topicalized." Thus, alongside (26) we also find (27) on the facing page.



The generalization about German word-order can be described this way: any phrase may be initial, but exactly one must be. German is sometimes described as having “Verb Second” word-order for this reason. The account we’ve just sketched of this captures Verb Second by way of a rule that moves the verbs that have moved into finite I^0 into C^0 , and by moving anything, but something, into the single position that exists to the left of C^0 .

That the verb in these cases has in fact moved into C^0 is further substantiated by cases where it V2 word-order is found in embedded clauses. Though there is considerable dialectal variation here, in the standard dialects, V2 is possible in embedded clauses just in those cases where a complementizer may go missing. As in English, it is possible in German to unexpress the complementizer when the CP it heads is the complement to a certain class of verbs. The verb *say*, for example, can go without a complementizer — as in (28) — and when it does so in German we find V2 word-order, cf. (29).

(28) Jerry said (that) Sally has a dime.

- (29) a. Er sagt, daß die Kinder diesen Film gesehen haben.
 He says that the kids this film seen have
 ‘He says that the kids have seen the film.’
 b. Er sagt, diesen Film haben die Kinder gesehen.
 he says this film have the kids seen
 ‘He says that, this film, have the kids seen.’

Despite the similarities in verb movement that German and English have, note that one place where they differ is whether main verbs fall under the scope of verb movement. Unlike English, German main verbs can undergo movement. This is seen by their ability to move into C^0 in main clauses, as in (25a).

There are a host of mysteries about the grammar of verb placement in German that we will not attempt to solve here. Like what, for instance, is responsible for forcing these extra movements in German root clauses. And why isn't the movement of a phrase into Specifier of CP allowed in embedded clauses, as in (30)?

- (30) * Ich sagte [_{CP} das Buch [daß [_{IP} Hans [_{VP} e gekauft] hat]]].

Vikner, in the chapter I have assigned as reading, discusses some of the ideas that have been offered for answers to these questions. We will not explore them here, however.

There are some things that we might notice about the processes in German that make use of Verb Movement which are familiar to the processes that move verbs in English. For example, the Head Movement Constraint is preserved in the grammar of verb placement in German. Sentences like the following, where *haben* ('have') moves past the modal are ungrammatical in German just as they are in English.

- (31) * Hans haben das Buch gekauft muß.

Further, note that in the cases of verb movement in German, just as in the similar processes in English, the verb (or I^0) is always landing in another head position. This appears to be something that is the same across German and English, and so it constitutes a candidate for a universal. One of the first proposals of this kind for this constraint is found in Baltin (1982), who argues that this is part of a more general constraint on movement rules. He proposes that the position to which a phrase or head is moved is always the same as the phrase or head being moved. So a head adjoins to another head position only, and a maximal projection can only move to another maximal projection position, and so on. Let's call this the LIKE- ATTRACTS-LIKES Condition:

- (32) LIKES ATTRACTS LIKES

An X^n may only adjoin to, or substitute into, a position that is also an X^n .

Finally, note that with the sole exception of Affix Hopping, the movement rules we have examined in these two languages all have the feature that the verb or I^0 that has moved has moved up. We don't find cases where Verb Movement has relocated the verb downwards, as in examples like:

- (33) a. * ... daß Hans das Buch gehabt kauft.
 ... that Hans the book had buys
 b. * John must had buy the book.

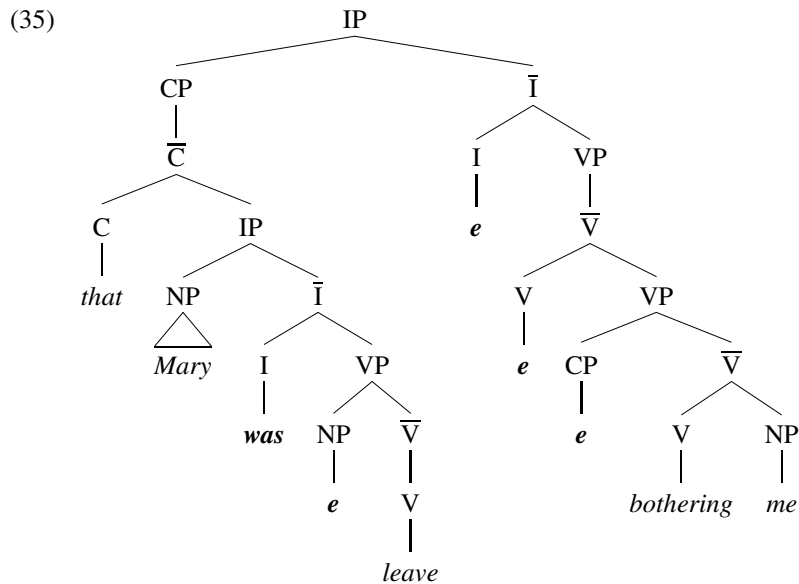
We have already found that this is a feature of the Argument Movement rule — it also relocates terms only to a position higher in the phrase marker.

Actually, we found that the constraint was more specific than this; it required that the moved term relocated to a c-commanding position. Let us formulate this constraint, shared by both German and English Verb Movement rules, as follows.

- (34) UPWARDS CONSTRAINT

α can move to position β only if β c-commands α .

That this narrower condition is required for verb movement is shown by cases like (35), which would be possible if verbs could move to non-c-commanding positions.³



Now, we find a situation similar to German in the Scandinavian languages, as Vikner reviews in the reading. Using Danish as an example, note that the same sensitivity to embedding is found in the placement of the finite verb in these languages. As the contrast in (36) indicates, Danish is like German in allowing any constituent, but only one, to precede the finite verb in independent clauses. That is, it shares with German the trait of being “Verb Second.”

- (36)
- a. Børnene har set denne film
kids-the have seen this film
‘The kids have seen this film.’
 - b. Denne film har børnene set.
this film have kids-the seen
‘The kids have seen this film.’
 - c. * Denne film børnene har set.
this film kids-the have seen
‘The kids have seen this film.’

But, as (37) shows, these traits are not shared by dependent clauses, where, instead, the subject must precede the finite verb.

³This parse does not reflect the topicalization that we’ve seen evidence for — recall, there is reason to believe that finite “subject” clause has been topicalized to IP.

- (37) a. *Jeg ved ikke hvor i går har koen stået.
 I know not where yesterday has cow-the stood
 'I don't know where the cow stood yesterday.'
- b. Jeg ved ikke hvor koen har i går stået.
 I know not where cow-the has yesterday stood
 'I don't know where the cow stood yesterday.'

Moreover, as the contrasts in (38) show, the placement of the finite verb relative to negation is sensitive to the embedded/non-embedded context.

- (38) a. Børnen har ikke set denne film.
 kids-the have not seen this film
 'The kids haven't seen this movie.'
- b. *Børnen ikke har set denne film
 kids-the not have seen this film
 'The kids haven't seen this movie.'
- c. Jeg ved at børnen ikke har set denne film
 I know that kids-the not have seen this film
 'I know the kids haven't seen this movie.'
- d. *Jeg ved at børnen har ikke set denne film
 I know that kids-the have not seen this film
 'I know the kids haven't seen this movie.'

This indicates that the finite verb has moved out of VP, past negation, into C^0 in independent clauses, just as it does in German.

It looks, therefore, like Danish has the word-order of English — placing verbs before their objects — but the syntax of verb movement like German. Note in particular that main verbs are able to move as well as auxiliary verbs, as we see partly in (37) (look at the root clause word-order). If we were to look further, we would see that Danish (and the remainder of Germanic) also have the range of constraints we have seen on verb movement operations.

There are a couple differences, however. One at least potential difference between Danish and German is the existence of V^0 -to- I^0 . Whereas we cannot easily discern whether such an operation exists in German, it apparently does not in Danish (nor does it in standard Norwegian and Swedish), since when the verb has not relocated into C^0 , it remains to the right of negation. (Of course, we have made the perhaps incorrect assumption that negation occupies the same place in Danish as it does in English.)

Another difference concerns the range of contexts where V2 is found in embedded contexts. Recall that in German the verb moves into C^0 in embedded clauses only in those contexts where a complementizer is not required to fill that C^0 . But in Danish (as in the other Scandinavian languages), V2 word-order is possible even in those embedded clauses which have a complementizer associated with them, as in (39).

- (39) Vi ved at denne bog har Bo ikke læst.
 we know that this book has Bo not read
 'We know that Bo hasn't read this book.'

A variety of hypotheses about this situation have been offered — the chapter following — the assigned one in Vikner gives a good overview; let us adopt for concreteness the hypothesis that there is a CP “shell” that can be embedded within a regular CP. This CP shell provides the C^0 into which verbs move in embedded clauses in Scandinavian.

So, let’s see what we’ve got so far. If we factor out the parts to the set of rules in German, Dutch and English that are common to the syntax of verb placement, we have the following:

- (40) UNIVERSALS
- a. Likes Attracts Likes
 - b. Upwards Constraint
 - c. The Head Movement Constraint
 - d. Stray Affix Filter
 - e. Word Criterion

The differences that are left can be boiled down to the following four parts.

- (41)
- a. German VPs (and IPs?) are head final.
 - b. English and Danish VPs (and IPs?) are head initial.
 - c. German and Danish: verbs (and I^0 ?) are movable.
 - d. English: auxiliary verbs and I^0 are movable.
 - e. German and Danish: Root C^0 and Specifier of CP must be filled. (i.e., V2)
 - f. English: Only wh-phrases move to Specifier of CP, and only questions trigger I^0 -to- C^0 movement.
 - g. Some differences in the nature of I^0 , yet to be explored.
 - h. English has affix hopping.

Putting this in the terms of language acquisition which we started with, the differences in (41) constitute what the child must learn in order to determine whether the grammar he or she is acquiring is German, Danish or English.

But these are not the sorts of things that go into “grammar construction,” in anything like the sense of that process that is familiar to linguists. And so the evaluation metric it hints of is not easily characterized in terms of “simplicity.” If (41) are the kinds of choices the child is making, then it doesn’t look at all like he or she is building up a grammar from the simplest rules possible. Rather, it looks like he or she is selecting from a menu of possibilities just those ones that fit the grammar to be acquired. This conception of the language acquisition process, and the kinds of grammars that have this arrangement, is called a “Principles and Parameters” model. Under it, there are an assortment of universal principles and rules with a variety of open parameters which, when set, fully determine a particular grammar. Much of what we will see in this course reflects this conception of how grammars, and the language acquisition process, are designed.

Let us now add to our family of languages, French. In French we also find evidence of rules that move verbs. Just as in English, French verbs distribute themselves relative to polarity items, like sentential negation, on the basis of their inflectional

class. Thus, for example, we find that finite verbs in French, just like their counterparts in English, only appear before negation:

- (42) a. Jean n'a pas lu livres.
 John n'have not read books
 'John hasn't read books.'
 b. *Jean ne pas a lu livres.

A word about negation in French is needed before the contrast in (42) will be interpretable. Unlike English, in French sentential negation is expressed by way of two words, one (*ne*) appears to be in a position like that we have put inflectional morphemes in English. We might imagine that like the contracted form of *not* in English (*n't*) it is contracted onto whatever word has moved into I^0 (as in (42a)). The second word (*pas*) is the one that has a distribution more like that of English *not*. And, as (42) illustrates, the finite form of *have* (*ha*) must precede this part of sentential negation.

French differs from English (but is like German) in placing finite forms of main verbs to the left of sentential negation too, as (43) demonstrates.

- (43) a. Jean n'aime pas Marie.
 John ne'love not Mary
 'John doesn't love Mary.'
 b. *Jean ne pas aime Marie.

As the contrast between the position of main verbs in (43a) and (42a) indicates, it is the inflectional class of the verb that determines its position. That is, just as in English, there is a correlation between the inflectional class of the verb and its syntactic position — a correlation that is captured by fixing the position of inflectional morphemes with the phrase structure rules and driving the verbs to these positions with the verb movement operation.

That main verbs may move is also responsible for another difference between French and English. This derives from the ordering that the Projection Principle places on arguments and non-arguments. As we have had occasion to discuss several times, now, the Projection Principle will force arguments to be closer to their θ -marking head than it will allow non-arguments. In verb initial languages, like French and English, this means that complements should come before non-complements. This is roughly true for English, as we have seen; but it is not the case for French.

- (44) a. Jean embrasse souvent Marie.
 b. *John kisses often Mary.

This apparent difference in the function of the Projection Principle can actually be seen as a product of verb movement. Since we already know that main verbs in French but not English move into I^0 , it makes sense that the main verbs in French, but not English, should be able to be separated from their objects by all sorts of material, including adverbs.

But now consider what happens in French non-finite clauses.⁴

⁴These data all from Pollock (1989).

- (45) a. **Comprendre à peine** l'italien après cinq ans d'étude dénote un
understand barely the-Italian after five years of study shows a
manque de don pour les langues.
lack of gift for the languages
'To barely understand Italian after five years of study shows a lack of
talent for languages.'
- b. **Perdre complètement** la tête pour les belles étudiantes c'est
lose completely the head for the pretty students it is
dangereux.
dangerous
'To completely lose your head for pretty students is dangerous.'
- (46) a. ne *pas* **sembler** heureux est une condition pour écrire des romans
ne not seem happy is a condition for writing novels
'To not seem happy is a (pre?)condition for writing novels.'
- b. *ne **sembler** *pas* heureux est une condition pour écrire des romans.
ne seem not happy is a condition for writing novels
'To not seem happy is a (pre?)condition for writing.'

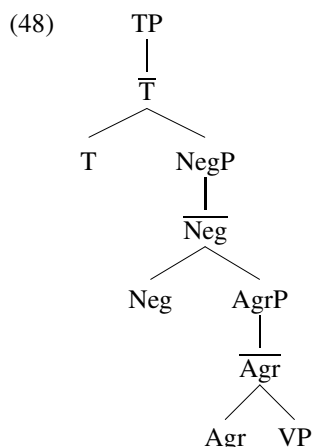
Here it appears that main verbs may move past adverbs but not negation. (The infinitival verbs are in bold face, in these examples, and the term that separates them from their objects is in italics.) Auxiliary verbs behave differently.

- (47) a. ne *pas* **être** heureux est une condition pour écrire des romans
ne not be happy is a condition for writing novels
'To not be happy is a condition for writing novels.'
- b. N'**être** *pas* heureux est une condition pour écrire des romans.
ne'be not happy is a condition for writing novels
'To not be happy is a condition for writing novels.'

They may optionally move past negation.

We learn two important things from these facts. First, that the contrast between auxiliary and main verbs that seems to distinguish English from German/Danish is actually not a distinction in languages. Instead, it is a difference which is found within a single language: French. That is, this is not a parameter along which languages vary, then, but rather a parameter along which clause-types vary. We need to express the distinction between English, and these other languages, in terms of the clause types these languages host, not in terms of the targets for their verb movement rules. So, we're looking for something that distinguishes English finite clauses and French non-finite clauses on the one hand from French finite clauses on the other.

The second thing lesson of this paradigm is the main point of Pollock's paper. He suggests that the two different positions that main verbs may occupy across finite and non-finite clauses warrants giving a structure like that in (48) to clauses, where the higher I^0 is equated with Tense and the lower one with Agreement. This answers to the correlation that appears to hold for main verbs between whether they bear tense morphology or not: in finite clauses they do, and in non-finite clauses they do not.



Pollock also introduces the idea that Neg heads a phrase which stands between the two other Inflectional phrases, as indicated in this parse. There are several reasons for wanting to treat negation differently from other adverbials. One is that the syntactic position of negation is comparatively rigid when compared to other adverbs. Another is that only negation blocks Affix Hopping, as we saw earlier; other adverbs don't. Pollock suggests distinguishing negation from adverbs structurally, and then making reference to this structural difference to capture these ways in which negation behaves uniquely. I won't examine in detail how this idea works, partly because we will eventually go in a slightly different direction than does Pollock. But let us nonetheless adopt — at least as an intermediary hypothesis — the thesis that negation does head a phrase as shown in (48). It should be noted, however, that this introduces a problem: how is it that Agr⁰ can move to T⁰ past negation without violating the Head Movement Constraint. Pollock offers a solution to this problem that I will come to soon.

Note that though the difference between main and auxiliary verbs in French non-finite clauses that we've just reviewed speaks on behalf of two head positions to which verbs may move, it doesn't really indicate what the identity of these two head positions might be. While most of Pollock's proposal involves examining how the hypothesis that there are two inflectional positions within a sentence can be used to explain the differences in verb position across clause types, he also assigns values to these inflectional phrases. But, in fact, it is extremely difficult to tell what the value of these heads is. Pollock decides in favor of giving the higher I⁰ the value of Tense and the lower one the value of Agreement. This involves two decisions: first that these positions should be equated with inflectional categories, and second that they should be equated with these particular ones. The first decision, note, is a bit of a leap here as there is no convergent evidence of the headedness sort to anchor this conclusion. That is, part of the reason we decided that the landing site of verb is I⁰ came in response to the evidence that I⁰ was a head.

There is, however, a sort of reasoning which suggests that the first decision is correct, but that the second isn't. Consider the following verbs, drawn from a variety of Indo-European languages.⁵

⁵The argument I'm reporting here comes from Belletti (1990).

- (49) a. Legg - eva - no
read - Imperfect - 3plur
- b. Parl - er - ò
speak - Future - 1sing
- c. Buhl - t - en
woo - past - plur

Note that in each of these cases, the morpheme which encodes tense information precedes that which encodes agreement. Is there a way of determine the syntactic arrangement of inflectional terms from their morphological arrangement? Mark Baker has made famous an argument for thinking that the answer to this question is yes. Following proposals by Pieter Muysken and Donna Gerdts, Baker argues that there are correlations between the order that inflectional affixes have relative to the verbal stem and the order of the syntactic operations that these affixes encode. Thus, for example, in Chamorro, the passive morpheme necessarily comes closer to the verbal stem than does the subject agreement affix, as (50) shows.⁶

- (50) Para.u.fan - s - in - aolak l famgu'un gi as tat-n-niha
irr.3pS - pl - pass - spank the children obl father.their
'The children are going to be spanked by their father.'

This ordering on affixation mimics the order in which the syntactic operations of Passive and subject agreement take place — the passive must bring the underlying object into the subject relation before the agreement process can apply.

Baker calls correlations of this type, which he suggests are widespread enough to be considered law-like, the "Mirror Principle."

- (51) MIRROR PRINCIPLE
Morphological derivations must directly reflect syntactic derivations (and vice versa).
(Baker 1985, p. 375)

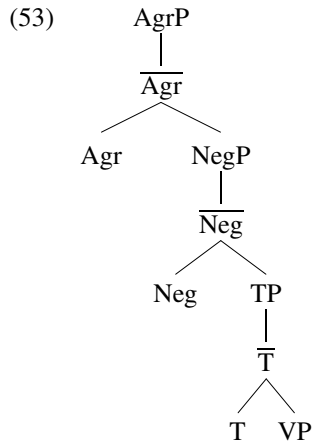
Baker argues that the Mirror Principle will emerge once the correct theory of the syntax-morphology interface is found. In particular, (51) can be derived on models that posit for each of the agreement, passive, and other such operations one rule underlying both the morphological and syntactic effects. Baker tentatively concludes that these metarules apply in the Syntax — that is, after the verbal roots are inserted into phrase-markers. The resulting picture is sketched in (52).

- (52)
- | | | |
|--------------|---------------------------------------|----------------------------|
| D-structure: | V_{root} | |
| | ↓ | |
| | $V_{root} + Af_{\alpha}$ | ← <i>Metarule</i> α |
| | ↓ | |
| S-structure: | $V_{root} + Af_{\alpha} + Af_{\beta}$ | ← <i>Metarule</i> β |

⁶These data come from Gibson (1980). The principle formulated in Baker (1985) has its roots in many earlier works, among which are: Muysken (1979, 1981) and Gerdts (1981).

Each metarule brings about the relevant syntactic operation and adds to the verbal root the corresponding affix. Note in particular that this picture correlates syntactic operations with affixes; the correspondences that Baker summarizes with the Mirror Principle only concern the relation between affixal orderings and the orderings of syntactic operations. Indeed, Baker's argument leads to the conclusion that the metarules in (52) necessarily involve affix-syntactic operation pairs. It is the order of affixes that correlates with the relevant syntactic operations, and not some more abstract information, such as morphological features or the like. We shall have an opportunity to revisit this point.

Now if Baker's conclusions from the Mirror Principle are imported into the domain of Verb Movement, then the arrangement of morphemes in (49) suggests that the higher of Pollock's positions should be associated with agreement morphology, and the lower with tense morphology. Then, the fact that tense morphology comes closer to the verbal stem will follow from the Head Movement Constraint. So we should adopt, perhaps, a picture like the following:



We can reach this conclusion, perhaps, through a different route. If we compare Icelandic with Danish, we see a difference in verb placement that suggests that only Icelandic has movement of verbs into I^0 . In embedded Icelandic clauses the finite verb must precede negation:

- (54) a. *Ég spurði af hverju Helgi ekki hefði lesið þessa bók
 I asked whether Helgi not had read this book
 'I asked whether Helgi hadn't read this book.'
- b. Ég spurði af hverju Helgi hefði ekki lesið þessa bók
 I asked whether Helgi had not read this book
 'I asked whether Helgi hadn't read the book.'

But as we have seen, Danish finite verbs cannot precede negation. Another example illustrating this fact is (55).

- (55) a. Jeg spurgte hvorfor Peter ikke havde læst den.
 I asked why Peter not had read it
 ‘I asked why Peter hadn’t read it.’
- b. *Jeg spurgte hvorfor Peter havde ikke læst den.
 I asked why Peter had not read it
 ‘I asked why Peter hadn’t read it.’

If we assume that the position of negation is constant across these languages, then this indicates that the verb has moved past this position in Icelandic, but not Danish. Now, interestingly, Icelandic has a full paradigm of subject agreement, but Danish has NO subject agreement. Comparing Icelandic with Danish, then, what we find is a correlation between subject agreement and position to the left of Neg⁰. Indeed, this correlation holds throughout the Scandinavian languages.⁷ Standard Swedish and Norwegian are like Danish in lacking subject agreement morphology, and they are like Danish as well in placing their finite verbs to the right of negation. This is just the correlation that (53) predicts.

Adopting this modified Pollockian parse for clauses, we are left with a series of questions about how verbs distribute themselves over the two head positions. Why can main verbs not move to the “agreement” X⁰ in non-finite clauses in French, for instance; and why are they unable to occupy this position even in finite clauses in English. I don’t think we have entirely satisfactory answers to these questions. Pollock relies on an observation due originally to Ian Roberts: The X⁰s in which main verbs cannot surface are devoid or close to devoid of morphological content. English agreement is much less “robust,” somehow, than is French agreement. And subject agreement in French non-finite clauses is considerably less pronounced than it is in French finite clauses. (In fact, it is thoroughly absent.) Now, this difference in agreement morphology Pollock exploits to control whether or not a verb can move into the lower of the two inflectional X⁰s — this is because Pollock assigned agreement to this lower head. But we can copy his idea into the assignment we have given to these positions. Tense morphology is also richer in French than it is in English, in both finite and non-finite clauses; and agreement is richer in finite clauses than it is in non-finite clauses too. In fact, there is no evidence of agreement at all in infinitival clauses.

Even if this “robustness” in the morphology associated with the T⁰ and Agr⁰ positions is not what is relevant (and further cases suggests that it isn’t), let’s continue to exploit the language that Pollock adopts for whatever the relevant difference turns out to be. Those heads that are associated with morphology that allows main verbs to occupy them, let’s call “strong.” And those head positions that are not associated with morphology that allows main verbs to sit in them, let’s call “weak.” We can now describe in these terms what we’ve seen so far. In particular, we will ascribe to the Tense and Agreement positions different “strengths,” and in this way describe how they control the Move X⁰ operation. For English and French, we might assign the

⁷More or less — there are some cases in which agreement morphology exists, though in a very impoverished form, and verb movement past negation nonetheless appears to be blocked. (This happens in certain dialects of Faroese, in a register of Norwegian, and in a Swedish dialect heavily influenced by Finnish.) Some of these facts are reported in the Vikner book, where references to the relevant literature can also be found.

values as shown in (56) and (57) below.

(56) **English:**

- a. Finite:
Tense: weak
Agreement: weak
- b. Infinitive:
Tense: *to*
Agreement: empty?

(57) **French:**

- a. Finite:
Tense: can't tell
Agreement: strong
- b. Infinitive:
Tense: strong
Agreement: empty

In these value assignments, I've decided to treat the infinitival marker in English, i.e., *to*, as something that can optionally move from T^0 to Agr^0 . On this view it is rather like an auxiliary verb in French, and should behave similarly, therefore, in infinitival clauses. For some speakers of English there is, in fact, an alternation that resembles that we've seen for French auxiliary verbs:

- (58) a. Jerry tried [not to eat candy].
b. Jerry tried [to not eat candy].

Note that we cannot determine whether finite T^0 is strong or weak in French because verbs are always driven up to Agr^0 in this context.

Turning now to Scandinavian, we have something like:

(59) **Danish, Norwegian and Swedish:**

- a. Finite:
Tense: weak
Agreement: not there
- b. Infinitive:
haven't looked yet (it varies across the group, as it turns out)

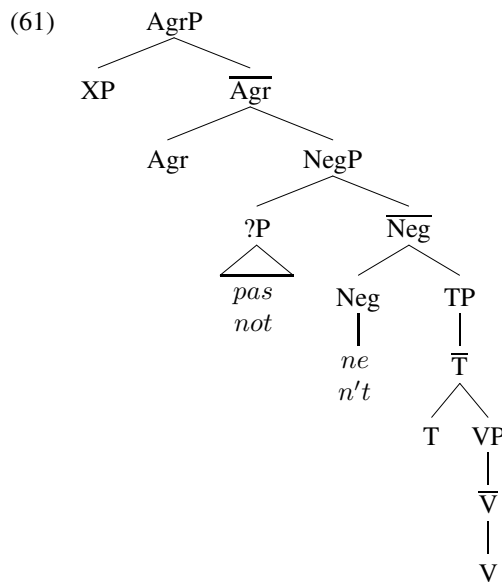
(60) **Icelandic:**

- a. Finite:
Tense: can't tell
Agreement: strong
- b. Infinitive:
haven't looked yet

My decision to classify the Mainland group as having weak tense morphology is based on the fact that this language group behaves like English, and not like French, with

respect to the positioning of main verbs relative to their complements. That is, in this group, non-complements cannot intervene between main verb and NP objects (except in contexts where the verb has relocated into C^0).

Okay, this is one part of the story. Let's now address the problem that treating negation as a head poses. Recall that we should not expect the Head Movement Constraint to allow movement of T^0 to Agr^0 past negation, if negation is a head that lies between them. Pollock's suggestion is that *not* and *pas* are in the Specifier of NegP, and that *ne* is what heads NegP in French. In English, Kayne has suggested that *n't* heads NegP.⁸ Thus, the correct way of parsing sentences with sentential negation in them is as in (61).

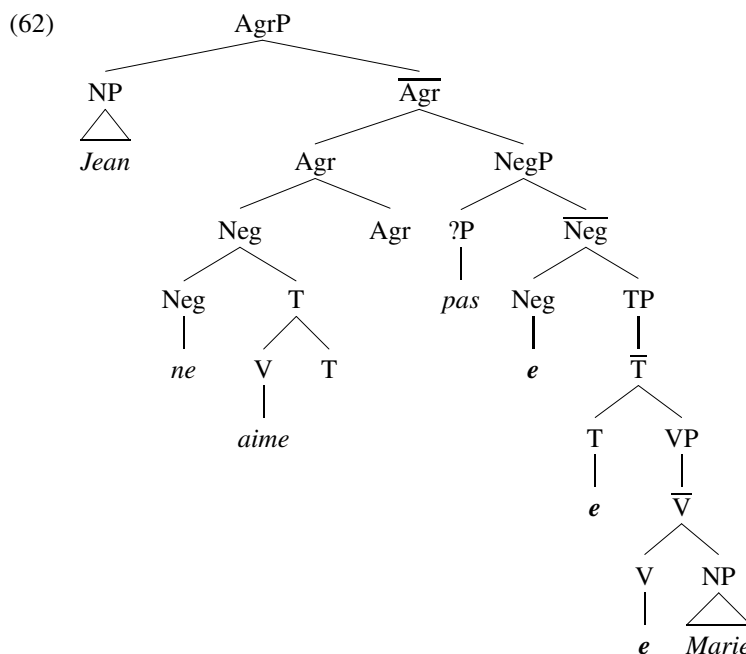


On this view, then, when T^0 moves to Agr^0 it doesn't have to skip Neg^0 , it can move through this position and pick up the head (if there is one) on the way.⁹ In a finite French clause, then, the surface parse might look roughly as in (62) on the following page. As this parse indicates, *ne* gets carried to the position before *pas* by the tensed verb as it moves into Agr^0 .¹⁰ Note that we have to let something force the bound morphemes, that is, the tense and agreement suffixes, to linearize to the right of the term that has adjoined to them, but force *ne* to linearize to the left of the term that has adjoined to it.

⁸See Kayne (2000).

⁹This predicts that it should be possible to get both the contracted *n't* and the full *not* in one sentence in English. Perhaps things like "Mary couldn't not do her homework" are such cases.

¹⁰The S-structure in (62) is derived as follows (ignoring the movement of the subject). The main verb, *aime*, is moved into T^0 , where it combines with the tense morpheme. The resulting tensed verb (that is, T^0) moves and adjoins to *ne*. the resulting negated verb (that is, Neg^0) moves into Agr^0 , where the agreement morpheme combines with the verb. This surface parse, then, claims that negation — *ne* — is closer to the verbal stem than is the agreement morpheme, and this is somewhat surprising.



This can't be the only way that *ne* manages to get to the left of *pas*, because *ne* shows up before *pas* in infinitival clauses as well, even when no verb has moved into Agr⁰. In these situations, we might imagine that *ne* moves on its own into Agr⁰. This, in fact, is what Pollock suggests. He hypothesizes that *ne* adjoins to Agr⁰ on its own in much the same way that so-called “clitics” do. “Clitics” are prosodically weak pronouns that, in the Romance languages, surface in construction with a verb; more particularly, they form a prosodic word with a verb and are therefore adjacent to that verb (or adjacent to something else that is part of the verb). In finite clauses, it is the finite verb that clitics appear with, and in infinitival clauses, it is typically the infinitival verb that the clitic appears on. One way of capturing this distribution is let clitics adjoin to Agr⁰, where they will come close enough to the finite or infinitival verb to be parsed with that verb prosodically.

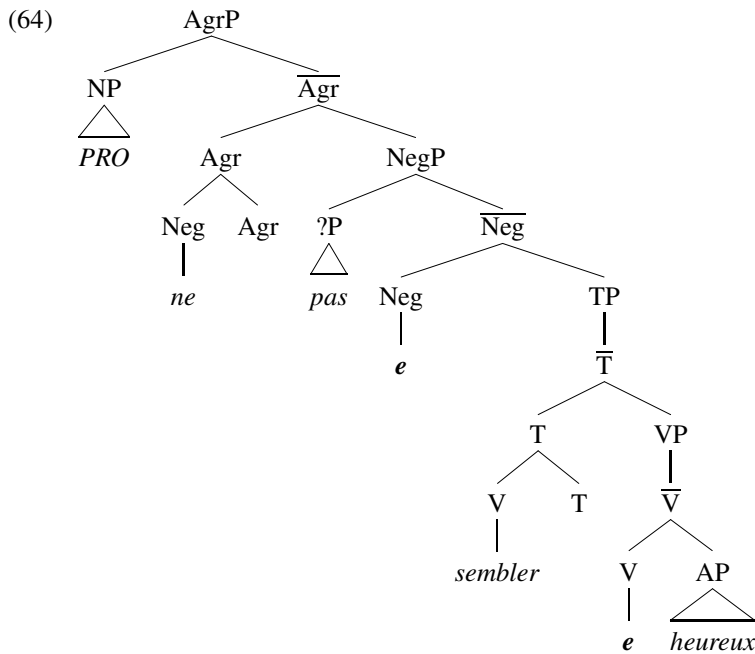
Clitics, then, instantiate another instance of spread. Object clitics, for instance, satisfy the Theta Criterion and Projection Principle by way of occupying a position that is sister to the verb they are arguments of, but surface in Agr⁰ position. This type of spread is called “Cliticization,” and it is commonly modeled derivationally just as we have done with Argument and Verb spread. Cliticization doesn't have the same properties that Head Movement has — it isn't subject to the Head Movement Constraint, for instance — and therefore we can treat it as being independent of the rules we are investigating here.

If we adopt Pollock's suggestion, that *ne* cliticizes to Agr⁰, we might be better poised to understand why it is linearized to the left of the verb+tense+agreement com-

plex. Indeed, we could impose a general constraint on adjunction that would linearize the moved item to the left of the phrase it adjoins to. Something along the lines of (63), perhaps.

- (63) ADJUNCTION LINEARIZATION
If α adjoins to β , then α precedes β .

With this assumption about *ne*, then, an infinitive that doesn't have verb movement to Agr^0 , but is negated, might have an S-structure like (64) below.



So, now what we would like to do is see if we can't get a better understanding of what it is about this “strong”/“weak” contrast that makes main verbs susceptible to it, but not auxiliary verbs. Pollock makes a proposal in this direction. He suggests that we think of main verbs as differing from auxiliary verbs with respect to their ability to assign θ -roles. In particular, he suggests that we define main verbs as ones that have a θ -role to assign, and auxiliary verbs as ones that don't. He then crafts a proposal that would prevent verbs from assigning θ -roles from a position associated with a “weak” inflectional head. And this will essentially prevent main verbs from both moving to and moving through the positions we have assigned the “weak” value. His idea goes something like this:

- (65) a. Assume that movement operations leave a “trace” of the moved X^0 or XP. This amounts to treating the “e”s in my abbreviations of derivations as real terms in the parses that movement operations create.
b. Let the Theta Criterion be enforced at S-structure.

- c. Let the trace of a verb be capable of assigning the verb's θ -role only if the verb has not adjoined to a "weak" X^0 .

There are some immediate problems with this idea that Pollock himself addresses. It wouldn't seem able to account for the British (67), nor for the contrast in (66).

(66) Mary hasn't a dime.

- (67) a. Mary isn't unhappy.
b. *Mary seemsn't unhappy.

Why doesn't the *be* in (66a) have the same number of θ -roles as (66b)? And it sure looks like *has* in (67) is capable of moving past negation, though it appears equally like it assigns θ -roles. Pollock suggests that the θ -role assigned in (67) comes not from the verb, but from some hidden term. Something similar might be attempted in distinguishing the cases in (66).

But there is another kind of problem, specifically for the prediction that movement through a weak X^0 is unavailable to main verbs, which seems to me insurmountable. Danish, as we have seen, and the other Mainland Scandinavian languages, show Verb Second word-order; (68) is an example, for instance, where the main verb has relocated to C^0 position.

- (68) Købte han bogen
bought he book-the
'Did he buy the book?'

But we have also just discovered that Danish has weak tense morphology. Thus, if we preserve the Head Movement Constraint, this entails that Danish main verbs are capable of moving into T^0 and from there moving into C^0 . If Danish T^0 is weak, this means, in turn, that moving through a weak X^0 is possible for main verbs.

For this reason, I will interpret the weak/strong typology to determine only whether a main verb may surface in the relevant X^0 . This seems to be the way these values are interpreted most commonly in the literature. It is, in fact, on this basis that I gave the assignments of "weak" and "strong" to Agr^0 and T^0 in English, French and the Scandinavian languages in (56), (57), (59) and (60). In these assignments, then, I have given a "weak" value to those X^0 positions for which there is evidence that main verbs cannot surface, and to those X^0 positions where there is evidence that main verbs can surface I have given the "strong" value. I have also indicated certain positions — namely, French and English infinitival Agr^0 — as empty positions. This is my, somewhat idiosyncratic, way of expressing the optionality of movement to these spots. Because they have no morpheme in them, these infinitival Agr^0 s do not cause the Word Criterion force a verb stem to move into this position. But, because it is present, it is still a possible landing site for X^0 movement. It's in this respect, then, that the infinitival Agr^0 s of these languages are different from the finite Agr^0 of, say, Danish. Movement of a verb to the position we've identified for finite Agr^0 is not possible in Danish — neither auxiliary nor main verbs may surface in this position — and I've coded that here by making Agr^0 completely absent in Danish.

Under this typology, X^0 positions identified with inflectional morphemes come in four strengths. In their weakest manifestation, they are completely absent, and nothing, as a consequence, may surface in their position. In their next strongest manifestation, they are present but hold no morpheme. In this guise we find optional movement of the infinitival marker *to* and auxiliary verbs, but main verbs are still not capable of surfacing in them. (This is what we see for the infinitival Agr^0 in French.) In their next strongest manifestation, they hold a “weak” morpheme. At this level of strength, the Word Criterion kicks in and requires that something join with the morpheme by S-structure; the “weak” property, however, prevents main verbs from doing this by adjoining to the morpheme. Finally, there are the X^0 positions occupied by a “strong” morpheme. These positions allow either auxiliary or main verb to surface in them, and the Word Criterion forces one or the other to. Because we want to prevent main verbs from surfacing in all but the last of these positions, the constraint that controls the auxiliary/main verb distinction should be stated as (69).

- (69) A main verb adjoined to α at S-structure creates ungrammaticality unless there is a “strong” morpheme in α .¹¹

This condition conflicts with the Word Criterion, then, in those cases where the affix should be rescued by a main verb stem but is “weak.” Those are the situations that we instead see Affix Hopping apply in. The two cases where we might expect to find this that we have discovered so far are: English main verbs, both finite and non-finite; and the Mainland Scandinavian group, in which, recall, main verbs get inflected with tense without moving from their underlying position.

Let’s take a closer look at Affix Hopping, now, and try to devolve this operation into the schema of Move X^0 and feature alchemy that we have converted verb movement into.

One important property of Affix Hopping is that it is apparently subject to the same kind of bounding constraint that we see in V^0 and I^0 Movement, but in reverse. So whereas Move V^0 and I^0 moves things always upwards and never past an X^0 , Affix Hopping moves the I^0 always downwards and never past an X^0 . This is indicated by the ungrammaticality of the cases in (70) and (71).

- (70) a. Jerry -s make [_{VP} Sally eat].
 b. * Jerry make [_{VP} Sally eats].
- (71) a. Jerry should remember [_{IP} Sally -ing eat].
 b. * Jerry should remembering [_{IP} Sally eat].

In (70), the present tense, third person agreement morpheme (-s) moves past *make* to attach to the embedded verb, *eat*. (Recall that *make* selects a VP small clause — a clause that has no inflectional projections in it.) This satisfies the Word Criterion, and the other constraints we have seen, but it is still ungrammatical. It, and all examples like it in which the affix has moved to a verbal stem past another, are ungrammatical.

¹¹This is going to have certain consequences for the syntax of sentence in which the verb has moved into C^0 . In these cases — questions and, in the Germanic languages other than English, root finite clauses — we must imagine that there is a strong morpheme in C^0 .

In (71), the inflectional suffix *-ing* has moved up, rather than down, to attach to a higher verb. Like all examples in which the affix moves to a position it doesn't c-command, this example is ungrammatical.

In general, then, (72) seems to be true:

- (72) When an affix, α , joins with a verb, β , by way of Affix Hopping, it must be that β could have undergone Move X^0 to α .

Because Move X^0 is subject to the Upwards Constraint (that is, it can move a verb only to a c-commanding position) and the Head Movement Constraint (that is, it can't move a head past another), (72) spreads the inverse of these constraints onto Affix Hopping. As (72) makes clear, then: Affix Hopping is somehow related to Verb Movement. It is Verb Movement in reverse.

There's another way in which Verb Movement and Affix Hopping are connected. The places where Verb Movement is responsible for rescuing an affix from the Stray Affix Filter, Affix Hopping is prevented from applying. In the context of finite clauses, for instance, auxiliary verbs in English are moved into the tense and agreement positions rather than the tense and agreement morphemes hopping down onto the auxiliary verbs. This is reflected in the fact that finite auxiliary verbs are driven into the position above negation, and can't, like main verbs, remain in their underlying position following negation. In fact, the only place in English where Affix Hopping occurs is just where Verb Movement cannot rescue an affix from the Word Criterion: in finite clauses that have no auxiliary verbs.

We see something similar when we look cross-linguistically. In those languages where the main verbs are capable of moving to join with affixes, that is what is done to satisfy the Word Criterion. In French, for instance, main verbs are capable of moving through the tense position into the agreement position, and this, not Affix Hopping, is how these morphemes are brought together. Again, this can be seen by the fact that the finite main verbs of French appear to the left of negation.

We need, then, to explain the dependency Affix Hopping has on Verb Movement; a dependency can be expressed with (73).

- (73) If Move X^0 can bring α and β together, then Affix Hopping can't.

A successful method of assimilating Affix Hopping into the framework we've developed for inflecting verbs should explain these properties. It should also explain why it appears to violate the Upwards Constraint, that is, why it relocates the affix to a non-c-commanding position. Though, as we've seen, it does obey what looks like the Upwards Constraint in reverse. And finally, it should explain why Affix Hopping is blocked by the presence of negation. This is one respect in which Affix Hopping and Verb Movement are dissimilar. As we have seen before, the presence of not completely destroys a finite clause without an auxiliary verb:

- (74) *Sally not eats natto.

For some reason, the presence of *not* blocks the only recourse available for satisfying the Word Criterion in this context: Affix Hopping.

We're going to look at one attempt to assimilate Affix Hopping into this framework that Noam Chomsky proposes.¹² His proposal doesn't capture all of the properties of Affix Hopping that we need to explain — it leaves out what makes (74) bad, for example — but it does point in a direction that has been influential, and that does a good job of explaining (72) and (73). The leading idea behind this proposal is that Head Movement — or Verb Movement in these particular cases — and Affix Hopping are different expressions of the same thing. It's for this reason that they compete with each other in the way that (73) expresses. And it's also the reason that the Upwards Constraint and the Head Movement Constraint are found governing both processes, though, oddly, in reverse for Affix Hopping. So, we can express, in vague terms, the idea as follows:

- (75) If α is an affix and β its verbal stem, then the Word Criterion is satisfied only if β Head Moves to α , forming an inflected verb, $\beta+\alpha$. In this situation, $\beta+\alpha$ is pronounced in the position of β , unless it can be pronounced in the position of α .

This way of putting the idea divorces the position things move to from the position they are pronounced in, and that isn't currently something that we've encountered yet.

What we want to do, then, is find a way of expressing movement in such a way that it doesn't effect an overt position change in the term that is being moved. Let's call these cases of movement: "covert movement." If we can find a way of expressing covert movement, then (75) would be a way of explaining the similarities that Affix Hopping has with Verb Movement because it essentially makes Affix Hopping a particular way Verb Movement can be expressed. Thus, the leading idea of Chomsky's proposal is that Affix Hopping amounts to covert Verb Movement.

So, what's left to be done is see how we can make sense of covert Verb Movement under the notion of "derivation" that we have adopted. This is what the specifics of Chomsky's proposal address. He suggests that we abandon the idea that the actual morphemes expressing tense and agreement are housed in the head positions we have seen verbs surfacing in. Instead, he suggests that these morphemes are always present on the verbs themselves. That is, he suggests that we insert verbs into the syntactic representation in their inflected form. The head positions that we have formerly assigned the tense and agreement morphemes to, he suggests instead have "features" which draw verbs to them. In particular, suppose that T^0 has "tense features" which attract verbs with tense morphology, and, likewise, that Agr^0 has "agreement features" which attract verbs with agreement morphology on them. This will give us a way of expressing the correlation between inflectional class and syntactic position that underlies the idea of Verb Movement without actually letting the syntax inflect verbs. As a consequence, it is now possible to see an inflected verb in the V^0 position; unlike our earlier model, inflected verbs don't only exist after they have moved into T^0 or Agr^0 , or an affix has lowered.

What this change means, though, is that we can no longer force a verb to move into a position in which a bound morpheme stands by way of the Word Criterion. Instead we're going to have to formulate a condition that requires a head with a tense

¹²What I will present here is a pared down version of his proposals in chapter 3 of Chomsky (1995).

or agreement feature to be paired with a verb bearing tense or agreement morphology. Chomsky suggests doing this with the following two principles:

- (76) An inflectional feature must be deleted.¹³
- (77) An inflectional feature is deleted when it shares an X^0 with a term that bears matching morphology.¹⁴

Like the Stray Affix Filter, or Word Criteria, this is going to guarantee that a T^0 or Agr^0 that is equipped with inflectional features comes to share an X^0 with a verb. As before, this is going to be satisfied by virtue of Verb Movement, which will bring the verb, and the morphemes it bears, into the T^0 and Agr^0 positions whose features need to be deleted.

What this part of the proposal does, then, is allow a verb to be in its underlying position and still be inflected. That's partly what we needed to express the idea in (75). Now what we need is a way of allowing Verb Movement to put the verb in the position it starts out in, not the position it moves to; that is, we need a way of enacting covert Verb Movement. For this, Chomsky relies on an idea that was developed for a different, but very similar, purpose. As we shall see, there are situations where an NP appears to be semantically interpreted in a position different from the one in which it is actually spoken. One approach to these facts has been to allow these NPs to "covertly" move into the position in which they are interpreted. So, just as in the situations we are grappling with now, these are instances where an NP surfaces in a position different from where it seems to "actually" be. The proposal made for these cases, and which Chomsky adopts for covert verb movement, is a species of the transformational, or derivational, solution to problems of these sorts. Just as we used movement transformations to solve the dilemma that arises from the observation that the position an NP gets its θ -role from and the position it gets its Case from are often different, so also are movement transformations used in solving these similar dilemmas.

Here's how the idea goes.

First, consider what derivations, as we have been using them up to now, look like. Presently, a derivation for a sentence consists of a set of parses, \mathcal{P} , one of which satisfies the Theta Criterion and Projection Principle (this is the D-structure), one of which satisfies the Case Filter and the condition on PRO, one, or more, of which satisfy the Extension to the Projection Principle, and, finally, one of which is spoken, the S-structure. These sets of parses are formed by taking a parse that is formed in accordance with the \bar{X} Skeleton, and the conditions that steer inserting lexical items into that phrase marker — the Theta Criterion, Projection Principle, Modification Rule, and so on. From this parse, the D-structure, the other members of the derivation are formed by applying rules such as Argument Movement and Head Movement. It is customary to constrain the action of this procedure so that the rules that generate members of a derivation always act sequentially, taking the last member of the derivation as input and defining a new "last element" as output.¹⁵ On this conception, derivations are the

¹³That is: there must be a phrase marker within a derivation that does not contain inflectional features.

¹⁴A morpheme *matches* an inflectional feature when the inflectional information associated with the feature is the same as the inflectional information carried by the morpheme. So, a present tense suffix matches, in this sense, a "present tense feature" borne to T^0 , and so on.

¹⁵We'll encounter an empirical consequence of this decision shortly.

ordered n -tuples in (78), where P_1 is the D-structure and each P_{i+1} is related to P_i by way of some (transformational) rule.

$$(78) \quad \mathcal{P} = (P_1, P_2, P_3, \dots, P_n)$$

We won't examine this in detail here, but as it happens, for the sorts of derivations we have been discussing, P_n , the final member of every derivation, has the form that is desired of an S-structure. This emerges as a consequence of the way that the conditions which derivations are invoked to satisfy are stated when mixed with the constraints on the rules that form derivations. For instance, with respect to Case Filter: when Argument Movement is called to create a derivation to satisfy this constraint, it will move some NP until it reaches a Case marked position, and then is prevented from moving it any farther. As a consequence, P_n of every derivation will have overt NPs in Case marked positions, and this is true of the S-structures for every derivation as well. Similarly, if Argument Movement is invoked to create a parse that satisfies the condition on PRO, it will generate a derivation that terminates in a parse in which PRO is in a position that is not c-commanded by any klose word; it will not be able to move PRO from this position because of how the constraints on movement interact with the environments in which the conditions on PRO are satisfied.¹⁶ Similar remarks hold for Head Movement. As a consequence of this, it is possible to define S-structure as the final member of a derivation.

The method of "covert movement" that Chomsky invokes in these contexts is to define S-structure in such a way that it is not necessarily the terminal item in a derivation. This makes it possible for a rule to add a member to the derivation without affecting the S-structure, and consequently gives the desired effect of "covert" movement. With this addition, we can think of derivations as consisting of two parts, the first of which terminates in an S-structure and the second of which starts with that S-structure:

$$(79) \quad \mathcal{P} = (P_1, P_2, \dots, P_n, P_{n+1}, P_{n+2}, \dots, P_{n+m})$$

As we'll see later in this course, one proposal has been to view the portion of the derivation that starts with an S-structure and produces the terminal parse as being responsible for creating a syntactic representation that the semantic component interprets. Under this view, the terminal parse is usually taken to be the one that is semantically interpreted, and is called *Logical Form* (or, LF), as a consequence. Somewhat confusingly, the portion of the derivation that stretches from S-structure to LF is also sometimes called Logical Form. Chomsky's proposal, then, is to let Verb Movement happen in the LF component.

If we adopt a view of derivations like this one, it becomes necessary to find a way of defining the S-structure. We want to preserve the effect of the earlier arrangement that S-structures place overt NPs in their Case marked positions, and PRO

¹⁶As we've seen, PRO is allowed in the Specifiers of phrases in adjunct or Specifier position, and, as it happens, there are constraints on movement we have not yet encountered that will prevent movement from phrases in adjunct of Specifier position. The only other place PRO is allowed is in the Specifier of a non-finite IP that is embedded within CP, and Argument Movement cannot move things from CP. As it happens, then, the positions in which PRO is allowed are also positions from which Argument Movement cannot move PRO.

in the positions they are licensed in; so we might take these as our criteria for being an S-structure. Putting this all together, we have (80).

- (80) a. Let \mathcal{R} be a transformational rule: that is, a function $\mathcal{R}(P_i) = P_j$, where P_i and P_j are phrase markers.
- b. Let *D-structure* be a phrase marker with lexical items that satisfies the \bar{X} Skeleton, the Projection Principle, the Theta Criterion and the Modification rule.
- c. Let *S-structure* be the phrase marker that satisfies the Case Filter and the PRO restriction and is phonologically interpreted.
- d. A *Derivation* =_{def} an ordered n-tuple of phrase markers, $\mathcal{P} = (P_1, P_2, \dots, P_n)$, such that:
- i. Each $P_i = \mathcal{R}(P_{i-1})$, and
 - ii. The first member is a D-structure, and
 - iii. The final member is semantically interpreted (i.e., an LF), and
 - iv. There is exactly one member that is an S-structure, and
 - v. For every IP, α , in the D-structure, there is some phrase marker which has a phrase in Specifier of α .¹⁷

Let's consider, now, how to steer the action of Verb movement in these derivations. If we adopt Chomsky's proposal for what makes Verb Movement "covert," we must let it happen in the LF portion of the derivation. This means that it is possible for (76) to be satisfied when the verb moves into an X^0 bearing features that need to be deleted "after" S-structure. The inflected verb remains in its underlying position at S-structure, but moves to occupy the X^0 with the feature at LF. This is the sort of derivation we want to happen just in those contexts in which we find the "Affix Hopping" phenomenon.

Now what we have to do is place controls on overt and covert Verb Movement so that we get the right one depending on the context. Our present description of when you get Verb Movement and when you don't is expressed in terms of morpheme "strength": strong morphemes allow all verbs to move into them, and weak morphemes allow only auxiliary verbs to move into them. Using the same strong/weak terminology, but applied to features now, Chomsky proposes to determine when main or auxiliary verbs move with the following:

- (81) a. Strong features kill S-structures.¹⁸
- b. Auxiliary verbs are not allowed in the LF component.¹⁹

This can be put into the terms (80) uses to define derivations with (82) and (83).

¹⁷This is the Extension to the Projection Principle.

¹⁸Chomsky actually proposes that strong features kill Phonetic Form (PF), a level of representation that is derived from S-structure, and feeds the phonetic interpretative component. Because there are no operations that will make relevant changes to S-structure for us, we can ignore the PF versus S-structure difference.

¹⁹Chomsky suggests that this is because auxiliary verbs have no semantic content, and that there is a constraint that requires everything in the LF component to have a semantic interpretation.

- (82) Let an *S-structure* be a phrase marker that satisfies the Case Filter, the PRO restriction, and **has no strong features in it** and is phonologically interpreted.
- (83) DELETE AUXILIARY
Delete all auxiliary verbs.
Condition: Must apply to the S-structure.

What (82) does is, essentially, enforce the Word Criterion at S-structure for those X^0 's that have a strong feature. Thus, (82) prevents covert Verb Movement from bringing a verb with the matching morphemes into an X^0 with strong features; instead, only Verb Movement that feeds S-structure, or overt Verb Movement, can be used. Thus, the strong features in the Agr^0 and T^0 position of finite clauses in French, or Icelandic, will require a verb with matching morphology to relocate into these positions by S-structure. A consequence of (82), then, is that the only time covert Verb Movement can be used to satisfy (76) is when an X^0 has a weak feature.

What (83) does is prevent auxiliary verbs from moving covertly. The condition on this rule forces the first part in the LF component to lack all auxiliary verbs. As a result, if an auxiliary verb has a morpheme that is needed to check off a feature on some X^0 , then the auxiliary verb must move into that X^0 by S-structure. Thus, whether the feature to be deleted is strong or weak, the only way an auxiliary verb can be used to satisfy (76) is if it moves overtly. This is how Chomsky proposes to explain the fact that in English finite clauses, auxiliary verbs are forced to overtly move into the Agr^0 bearing a weak feature.

Together, then, (82) and (83) correctly force overt verb movement in cases where strong features are found, or when auxiliary verbs are involved. Thus, they allow covert Verb Movement to occur only when a main verb is checking off a weak feature. Now, all we have to do is find a way of making sure that in this context — main verb checking off a weak feature — not only are we allowed to use covert movement, but we are prevented from using overt movement. Chomsky suggests we do this with a principle that favors covertly moving terms over overtly moving them. Of the various ways of expressing this idea, we can adopt:

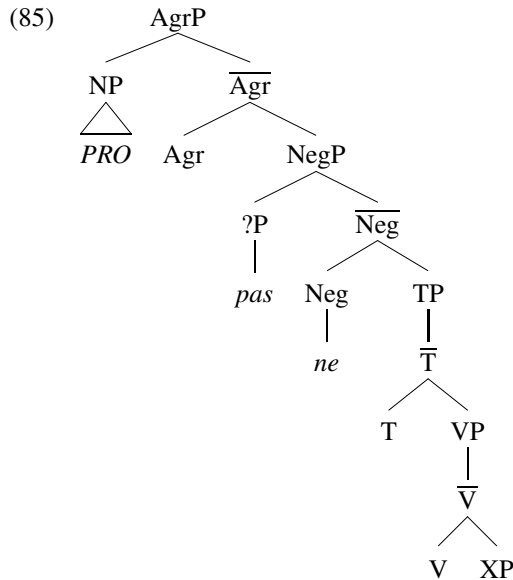
- (84) ECONOMIZE
If two derivations equally satisfy the same constraints, choose the one that minimizes the D-structure to S-structure portion of the derivation.

Because main verbs are able to move into an X^0 that bears a weak feature in the S-structure to LF portion of the derivation, they are forced to do so by Economize. In other words, when main verbs check off weak features, they must do so by covert Verb Movement.

This system works differently than the one we started with, but it has many of the same consequences. It will, as we have seen, force overt Verb Movement in finite English clauses when an auxiliary verb is involved, and it will similarly force overt Verb Movement in the finite clauses of French and Icelandic. Moreover, it will force covert Movement of main verbs in English finite clauses, and in the finite clauses of Danish, Swedish and Norwegian too. (The difference here being that English main verbs covertly move through T^0 (bearing a weak tense feature) and into Agr^0 (which we must now assign weak features to), whereas the mainland Scandinavian group doesn't

have Agr^0 , and so their main verbs move covertly just to T^0 .) Where this system and the one we started with have different consequences is in infinitival clauses, so let's look back at these cases.

Consider first French infinitives, which have a structure like that in (85) before Head Movement has applied.



What we desire is to let the verb move to T^0 , and to allow auxiliary, but not main, verbs to move on to Agr^0 . Moreover, we want the movement from T^0 to Agr^0 that auxiliary verbs make to be optional. We've not seen any evidence that determines whether the verb's movement to T^0 is optional or not, but it's consistent with the facts to take this movement to be obligatory.²⁰ If we do that, then we want to let T^0 have strong features in French infinitives, forcing overt movement of both auxiliary and main verbs into T^0 .

Now, what about Agr^0 ? In my earlier characterization of French infinitival Agr^0 , I suggested that it was empty. If we continued to characterize it this way, then Agr^0 would presumably lack features altogether; and this, of course, would mean that

²⁰Recall that the reason Pollock gives us for believing that verbs move to T^0 in French infinitival clauses is that they can be separated from their complements by a non-argument (an adverb, for instance). This is explained, recall, by letting the main verb move leftwards across the non-argument. By this criterion, there is, in fact, some evidence that the movement to T^0 is optional, since in cases such as (1) it is possible for an adverb to fall either either between the verb and its object, or to the left of the verb

- (1) a. Ne pas lire souvent Joyce est compréhensible.
 not to read often Joyce is comprehensible.
 'To not read Joyce is understandable'
- b. Ne pas souvent lire Joyce est compréhensible
 not often to read Joyce is comprehensible
 'To not read Joyce is understandable.'

It could be, however, that this alternation is due to the relative freedom of adverb placement — nothing would prevent *souvent* from being able to occupy a position to the left of T^0 .

it does not have strong features. As a consequence, (82) will not force verbs to overtly move to Agr⁰ in French infinitives. When verbs aren't forced to move overtly, then Economize is going to force them to move covertly. That means that neither main nor auxiliary verbs should be able to overtly move to infinitival Agr⁰ in French. That's not the result we want, because, as we've just reviewed, it's only main verbs that cannot move to Agr⁰ in French infinitives.

We'll face a very similar problem in English infinitives. Recall that under my characterization of English infinitives, the fact that the infinitival marker *to* can appear either before or after negation corresponds to the fact that *to* is generated in T⁰, and moves, optionally, into Agr⁰. But as before, what can the features on Agr⁰ be that will allow optional overt movement of *to*? If they are strong features, then overt movement will be forced (by (83)), and if they are weak features, then Economize will force movement to be covert.

In fact, as these two cases make clear, one of the consequences of Chomsky's Economize proposal is that it will make overt movement impossible unless it is forced. Or, to put it differently, it amounts to predicting that there is no such thing as "optional" overt movement. That's why these two cases are problematic, because they seem to involve precisely that: optional overt movement.

We'll have to reanalyze these cases if we are to maintain Chomsky's proposal, then. Let me make the following suggestions. For English, let us adopt the hypothesis that *to* can be generated in either T⁰ or Agr⁰. This amounts to the claim that we have two *tos* in English, one that is an Agr⁰ and another that is a T⁰. On this view, we will not have to move to into Agr⁰ from T⁰, and so there is no consequent optional movement. For French, I suggest that we take the occurrence of Agr⁰ in infinitival clauses to be optional. When French infinitives don't have Agr⁰, then they are essentially like finite clauses in Danish (Swedish and Norwegian) and movement of auxiliary and main verbs past negation is blocked (since there is no position to move to). When Agr⁰ is present in French infinitival clauses, then we will need to find a way of preventing main verbs from moving to it, but allowing auxiliary verbs to move. Under Chomsky's proposal, this can be done by giving Agr⁰ weak features. Because the features are weak, main verbs won't be forced to move overtly to Agr⁰, and Economize will then force the main verb to move covertly. But, when the verb that would check off these features is an auxiliary, (83) is going to prevent the auxiliary from moving covertly to Agr⁰, and as a result it will be forced to move overtly. Thus, when Agr⁰ is present in French infinitives, it will drive an auxiliary verb into them overtly but still prevent a main verb from moving overtly.²¹ And when they are absent, we get movement of neither main nor auxiliary verb.

²¹One problem with this solution is that it does not mix successfully with the account of *ne pas* in infinitives that we adopted. Recall that in infinitives we find that *ne* appears to the left of *pas*, even when no verb has moved into Agr⁰. This is unexpected if *ne* heads NegP, and *pas* is in its Specifier; they should be in the reverse order. Pollock's suggestion, which we adopted, was that *ne* instead "cliticizes" onto Agr⁰. But according to the way I have just fit Chomsky's proposal to the case of French infinitives, whenever Agr⁰ is present, it forces movement of auxiliary verbs. Thus, when Agr⁰ isn't present in infinitival clauses, both *ne* and an auxiliary verb should surface to the right of *pas*; or when Agr⁰ is present, both *ne* and the auxiliary verb should surface to the left of *pas*. What happens, by contrast, is that *ne* always appears to the left of *pas*, even when the auxiliary verb surfaces to the right of *pas*. Something is wrong.

Let's adopt this method of expressing the controls on verb movement, and in this way assimilate Affix Hopping to the general schema that Pollock's framework affords. It successfully accounts for the odd relationship between Affix Hopping and Verb Movement, and does so in a way that is consistent with Pollock's description of the main verb/auxiliary verb contrast in terms of morpheme "strength." What it doesn't do, it should be noted, is provide an explanation for the fact that negation destroys Affix Hopping, or as we would describe it now, covert Verb Movement. And, because it no longer says that verbs and their inflectionally morphology are brought together by Head Movement, it is not longer equipped to express Baker's Mirror Principle.

Chapter 6

Determiner Phrases and Noun Movement

One of the puzzles we stumbled over in introducing phrase structure rules involved the internal shape of determiner phrases. I noted that the set of strings that constitute DPs is miserably anemic. There are very few examples that offer themselves as possible determiner phrases; typically, a determiner phrase appears to be constituted of nothing more than its head. The sorts of examples I offered as candidates for this family of strings were things like the bold-faced material in (1).

- (1) a. **all but three** determiners
b. **more than six** children
c. **two dozen** eggs

But there are reasons to think that these examples don't have parallel parses, and that, in fact, none of them fit to a DP string in the desired way. It's probable that *dozen*, in (1c) is an adjective; this can be seen by observing that it can follow other adjectives (something determiners aren't capable of doing):

- (2) an expensive dozen eggs

(1a) involves a coordinator, *but*, which will invoke the sorts of structures we have encountered before with coordinations. (1b) involves what is known as a "comparative construction," whose syntax, like that with coordinations, invokes larger joined structures. We won't examine these cases in any detail here, but let me offer as a way of thinking about the syntax of these cases that makes their semantics transparent, something along the lines of (3), where the ~~struck-out~~ material should be understood as syntactically present, but phonetically absent.

- (3) a. [all ~~of the determiners~~] but [three determiners]
b. [more ~~of the children~~] than [six children]

Imagine, that is, that these cases involve bringing two full NPs together, and that a process of ellipsis removes the \bar{N} from the first NP and, moreover, this \bar{N} is understood to refer to the same set of individuals that the \bar{N} in the other NP refers to.

If these cases don't involve strings that have the same distribution as determiners, then where are these strings? Why, then, are there so few clear examples of determiner phrases?

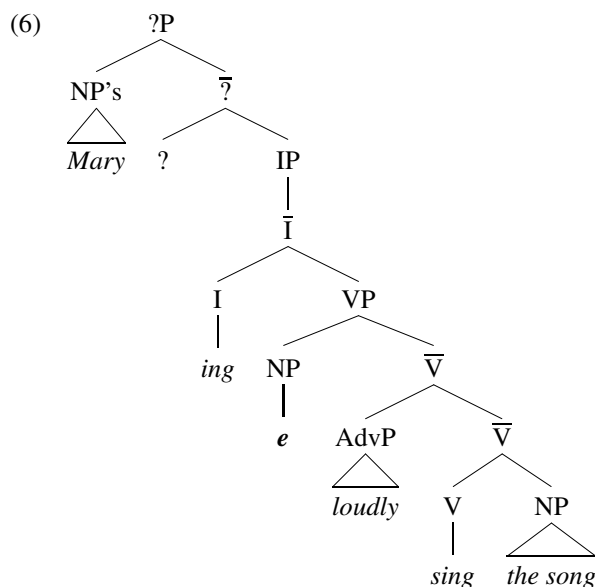
This problem can be related to another, worse, problem. Remember that determiner phrases compete with genitive NPs for Specifier of NP:

- (4) a. Mary's lamp
b. the lamp
c. * the Mary's lamp

We adopted a view of NPs that embraced a constraint — yet to be found — that limited DPs and genitive NPs to their Specifier position. That is, this fact was encoded in the Phrase Structure rules we began with as a fact about how NPs are expanded; but when we transitioned to a category-free set of phrase structure rules, this fact has to be encoded in the general principles that govern where modifiers are allowed. In fact, the method of positioning modifiers that we adopted doesn't have this consequence, and so this is business left unfinished in our transit from grammars that are free of phrase structure rules. The problem is that there are some phrases that look rather like clauses, but which nonetheless have genitive subjects:

- (5) a. [Mary's loudly singing the song] bothered us.
b. I recalled [Mary's fixing the car].
c. [Mary's having talked to John] wasn't widely known.

This suggests that these phrases have an organization something like (6).



But what's the “?” in this graph?

There is some evidence that ?P has the same distribution as NPs. Recall that NPs are subject to the Case Filter, and as such, are unable to stand after adjectives,

which apparently are incapable of assigning Case. The same is true for these sorts of expressions:

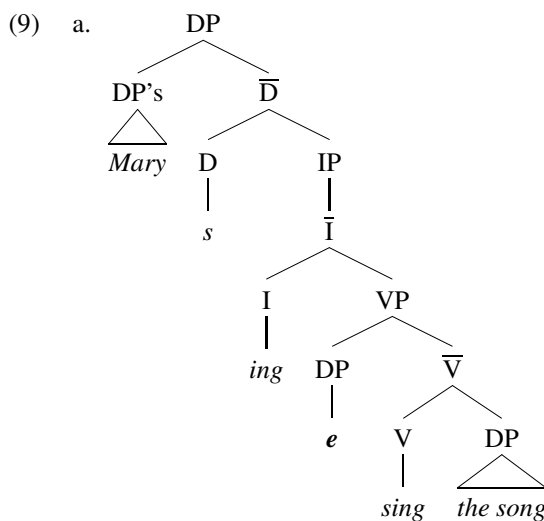
- (7) a. I was happy with Mary's singing the song.
 b. *I was happy Mary's singing the song.
 (compare: "I was happy that Mary sang the song.")

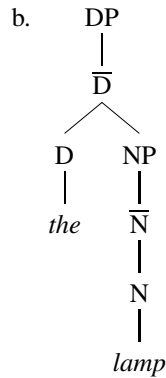
And these expressions can be conjoined with NPs, which, if we've got the constraints on Coördination correct, also indicates that they are NPs.

- (8) [Mary's singing the song] and [my subsequent departure] enraged the organizers.
 (compare: "[*Mary's singing the song] and [that I subsequently departed] enraged the organizers.")

But if ?P is a noun phrase, then the law of endocentricity demands that ? be a noun, contrary to fact. Something's amiss.

One way of thinking about this problem goes as follows. What makes the distribution of ?P look like that of noun phrases is the presence of the genitive NP. So, maybe we should call "?" the head that determines genitive Case on its Specifier. Maybe, actually, it is the genitive *s* itself. This would mean that the distribution of Genitive Phrases is the same as NPs (perhaps). And since Genitive NPs are in complementary distribution with determiners, maybe we should rethink how we earlier characterized the phrases that we called NPs. Maybe they are in fact determiner phrases, as in (9).





If this is correct, it would also answer the problem we began with. The reason DPs look so anemic is because they're considerably larger than we thought.

The two considerations I've just adduced in favor of reanalyzing NPs as DPs with NPs inside them can be found in Abney's dissertation — the first two chapters of which I've assigned as reading.¹ He gives another, smaller, argument on behalf of this reanalysis that relies on a mystery concerning the expression of Adjective Phrases in English. The mystery is that there is a constraint on Adjective Phrases in English nominals which determines how large they can be depending on whether they precede or follow the noun. As (10) shows, when an AP has nothing but its head in it, it prefers preceding the noun, whereas if it contains material following the A⁰, it prefers following the head.

- (10) a. some angry children
 b. *some children angry
- (11) a. *some angry at Bill children
 b. some children angry at Bill

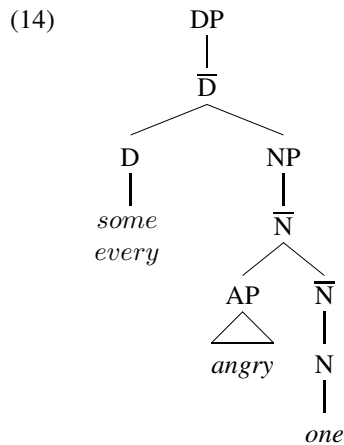
There are two exceptions to this, and these are the expressions *everyone* and *someone*:

- (12) a. someone angry
 b. *angry someone
- (13) a. everyone angry
 b. *angry everyone

Abney² suggests that an analysis of this exception should not make it accidental that the determiners *every* and *some*, and the noun *one* are involved. More particularly, it should not be accidental that the only expressions in English which seem to be made up of a determiner and noun sequence should be the very expressions which seem to violate this generalization. He recommends that we see these cases as coming about through movement of *one* onto the determiner; that is, he suggests that (12) be involved in a derivation that includes the parses in (14).

¹I've changed slightly his discussion of cases like "Mary's singing the song" — but the spirit of the argument is his.

²Who is here following a suggestion of Richard Kayne's, who in turn is building on ideas in Postal (1969).



Now Head Movement can combine one with D^0 to form the DPs in (12).

Further, to the extent that combining *one* with *some/every* really involves Head Movement, we have an argument for the reanalysis of NPs into DPs. This is because the Upwards Constraint and the Likes Attracts Likes constraint combine to allow *one* to adjoin only to a head that c-commands its original position. So, if the *some* and *every* parts of *someone* and *everyone* are determiners, and the *one* part is a noun that has Head Moved into these determiners, then it must be that D^0 c-commands NP.

If these considerations are on the right track, it demands that we change our way of talking about nominals altogether. Everything we once thought to be true of NPs, is now true of DPs instead. So, for instance, the Case Filter is something that fixes the position of DPs, not NPs. NPs are now found pretty much only inside DPs and not, as we previously thought, in subject and object positions. From this point forwards, then, let everything that we have credited to NPs hold of DPs instead, and let NPs be selected only by determiners, thereby fixing their position within DPs.³

We have also seen, faintly, evidence that nouns move internal to DPs in a way somewhat like the movement of verbs internal to CPs. Indeed, there is a variety of interesting evidence that Noun Movement exists to a larger degree than just that found in the *someone* and *everyone* cases mentioned above. Moreover, there is some evidence that this movement relocates a noun to a head associated with inflectional morphology, much like the situations we have viewed involving verbs and inflectional morphology. In gross terms, then, DPs and IPs have a variety of parallels; it is this parallelism that Abney focuses on in the chapters that we are reading.⁴

In this chapter, we will examine a paradigm of facts which focus on the noun movement part of this parallelism.

One paradigm of facts that has been accounted for in terms of noun movement concerns a difference in the position of a noun's "Subject," which is how we might

³We will quickly see that there is evidence that there are phrases that DPs embed that in turn embed NPs, and, consequently, there is reason to believe that D^0 does not select NP. If this evidence is correct, we'll need instead to let D^0 s select the phrase whose head selects NP. In general, what we'll want is to guarantee that the terms which select NPs are always found within DPs (unless, of course, we discover that NPs can be found elsewhere).

⁴The first two chapters in Abney (1987).

characterize the terms that appear as genitives in English. In Romance (by which I will mean here Catalan, standard French and Italian), the “subject” argument can appear between the noun and its complements.

- (15) a. L'opinione di Maria di Gianni (Spanish)
 the opinion of Mary of John
 ‘Mary’s opinion of John’
 b. les novel·les d’en Pere de Maria (Catalan)
 the novel of Pere of Maria
 ‘Pere’s novel of Mary’
 c. le portrait de chaque peintre étranger de son enfant (French)
 the portrait of each painter foreign of his child
 ‘the picture by each foreign painter of his child’

This could be made sense of if we adopt the Derived Subjects Hypothesis for nominals as well as for clauses, and suppose that there is N^0 movement in Romance but not English at S-structure. If we understand the Derived Subjects Hypothesis to claim that it is the highest \bar{N} which assigns the θ -role that “subjects” in DPs receive, then this will put these subjects in Specifier of NP underlyingly. If nouns then move leftwards in Romance, and the subjects of DPs can remain in their underlying position, this will have the desired consequence of placing nouns to the left of the subjects.

I think the first argument of this sort comes from Cinque,⁵ who makes the argument based on the position of “ethnic” adjectives, which also can be found post-nominally in Romance.

- (16) L'invasione tedesca dell'Austria.
 the invasion german of Austria

Ethnic adjectives seem able to bear a subject θ -role assigned by a noun. So, consider, for example, the contrast in (17)

- (17) a. the American car in the showroom
 b. the American waltz on the radio
 c. the American opinion of the blockade
 d. the American discussion of trade barriers

In (17a), *American* serves as a genuine modifier, merely attributing to the referent involved that it has the property of being American. But in (17b), *American* refers to an abstract entity that is constituted of the American people or the American government.⁶ That this is a function of these adjectives bearing the “subject” θ -role which *opinion* and *discussion* assign is indicated by the fact that this meaning is lost when there is another external θ -role bearer in the nominal.

- (18) a. Uganda’s American opinion of the blockade
 b. Morocco’s American discussion of trade barriers

⁵In an unpublished talk delivered at the Università di Venezia in 1990. See his GLOW abstract from the 1992 Lisbon meeting.

⁶This quality of ethnic adjectives is first discussed, to my knowledge, in (Kayne 1984a, Chapter 7).

Like other external θ -role bearers, then, this one shows up postnominally in Italian (and the other Romance languages), which can be accounted for if we let these adjectives be underlyingly placed in Specifier of NP, and then make nouns move leftwards past them.

There is an alternative method of generating these word orders. Giorgi and Longobardi (1991) suggest that there is a difference in the direction that Specifiers can branch in Romance and Germanic, and that this is responsible for the fact that the phrases that appear in these Specifiers, i.e., “subjects,” show up following the noun in Romance but not Germanic. This alternative account predicts that postnominal “subjects” can follow the complements, and this is generally possible too. Thus the *dil/de* phrases in (15) are actually ambiguous; either of them can have the subject or object reading. Giorgi and Longobardi suggest that this word-order alternation arises by virtue of a rule that moves the “object” past the right-branching subject position. The Noun Movement account would have to claim that the subject can move rightwards past the object.

There are reasons for doubting that the Giorgi and Longobardi account is correct, and this direction has largely been abandoned in the literature. One of these is that, as Valois (1991a) and Bernstein (1993) note, “ethnic” adjectives cannot follow complements in Romance.

- (19) *L’invasione dell’Austria tedesca
the invasion of Austria german

(Valois 1991a, p. 374)

This can be related to the fact that ethnic adjectives seem unable to move. There is evidence in English for this which comes from the fact that ethnic adjectives are unable to undergo the passive-like operation that nominals support in examples like (20), compare (21).

- (20) a. Iran’s bombardment by Russia took several weeks.
b. Uganda’s invasion by Tanzania grinds slowly on.
(21) a. *The Iranian bombardment by Russia took weeks.
b. *The Ugandan invasion by Tanzania grinds slowly on.
(with an object interpretation for the adjective

(basically Kayne’s 1984, (32) and (33), p. 139)

As we shall have occasion to see, there is evidence that the genitives in (20) have undergone A Movement from a position to the right of the noun, where they receive their θ -role. Ethnic adjectives, apparently, are unable to move from this position. Instead, they are stuck in the position from which they get their θ -role. Thus, the fact that they appear in Romance between the noun and the noun’s complements suggests that the underlying position to which the external θ -role is assigned in nominals is to the left of the complement. This follows if the Specifier of NP branches to the left rather than the right. This fact, then, fits the model of Romance nominals which has the noun moving leftwards past the subject.

So, now, where are the nouns moving in these cases? One possibility, explored in a variety of places, is that the intermediary position is where number morphology is associated. There is some *prima facie* reason for thinking that number morphology

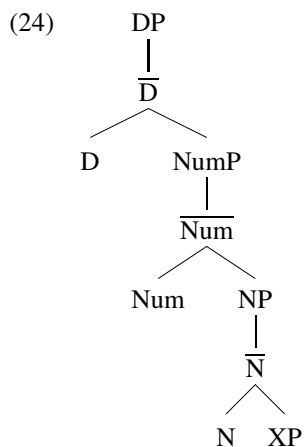
heads a syntactic phrase: Cross-linguistically this is common, as Dryer (1989) shows. Thus, in Yapese, for example, the plural/singular/dual categories are expressed with separate morphemes.

- (22) a. ea rea kaarroo neey
sing car this
b. ea gal kaarroo neey
dual car this
c. ea pi kaarroo neey
plural car this

This at least suggests that Universal Grammar makes projecting a syntactic phrase above Number a possibility. Further, Dryer shows that the relative order of Num⁰ and Noun correlates with Verb-Object word order. This would be explained, on standard Theory, if Num⁰ is in a head complement relation with Nouns. Further, Dryer finds that most times there is a number word, it falls more embedded in the nominal than do determiners, but above adjectives and the noun. He provides examples like the following.

- (23) a. ha ongo puha'e ua
art dual box two
(Tongan)
b. do mamu ragha
tree big plural
(Kimaghama)
c. me-ria rabiri
plur-new paddle
(Cayuvava)

There are exceptions, but this can be said to be the “basic” order among these terms. If so, the pattern that emerges can be sketched in (24).



This is what the statistical study yields.

Of course, if Dryer's conclusion that adjectives come between Num⁰ and N⁰ is valid for English, then English nouns must combine with this morphology in one of the ways we have seen possible in the verb/inflection cases. One possibility is that nouns overtly move to Number, but that this doesn't bring the noun to the left of the possessor in English because possessors are driven into Specifier of DP, an even higher position. Or, we might imagine that the noun undergoes covert movement to Number head.

So now what we want to determine is: Is there language internal evidence for the picture that Dryer gives us statistically? And, is there evidence that bears on whether English differs from other languages in not enjoying overt N⁰ movement to Number?

Let's tackle the second question first.

We have seen evidence for the movability of nouns in Universal Grammar. Is there language internal evidence that the site of this movement is Num⁰. The best argument I know for this in the literature is found in Bernstein's paper, who manufactures a Correlation Argument. She claims that there is reason to believe that the position of nouns relative to adjectives correlates with the presence of number morphology on the noun. Her evidence comes chiefly from a comparison of Walloon and standard French. The contrast she describes is very like one that holds between English and French, however, so I will begin with an illustration of this difference.

In French, but not English (with the exception we've already noted), it is possible for single adjectives to follow the noun they modify.

- | | | | |
|------|----|--------------------|-----------|
| (25) | a. | dès bêtes malades | (French) |
| | b. | some sick animals | (English) |
| | | *some animals sick | |

It is also possible to find prenominal single adjectives in French, as in the following example.

- | | | |
|------|----|------------------|
| (26) | a. | une large vallée |
| | a | large valley |
| | b. | une vallée large |
| | a | valley large |

But here Bernstein notes that there is a difference in meaning: in (26a), the nominal refers to an individual drawn from the set of things that are large valleys. In (26b), by contrast, a "member of a class of valleys which happens to be large" is denoted. In Giorgi and Longobardi's study of this phenomenon in Italian, they suggest that the prenominal depictive adjective can only get an appositive interpretation, whereas the postnominal one can have either an appositive or restrictive reading. The difference between an "appositive" and a "restrictive" reading is subtle. Roughly speaking, appositive modifiers contribute their meaning to the expression they are attached to in a way that is reminiscent of conjunction. So, for instance, in (27a) the PP *from Duluth* stands in the same relation to *Mary* as it does in (27b).

- (27) a. Mary, from Duluth, has arrived.
 b. Mary has arrived and she is from Duluth.

In an example such as (28), by contrast, *from Duluth* plays a more direct role in determining the reference of the DP it is attached to.

- (28) Jill likes women from Duluth.

In this case, *from Duluth* restricts the reference of women to just those that have an attribute that Jill values: being from Duluth. One could not capture the meaning conveyed by (28) with a circumlocution, parallel to (27b), like:

- (29) Jill likes women, and they are from Duluth.

Perhaps it is this sort of difference in meaning that correlates with the pre-nominal/post-nominal position of adjectives.

If so, it doesn't appear to always be the case, however. There are some examples where the alternation between Adj+N and N+Adj order doesn't appear to invoke any meaning difference. Valois (1991b) provides some examples in nominals with a deverbal noun.

- (30) a. La probable invasion de Jupiter
 the probable invasion of Jupiter
 La invasion probable de Jupiter
 the invasion probable of Jupiter
 b. La fréquente invasion de Jupiter
 the frequent invasion of Jupiter
 La invasion fréquente de Jupiter
 the invasion frequent of Jupiter

(Valois 1991, p. 374)

Valois claims that there is no difference in meaning attendant with these word orders. What's going on here will have to await a better understanding of the syntax-to-semantics mapping of modification.

What is the source of the difference between French and English with respect to placement of these single adjectives. Why can they appear after the noun in French but not in English?

One possibility would be to blame whatever it is that prohibits bare adjectives from being right-adjoined to the nominal projection as the cause. Maybe this constraint is not present in French? Actually, however, there is evidence that this constraint is also present in French. We've seen that bare adjectives can follow the noun, but they cannot follow the noun's complement, as in the following example.

- (31) *L'invasion de Jupiter complète

If bare adjectives could right-adjoin to a projection of a noun, there would be no reason for this asymmetry — they should be able to follow everything that is in an NP. But, on the other hand, if we assume that this constraint operates in French just in the same way

that it does in English, then (31) will be ungrammatical for the same reason that the English version of this DP is. And, the fact that bare adjectives can follow single nouns, on the other hand, can be explained if we allow nouns to Head Move past adjectives in French, but not English.

Now, interestingly, Bernstein shows that Walloon patterns with English, and not with the other Romance languages, with regard to adjective placement.

- (32) a. dés malâtès bièsses
the sick beasts
b. *dés bièsses malâtès
the beasts sick

(Walloon)

She suggests that the difference between French and Walloon is that the noun moves past the adjective in French but not in English or Walloon. And, further, she argues that the contrast between Walloon and French correlates with a difference in the way that number morphology is expressed on the noun. This, she argues, suggests that the position the nouns are moving to in French is a position associated with number morphology.

Her argument that number morphology is centrally implicated rests on several observations. First, she points out that neither colloquial French nor Walloon show (phonetically) the plural morphology that is orthographically present on nouns. Thus the bracketed portions of the following nouns are not pronounced.

- (33) a. dès r'tchâfés cromptîre[s]
the reheated potatoes
b. des petites fille[s]
the small girls

(Walloon)

(French)

These plural affixes are not present phonetically even when the context for liaison is provided.

However, she notes that in French there is a certain class of suppletive forms where the nouns do show a morphological mark for plurality. Examples are given below.

- (34) a. un mal
an evil
b. des maux
evil(s)
(35) a. un oeil rouge
a red eye
b. des yeux rouges
red eye(s)
(36) a. quel cheval
which horse

- b. quels chevaux
 which horses

In Walloon, however, these forms always appear just in their singular form.

- (37) a. on mâ
 evil

- b. dès mâ[s]
 evils

- (38) a. on rothch[e]-oûy
 red eye

- b. dès rodj[e]-z-oûy
 red eyes

- (39) a. [ke: dzva:]
 which horse

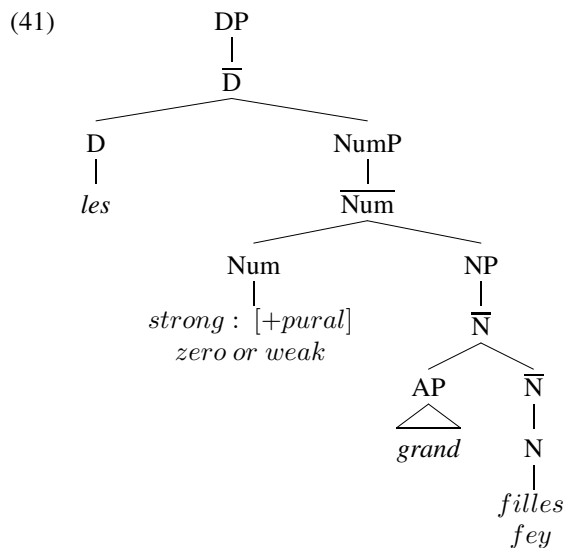
- b. [ke: dzva:]
 which horses

Further, She notes that liaison is possible in literary French, but never in Walloon.⁷ Thus, the orthographically present, but normally not phonetically manifest, plural suffix is pronounced in literary French in contexts like the following.

- (40) Les train-z-ont du retard.
 the train-s-are running late

From these data, she concludes that the plural affix in French is *-es*, or a suppletive trigger, and that it is absent altogether in Walloon (well, almost we'll revise this immediately). Thus, using the inflectional feature model we adapted from Chomsky in connection with verbal syntax, this gives to French/Walloon nominals the following D-structure.

⁷but it's not that liaison is completely absent in Walloon — Bernstein notes that it is still present after plural determiners and pronouns (for this, see (Bernstein 1991, note 7, p. 107).



In French, the head noun is driven overtly into the Num⁰ position to delete the strong plural feature residing there, and this will bring it past the adjective. This movement is blocked in Walloon because either there is no NumP, or its head is associated with a weak feature. Thus the availability of plural forms in French is correlated with its ability to appear before single adjectives. This achieves the desired correlation between presence of number morphology and N+Adj word order, and also supports the idea that number morphology is associated with an inflectional category that projects its own phrase.

Is it possible to tell whether Walloon has a NumP, or whether it is absent altogether? Bernstein suggests that there are reasons for assuming that Walloon does have NumP and, moreover, there is some reason to believe that it is actually filled with morphology. If this is correct, the crucial difference between Walloon and French is not whether NumP is present or not, but instead how it combines with the noun that follows. Interestingly, Bernstein argues that it combines in a way that we would not have expected from our examination of verbal syntax. She argues that the plural morpheme in Walloon begins in Num⁰ and attaches to the left edge of the following noun; but, somewhat surprisingly, it shows up orthographically as the final syllable of an adjective which precedes the noun. Let's briefly examine how she arrives at this conclusion.

One fact, due to ?, that leads her in this direction is that liaison between prenominal adjectives and a following noun is absent in Walloon, though present in French.

- (42) a. un gro-z -arbre
a big tree
b. une peti-t -enfant
a little child

(French)

- (43) a. on gro[s] abe
a big tree

(Liège Walloon)

- b. on peti[t] èfant
a small child

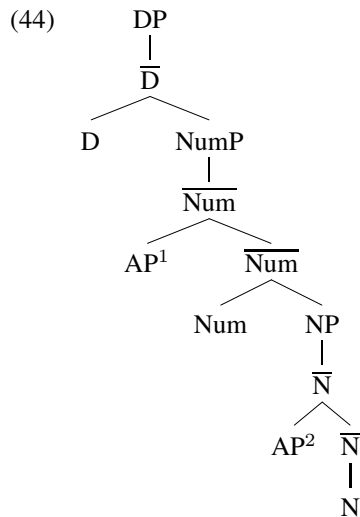
(Gondecourt Picard)

She suggests that the account offered here would provide an immediate explanation for this, if in Walloon there is a Num⁰ that lies between the prenominal adjective and the following noun. This intervening category might then be responsible for blocking liaison in Walloon. In French, by contrast, the noun has moved into Num⁰, and is therefore in a position to trigger liaison with a preceding adjective. For this reason, she suggests that Num⁰ should be present in Walloon, but not able to attract N⁰s to it.

Note how this account presupposes that adjectives cannot come between Num⁰ and N⁰; if they could, then an intervening Num⁰ cannot be blamed on the absence of liaison. This, however, is incompatible with the proposal that adjectives follow nouns (when they do) because of movement to Num⁰.⁸ Thus, either Bernstein must posit two differences between French and Walloon — adding that in addition to the differing expressions of noun movement they also differ in placement of adjectives — or something is being missed here. Note in particular that adopting a view which holds that there are two independent differences between French and Walloon will destroy the basis for the Correlation Argument.

In fact, we have another kind of problem that arises as soon as we adopt the view that it is movement of Nouns to Num⁰ that is responsible for the N+Adj word order. In the system we have adopted from Chomsky, there is no way of making head movement systematically optional. And, as we have seen, in the Romance languages which allow the N+Adj word order, the Adj+N word order is also possible. Moreover, recall that with some adjectives, there is a difference in meaning that correlates with these two orders. We need to find a way of fitting these facts to our goal of correlating the availability of the N+Adj word order with overt movement to Num⁰. One way we could do this is to imagine that adjectives can be placed either above or below Num⁰, as indicated in (44).

⁸It also runs counter the intuition that adjectives are modifying nouns, or their projections, and how we have expressed the syntactic relationship that holds between modifiers and the things they modify.



Once the noun has moved into Num⁰, there is still, according to this model, a place for adjectives to the left of the Num⁰+N⁰ pair. In order to account for the meaning difference that (sometimes) arises, we might imagine that adjectives in the AP² position get a different interpretation (maybe restrictive) than do adjectives in the AP¹ position. We might seek an account for this difference from the fact that these adjectives are modifying different things: a Num⁰ in one case and a N⁰ in the other. This way of modeling the meaning difference, then, would predict that, with respect to those adjectives that show the difference, whether the adjective appears to the left or right of the noun will completely disambiguate its meaning. Thus, for example, if an adjective shows up to the left of the noun, it'll have to be non-restrictive; whereas if it appears to the right of the noun, it'll have to be restrictive. I don't know if this is a correct outcome.

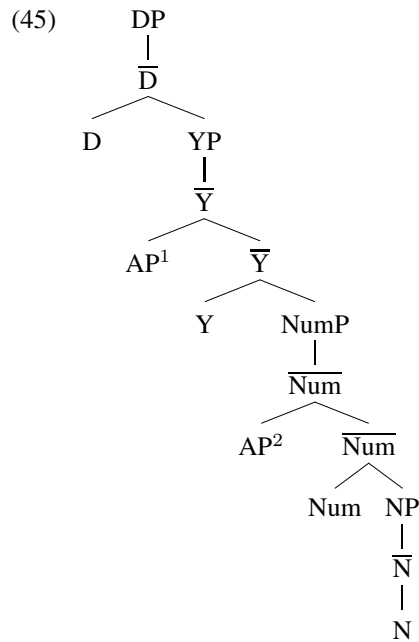
A different way of modeling the N+Adj/Adj+N word order, that still correlates the availability of the N+Adj order with overt Noun Movement to Num⁰ and would also address the problem we encountered with Bernstein's explanation for the contrast in liaison between Walloon and French, would be to hypothesize an optional projection above NumP. Then we could place adjectives on either side of this optional projection, and let the noun move into its head when it is present. I don't know what this optional phrase is, so I will simply call it "Y" here. The idea, then, would be to give to DPs the shape in (46).⁹

⁹Relevant to this decision is that ethnic adjectives can't appear prenominally in (standard) Italian or French.

- (1) *la tedesca invasione dell'Austria
the german invasion of Austria
- (2) quel tedeschissimo comportamento
that very German behavior

(Valois 1991, p. 374)

To the extent, then, that ethnic adjectives show us where the D-structure position of external θ -role bearers are in nominals, this fact suggests that these external θ -role bearers are positioned before objects in Romance nominals.



Now, as before, let nouns be driven into Num⁰ by a strong feature in Romance. When YP is absent, then nouns will surface to the right of adjectives, both those in AP¹ and those in AP² position. If, as before, we associate these two positions with the two interpretations that these adjectives can get, we will, in this situation, allow prenominal adjectives to have either interpretation. When YP is present, assume that Y⁰ has a strong feature, and the Num⁰+N⁰ pair will be driven into Y⁰. In that situation, the noun will surface to the left of adjectives in AP², thus appearing to the left of adjectives with a restrictive interpretation, and will still remain to the right of adjectives in AP², presumably those with a non-restrictive interpretation.

Consider, by contrast, a language which does not allow Nouns to move into Num⁰, presumably Walloon for instance. In these languages, movement into Y⁰ will be blocked by the Head Movement Constraint. That is, if we could find something that not only prevented Nouns from surfacing in Num⁰, but also prevented them from moving through Num⁰, then we would have a way of correlating access to Y⁰ with access to Num⁰.

This, in fact, is what Bernstein does. She argues that the plural morpheme in Walloon is placed in Num⁰, and this blocks movement of the noun in Walloon.

There are two plural morphemes in Walloon, one for feminine nouns and one for masculine nouns. Both are expressed orthographically on the prenominal adjective. The “feminine plural morpheme” is realized before consonant initial words as an unstressed vowel and before vowel initial words as [ɛz]. Examples follow.

- (46) a. les belè[s] feyes
the pretty girls
b. dès neûrè-z -amonnes
some black berries

Compare:

- (47) li neûr sipène
the black thorn

The masculine plural morpheme (-s) shows a similar pattern, though it is phonetically manifest only in contexts of liaison.

- (48) a. dès deûr[s] tchivès
the black hair
b. dès neûr-z -ouy
the black eyes

She argues against composing the feminine plural marking of a gender morpheme and a number morpheme because this would assign to feminine the suffix *-e*, and this doesn't show up in singular nominals.

- (49) li neûr sipène
the black thorn

So she supposes that there is only one morpheme, a plural one, that is to be found here. And, she conjectures that this morpheme is portmanteau with gender, or what she calls a word-marker, following work by Jim Harris. From now on we will illustrate this morpheme with the phonetically more salient feminine one.

The evidence that these morphemes are actually attached to the noun that follows them is as follows. First, only prenominal adjectives show this morphology.

- (50) a. Èle sont neûr.
they are black
b. Èle sont tot[es] petit[es].
the are very little

And when the nominal that the adjective precedes is absent, this morpheme does not appear, as in these copular constructions:

- (51) a. C'è dès bèl[es].
those are good
b. *C'è dès bèle[s].
those are good

Second, only one of these morphemes appears when two prenominal adjectives are conjoined.

- (52) dès bèl[es] èt bounè[s] bièsses
some nice and good animals

This, at least, is the case in one dialect of Walloon (Boulogne Picard). In another, Liège Walloon, it is possible to find the plural morpheme on all adjectives in the series.

- (53) *dès bèlè[s] gradè[s] djônè[s] fèy[es]*
 some nice and strong young girls

She suggests that in these cases, the adjectives aren't actually stacked, but are instead conjoined. She notes that the conjunction is *es* in Walloon, and therefore homophonous with the plural morpheme.

Third, there is phonological evidence that this morpheme is a proclitic on the following noun and not suffixed onto the preceding adjective. First, there is a widespread process of final obstruent devoicing in Walloon, that Bernstein illustrates with the following pair.

- (54) a. *grandeûr*
 big
 b. *grande amice [grât amis]*
 good friend

When adjectives are followed by the plural morpheme, they show obstruent final devoicing, as the contrast below illustrates.

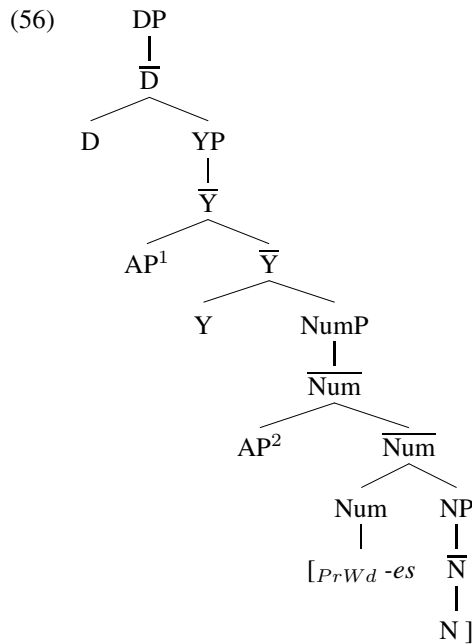
- (55) a. **grandè[s] fèyes*
 big girls
 b. *grantè[s] fèyes*
 good girls

A second phonological reason for thinking that the plural affix is not part of the preceding adjective is that it is unstressed. She cites Morin who argues that all words in Walloon have stressed final syllables. Finally, again following Morin, she points to the fact that in Gondecourt Picard, the plural morpheme, *ès*, triggers harmony on the following noun. She follows Morin and adopts the proposition that harmony is restricted to words in Walloon, which leads to the conclusion that *ès* is part of the following noun, not the preceding adjective.

This pattern of data all makes sense, Bernstein concludes, if the Walloon plural suffix combines with the following noun not by way of N^0 movement, but instead, by procliticizing onto the unmoved, following N^0 , as indicated on the facing page. (Understand the material enclosed within “[*PrWd*]” to form a phonological word.) As noted above, this will explain why Walloon nouns surface always to the right of adjectives, because they will not be able to move through Num^0 into Y^0 .

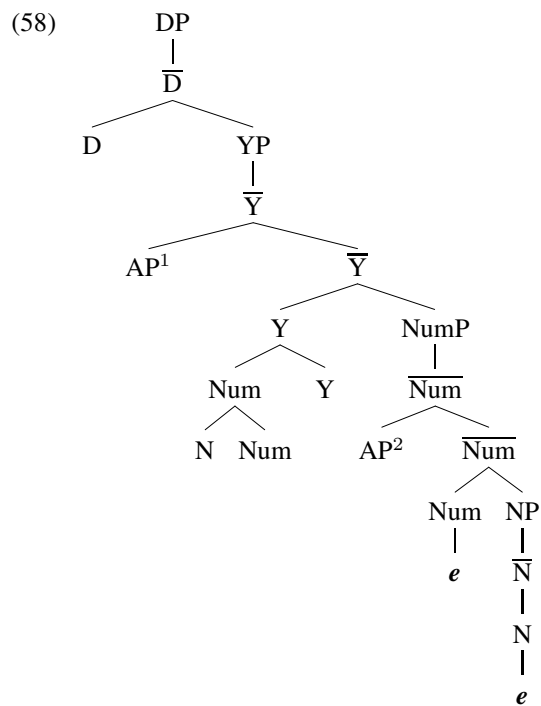
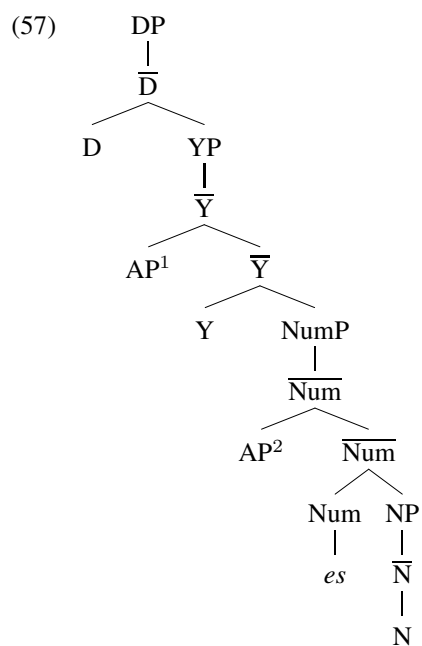
Now, if we want to transfer this account of the difference in adjective placement between Walloon and French/Spanish/Catalan to the similar difference between English and French/Spanish/Catalan, we will have to find a reason why Nouns are prevented from moving through Num^0 in English. But it is unlikely that this movement is prevented in English for the same reason that it is in Walloon. At least, there is no evidence that plural morphemes in English are proclitics on the noun, no evidence, that is, of the sort we have just witnessed for Walloon.

Given the tools developed here, perhaps the most straightforward way of modeling the difference between English and Romance is to credit the mysterious Y^0 with



strong features in French, Catalan and Spanish, but not in English or Walloon. This will force nouns in French, Catalan and Spanish to move through Num⁰ and surface in Y⁰, thereby bringing them to the left of (some) adjectives. In Walloon, Y⁰ must either have no features or weak features, because nouns remain to the right of Num⁰ in this language and consequently don't move into Y⁰ by S-structure. This is the outcome we also want for English, so in English too, Y⁰ must not have strong features. The difference between English and Walloon will reside in how these languages express number morphology. In Walloon, as we've seen, the number morpheme remains in Num⁰ and is procliticized onto the following noun. Let's speculate that this process requires nothing more of the syntax than for it to put the number morphology and the noun right next to each other. Perhaps, for instance, there is a rule of prosody that brings a number morpheme and an adjacent noun together as a single word, and so there is nothing that demands that the noun move into Num⁰ in Walloon.¹⁰ We might go so far as to find a way of preventing overt movement of the noun into Num⁰ in Walloon — this could be done by giving Num⁰ weak features, or by treating the number morpheme in Walloon as a word, and invoking the Word Criterion to prevent it from sharing Num⁰ with a noun. So, Walloon S-structures will comport DPs as in (57), while in the remaining Romance languages, nouns will surface in Y⁰ as in (58) on the next page.

¹⁰The failure of the number morpheme to materialize in situations where it is not followed by a noun might be credited to this rule failing to parse the number morpheme unless the conditions for putting it in a word with a noun are met.



Postnominal subjects don't arise in English. But this could be because the other ingredient necessary to getting postnominal subjects is missing from English:

perhaps the subjects cannot remain in Specifier of NP position. We might characterize this difference between English and Romance in terms of the positions that Case is assigned to within DPs. Let the Specifier of DP be assigned Case in both English and Romance, but let only Romance assign Case to Specifier of NP. Not that this Case is manifest in what appears to be a preposition — *di* or *de*, depending on the language. Let's call this Case, the one expressed by a preposition, "Nominal Case." On this view, then, the difference in availability of postnominal subjects between English and Romance boils down to the availability of Nominal Case in Specifier of NP. Indeed, the "subject" arguments of DPs uniformly appear with the genitive Case in English, and this is a position, as we've seen, associated Specifier of DP. Thus, no matter what its position relative to the noun, the subject of a DP cannot appear marked with Nominal Case:

- (60) a. * the discussion of Jill of the problem
 b. * the discussion of the problem of Jill
 c. * the of Jill discussion of the problem
 d. * the placement of Mark of the sofa
 e. * the placement of the sofa of Mark
 f. * the of Mark placement of the sofa

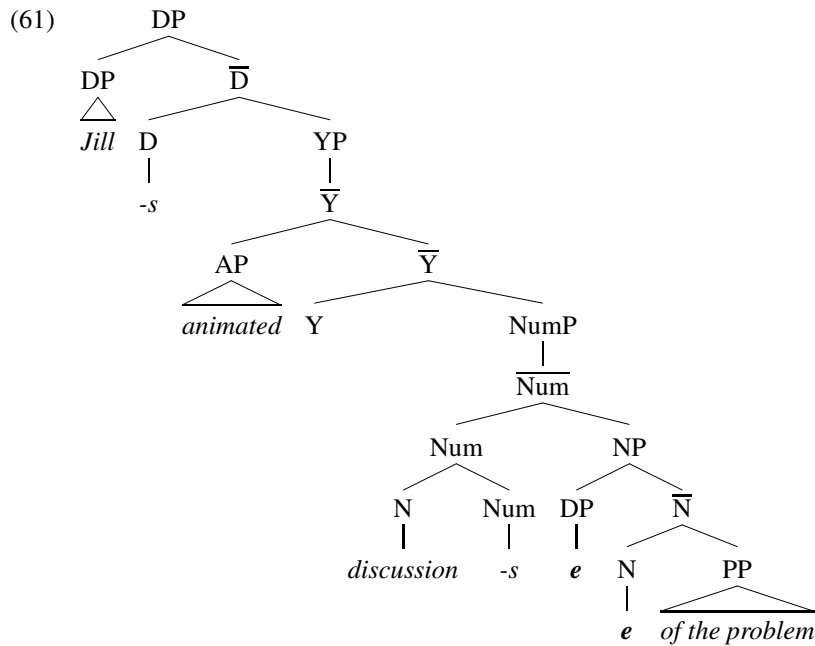
It would be reasonable, therefore, to expect the general absence of postnominal subjects in English DPs to be caused by the absence of Nominal Case, and not by the absence of noun movement in English. This hypothesis, then, would give an example like "Jill's animated discussions of the problem" a representation like that in (61) on the next page.

It's not, incidentally, that Nominal Case is completely absent in English. It is possible for Nominal Case to be found on the arguments of nouns that derive from unaccuative verbs, as in (62).

- (62) a. the death of her
 b. the arrival of it
 c. the appearance of Jill
 d. the sinking of the ship

Though it's not possible on the arguments of nouns derived from transitive verbs, even if these nouns do not express their object argument. Leaving the objects unexpressed in the examples in (60), for example, does not improve them.

- (63) a. * the discussion of Jill
 b. * the discussion of Jill
 c. * the of Jill discussion
 d. * the placement of Mark
 e. * the placement of Mark
 f. * the of Mark placement



With nouns derived from unergative verbs, the situation is somewhat intermediate.¹¹

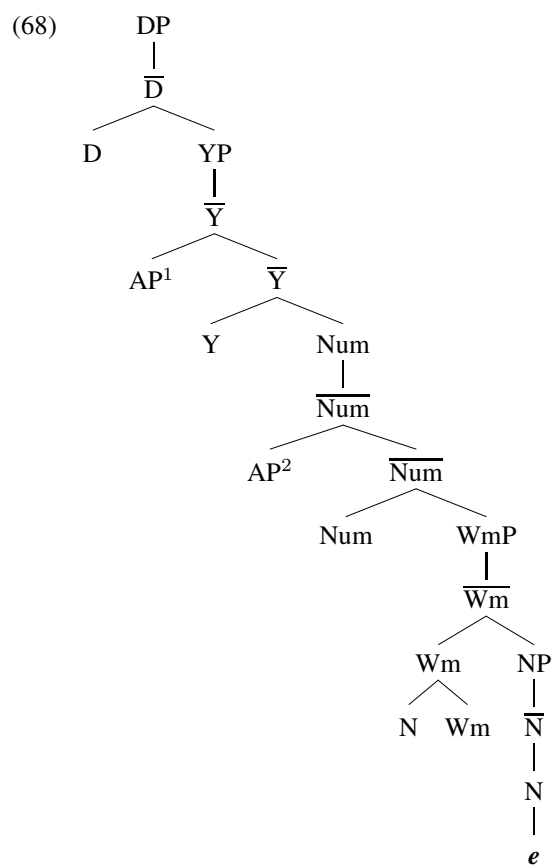
- (64)
- a. ?* the running of her
 - b. * the talking of him
 - c. ?? the dancing of Jill
 - d. ?* the speaking of the woman
 - e. ?* the sitting of Mark

If we interpret these facts as indicating that there is a distinction between the “unaccusative” nouns and the others — that is if we set aside the cause for the intermediate status of the “unergative” nouns — then this pattern can be described with (65).

- (65) NOMINAL CASE ASSIGNMENT: ENGLISH
Nominal Case is assigned to the “object position” of nouns.

We’ll set to defining what “object position” means later; but, importantly, it can’t have the same sort of definition we’ve given to the positions that verbs assign their “object” Case to if we adopt the view that nouns move overtly to Num⁰ in English. Object Case is assigned by verbs to positions they govern. If we let Nominal Case be assigned by nouns to positions they govern, then once a noun has moved to Num⁰ it should be able to assign its Case to a DP within Specifier of NP: just the effect we are hoping to avoid.

¹¹See Grimshaw (1990) for a discussion of these facts and an argument that nouns divide into the unaccusative and unergative classes.



Chapter 7

Complement Structure

We've seen now that the syntax of nominals has features similar to the syntax of clauses. Like CPs, DPs have embedded within them at least one inflectional phrase which, in turn, embeds a phrase holding the θ -role assigner. In the case of clauses, the inflectional phrases are Agreement and Tense, and the θ -role assigner is a verb and its projection. In the case of DPs, the inflectional phrases hold, perhaps, number morphology and something else, and the θ -role assigner is a noun and its projection.

There are other ways in which nominals have a syntax like that we discovered in clauses. They allow for processes which arrange the arguments to the nouns in ways similar to the process which arrange the arguments to verbs. We have seen, for example, that there is an operation that looks something like the Passive, as we briefly noted. This is clearest in cases where a DP contains an N which has been formed from a verb, as in (1).

- (1) a. The Iberians' construction of Rome.
- b. Barry's enlargement of the slide.

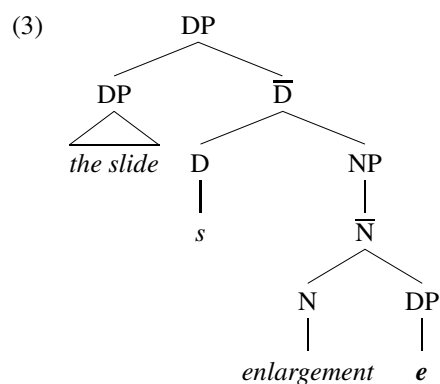
From (1) can be formed nominals like:

- (2) a. Rome's construction (by the Iberians)
- b. the slide's enlargement (by Barry)

which bear a superficial similarity to the Passive. One might imagine that, as in the passive, there is an operation which "moves" the object of a noun into Specifier of DP position in such cases. These nominals would therefore have a representation something like that in (3) on the following page, once the "object" has moved.

There are, of course, many ways in which nominals are unlike clauses, in both their internal organization and the way argument structure is expressed. Many of these, especially those having to do with argument structure, are explored in a famous paper by John Ross called "Nouniness."¹ One of these differences concerns the typology of infinitival clauses that we've seen can be the complements to verbs. Recall that these infinitives partition into two classes: transparent and opaque. The opaque ones can

¹See Ross (1974).



house PRO, prevent A chains from spanning them and don't allow Case assignment into their "subject position." Transparent infinitives have the opposite mix of properties: they cannot house PRO, the do permit A chains to span them, and they also allow Case assignment to penetrate them from outside. We characterized the difference between these clauses in terms of the status of their projection: opaque infinitives are CPs and transparent ones are AgrPs

Interestingly, nouns do not seem to permit transparent infinitives. This can be appreciated by noting that verbs which select opaque infinitives continue to do so when they are nominalized, as in (4).

- (4)
- a. I attempted [to read all of *Aspects*].
 - b. my attempt [to read all of *Aspects*]

 - c. I promised [to start reading *LSLT*].
 - d. my promise [to start reading *LSLT*]

 - e. I persuaded Sal [to read *Connectedness*].
 - f. my persuasion of Sal [to read *Connectedness*]

But, so far as I know, no verb which selects a transparent infinitive can do so once it's been nominalized. Thus, verbs which assign accusative Case to the subjects of their complement infinitive are ungrammatical when nominalized, as in (5).

- (5)
- a. I believe [Sal to be happy].
 - b. * my belief [of Sal to be happy]

 - c. I consider [Sandy to be intelligent].
 - d. * my consideration [of Sandy to be intelligent]

 - e. I find [Jeri to be a good actress].
 - f. * my finding [of Jeri to be a good actress]

And verbs whose complement infinitives have an argument that moves out of that infinitive, are also ungrammatical when nominalized.

- (6) a. I appear [to be confused].
- b. *my appearance [to be confused]

- c. This proposal is likely [to be incomplete].
- d. *this proposal's likelihood [to be incomplete]

- e. She seems [to be tired].
- f. *her semblance [to be tired]

The absence of this last class of transparent infinitive in nominals can be seen too by observing that the “passive” guise of the nominals in (5) are also ungrammatical.

- (7) a. *Sal's belief [to be happy] (by me)
- b. *Sandy's consideration [to be intelligent] (by John)
- c. *Jeri's finding [to be a good actress] (by Sal)

The “verbal” versions of these passives is perfectly grammatical, of course:

- (8) a. Sal is believed to be happy by me.
- b. Sandy is considered to be intelligent by John.
- c. Jeri was found to be a good actress by Sal.

As with (8), there is something about being the complement to a noun that makes this process ungrammatical.

What is it about transparent infinitives, then, that makes nominals lethal for them?

In fact, the absence in nominals of this second class of transparent infinitive is expected. Another discovery of John Ross's was that, in general, movement out of the clausal complements to nouns is blocked. We have not yet encountered this process, but there is a mechanism in English syntax that lets certain interrogative phrases be moved to the front of a sentence in forming questions. Thus, for instance, the expression *which woman* is moved in (9) from the position marked with *e*.

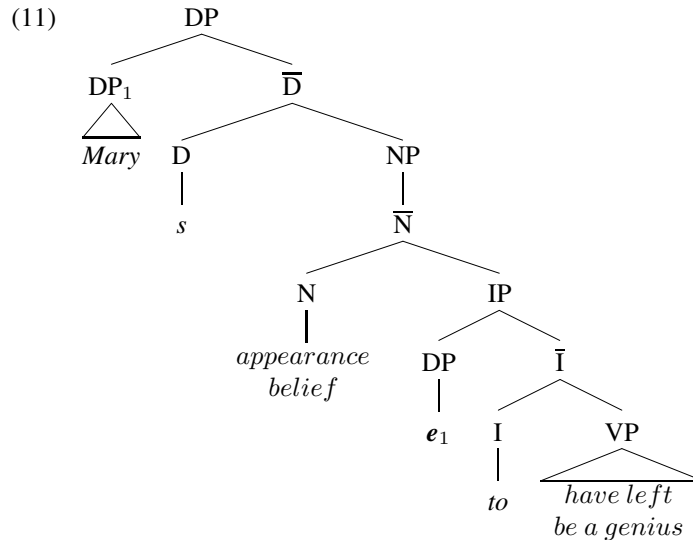
- (9) Which woman did you remember [that Sam talked to *e*]?

Note that *which woman* has moved out of the finite clause *that Sam talked to*. When this clause is embedded as the complement to a noun, however, the movement is blocked:

- (10) *Which woman did you remember [the rumor [that Sam talked to *e*]]?

Clausal complements to nouns are said to be “islands” for movement. The islandhood of clausal complements to nouns, then, will be responsible for the ungrammaticality of

(6) and (7). In these examples, as (11) shows, it will be necessary to move an argument out of the clausal island.²



Indeed, we might view the absence of transparent infinitives of this sort in nominals as evidence for the movement account we adopted for these cases. From a purely descriptive point of view, these infinitives exemplify situations in which an argument is outside the clause from which its θ -role derives. That this is correctly modeled by letting these arguments move out of the infinitives is supported here, then, by the fact that in contexts where these clauses are islands for other forms of movement operations, they are also incapable of supporting distant arguments.³

This leaves us still with the task of explaining why the other form of transparent infinitive cannot be found as complements to nouns. There is something about nouns which prevents them from supporting Case assignment into their complement infinitives, it would seem. If we treat the *of* preposition that shows up commonly on the objects to nouns as the nominal parallel to accusative Case, then what we want to know is why of assignment cannot go across infinitival complements.

Let's postpone addressing this part of the problem for a while, and turn first to consider what can be learned about other complement structures from the islandhood of complements to nouns.

²In this representation, I ignore NumP and the other inflectional phrases that might be embedded within DPs, and I similarly ignore the AgrP/TP distinction in parsing the infinitival clause.

From here on out, I will abbreviate derivations by use of the "e" notation, giving numerical subscripts to the phrases and es that movement relates. Thus, (11) abbreviates two parses: one in which *Mary* is in the position shown, and another in which it is in the position that "e" stands in. These are not the only phrase markers in this sentence's derivation — there is at least one other in which *Mary* stands in its θ -marked position. But it is sufficient to see the point I am making here with just the two parses shown. In general, this shall be my convention from now on. I will only show those parses in a sentence's derivation that are necessary to see the point at hand.

³That is, these phenomena support treating the long-distance dependency that (9) illustrates arise in questions in English with the same mechanism that is used to treat "Argument spread."

There are other systematic absences in the complement types to nouns. One of these is that “small clauses” cannot be found as complements to nouns. There are no instances where a verb that selects a small clause complement can continue to do so once it has been nominalized. The pairs in (12) illustrate this generalization.

- (12) a. I believe this fact interesting.
 b. * this fact’s belief interesting
 c. I consider this sentence instructive.
 d. * this sentence’s consideration instructive
 e. Some find these facts remarkable.
 f. * these facts’ finding remarkable

In each case, I’ve constructed a nominalized version of a verb that selects a small clause, and moved the subject argument of that small clause into the genitive Case-marked position. Doing this when the small clauses are complements to verbs is perfectly fine — witness (13).

- (13) a. This fact is believed interesting.
 b. This sentence was considered instructive.
 c. These facts were found remarkable.

It’s being the complement to a noun that’s ruining things in (12).

These cases too can be credited to the island nature of complements to nouns. If small clauses, like “big” clauses, are islands for movement, then the ungrammaticality of the cases in (12) is expected. As might be expected, it’s also true that nouns can’t assign their of Case to the subjects of small clauses. So examples like (14) are also ungrammatical.

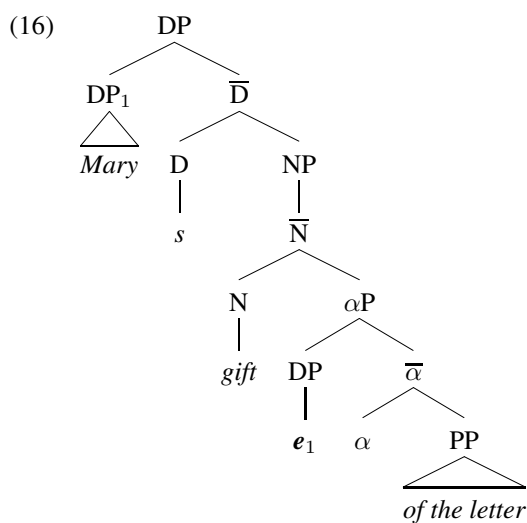
- (14) a. * my belief of this fact interesting
 b. * my consideration of this sentence instructive
 c. * my finding of these facts remarkable.

Small clauses, then, pattern with the transparent infinitives, and when we find the reason Case assignment by nouns is blocked into transparent infinitives, we can spread that account to the cases in (14).

Another place where nouns don’t seem capable of supporting the complement structure that verbs can, is found with the double object construction. We have put off understanding how these constructions can be squared against Stowell’s hypothesis that what forces DPs to be initial among complements is their need to be adjacent to the Case assigning verb. So, these stand presently as puzzles. The fact that double objects are not possible as complements to nouns, as (15) illustrates, perhaps offers information that might help us understand how to treat them.

- (15)
- a. Her teacher gave Mary the letter.
 - b. * Her teacher's gift of Mary of the letter
 - c. * Mary's gift of the letter (by her teacher)
 - d. Her classmate offered her a crayon.
 - e. * her offer of a crayon by her classmate
 - f. * her classmate's offer of her of a crayon
 - g. This salesman sold us that defective natto pot.
 - h. * our sale of that defective natto pot (by this salesman)
 - i. * this salesman's sale of us of that defective natto pot
 - j. The coach tossed him the ball.
 - k. * his toss of the ball (by the coach)
 - l. * the coaches toss of him of the ball
 - m. I rented John the office.
 - n. * John's rental of the office (by me)
 - o. * my rental of John of the office

In a pattern that is now familiar, the double object construction appears to be a sort of transparent clause. In particular, the first object of the double object can neither move in a passive like way (as indicated by the first of the bad DPs in each triple), nor can it get the of Case that nouns assign (as indicated by the second of the bad DPs in each triple). Concentrating on the cases in which the first object has moved, the ungrammaticality of these examples follows from the islandhood of complements to nouns, if there is a hidden clause in double object constructions, as indicated in (16).



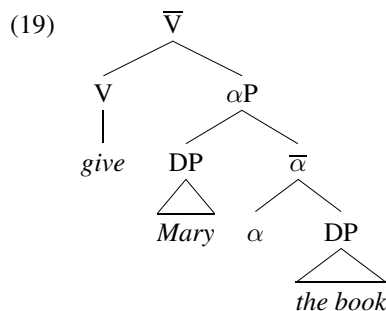
As in the other comparisons we've made, the absence of argument movement in these cases is something that distinguishes the situation that arises when the double objects are complements to nouns from when they are complements to verbs, as in (17).

- (17) a. Mary was given the letter by her teacher.
 b. She was offered a crayon by her classmate.
 c. We were sold that defective natto pot by this salesman.
 d. John was rented the office by me.

And it's also the case that it's just the double object construction that has this restriction in nominals. Other two complement constructions are permitted as complements to nouns:

- (18) a. my toss of the ball to Sally
 the ball's toss to Sally (by me)
 b. my placement of the item on the table
 the item's placement on the table (by me)
 c. my forfeiture of the game to Shawn
 the game's forfeiture to Shawn (by me)
 d. my sale of the car to you
 my car's sale to you (by me)
 e. my rental of the office to Hugh
 the offices rental to Hugh (by me)
 f. Jane's explanation of the problem to Bill
 the problem's explanation to Bill (by Jane)

These data, then, point to giving the double object construction an organization like that in (19).



When this complement structure finds itself embedded as the complement to a noun, the familiar island effects will prevent movement of the first object out of the construction.

This is the argument made in Kayne (1984b). Extending the treatment of the absence of transparent infinitives in nominals that relies on the islandhood of clausal complements to nouns to the double object construction leads to the conclusion that double objects are embedded within a hidden small clause.

There is some independent evidence that this is the right way of organizing the two DPs in a double object construction. One such piece of evidence concerns “scope.” Scope is a relationship that holds for a variety of semantic processes; it is computed on the basis of the syntactic relationship the terms involved in the process have to each other. It will emerge as we go along — and I will come back to a specific argument on behalf of this in the section on Binding Theory — that c-command is the relevant syntactic notion for computing scope. Thus:

- (20) α is in the scope of β iff α c-commands β .

The examples in (21) The first point to be made is that there appears to be evidence that the VP that contains these arguments is not flat. This can be seen, for example, by noting the scope relations between the two arguments.⁴

- (21) a. Mary showed the boys₁ each other's₁ parents.
 b. *Mary showed each other's₁ parents the boys₁.
 c. Mary gave each boy₁ his₁ toy.
 d. *Mary gave its₁ owner each toy₁.
 e. *Mary baked her₁ Sally's₁ recipe.
 f. Mary baked Rover's₁ owner Rover₁/the dog₁.

This suggests either that scope can be computed on the basis of something other than hierarchic arrangement, or that the first object is asymmetrically higher than the second.

Larson (1988) defends the parse for double objects in (19) and provides, in addition, a suggestion as to the identity of “ α .” He suggests that α is, in fact, the verb, and that the VP which contains α P — a VP he calls a “VP Shell” — is headed by an empty V^0 position into which the verb moves.

Note how this gives us a way of shoring up the double object construction with Stowell's adjacency condition on Case assignment. If we suppose that both V^0 s are able to assign accusative Case, then we can see how each of these two DPs can meet the Case Filter. That only the first DP gets structural Case, the one that is lost in the passive, Larson suggests is due to the proximity to T^0 that the verb it gets Case from has. In particular, he suggests that a V^0 that is governed by T^0 assigns structural Case, and that all other V^0 s assign inherent Case.⁵

Actually, Larson's proposals are more far reaching, because he also proposes that there is a transformational relationship between the double object construction and the alternative DP+PP frame. He gives two reasons for adopting this view (see the discussion in his paper on pp. 350-351). First, he suggests that the productive relationship between these two word-orders in some languages (the applicative of Bantu, for example) makes it a candidate for a transformational rule. The alternation is not uniformly

⁴The use of numerical subscripts in these examples is meant to indicate referential dependence. In (21a,b) and (21e,f), take the DPs which have the same subscript to be used in such a way that they refer to the same individual, or set of individuals. In (21c,d), understand the pronoun to get its referential value based on how the quantificational DP it shares a subscript with gets its referential value. This catalogue of scope asymmetries is explored in Barss and Lasnik (1986).

⁵Larson's paper predates the Pollockian revolution to clausal structure, and so he identifies T^0 with I^0 .

“productive” in English, but he suggests that this is a special property of English and makes some suggestions as to what this property is.

I don’t think this can be an argument for deriving the double object frame from the DP+PP frame syntactically, however. Productivity is a mark of a rule, nothing more. It cannot help us decide what this is a rule of.

Larson’s second reason for thinking that the two frames are transformationally related is Baker’s Uniformity of Theta Role Assignment Hypothesis (UTAH), which states:⁶

(22) UNIFORMITY OF Θ ROLE ASSIGNMENT HYPOTHESIS (UTAH)

Identical thematic relationships are represented by identical structural relations at the level of D-structure.

This reason, then, is as valid as UTAH is, and the guess that the θ -roles involved in both frames is the same.

Now to get a transformational relation between these two frames, Larson takes the standard approach and supposes that the double object word order is derived from the DP+PP word order. But in order to implement this idea he must employ a novel approach to the DP+PP frame. He adopts a Verb Movement based version of Bach/Dowty’s “Right Wrap” theory, which claims that the PP forms a constituent with the verb which excludes the DP. He offers two reasons for thinking this might be right.

First, there is an idiomatic relationship that holds between the verb and indirect object which can, in certain cases, influence the θ -role borne by the direct object, as in (23).

- (23) a. Mary took Felix to the cleaners.
b. Felix threw Oscar to the wolves.
c. Beethoven gave the fifth symphony to the world.

This is very much like the situation we saw holding of a verb and its object with respect to the θ -role assigned to the subject. We concluded in the case of the subject that the θ -role assigned came from a phrase made up of the verb and its complement. In this situation, then, we might want to make the same conclusion: the verb and indirect object form a phrase that together θ -marks the direct object. This suggests, then, that there is a phrase that holds the verb and indirect object but excludes the direct object.

The same conclusion is reached by looking at the strings that can be targeted by Gapping — an operation very much like VP Ellipsis. Larson (1990) points out that Gapping can affect strings in cases like the (24) which are discontinuous.

- (24) Some gave books to Sam, and others, magazines.

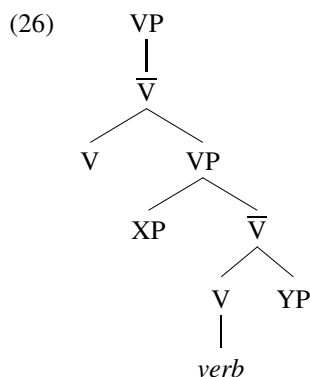
To the extent that syntactic operations like Gapping are constrained to apply only to constituents, we should conclude from cases such as these, then, that there is a constituent made up of the verb and indirect object that excludes the direct object. This is just what Larson’s proposal would provide.

⁶See Baker (1988).

Further, as Jackendoff reminds in a rejoinder to Larson's paper,⁷ there are scope asymmetries of the sort we've seen in double object constructions for all double complement verbs.

- (25) a. Mary showed every paper₁ to its₁ author.
 b. * Mary showed her₁ problem to every author₁.
 c. Mary introduced the girls₁ to each other's₁ parents.
 d. * Mary introduced eather other's₁ parents to the girls₁.
 e. Mary told the fink₁ about her₁ shortcomings.
 f. * Mary told her₁ about the fink's₁ shortcomings.

For these reasons, then (and one more we come to immediately), Larson suggests that the structure for ALL double complement verbs is as follows:



The verb will start in the lower position and move into the empty V^0 , bringing it before both of its complements.

This presents a problem for the paradigm of facts that we saw in Kayne's work. Recall that in nominals we find a crucial difference between double object constructions and other two complement constructions, a difference that suggests that only the double object construction has a hidden small clause in it. We will have to reconcile Larson's decision to treat all two complement constructions as involving an embedded clause with this contrast. More work for later.

Notice also that Larson's schema departs from Stowell's program in what it gives the responsibility for ordering complements. On Stowell's program, this is achieved by the Case Filter, which distinguished complements by their categorial status: DPs had to be in the Case marked position, and CPs had to move from it. But Larson's program orders complements according to their θ -roles. Indeed, Larson's idea is that θ -roles arrange themselves in a hierarchy that governs the position in the tree that they will be assigned to. The higher the θ -role on the hierarchy, the higher in the phrase marker that θ -role will be assigned. In particular, he suggests that AGENT is higher than THEME which is higher than GOAL/BENEFACTIVE which is higher

⁷See Jackendoff (1990).

than LOCATIVE, and so on.⁸

The final reason Larson offers for accepting this view of the arrangement of complements is that it provides a method for characterizing Ross's "NP Shift" that helps explain some of its puzzling properties. One puzzle is why leftward movement, but not "NP Shift," can strand prepositions, as the contrast in (27) indicates.

- (27) a. The boys, you talked to.
b. * You talked to yesterday the boys.

And it also appears that leftward movement of the first of the two objects in the double object construction is possible, but that "NP Shift" is not:

- (28) a. Sally, I gave the pictures.
b. * I gave the pictures all my relatives in Duluth.

Larson's proposal allows for a reanalysis of "NP Shift" that explains its differences from leftward movement. Indeed, instead of seeing "NP Shift" as the result of rightward movement of the DP, Larson suggests seeing it as leftward movement of the \bar{V} that contains the verb and the secondary complement. Larson's particular way of expressing this option is by positing a rule that "reanalyzes" as a V^0 along the following lines:

- (29) When \bar{V} contains one undischarged internal θ -role, it may be reanalyzed as a V^0 .

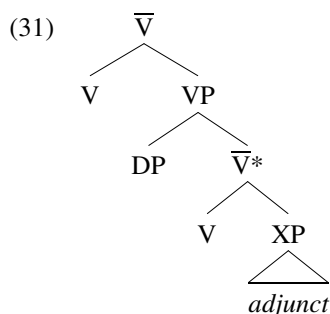
This straightforwardly explains the inability of "NP Shift" to strand prepositions because it no longer requires that NP Shift be seen as movement of the DP rightwards; instead, everything else moves to the left of the DP. Hence, NP Shift does not strand prepositions because it never targets the complement of a preposition.

If Larson's characterization of "NP Shift" is correct, then cases like the following suggest that we've got the position of adverbs screwed up.

- (30) Sam visited yesterday all his relatives in Duluth.

On Larson's characterization of NP Shift, the verb and adverb (*yesterday* in this example) have moved together past the object (*all his relatives in Duluth*). This requires that the verb and adverb form a constituent that excludes the object, and this requires that we find a way of finding a constituent in which the adverb is closer to the verb than the object is. Indeed, Larson suggests that adverbs are base-generated as sisters to the verbs they bear a relation to the verb that is simply very low on the Θ Hierarchy, Larson suggests. So he gives a parse like that in (31) for cases where a verb combines with a DP complement and a non-argument.

⁸What's wanting, then, is an explanation for why this particular arrangement of θ -roles should map onto the categories that it seems to.



His reanalysis rule will be able to affect \bar{V}^* , and thereby move verb and adjunct past the heavy DP in Specifier of VP.

Larson's view is a genuine revolution to the standard way of conceiving of the projection of argument structure. Recall that we have seen evidence from VP Ellipsis, coördination, *do so* Replacement and other constituency tests that complements are closer to the verb at D-structure than are non-arguments. It was for this reason that we formulated the Projection Principle, which forces θ -role bearers closer to their θ -role assigners than non θ -role bearers. So Larson's conjecture would require that we abandon the Projection Principle, and it seems at odds with the constituency evidence that we examined earlier.

This might be considered reason enough to abandon Larson's treatment of NP Shift, but in Larson and May (1990), he provides evidence from scope phenomena for his unconventional view about the relationship between complements and adjuncts.

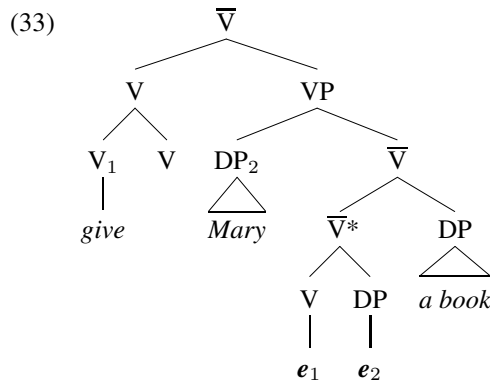
- (32) a. John visited few friends any day this week. (Larson 1990, (52):621)
- b. * John filed them₁ without reading [Mary's articles]₁
(from Contreras (1984), in Larson 1990 (53a):622)
- c. I will visit every child₁ before his₁ mother.

In these examples, it looks as if the direct object has non-arguments within its scope, which is consistent with the parse that Larson would give to these cases, but not consistent with the more conventional parses that the Projection Principle is designed to guarantee. So here, both Larson's interpretation of the "NP Shift" operation and scope fact seem to give the same picture of constituency. There appears to be a genuine conflict, then, between the picture of constituency that phenomena like coördination and anaphora reveal and the picture of constituency that scope gives.

Let's set this paradoxical situation aside for the moment; we return to it immediately.

So, we've seen how Larson's treatment of "NP Shift" accounts for its apparently puzzling boundedness. Let's now turn to its inability to affect the first of the two objects in a double object construction. Larson's solution to this problem relies in part on the mechanism he proposes is responsible for deriving the DP+DP frame from the DP+PP frame, and in part due to the condition he places on the reanalysis process that gives rise to "NP Shift" configurations. So we need to first examine his proposals for the dative transformation he proposes.

He suggests that there is a grammatical function changing operation that “demotes” the THEME argument and absorbs the Case morpheme to, forcing the GOAL argument to move into the structurally Case-marked Spec VP position. (Note, then, that this operation must rob the verb only of its inherent Case.) Thus, this operation forms the following from the DP+PP frame.



How does the demoted DP get Case? By reanalyzing the lower \bar{V} , i.e., \bar{V}^* into a V^0 , and then using the assumption that all verbs can assign inherent Accusative, but only those governed by I^0 can assign structural Case. This difference in how Case is assigned can be put to the service of untangling why Passive robs the Case of the first object only in these constructions. Assume that Passive destroys on the structural Cases that a verb assigns, and in the context of double object constructions, this will mean only the Case assigned by the higher V^0 will be lost.

It is this process of Case robbing that Larson suggests might be responsible for the fact that the double object frame is not always related to a DP+PP frame. Suppose, he suggests, that *to* assigns a very particular GOAL θ -role: “Goal of motion along some Path.” Further, imagine that it can be deleted by this process only when the θ -role assigned by the verb is identical. Thus, he supposes, cases which allow only the DP+PP frame, as in (34), are those where the verb does not assign an identical θ -role.

- (34)
- a. I donated money to charity.
 - b. * I donated charity money.
 - c. I distributed apples to the children.
 - d. * I distributed the children apples.
 - e. I contributed my time to the auction.
 - f. * I contributed the auction my time

And cases where only the DP+DP frame is possible, as in (35), he suggests arise because the θ -role that *to* would assign is not appropriate:

- (35) a. The judge spared John the ordeal.
 b. * The judge spared the ordeal to John.
 c. They denied him tenure.
 d. * They denied tenure to him.

But this, as far as I can see, is incompatible with his views about Case assignment: why would the GOAL have to get Case from *to* when the verb can assign it inherent Case? Since Larson allows inherent Case to be assigned to a DP whenever that DP is adjacent to a V^0 (or a \bar{V} that can be reanalyzed as a V^0), what would force any other Case assigning mechanism? That is, Larson has to say both that a reanalyzed \bar{V} can assign inherent Case to the demoted “subject” in a double object construction, thus to an argument that is not an underlying complement to that V^0 , and that the inherent Case which a reanalyzed \bar{V} assigns is finicky enough to care about the kind of Goal θ -role that is assigned. I can’t see how to resolve these needs: how would we guarantee that the \bar{V} *donate* heads is incapable of assigning its inherent accusative Case to *money* because that is not the right kind of Goal θ -role, but at the same time allow the \bar{V} s which reanalyze to, in general, assign their accusative Case to whatever happens to land within their government domain (hence, the demoted subjects of double object constructions generally). And it’s also unclear that his account of the other cases can go through: One wonders, for example, on what basis we can ascribe to *distribute* a different θ -role than *give*.

I think we should abandon this attempt at explaining the restrictions. So far as I know, this is still an open question; with a variety of factors appearing to be relevant.

Now we are prepared to see how Larson explains the failure of “NP Shift” to relocate the initial DP of the double object construction. As it is formulated above, the \bar{V} to V^0 reanalysis process is restricted to apply only to \bar{V} s that have (at least) one θ -role unassigned. But this is not true of the \bar{V} that would have to reanalyze to bring the second DP and verb to the left of the first DP in the double object construction. Look back at the parse of a double object construction in (33). The \bar{V} that would have to reanalyze to form a NP Shift construction of the first object is labeled \bar{V}^* in this tree. But this does not have one unassigned θ -role in it; all of the arguments of the verb heading this \bar{V} are inside the \bar{V} . The verb’s GOAL argument gets its θ -role from the trace in sister of V^0 position, and the THEME θ -role is borne by the DP adjoined to \bar{V} . So the restriction in (29) would prevent reanalysis of \bar{V}^* .

Let’s summarize some of Larson’s claims, and the arguments we’ve seen for them.

- (36) a. DP+DP is transformationally derived from DP+PP.
 – based on UTAH
 b. V^0 assigns inherent Case, and a V^0 governed by I^0 assigns structural Case.
 – accounts for the way Passive applies in double object constructions
 c. VP Shell for double object construction.
 – scope facts

- Kayne’s nominalization argument
- gives a good account of NP Shift
- d. VP Shell for every VP with two things in it.
 - scope facts
 - new account of NP Shift
 - Gapping facts
 - idiomatic θ -role to direct object

Now let’s begin deciding which parts of this to accept.

There are reasons for doubting that the double object frame is transformationally derived from the DP+PP frame. UTAH doesn’t argue for it, I think, because there are subtle differences in the θ -roles these two structures license. Consider, by way of illustration, the examples below.

- (37)
- a. This book taught Mary French.
 - b. * This book taught French to Mary.
 - c. Your article showed Henry a problem.
 - d. * Your article showed a problem to Henry.
 - e. The manual told Susan everything.
 - f. * The manual told everything to Susan.
 - g. The TV gave Gary the bad news.
 - h. * The TV gave the bad news to Gary.
- (38)
- a. John bought a book for Bill.
 - b. John bought Bill a book.
 - c. John taught French to Mary.
 - d. John taught Mary French.
 - e. John threw a ball to first-base.
 - f. * John threw first-base a ball.
 - g. Mary sent a letter to New York.
 - h. * Mary set New York a letter.
 - i. Kill a commie for Christ.
 - j. * Kill Christ a commie.

There is a systematic difference in meaning between the DP+DP and DP+PP frames that is illustrated in (38). Green (1974), from whom many of these examples are taken, notes that the DP+DP frame systematically involves asserting a possession relation between the first and second of the two objects. Thus, in “John taught Mary French,” that Mary came to “have,” that is learn, French is entailed. But not so in “John taught French to Mary,” where nothing about the success of the lessons is entailed. And the oddness of “John threw first-base a ball” is related to the fact that first-base is not the kind of thing that can possess a ball. It looks from these examples, then, that there is a constant element to the meaning of the double object frame that is not always part of the other complementation structure. In so far as this difference between the two has

to do with argument structure, we should conclude that they do not share a D-structure — at least not if UTAH is correct.

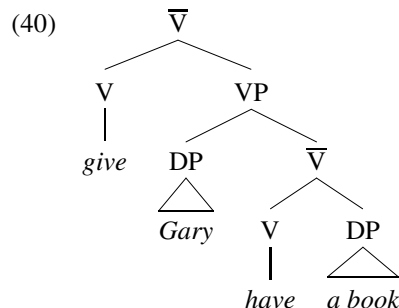
The examples in (37) lead to the same conclusion. It looks like the external θ -role for double object verbs is CAUSE, whereas that for verbs when they take the DP+PP frame is narrower, maybe AGENT. This looks like a straightforward difference in the θ -roles these structures license, and so again, UTAH would lead us away from giving them a common D-structure.

So let's abandon the view that these two are transformationally related. There is also some trouble for the "NP Shift" story. It seems unable to distinguish the following two sentences, as Jeff Runner points out.⁹

- (39) a. I gave a call yesterday to all my friends in Duluth.
 b. *I gave yesterday a book to all my friends in Duluth.
 c. *I gave yesterday Mary a book.

Larson's particular mechanism for generating NP Shift structures would allow a verb and adverb to move together past any number of complements. But, in fact, it looks like only one complement can be "NP Shifted," as the contrast above shows. This might be a problem for Larson's view about where adverbs are placed which we've seen is already problematic given the Projection Principle — or it might simply be that the particular mechanism that Larson proposes for "NP Shift" is faulty. Until we can see a successful treatment of "NP Shift" facts, let's hold off using it as a diagnostic for the structure of VP.

If we agree that we are not forced by UTAH to give the double object construction and the DP+PP frames the same underlying structure, we should ask how they are different. The conclusion that the double object construction has a hidden phrase in it, in which the two "objects" are contained, is reasonably strong. Kayne's argument from nominalizations suggests it, as do the asymmetric scope facts. Moreover, we have not encountered any evidence that suggests that it is wrong. Let's conclude, then, that the double object construction does have the parse that Kayne, Larson and others envision. But, departing from Larson, let's imagine that the hidden phrase is headed by a term whose meaning contributes the "possession" entailment that Green argues is a constant element to the meaning of the double object construction. So, on this view, the double object construction has a representation like that in (40), in which the *have* verb is unspoken.



⁹See Runner (1995).

Now what about the other two complement structures? Should we employ a VP Shell arrangement for these constructions as well? If the double object construction does not have the same D-structure as the DP+PP frames, we are not forced to give DP+PP frames, and other two complement structures, a parse like that in (40). “NP Shift” no longer supports this parse either. But the scope facts do: When a verb is followed by both an argument DP and an argument PP, the DP has the PP asymmetrically within its scope. This is just exactly the scope relation we find between the two objects of the double object construction, and which is now explained for these constructions by the parse in (40). So we might be led by the scope facts to give both the double object frame and the DP+PP frames the same structure.

But there are also reasons for thinking this might be the wrong way to capture the scope facts. One that we’ve seen already is Kayne’s nominalization argument, which, recall, argues that we want to give double objects a different D-structure parse than other double complement structures. Another is that ordering complements the way that VP Shells does wouldn’t capture an important fact about the linear order of terms. Remember that, as Stowell’s framework accurately reflected, the ordering of complements is determined on the basis of the categorial status of the complements; consider:

- (41) a. Mary explained [_{DP}the brats] [_{PP}to their parents].
 b. ?? Mary explained [_{PP}to their parents] [_{DP}the brats].
 c. ?? Mary explained [_{CP}that the brats needed help] [_{PP}to their parents].
 d. Mary explained [_{PP}to their parents] [_{CP}that the brats needed help].

Here the θ -roles borne by the two arguments following *explain* are the same; but whether the PP comes first or last depends on the category of the other θ -role bearer. It’s paradigms like this that suggest that the order of complements in English is determined by the category of the phrases involved, and not their θ -roles. The θ -role that *explain* assigns in these examples are the same, and yet when “direct object” is a DP it precedes the PP argument, and when it is a CP it follows that argument. In general, as we’ve seen, no matter what the verb is, or its θ -roles are, DPs tend to precede other complements. Larson’s method of ordering complements makes this information that each verb carries. But that wrongly implies that the ordering of complements can vary from verb to verb, and this isn’t what we see.

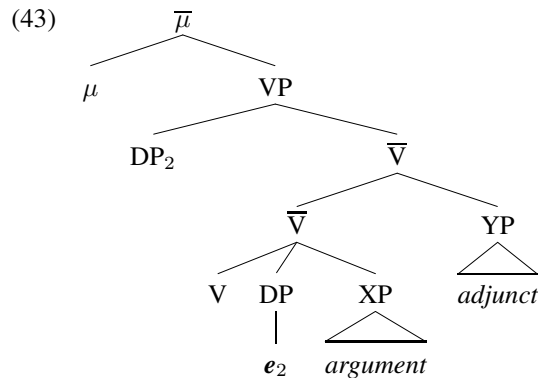
So we’re looking for a way of preserving the scope facts that meshes with both this aspect of the ordering facts and the nominalization facts. Finally, we need a way of preserving the difference in the kinds of Case that show up on the double objects. It is only the first object of a double object construction that loses its Case under the Passive, as (42) illustrates.

- (42) a. Mary was shown a book.
 b. * A book was shown Mary.

This suggests that it is only the first object which bears “structural” Case, which is the kind of Case that Passive robs. Let’s define structural Case as the Case assigned by the functional head above VP.¹⁰

¹⁰This is just a strengthening of Larson’s definition of structural Case. He proposed that a verb assigned

This is the proposal that I make in the “Object Positions” paper, and which Chomsky also makes in Chomsky (1991). Chomsky calls this functional head Object Agreement (AgrO), and I, following a suggestion of David Pesetsky’s, call it μ . If this head is where structural accusative Case comes from, then we can see the fact that DPs precede other complements (and non-complements) as a response to their need to move into proximity to μ . And, if this movement is upwards, this will also account for the fact that DPs appear to be higher than other VP internal material, i.e. it will give us a handle on the scope facts. So this vision of how complements are ordered might give a parse like that in (43) to cases where a verb is followed by two complements.



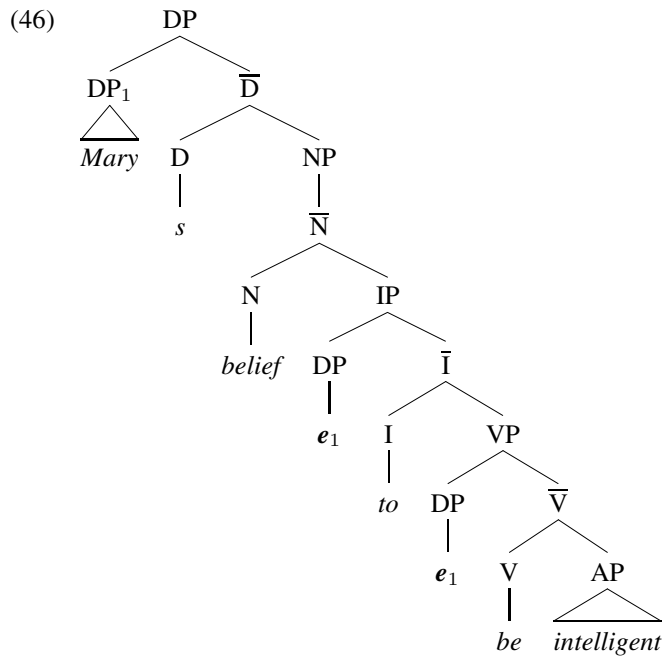
In (43), the complement DP has moved into a position that is governed, hence Case-marked, by μ . Suppose this position is Specifier of VP. From here, the DP will c-command the rest of the material within VP, both non-arguments and other arguments, and this accounts for how DP arguments are able to c-command non-arguments without abandoning the Projection Principle and the facts that the Projection Principle explains. In (43), there is a level of representation — D-structure — where arguments are closer to the verb than are non-arguments. This is in accordance with the Projection Principle. It's only at S-structure, after the DP argument has relocated into Specifier of VP, that the scope facts are reflected. In this way, the scope facts are captured without leading to the paradox that Larson's proposal creates.

Further, because only DPs are subject to the Case Filter,¹¹ we account for the fact that it's DPs, irrespective of the θ -roles involved, that precede other complements. Note also that once the DP has moved into Specifier of VP, there is a constituent that includes the verb and secondary complements but which excludes the DP complement. This constituent can be seen as the target for Gapping in cases like those we saw earlier.

There is perhaps another way in which this model of the scope and Gapping facts might be considered superior to Larson's. It provides a method for capturing the part of Kayne's nominalization argument that was left out. Recall, that Kayne's argument suffered from not giving (an entirely satisfactory) account for the absence of cases like the following (44).

structural Case only when it is governed by the functional head governing VP — this was I^0 for him.

¹¹Note, then, that we will have to abandon Stowell's proposal that CPs are also subject to the Case filter.

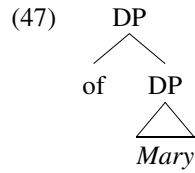


- (44)
- * John's belief of Mary to have left
 - * the editor's assumption of the article to contain several errors
 - * John's belief (of) Mary a genius
 - * their assumption of John dangerous
 - * her teacher's gift of Mary of the letter
 - * her rental of John of office space

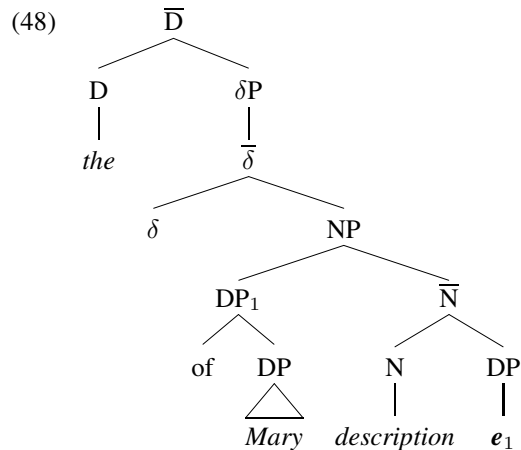
If we assume that *of* is a kind of structural Case assigned within nominals, then the method of assigning structural Case that Chomsky suggests, when extended to nominals, allows us to give the same explanation of the absence of these cases that Kayne gave for the passive nominals. Recall that Kayne's account of the absence for passive nominals of the sort that (45) illustrates is that it would involve movement of the argument out of a clausal complement to a noun. Clausal complements to nouns, however, are islands for extraction, and so this movement, illustrated in the tree in (46) below is blocked.

- (45) *Mary's belief to be intelligent

The ungrammaticality of (44) might be seen as having a similar source. Suppose, for concreteness, that *of* is a Case prefix, not a preposition. That is, imagine that the structure assigned to *of Mary* is something like:

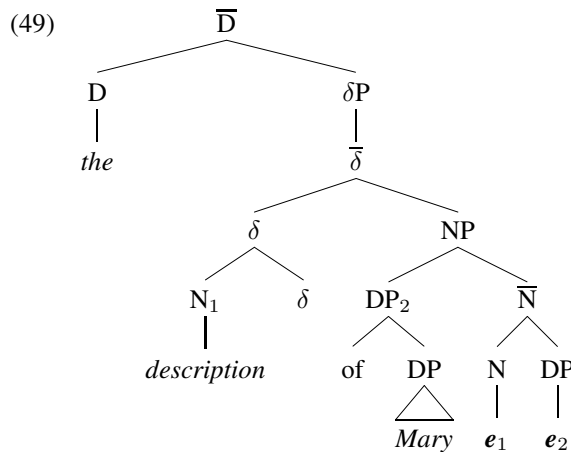


This is one way of making the assignment of *of* Case parallel to the assignment of accusative Case. We can think of *of* as being the structural Case that is assigned within nominals. Indeed, for those nouns which clearly have complements (typically nouns derived from verbs), the complement that would correspond to an accusative Case marked object in a VP appears with *of* marking in the NP. And, moreover, the canonical order of complements of a noun puts the *of*-marked complement before the others, just as, within the VP, it is the accusative Case marked complement that comes before the others. So, if we treat *of* marking as being fully parallel to accusative Case marking, and we adopt the view that accusative Case marking is done by an empty head, like μ , then we will want to do the same for the *of*-marker within NPs. Let's call the head that assigns *of*, δ . This would give a nominal like "the description of Mary" a representation something like (48) below. This isn't quite right, of course, since it



puts the *of* marked DP to the left of the noun that it is an argument of. In fact, that is a problem with what we have done with accusative Case marked complements too — the parse in (43) wrongly puts the object to the left of the verb. We're about to address this problem for the situation involving accusative Case, but for the situation in (48) let's adopt the solution that allows the noun to move to δ . This, as we shall see, is one solution to the parallel problem in (43). This would transform (48) into (49) on the next page. This gets the word-order right.

Finally we're prepared to consider what the source of the ungrammaticality of the examples in (44) might be. The relevant property of these examples is that the DP argument that is in need of the *of*-Case is embedded within a clausal complement to the noun. If the suggestions above are correct, this DP will need to move into proximity to



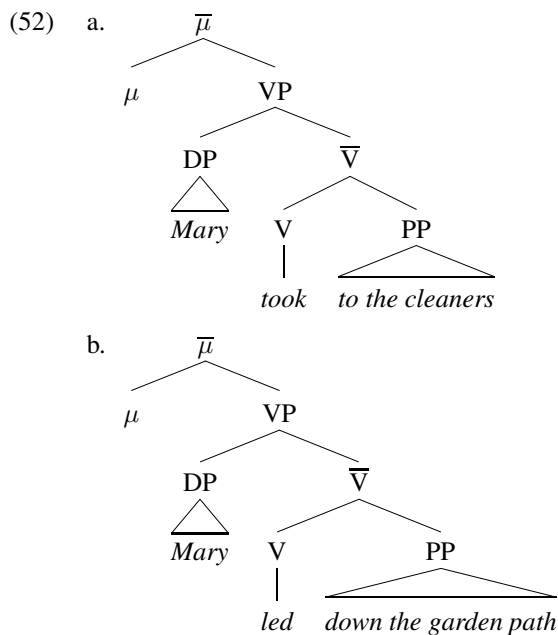
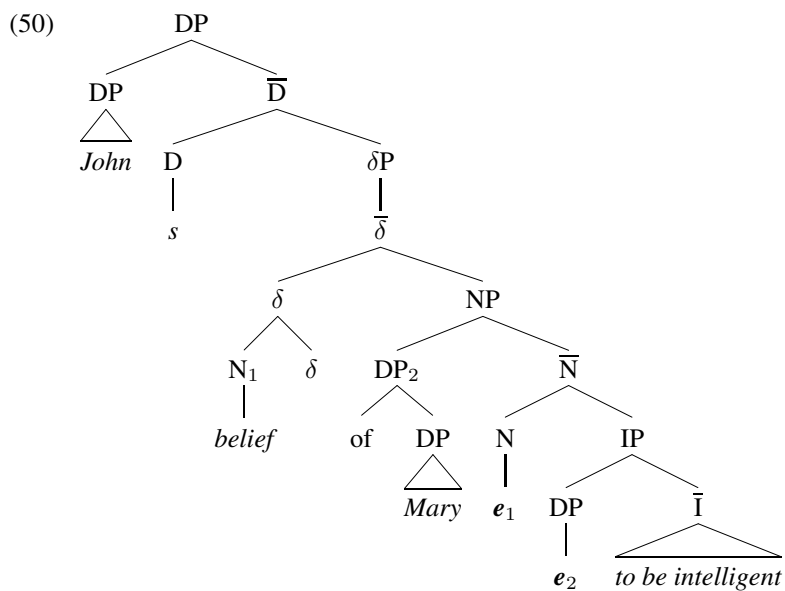
δ to receive Case. But this will require that this DP move out of the clausal complement, and, as we've seen, the clausal complements to nouns are islands for movement. Thus, these examples will be ungrammatical for precisely the same reason that (45) is. Just as in (46), the DP in (44) will have to illegally move out of the clause to get Case, as (50) on the following page indicates. The general idea here, then, is to preserve the parallelism between DPs and CPs/IPs that the analysis of "noun phrases" we get from Abney allows. Let the "object Case" that NPs and VPs support have a similar syntax, one that forces the term which receives that Case to move from its underlying position, and we will gain an account for the ungrammaticality of (44).

There are several reasons for adopting the thesis that accusative Case marked complements move leftwards to get Case in English, then. If this thesis is embedded in a structure like (43), then we preserve Larson's account for the scope facts, and for his Gapping data. But we are also able to preserve Stowell's idea that complements are ordered as a function of the Case filter, and we can preserve the idea that at D-structure non-arguments are higher than are arguments. And finally, as we've just seen, this hypothesis, if carried over to nominals, allows us to bring part of Kayne's paradigm into the account of the rest of his data.

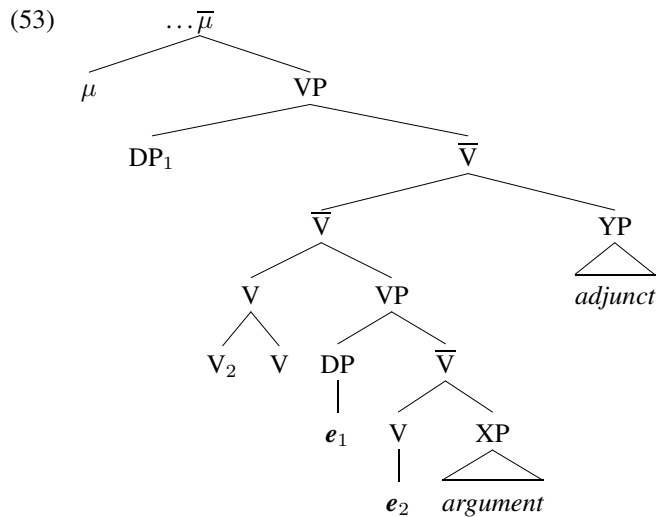
There is something of Larson's account that is lost by adopting (43), however. And that is his account for the idiom facts. Recall that he pointed to examples such as (51) and concluded that the first object must be able to get its interpretation — perhaps its θ -role — by combining with a phrase made up of the verb and the second object.

- (51) a. Jerry took Mary to the cleaners.
b. John led Mary down the garden path.

In examples such as these, the phrases *took to the cleaners* and *led down the garden path* get an idiomatic, fixed, interpretation. And, moreover, these expressions seem to take as one of their arguments *Mary*. If we hold constant our assumption that sisterhood is the structural relation which θ -roles are assigned, then this would seem to force us to see D-structure parses such as (52) in these examples.



Because we are taking θ -roles to be assigned at D-structure, we will want to let these examples have this sort of organization at D-structure. (43) doesn't do that. As a response to this problem, we could import some of Larson's structure into (43), maybe along the lines of (53).



This merely embeds Larsonian VP shells inside the functional projection that assigns Accusative Case, μ P.

There is some confirmation of this structure that comes from considering the scopal relations of complements when neither of them is a DP. In this circumstance, recall, the requirement that finite CPs come string finally will lead towards fixing the PP+CP order. And in contexts when both complements are PPs, we expect either order to be natural. So, we find paradigms like (54) and (55).

- (54) a. Jerry explained to Sally that syntax is gripping stuff.
 b. ?? Jerry explained that syntax is gripping stuff to Sally.
- (55) a. Jerry talked to Sally about Bill.
 b. Jerry talked about Bill to Sally.

There is some reason for thinking that the ordering effect in (54) doesn't have anything to do with the category of CP, contra Stowell, but rather with its length. This is because a similar effect emerges with long DPs, as in (56).

- (56) a. ?? She explained [the problem that we're having] [to Sam].
 b. She explained [to Sam] [the problem that we're having].

We might want to conclude from this that the orderings are due to length, maybe reflecting the behavior of the parser, which might face an obstacle when trying to parse a long string when it falls between two things that need to be packaged together (as do the verb and its complement PP in these configurations). Let's conclude, then, that as far as the syntax is concerned, either order of complements is allowed.

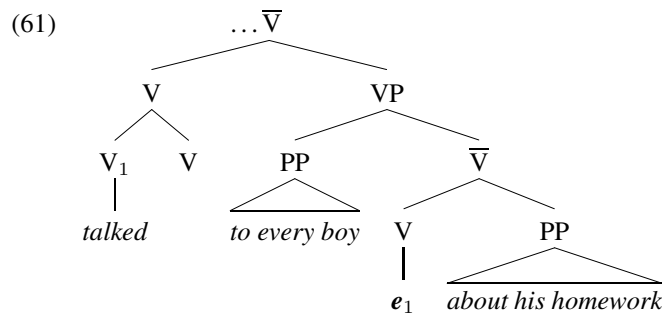
Now let us consider the scope facts which emerge with these terms. When a preposition phrase comes before a clause, it appears to have scope over that clause. The same thing appears to be true with respect to a PP that precedes another PP.

- (57) a. She explained to every boy₁ what he₁ would be fed.
 b. She carefully explained to the kids₁ when each other₁'s parents would pick them up.
 c. * She explained to the poor bastard₁ when Sam₁ would be paid.
- (58) a. She talked to every boy₁ about his₁ homework.
 b. She talked to the boys₁ about each other₁'s homework.
 c. * She talked to the poor bastard₁ about Sam₁'s homework.

But this is not true when the *to* phrase follows the other complement, as in the following cases.

- (59) a. ? She explained what he₁ would be fed to every boy₁.
 b. * She carefully explained when each other₁'s parents would arrive to the kids₁.
 c. She explained when Sam₁ would be paid to the poor bastard₁.
- (60) a. ? She talked about his₁ homework to every boy₁.
 b. * She talked about each other₁'s homework to the boys₁.
 c. She talked about Sam₁'s homework to the poor bastard₁.

In other words, in this context we are finding the scope asymmetry that we earlier discovered with regard to DP complements and other VP-internal terms. If we maintain the thesis that scope reflects c-command relations, then these facts call for the sort of parse that we find in (53), where the first phrase has the second in its scope. For example, in the PP+PP scenario, we might imagine a parse like (61).



Let's adopt, then, this aspect of Larson's proposal: verbs combine with their complements always in a way that leads to binary branching phrase markers. This will mean that we must revise the Projection Principle because, recall, this condition requires complements to be a sister to the head from which their θ -roles derive. In cases like these, where there are two complements, the Projection Principle would require that we have a ternary branching \bar{V} , containing the head and both complements. Let's change the Projection Principle so that it reads as in (62), then.

(62) THE PROJECTION PRINCIPLE

- a. Let α have c-selection requirements, and $\alpha^{n=1,2,\dots}$ be the set of \bar{X} projections of α . Then there must be an α^x such that every phrase β within α^x , $\beta \notin \alpha^n$, is c-selected by α .
- b. If α θ -marks β , then α and β are sisters.

The change I've made is in the first clause, which now no longer requires that all complements (i.e., things which are c-selected) be a sister to the head doing the c-selecting. Instead, the condition requires that the c-selected items be placed within the projection of the c-selecting head in such a way that they all appear closer to the head than do non-c-selected items. This preserves what the original version did: it places complements lower in the phrase than non-complements.

So, overall, the system we've adopted here combines complements with a verb in either order to form a binary branching projection of the verb. Modifiers, and other non-arguments may be adjoined to the verbal projection in higher positions. The DP complement to the verb, should there be one, will move leftwards into a Specifier that brings it close to the term responsible for assigning accusative Case, μ , in our phrase markers. This last step is responsible for bringing DP objects to an initial position, and for giving them exceptionally wide scope. It also, as we've reviewed, provides the necessary structure for Gapping, and, when transferred into a nominal context, explains why transparent infinitives are blocked as complements to nouns.

Although there are some apparent advantages to the thesis that accusative Case marked complements move to the left for their Case, there are quite a number of problems with it too. It wrongly places the object before the verb, as we've noted. And it also raises problems when we consider how the Internal Subjects Hypothesis is embedded within these structures. Recall that we have seen evidence which suggest that external θ -role bearers move from a VP internal position. And we have supposed that this puts subjects in Specifier of VP at D-structure. This is obviously at odds with placing the object in this position too. We'll turn to resolving this conflict in the next chapter. For the remainder of this chapter, then, let's suspend our earlier conclusion that the external θ -role is assigned to Specifier of VP (or other XP, as appropriate).

Another conflict with our earlier conclusions that crediting μ , rather than a verb, with the ability to assign accusative Case concerns characterizing the contrast between unaccusative and unergative verbs. If Case assignment is determined by something other than the verb, then how are we to account for the verb's apparent influence on the availability of accusative Case, which is one of the ways unaccusative verbs are distinguished from unergative verbs. This problem too we will take up in the chapter that follows.

For the remainder of this chapter, let's focus on the problem of fixing the order of the verb relative to the object DP. Chomsky suggests that we should see the movement as covert, by which he specifically suggests LF movement. But this would appear to lose the advantage this system has in describing the surface word-order of complements. That is, it doesn't directly address how to guarantee that DP complements precede the others.

An alternative, one that I suggest in the "Object Positions" paper is that, contrary to the Emonds/Pollock claim, main verbs in English do move. Recall Pollock's

claim: the contrast between the examples below is due to the difference in movability of main verbs in French and English.

- (63) a. * John kisses often Mary.
b. Jean embresse souvent Marie.

Note, however, that English patterns with French when the verb's complement is not a DP:

- (64) a. John talks often to Mary.
b. John described slowly how Mary had solved the problem.

This doesn't look like it could be derived through NP Shift, however this process is to be described, because it leaves the categorial contrast below puzzling.

- (65) a. * John kisses often her.
b. John talks often to her.

So this evidence leads to the conclusion that the contrast between French and English is not purely a difference in verb movement, as Pollock suggested.

Instead it is contingent on the categorial status of the complements involved. Indeed, it is contingent on the same DP/non-DP distinction that the ordering of complements appears to be sensitive to. This correlation can be captured on a "movement" theory for the DP-first ordering of complements, like the one we are exploring here. Still, we are in search of an account for the contrast in (63). We no longer can rely on Pollock's straightforward account in terms of verb position. Let's put off this question for a while.

Another consideration which favors the view that main verbs move in English comes from a consideration of the particle verb. These are constructions where the normal adjacency between verb and object seems to favor viewing the verb as having two parts, as illustrated in (66).¹²

- (66) Mary [_V handed out] the toys.

This is also suggested by morphological processes which typically apply to verbs; these seem to operate on the two-part verb+particle, suggesting that these amalgams are verbs.

- (67) a. The toys were handed out by Mary.
b. the handed out toys
c. These toys hand out easily.
d. the handing out of the toys
e. this handout

But, interestingly, it looks like there is room between the verb and particle, as shown by cases like (68).

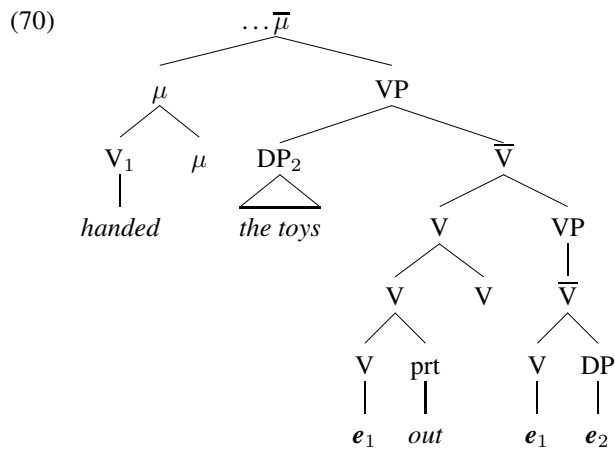
- (68) Mary handed the toys out.

¹²These data are parallel, then, to the separable prefix verbs we examined in homework.

We can resolve the separate syntactic positions of verb and particle with their apparent wordhood, if we let words (at least some minority of them) be amenable to being separated through syntactic means. So let particle verbs have the structure [_V V prt], and the hypothesis that main verbs can move in English will allow either the higher or the lower verb to move. This intersects with the “movement” account of the DP-first word order to account for why only DPs are able to intervene between verb and particle.

- (69) a. Sally tried out dancing the jitterbug.
 b. * Sally tried dancing the jitterbug out.
 c. Sally stepped out onto the platform.
 d. * Sally stepped onto the platform out.

This is expected if movement to the Case marked Specifier is allowed only of DPs. So, on this view, the parse of (68) is as in (70).

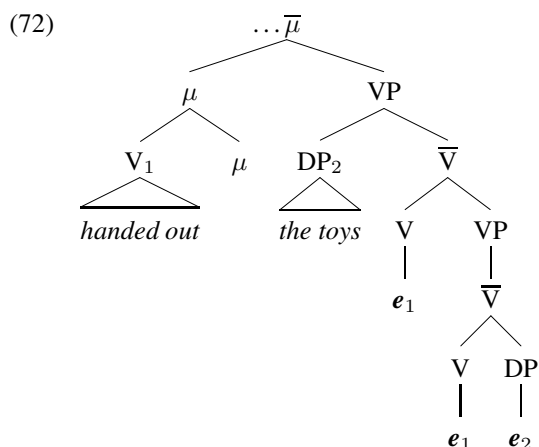


Because only DPs move to the Specifier of the higher VP, only DPs will be able to precede the particle, thereby deriving the contrasts in (69).

It's also possible for the particle to precede an object DP, however, as in (71).

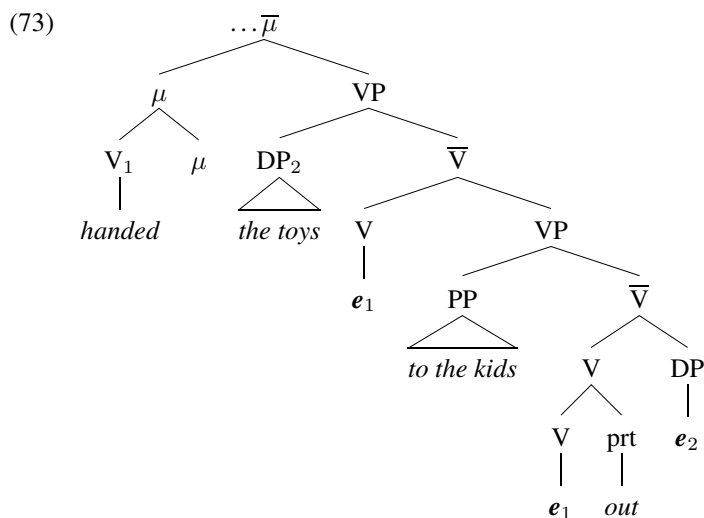
- (71) Mary handed out the toys.

This word-order can be captured if we allow Verb Movement the option of moving the verb and particle together, as in (72).



Particle constructions, then, also suggest that main verbs in English do in fact move past the position that object DPs are forced to surface in. So let us embrace this as our method of solving the problem that moving object DPs into Specifier of VP gives rise to — in particular, let us suppose that this is why verbs precede objects even though objects have relocated to this higher position.

It should be noted, however, that this analysis of particle verbs does not straightforwardly combine with Larson's method of positioning complements — at least not our version of his method. If complements combine with their verb in the binary branching way that (61) illustrates, and the order in which these complements combine is left open, then the analysis of particle constructions here should allow for parses such as (73).⁷



What's wrong with this parse, of course, is that something other than a DP appears to the left of the particle. In general, letting particles stay in the position that verbs begin in, and letting complements freely come on either side of this position has the

consequence of wrongly allowing complements other than DPs to precede the particle. Yet another problem that we'll have to come back to.

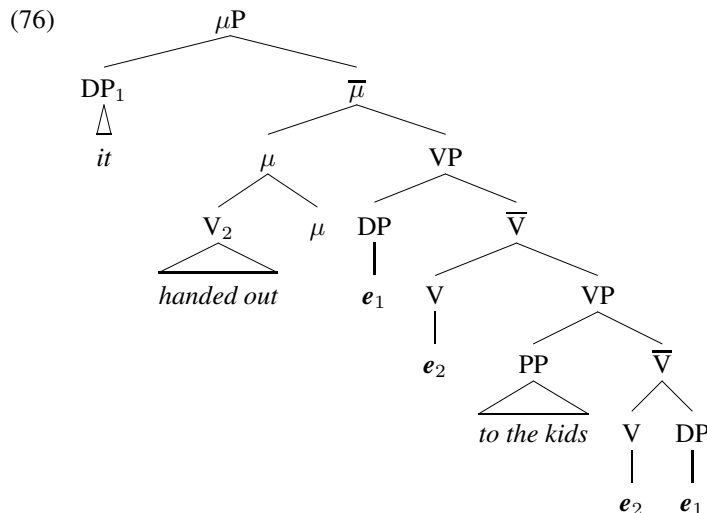
I took the view that verbs in English move farther than just μ in the "Object Positions" paper for the following reason. When the particle verb's complement is an unstressed pronoun, it must come between the verb and the particle.

- (74) a. * Sally handed out it.
b. Sally handed it out.

This can be related to an apparent cross-Germanic generalization to the effect that if a pronoun can move leftwards out of VP, it will. Thus, in German, for instance, we find contrasts like the following.¹³

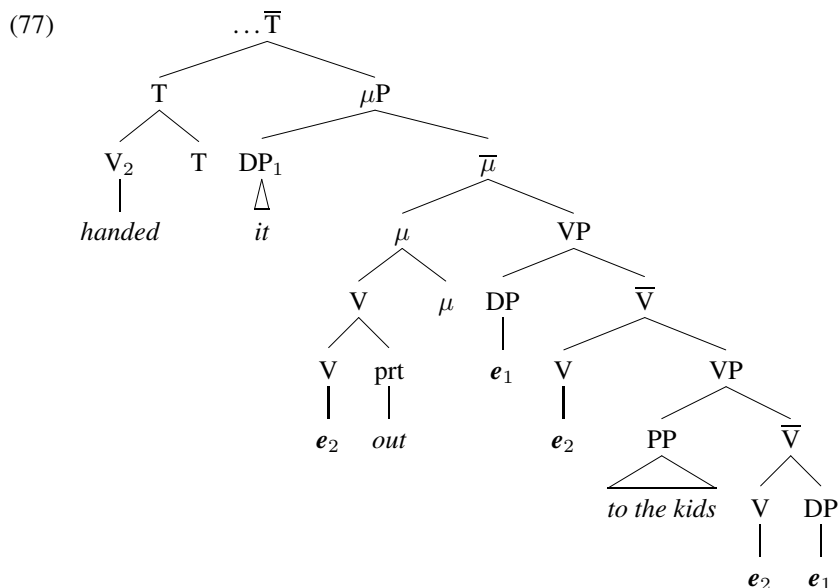
- (75) a. * ... daß Hans ja doch es gelesen hat.
...that Hans indeed it read has.
'...that Hans has indeed read it.'
b. ... daß Hans ja doch ein Buch gelesen hat.
...that Hans indeed a book read has.
'...that Hans has indeed read a book.'
c. ... daß Hans es ja doch gelesen hat.
...that Hans it indeed read has.
'...that Hans has indeed read it.'

There is evidence that *ja doch* is an adverb which sits to the left of the VP. So the contrasts in (75) suggest that pronouns, unlike "normal" objects, are forced to move out of VP in German. If we suppose that weak pronouns in English are likewise forced out of VP by S-structure, let's say into Specifier of μ P, then we can explain why they must precede the particle. But this will require that verbs move beyond μ P, because they always come to the left of objects, even pronominal ones. So, (74b), for instance, will receive a representation like that in (76).



¹³See Diesing (1992) for a discussion of these examples.

If this is correct, then, of course, the verb must move yet again to an even higher position; and, moreover, this movement must not be able to drag the particle along with it. I suggested in “Object Positions” that the position the verb moved to is the one that determines the inflection on the verb. In a simple clause, where there are no auxiliary verbs, this position will be what we’ve called T^0 . Thus, (76) will produce a surface representation like that in (77) below. I suggested that the fact that this functional head



is related to the inflectional morphology that the verb bears is responsible for the fact that the particle cannot be lifted into this position. The reason this might have an effect on the movement of the particle involves a peculiarity of particle verbs when it comes to attaching morphology to them.

When particle verbs combine with derivational or inflectional morphology, this morphology cannot reside outside the entire V+particle complex. Instead, it must reside on the interior verbal part; this is what the alternations in (78) indicate.

- (78)
- a. the handed out leaflet
 - b. * the hand outed leaflet
 - c. the looking up of the information
 - d. * the look uping of the information

Let’s describe this peculiarity as follows: the verb+particle verb is opaque to morphology. Morphological process don’t see this constituent. Now, let the feature checking algorithm be part of the battery of things that are “morphological processes,” and this will make the verb+particle constituent opaque to feature checking as well. Thus, if the strong morphological feature on T^0 is to be checked off by S-structure, a verb

with matching morphology will have to be adjoined to T^0 and not a verb+particle with matching morphology.¹⁴

¹⁴In fact, the idea expressed in Johnson (1991) presupposes a model in which the morphemes themselves reside in T^0 , Agr^0 and the like. I've recast this idea into present terms.

Chapter 8

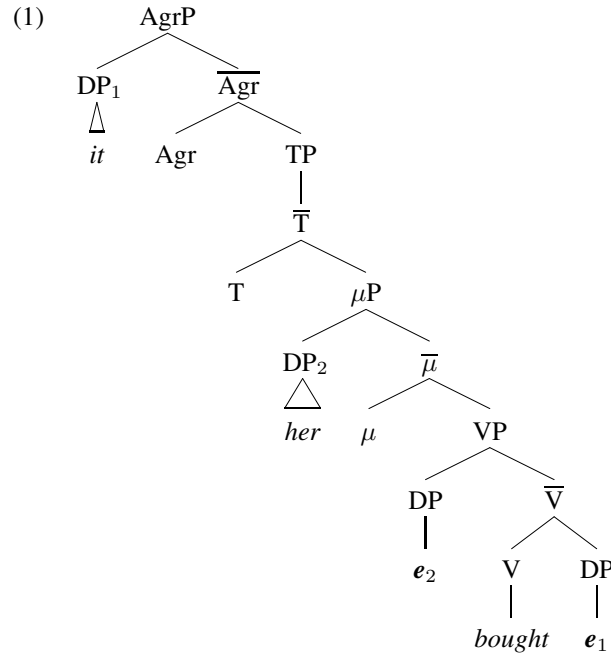
Subjects and Complex Predicates

The picture that emerges from the previous chapter is that complements are arranged in VPs with considerable more structure than originally thought. Moreover, at least on one version of how they are arranged within this structure, both verbs and “object” DPs overtly move. It's the movement of the object DP that is responsible for bringing this object into initial position, for giving it scope wider than the rest of the material in the VP, and for allowing it to precede the particle.

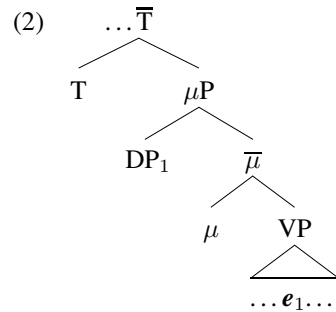
But, as touched on, this account runs into several incompatibilities with conclusions we have reached earlier. If the mechanism that is responsible for making “object” DPs move is that accusative Case is assigned by μ , and not verbs, then verbs are not distinguishable by virtue of their ability to assign accusative Case. But this is something that we have earlier suggested is possible. Unaccusative verbs differ from transitive verbs by virtue of having an accusative Case to assign. And, if objects are driving into Specifier of VP to be Case marked by μ , then where are subjects placed — up to now we have taken the view that external θ -roles are assigned to subjects in this position. Could Specifier of VP be occupied by both subject and object?

Let's focus on this second difficulty for a moment. There are a variety of related puzzles that come with trying to embed the view that subject arguments originate in Specifier of VP with the view that accusative Case is assigned by a functional head outside VP. We might wonder, for instance, why subjects aren't capable of remaining in Specifier of VP, bearing the accusative Case. Indeed, if it's not just subject DPs that are moving into their Case marked positions, but object DPs as well, then what prevents these DPs from surfacing with the wrong Cases? Why can't a sentence with the meaning “She bought it” surface with the form in (1) on the following page. If this general approach to complement structure is to be maintained along with the “Derived Subjects” hypothesis, then it's clear that we will need greater controls on how object and subject DPs move to their surface Case marked positions.

We can take an initial step towards addressing this problem by simplifying the scenarios in which objects move to just one. Assume that the object moves into



Specifier of μP always, and not into the Specifier of VP that μ governs. Assume that object DPs always are positioned in S-structures in the way that (2) indicates.



We might ensure this outcome by letting accusative Case be assigned by μ not under government, but to Specifier position. That is, we can reduce the various combinations of object movements that we explored in the previous chapter to just the one shown in (2) with the conjecture in (3).

- (3) Structural Accusative Case is assigned by μ to its Specifier position.

This simplified picture has some virtues. It straightforwardly answers the question why subjects don't get Accusative Case by virtue of their underlying position. Even if subjects start out beneath μP , under (3) they will have to move to get to Specifier of μP if they are to surface with Accusative Case. Thus, by changing the conditions under

which structural Accusative Case is assigned, we are put in the position of being able to prevent subjects from getting Accusative solely by virtue of conditions on movement, conditions which, hopefully, will steer subjects and objects into the correct Case marked positions.

Further, (3) opens the way for a unification of the conditions under which structural Case is assigned. Recall that, presently, our description of structural Case assignment (repeated in (4)) has an ugly bifurcation in it.

(4) CONDITIONS ON STRUCTURAL CASE ASSIGNMENT

- a. Specifier of finite AgrP is assigned Nominative Case.
- b. X^0 assigns its Case to the position α only if X^0 governs α .

If the two structural Cases are Nominative and Accusative, then Accusative is assigned under government.¹ One of the places, of course, which (4b) is designed for is structural Case assignment by a verb, and so if (3) can be maintained, this scenario will no longer fall under (3). The only other situation that (4b) is designed for is Case assignment by the complementizer *for* in certain infinitival clauses, like that in (5).

- (5) I prefer [_{CP} for [_{AgrP} Mary to be on time]].

However, some have suggested that this case has a different analysis — we'll examine one of these later in this chapter — and this clears the way towards a more uniform treatment of structural Case assignment; something along the lines of (6), for instance.

(6) Structural Case is assigned by α to its Specifier position

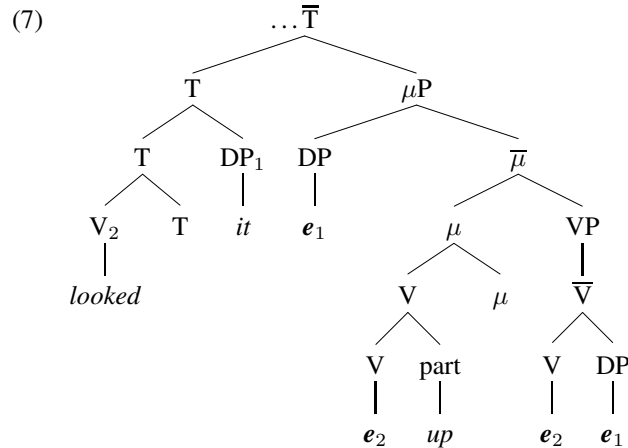
- a. finite Agr⁰ assigns Nominative.
- b. μ^0 assigns Accusative.

What's lost with by adopting (3) is being able to give object DPs two positions to surface in. This, recall, was instrumental in explaining why simple pronominal objects must show up to the left of the particle, whereas full DP complements can surface to the right of the particle as well. The guiding idea in that account was: (1) there are two positions that object DPs can surface in and (2) there is a pan-Germanic tendency for pronominal objects to surface in the leftmost of the positions available to objects. The two positions that the first of these conjectures employed were Specifier of VP and Specifier of μ P. If we are to adopt (3), we will have to find another way of expressing this account.

What is needed is a position that pronouns are driven to that brings them always to the left of a particle, while still leaving another position to the right of the particle available for non-pronoun objects. If the verb moves beyond μ P, and non-pronouns are in Specifier of μ P, this means that pronouns must have a position that is higher than Specifier of μ P. One way of capturing this effect would be to treat pronouns in English as sorts of clitics. Like the weak pronouns we have seen in Romance, they might be thought of as needing to prosodically become part of a verb. Imagine in particular, that they adjoin to the right edge of the closest verbal stem that c-commands

¹Recall that government is the relation defined to express when a Case assigner and Case bearer are sufficiently *klose* to each other. α GOVERNS β iff α c-commands β and there is no more than one phrase that contains β but not α .

them. If we let particles move with the verb out of μP , say as far as T, this would give us representations like:



Note that in (7) the particle has been stranded in μ . We might imagine that this is forced by the action of cliticization — perhaps, in a way similar to the suggestion made at the end of the last chapter — there is something special about particle verbs that makes them resist anything morphologically attaching to the outside of the verb+particle complex. Or, alternatively, we might imagine that the clitic must find itself attached to the verbal root — maybe for the same reason — and this will cause it to migrate within the verb+particle complex even should the particle be able to move with the verb into T^0 .

This is, in outline, the solution to the problem in Diesing and Jelinek (1995). It preserves the structure of the account developed in the previous chapter: there are two positions for accusative DPs, the leftmost one reserved for pronouns, and pronouns are driven by the pan-Germanic inclination for leftwardness into this position.

As our first step in solving the problem of distributing nominative and accusative Cases to the subject and object correctly, then, we've revised our model in a way that is consonant with (3), giving us just the scenario in (2) to consider. Now the problem reduces to finding a way of guaranteeing that the subject argument moves into Specifier of AgrP and the object DP moves into Specifier of μP , rather than the other way round.

There is a proposal that a variety of people have entertained about how subjects are introduced into structures that might be of assistance with this problem. The idea, in a nutshell, is that subjects are not arguments of the same predicates that objects are arguments of. Instead, the predicates that subjects and objects are arguments of are connected in such a way that they come together, giving the appearance, in normal circumstances, of belonging to the same predicate. This conclusion has been reached by different authors in different ways. I will focus here on two reasons that are salient in Kratzer (1996).

The first is simply that it is difficult to steer the rules of interpretation to derive the fact that the narrow content of the θ -role that a subject gets is a function of the appropriate \bar{V} . This, recall, was the reason that Marantz gave for letting the θ -role which subjects get be produced jointly by the verb and its complements. In our present

schema, this is described by letting one θ -role be designated to be assigned externally. But how would such a designation derive this consequence? It is not enough to just say that such a θ -role shall be assigned externally; we should find a way of representing this θ -role so that this consequence is derived. And, furthermore, why should this θ -role be special in this particular way? What is it about this θ -role that makes it different from the others? And why does it get assigned to the Specifier of VP — that is in the position that is sister to the highest \bar{V} — and nowhere else?

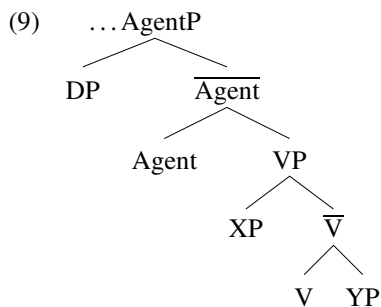
The second is the observation that subjects are not selected in the same way that objects are. Recall that the categorial kinds of objects are determined by the verb that selects them. Verbs “c-select” their complements. Subjects, by contrast, have their categorial nature determined completely by their θ -role. And why should these two particular properties hold of subjects: that they are θ -marked in a higher position than other arguments, and they are not c-selected?

If we concentrate on the second of these observations — that it is the θ -role assigned to a subject that determines its categorial nature — then we might express this in terms of selection. Perhaps the θ -role itself is c-selecting the subject argument. Recall that our revised Projection Principle, repeated in (8), requires in this case that the θ -role itself project a phrase in which the subject is found, and in which only other c-selected items are found.

(8) THE PROJECTION PRINCIPLE

- a. Let α have c-selection requirements, and α' be some projection of α . Then there must be an α' such that all the phrases within α' are c-selected by α .
- b. If α θ -marks β , then α and β are sisters.

One way of doing this, that will also place the subject argument higher than the objects, is with (9).



(9) represents a situation in which there is a subject bearing the Agent θ -role, and there are two internal arguments as well. This might be the underlying arrangement of arguments in “Jill pushed the book onto the table,” for instance.

This is essentially the proposal in Kratzer (1996), though she calls the phrase that embeds VP, and which Agent might be thought of as a special instance: Voice Phrase. She suggests that both of these phrases be thought of as predicates of events, and that they are semantically composed in such a way that they predicate of the same

event. For example, “Jill pushed the book onto the table” would be interpreted along the lines of (10), where “e” is a variable over events.

- (10) $\exists e [\text{Agent}(\text{Jill})(e) \ \& \ \text{Push-the book-onto the table}(e)]$

She suggests that the connection these two phrases have to each other — that *push* takes an Agent, rather than, say, an Experiencer θ -role — be expressed with the kinds of events available. Suppose, for instance, that events come in different kinds: actions, states, and so on, and that predicates select for them. Because the method of composing these two predicates will require that they take the same event argument, there will be constraints on which of these predicates can combine. If the typology of events makes the same number of distinctions that external θ -roles turn out to, then this method at least provides a way of ensuring that external θ -role and main predicate are matched up correctly. This account, then, makes the selection of external θ -roles part of the event structure of sentences. And this in turn, Kratzer suggests, should be part of an over arching theory of Voice, which will control when these external θ -role predicates are found.

There is another potential advantage to this approach to external θ -roles. It gives us a means of capturing the part of Burzio’s Generalization that cannot be derived from syntactic constraints. This is the statement that establishes a connection between external θ -roles and (structural) Accusative Case assignment:

- (11) BURZIO’S GENERALIZATION

If a verb assigns (structural) Accusative Case, then it assigns an external θ -role.

Recall that what this statement records is an unexpected gap in the logically possible outcomes, should the Case marking and θ -marking properties of verbs be independent, as the theory up to now has held. What (11) suggests, by contrast, is that Accusative Case assignment and external θ -role assignment are not independent. We have since adopted an account of Accusative Case assignment that credits it not to the verb, but to μ . If we also move to the view that external θ -roles are not assigned by verbs, but are instead assigned by Voice, then we can let the contingency expressed by (11) come about by way of selection. Let one of μ and Voice c-select the other, and (11) will emerge. This is a proposal of Chomsky (1995), who refers to the phrase that Kratzer calls Voice Phrase as a “light VP.”²

Finally, notice that this move allows a simplification of the Projection Principle. The only reason for stating separately that θ -roles are assigned under sisterhood is to capture the case of external θ -roles, which, up to now, were the only θ -roles assigned without reference to c-selection. But under the present hypothesis about how external θ -roles are assigned, they are linked to c-selected arguments like all other θ -roles; (8b) becomes superfluous. Instead, (8b) might fall into a larger generalization, one that does not single out θ -roles. This generalization, of which the modification rule could be a particular example, is that sisterhood steers the way in which syntactic representations are semantically interpreted. This generalization can be expressed as (12).

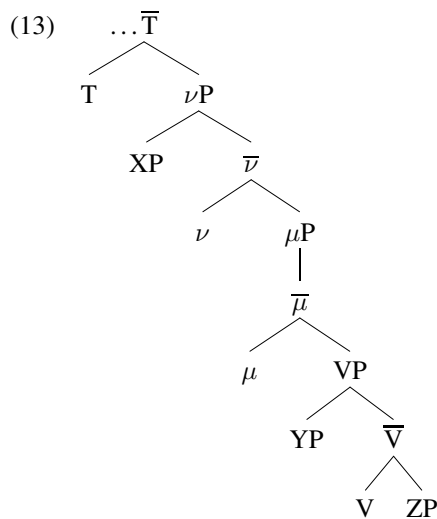
²Chomsky’s proposal, and his choice of terminology, is inspired by Hale and Keyser (1993), which sketches out a way of expressing the transitive/intransitive distinction in these terms. Ken Hale and Jay Keyser have since developed this work into a general theory of argument structure; see Hale and Keyser (2002). An early version of this approach is Bowers (1993), who argues for a “Predicate Phrase,” which might be seen as equivalent to μ P. Bowers’ work, in turn, builds on Williams (1980).

- (12) The denotation of α is always the result of combining the denotations of β and γ , where β and γ are α 's (immediate) daughters.

It's (12), and not the narrower (8b), perhaps, that derives Larson's observation that the first object's role is determined by the verb and second object together. In other words, Larsonian shells, when coupled with (12), will still derive the fact that the role *Mary* has in *took Mary to the cleaners* comes from syntactically combining with a phrase made up of *took to the cleaners*.

Whether (8b) should be abandoned for the more general (12) depends, of course, on whether (12) is true. This can only be determined by an exhaustive examination of the semantics of clauses and, naturally, this hasn't been completed. Nonetheless, a reasonably large fragment of English can be shown to conform to (12),³ and so there is some hope that it is true. Let's adopt it from here on out.

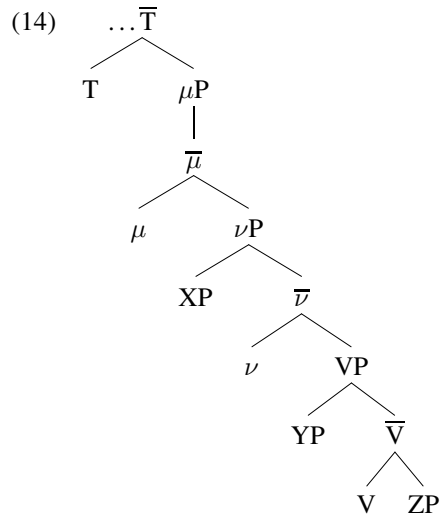
The reason this approach to external θ -roles might be helpful for our problem of delivering objects and subjects to the correct Case marked positions is that it allows the external θ -marked position to be removed from the main VP. It's possible, therefore, to place the underlying position of subjects higher than the position in which Accusative Case is assigned. The prohibition on moving downwards will then guarantee that subjects don't surface in the Accusative Case. If we use " ν " — which equivocates between Voice and "light verb" — to represent the phrase in which external θ -roles are determined, this thesis would produce the representation in (13).



On this syntax, then, Burzio's Generalization is expressed as a consequence of ν selecting μ P. Everything we have seen about the derived subjects hypothesis is consistent with (13). In particular, it correctly allows quantifiers to be floated beneath auxiliary verbs, but not below main verbs or their objects.

Interestingly, there is some evidence that this straightforward solution isn't correct. This evidence suggests that ν P is instead embedded within μ P, as in (14).

³See Heim and Kratzer (1998).



The evidence comes from German data, and is based on an interesting property of some adverbs, of which *wieder* ('again') is an example.

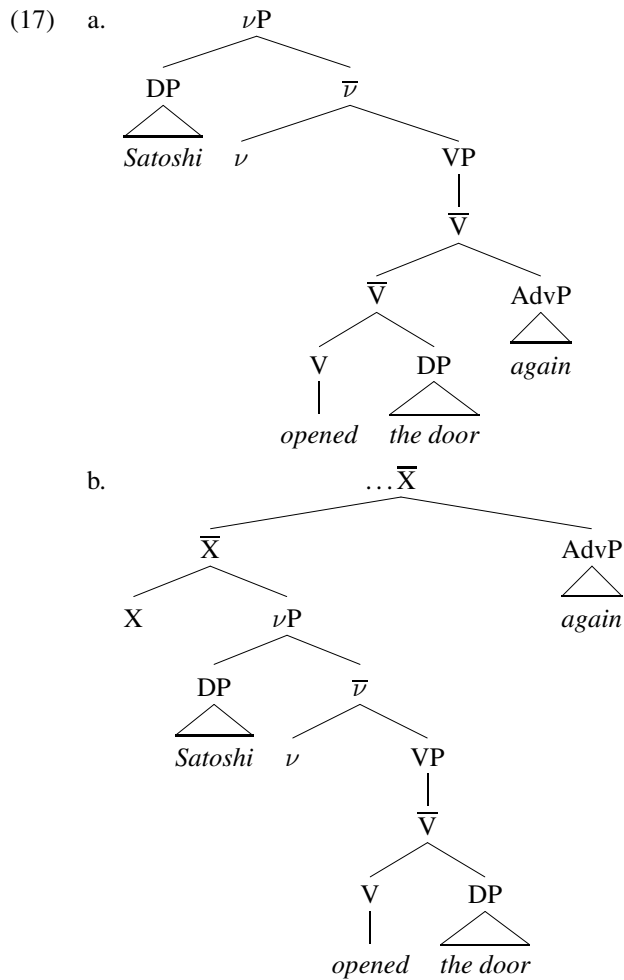
Sentences in both German and English which have this adverb in them are ambiguous in a way that suggests that *wieder/again* is capable of modifying two different things. Consider, by way of illustration, (15).

- (15) Satoshi opened the door again.

This sentence reports that Satoshi opened the door, and adds, by way of *again*, that something has been repeated by doing so. On one reading, what has been repeated is just Satoshi's opening of the door. On this reading, then, Satoshi opened the door on some previous occasion, and (15) reports that he has done so a second time. On another reading, what is repeated is the state of the door being open. On this reading, Satoshi may be opening the door for the first time, but by doing so he is restoring it to a state it was previously in, namely open. Let's call the first interpretation the "repetitive" reading of *again*, and the second the "restitutive" reading. So (15) has the two readings in (16).

- (16) a. Satoshi opened the door, and he had done that before.
(repetitive)
- b. Satoshi opened the door, and the door had been opened before.
(restitutive)

von Stechow (1996) argues that this ambiguity reflects a structural ambiguity. In the repetitive reading, *again* modifies the constituent we have identified with νP , as in (17a). In the restitutive reading, by contrast, *again* modifies the lower predicate, as in (17b). (These parses, both on the next page, don't reflect the movements necessary to achieve the surface representations.) Note that this account requires the meanings of νP and VP to be slightly different than those we described in connection with Kratzer's account. It's necessary here for the VP *open the door* to refer to a state — the state the door is in when it is open. The rest of the description of the event this sentence refers to



must come from the meaning of νP . Roughly, ν in this sentence must be something like “makes,” so that the parts come together to mean something like “Satoshi makes the door be in the open state.” There is some work to be done, then, in bringing Kratzer’s suggestions about νP together with the treatment of *again* that Stechow advocates.⁴ If νP and VP in (17) have the meanings just described, however, then (17a) has a meaning paraphrased by “Satoshi again made the door be in an open state,” and (17b) has a meaning paraphrased by “Satoshi made the door again be in an open state.”

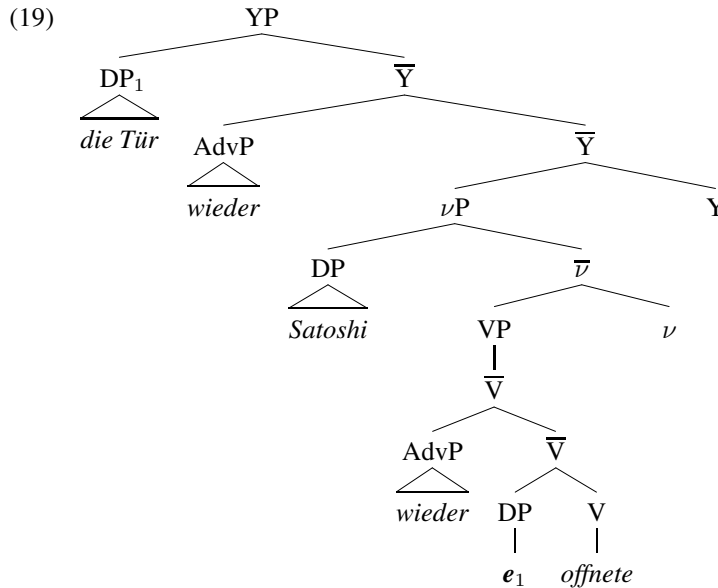
Stechow’s argument on behalf of this structural account is based on the fact that word order plays a role in disambiguating *again* in German. The German version

⁴von Stechow (1996) has some discussion of this, one that enriches the syntactic representation. For another approach, one that credits the additional “make” component of the meaning to a semantic rule of composition, see Beck and Snyder (2001).

of (15) is (18).

- (18) ... weil Satoshi die Tür wieder öffnete.
 ...since Satoshi the door again opened.
 ‘...since Satoshi opened the door again’

Like (15), (18) has both the repetitive and restitutive readings. On the structural account of this ambiguity that (17) sketches, this means that both νP and VP must be embedded beneath the position that the Accusative Case marked object, *die Tür*, surfaces in. If we let Specifier of XP be the position the object is surfacing in in this example, we could represent this ambiguity by letting *wieder* have either of the two positions indicated in (19) below. Interestingly, if *wieder* shows up to the left of the object, as in (20), only

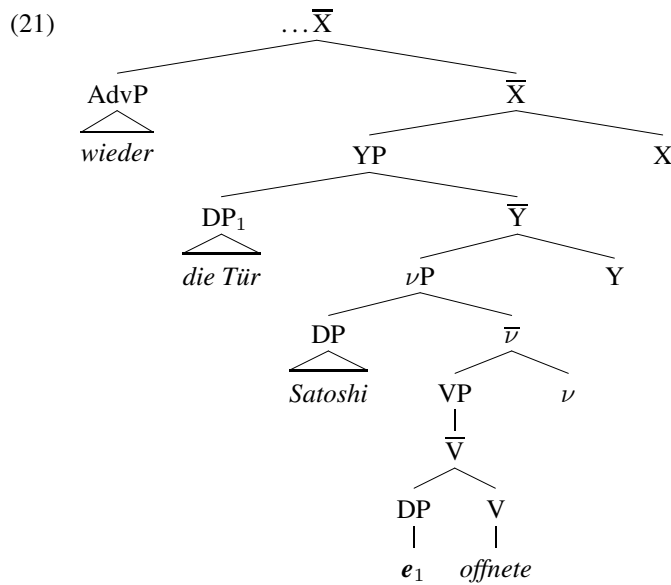


the repetitive reading is found.

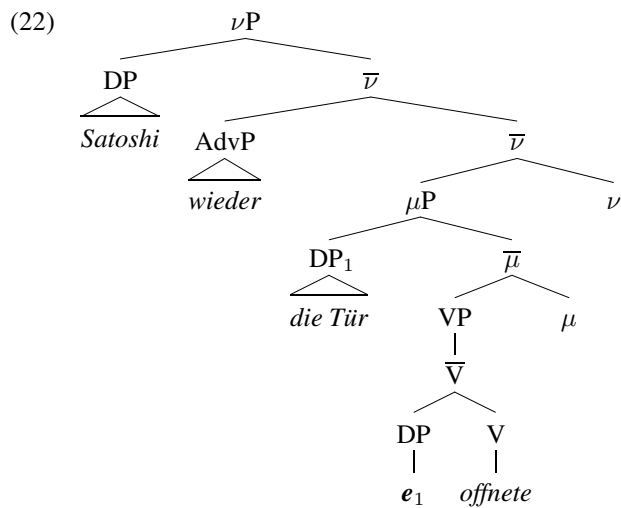
- (20) ... weil Satoshi wieder die Tür öffnete.
 ...since Satoshi again the door opened
 ‘...since Satoshi again opened the door.’

This follows from the structural account of this ambiguity, since, in this case, *wieder* will be forced into a higher position, as indicated in (21) on the next page. As a consequence, *wieder* will have to modify a constituent containing the “Satoshi makes” part of the meaning, and this is just the repetitive reading.

What is YP in (19) and (21)? It must be the phrase in whose Specifier accusative Case marked objects are forced to surface in German. If German is like English with regard to accusative Case marking, then this is μP . Here, then, is the reason for believing that the phrase that introduces the subject argument, i.e., νP , is



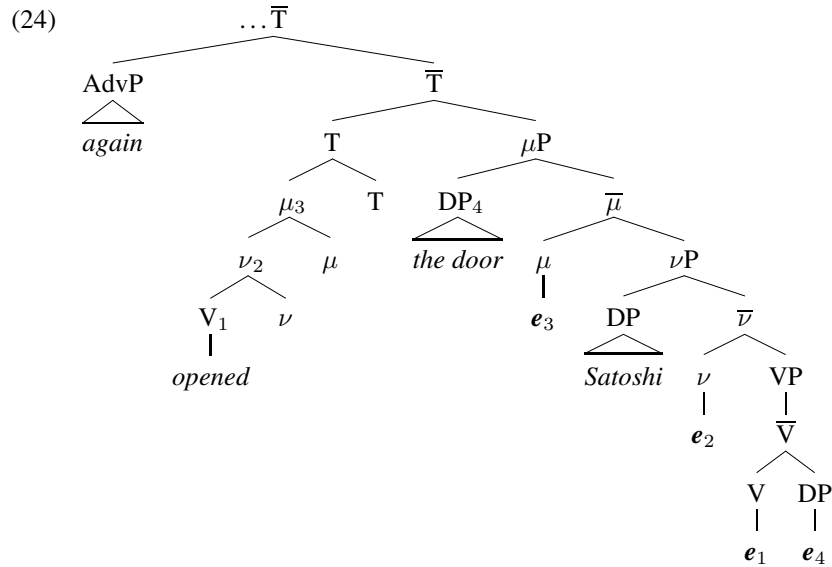
embedded within the position in which accusative Case is licensed, i.e., μ P. If ν P could be higher than μ P, then we should be able to give (20) the representation in (22), and this fails to capture the fact that *wieder* must modify a constituent that has *Satoshi* in it.



A parallel argument can be made for English. If *again* precedes the verb in English, as in (23), then only the repetitive reading is possible.

- (23) Satoshi again opened the door.
 = Satoshi again made the door be in an opened state.
 ≠ Satoshi made the door again be in an opened state.

Just as in the German (20), this word order is only possible if *again* is adjoined to some \bar{X} that contains the surface positions of the verb and object, as in (24) below. As a

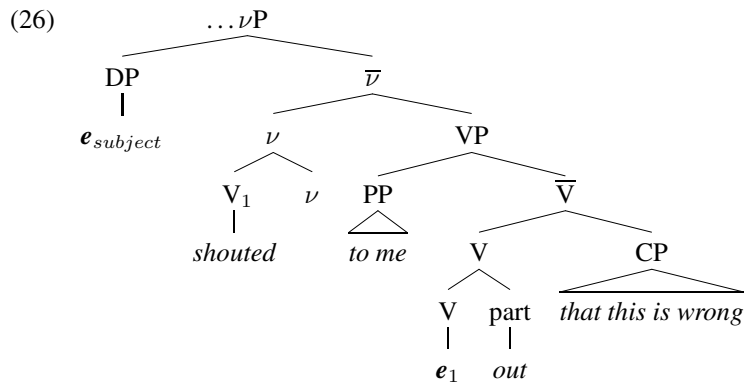


consequence, *again* will be forced to modify a constituent that contains the “make” part of the meaning, i.e. νP , but only if νP is necessarily embedded within μP .

For this reason, then, let’s abandon the solution to the problem of delivering the subject and object arguments into the proper Case marked positions that is based on putting νP higher than μP . Instead, we should place νP within μP , and find another solution to the Case problem. Indeed, we could equate the highest VP of Larson’s shells with νP .

But before turning to the task of finding a solution to the Case problem, let me wield what we have discovered about *again* to work on another of our open problems. There is, recall, a difficulty in putting the approach to particle verbs that Johnson advocates together with Larsonian shells. If verb movement is capable of stranding a particle, then we should expect it to be able to leave a particle in a position to the right of a non-DP argument, and this is not the correct outcome. That is, it should be able to produce the ungrammatical (25) with the representation in (26).

- (25) *Sally shouted to me out that this is wrong.
(compare: “Sally shouted out to me that this is wrong.”)

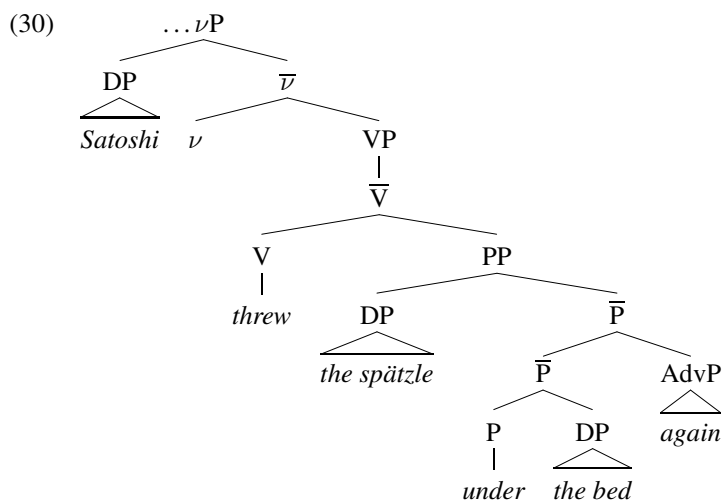


There is something, I suggested, that must be responsible for preventing particles from being stranded in their underlying position. But what could be responsible for this?

Consider now what we can learn from *again* about the structure of VPs with PPs in them. Focus on the restitutive readings in (27) through (29) (these are paraphrased for each case in the b-examples). In each case, the state that *again* says has

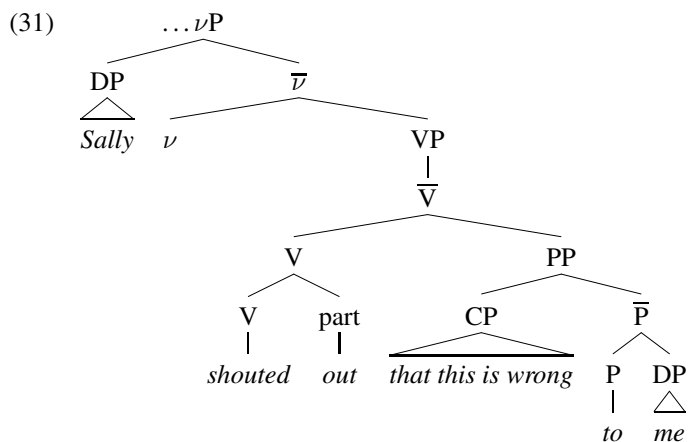
- (27) a. Satoshi threw the spätzle under the bed again.
 b. Satoshi threw the spätzle under the bed, and the spätzle had been under the bed before.
- (28) a. Satoshi threw the ball onto the field again.
 b. Satoshi threw the ball onto the field, and the ball had been on the field before.
- (29) a. Thilo threw the ball behind the fence again.
 b. Thilo threw the ball behind the fence, and the ball had been behind the fence before.

been restored would appear to be fully described by way of the meaning of the prepositional phrase. This suggests a structure like (30), in which the PP, “the spätzle under the bed” denotes something like “the spätzle is in the state of being under the bed.”



If this is the correct structure for Prepositional Phrases in all cases, then the problem posed by word order in (25) is solved. Indeed, Pesetsky (1995) has argued extensively, and on independent grounds, that (30) is quite representative of how PPs are positioned in English.

This solves the problem for particles because it has the consequence that PPs will always follow the underlying position of the verbs that select them. For example, the D-structure representation for (25) will be (31) below. Clearly, even if the particle is



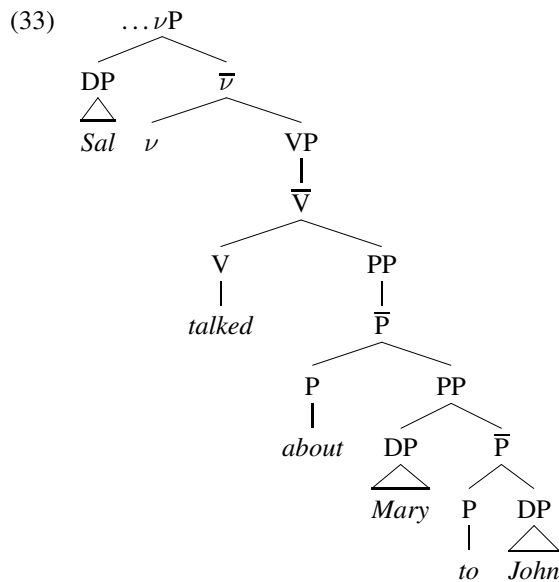
stranded in its underlying position in (31) it will remain to the left of the PP. Of course, (31) wrongly puts the CP to the left of *to me*. But we can rely on the ordering principle of English complements to ensure that this CP surfaces at the right edge. Perhaps this is achieved by moving the CP with NP Shift to the right edge, as on the Ross model that we began with. Or perhaps the Specifier of PP is linearized so that it follows its \bar{P}

when its contents are a CP.⁵

Perhaps the most difficult case to treat in this way are those verbs that take two PPs as complements, such as (32).

- (32) Sal talked about Mary to John.

Pesetsky argues for the representation in (33) below for such cases. One of the difficul-

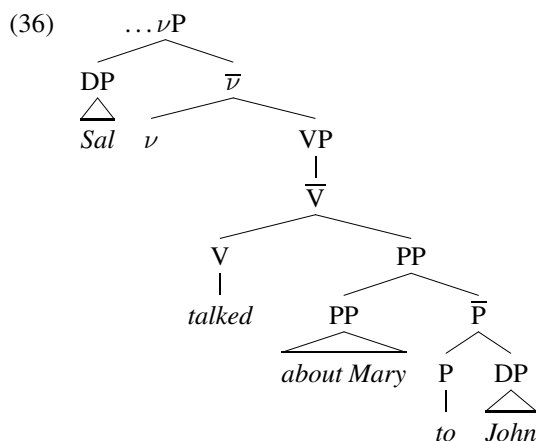


ties for this structure is that it will make it problematic to let the two PPs move about independently, something they are apparently capable of doing, as the topicalization sentences in (34) and indicate, as do the word orders in (35).

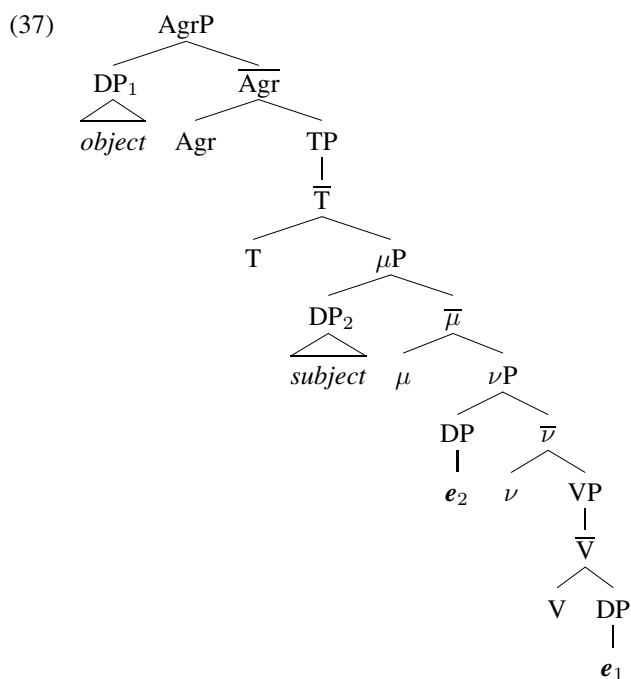
- (34) a. About John, Sal talked to Mary.
b. To Mary, Sal talked about John.
- (35) a. Sal talked to Mary yesterday about John.
b. Sal talked about John yesterday to Mary.

Pesetsky has a sophisticated proposal for these problems, one that would take us too far afield to explore. We might instead entertain the possibility that these cases get a representation like that in (36).

⁵What will be lost on these views is that the indirect object can have the clause that follows it in its scope. Pesetsky offers a solution to this problem that credits this case with a structure like the one for two PPs in (33).



With these preliminaries behind us, let's now face squarely the problem of putting the subject and object arguments into the proper Case marked positions. Our problem is how to prevent subject and object DPs from ending up in the surface positions indicated in (37) below. This remains something of an open problem. We will



eventually examine two proposals as to its solution. I will present one now that comes from Chomsky (1992).

This solution is built on the proposition that A movement is subject to a constraint we have not yet encountered. This constraint prevents some term from A moving

past a c-commanding subject. Chomsky (1977a,b) are the first systematic examinations of such a constraint, and there it is named the “Specified Subject Constraint.” We may adopt the fairly descriptive formulation of it in (38).

(38) SPECIFIED SUBJECT CONSTRAINT

If α A moves to β , then there must be no γ in a Specifier position that c-commands α but not β .

If such a constraint does exist, we are not going to be able to see it on the basis of English data, at least not readily. This is because the most easily recognizable Specifiers, past which things might A move, are subjects; and Burzio’s Generalization combines with the constraint that prohibits moving from a Case marked position to independently rule out most of these scenarios. Consider, by way of illustration, a situation like that diagrammed in (39).

(39) $[_{XP} \beta X^0 [_{YP} \gamma Y^0 \alpha]]$

If γ is a subject, then it must be Case marked. This means that either YP is a CP, or γ is getting Case from outside YP. If YP is a CP, then movement from α to β will violate the constraint against moving out of CPs. If γ is getting Case from outside YP, then it is either moving into the nominative or accusative Case marked position in XP. But if there is an Accusative Case marked position in XP, then Burzio’s Generalization will require that XP also have its own subject argument, and this will compete with α for the Nominative Case marked position. But if γ itself moves to the Nominative Case marked position, then it will also compete with α . In brief, the Case conditions on A movement will derive the consequences of (38), at least over a broad range of situations.

But there are versions of A movement, found outside English, which are not subject to the Case conditions on A movement. These movement operations do not move a term from a non-Case marked position into a Case-marked position. Instead, the trigger for these movement operations is not the Case filter, but some other surface requirement. The process of Cliticization that we have had occasion to touch on is one such operation. This movement operation, first systematically studied in Kayne (1975), moves prosodically weak pronouns into a designated position. In Romance, this designated position brings them into proximity to a verb, with which they become a prosodic unit. Cliticization is what is responsible for the different placement of objects in (40), for example.

- (40) a. Jean est fidèle à ses parents.
 John is faithful to his parents
 ‘John is faithful to his parents.’
 b. Jean leur est fidèle.
 Jean to-them is faithful
 ‘John is faithful to them.’

Cliticization has many of the same attributes that we have seen for Argument Movement. It obeys the Upwards Constraint and it is subject to something very like the same bounding constraints we’ve seen on A movement. In fact, from a descriptive point of

view, clitics cannot move out of CPs, just as arguments are unable to. The one difference of interest to us is that clitics are capable of moving from Case marked position into non-Case marked positions. This makes this kind of A movement unsusceptible to the constraints which confound the Specified Subject Constraint.

And in fact, as the contrast in the following French examples illustrates, Cliticization seems to be subject to something like the Specified Subject Constraint.

- (41) a. Marie *le* croyait heureux.
 Mary him believed happy
 'Mary believed him happy.'
- b. *Marie *lui* croyait le cadeau envoyé depuis longtemps.
 Mary to-him believed the gift sent a long time ago
 'Mary believed the gift (to be) sent to him a long time ago.'
 compare:
- c. Marie croyait le cadeau envoyé a Jean depuis longtemps
 Mary believed the gift sent to John a long time ago

In (41a), the accusative clitic, *le*, has moved out of the embedded adjectival small clause and attached to a position (presumably near Agr) into which the finite verb is driven. But a similar movement is blocked in (41b), where the dative clitic *lui* has tried to do precisely the same thing, but in so doing has moved past a c-commanding subject.

If Argument movement is subject to the Specified Subject Constraint, then this will prevent the derivation that leads to the ungrammatical outcome in (37). In particular, the movement of the object into Specifier of AgrP will pass the c-commanding subject in (37), in violation of the Specified Subject Constraint. Unfortunately, this constraint also similarly blocks the grammatical outcome, shown in (42). Movement of the object into Specifier of μ P will have passed the c-commanding subject, and, potentially movement of the subject into Specifier of AgrP will have similarly passed the c-commanding object in Specifier of μ P.

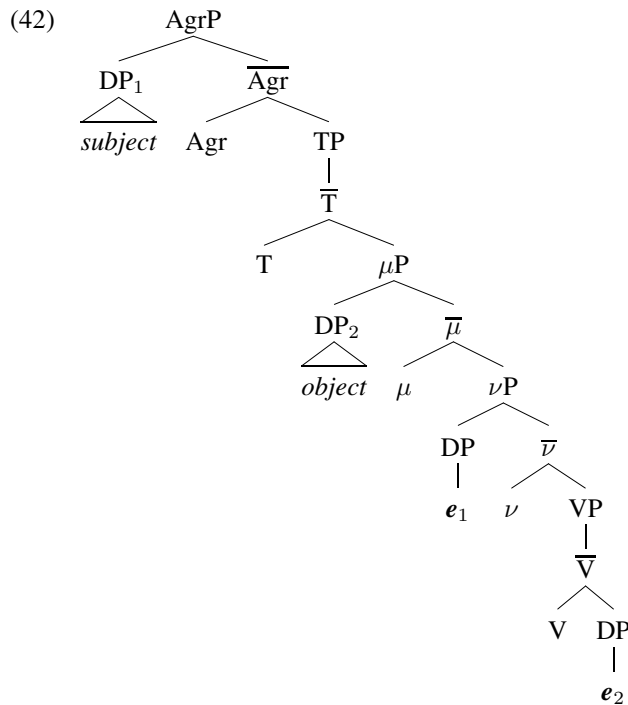
Chomsky's strategy is to find a loosening of the Specified Subject Constraint that lets (42) in, but continues to block (37). He is guided by other apparent violations of the Specified Subject Constraint found in causative constructions.⁶ Let's briefly consider these constructions and the way in which they are similar to the process we are confronted with in (37) and (42).

Consider, first, the way in which causatives are formed in Italian (which is very much like French and Spanish in the respects which concern us). We find in these contexts that the subject of the causativized verb must appear post-verbally.⁷

- (43) a. Faró scrivere Giovanni.
 I will make write Johnny
 'I will make Johnny write.'
- b. Faró lavorare alcuni prigionieri.
 I will make work a few prisoners
 'I will make a few prisoners work.'

⁶Chomsky (1992) cites the Government Transparency Corollary in Baker (1988) as his guide, one special instance of which is found in the causative construction.

⁷My description of Italian causatives, and all the examples, comes from Burzio (1986).



This is striking because subjects are otherwise quite capable of surfacing preverbally in Italian.

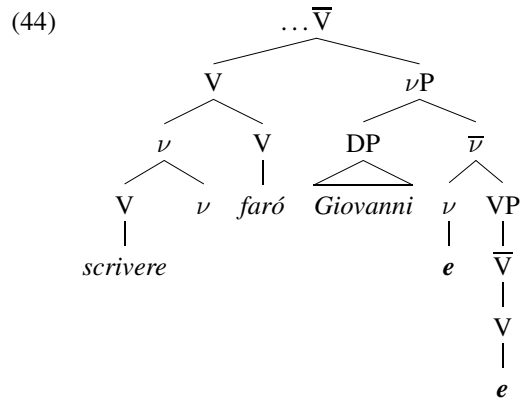
A widely-held account for this word order is that the lower verb in these contexts has moved and adjoined to the higher causativizing verb, as in (44) on the following page.⁸ Indeed, not only may the subject of the lower clause not intervene between the causative and lower verb, but very little else can.

There is evidence that suggests that, just as this treatment entails, the subject of the causative is still in “subject” position. One is the behavior of clitic placement. As we have seen, clitics cannot move past subjects. The ungrammaticality of the following example will follow as a consequence of this condition, if *Giovanni* is in “subject” position.

- (45) a. Faccio telefonare Giovanni a Maria.
I will make telephone Johnny to Mary
'I will make Johnny telephone Mary.'
- b. * Gli faccio telefonare Giovanni.
I to-him will make telephone Johnny
'I will make Johnny telephone him.'

Similarly, we find that anaphors which are stranded beyond the subject cannot be bound

⁸For an early proposal along these lines, see Aissen (1974).



by terms in the higher clause. This is illustrated by the contrast in the following French examples (46).

- (46) a. Elle aurait fait tirer les pauvres soldats₁ l'un sur l'autre₁.
 she would have made shoot the poor soldiers the one on the other
 'She would have made the soldiers shoot each other.'
- b. *Elles₁ auraient fait tirer le pauvre soldat l'une sur l'autre₁.
 they would have made shoot the poor soldier the one on the other
 'They would have made the soldier shoot each other.'

The reciprocal in these examples can be bound to the subject of the lower, causativized, verb. But it cannot accept as antecedent the subject of the higher clause. This follows if the subject argument of the embedded VP is in fact in a c-commanding "subject" position because, in general, anaphors cannot be separated from the things they corefer with by a c-commanding subject. (We will explore this description of the relationship between anaphors and their antecedents later in the course.)

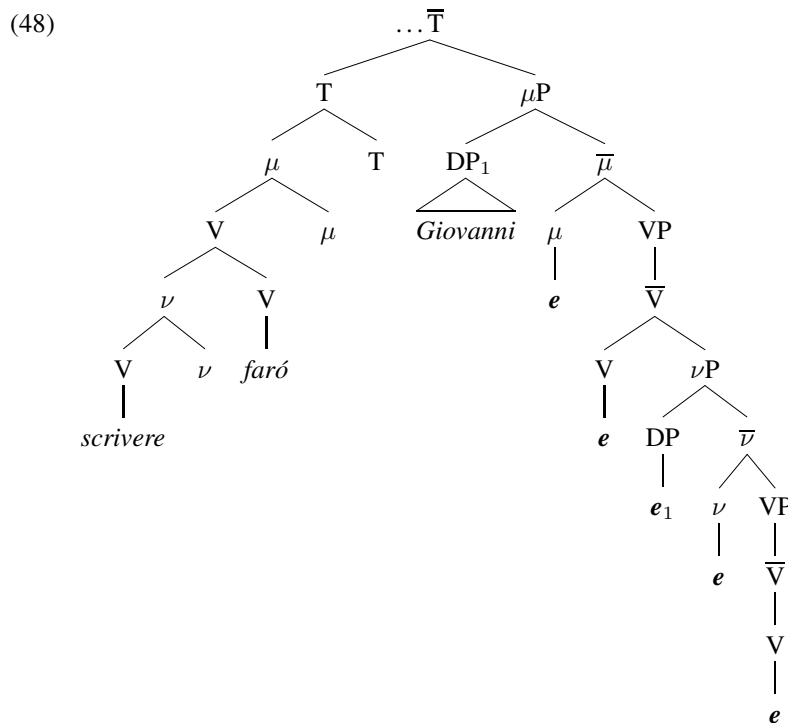
So, on the basis of this evidence, let us adopt a picture of causatives which involves verb movement, as in (44). We might imagine, for instance, that the causative verb in Romance is a bound morpheme with "strong" features. This will force overt verb movement, under the feature alchemy picture of Head Movement that we have adopted to account for verb placement.

Now, an interesting fact about these causatives is that the subject of the lower VP appears to be dependent on structural Case from something in the higher clause. This can be seen from the fact that when the causative verb bears Passive morphology, the accusative Case borne by the subject of the lower VP is lost:

- (47) Alcuni prigionieri furono fatti lavorare.
 a few prisoners were made work
 'A few prisoners were made to work.'

On our original view of how accusative Case is assigned, this would mean that the subject of the lower VP gets Case from the causative verb itself, because it is this Case

that has been robbed by the Passive. Under the present hypothesis, which credits μ with accusative Case, this will mean that the subject of the lower VP has moved into the Specifier of μ P in the higher clause to get Case. So, the tree in (44) should look instead something like (48) below. In this representation, the verb *scrivere* has adjoined



to *farO*, and then they have together adjoined to μ and then to T. Then the subject of the embedded VP, *Giovanni*, has moved out of Specifier of VP into its Case marked position in Specifier of μ P.

We have so far just considered what happens when a verb that has a subject argument and an optional indirect (i.e., PP) complement is embedded under a causative verb. Let's consider now what happens when a VP that has both a subject and a DP complement is causativized. In these situations, there will be two DPs in the lower clause, both of which will have to satisfy the Case Filter. What happens in these examples, interestingly enough, is that the subject argument shows up with a dative Case morpheme attached to it, and the object of the lower verb appears to get accusative Case from the higher clause. This, as a consequence, brings the object DP into a position preceding the subject DP, as in (49).

- (49) Maria ha fatto riparare la macchina a Giovanni
 Mary has made repair the car to John
 'Mary has made John repair the car.'

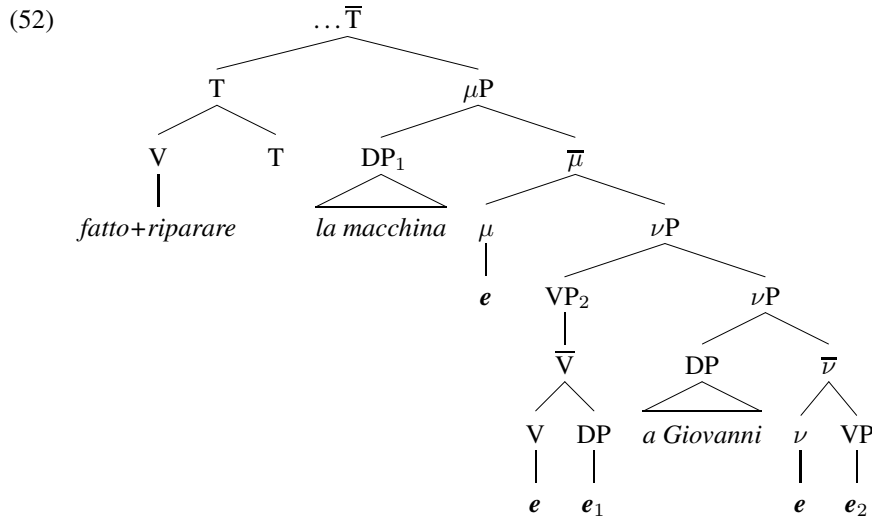
There are good reasons for thinking that the external θ -role bearer in this example, *a Giovanni*, is in fact in a “subject” position. It blocks Cliticization of phrases that follow it, as (50) indicates, and this is just what the Specified Subject Constraint predicts if it is in subject position.

- (50) * Gli faccio scrivere una lettera a Maria.
 I to him will make write a letter to Mary
 ‘I will make Mary write a letter to him.’

The reason for believing that the object, *la macchina*, is dependent on Case from something in the higher clause is that it’s the one that loses accusative Case marking when the causative verb passivizes, as (51) indicates.

- (51) La macchina fu fatta riparare a Giovanni
 the car was made repair to John
 ‘The car was made to be repaired by John.’

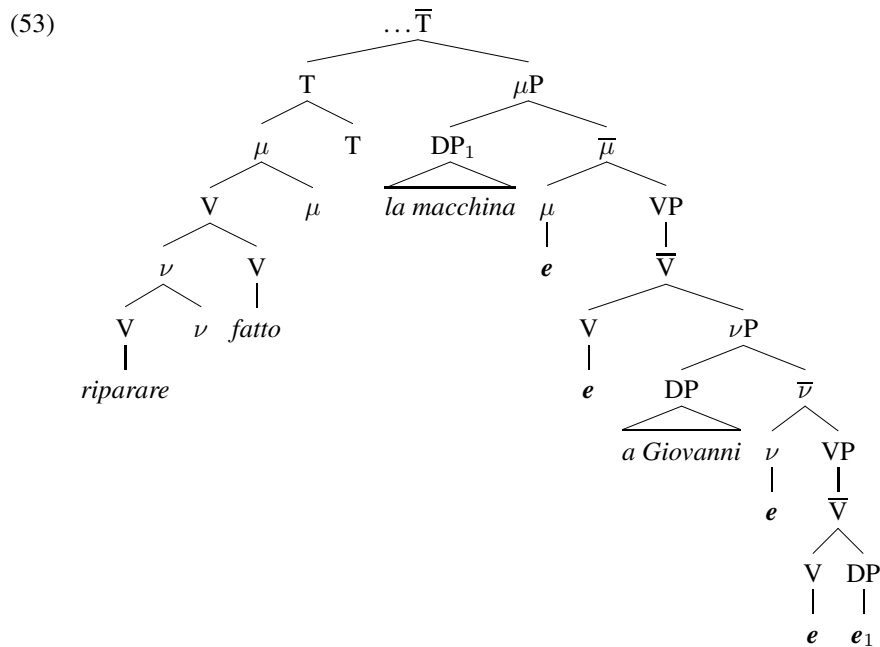
How can the object move past the subject in these cases? This is the interesting part of this construction for us, because it would seem to be a violation of the Specified Subject Constraint. The “classical” solution — one that goes back to Kayne’s work on French — is that a verbal projection has fronted in causatives, and this verbal projection has brought both the verb and its object past the subject. If this account were embedded within the current framework of Case assignment, it would give to examples like (49) a representation something like (52).



This would be a way of allowing the object to move past the subject without violating the Specified Subject Constraint, since the object is not, itself, moving past the subject, and presumably the movement that we see here is not a kind of Argument Movement and therefore not subject to the Specified Subject Constraint. The problem with this account, however, is that it doesn’t really explain why it’s only the “direct object” — that is the accusative Case marked complement — which seems enabled to pass the

subject. Dative arguments to the verb, and other material that we might expect to be part of the VP, are unable to move past the subject. This is indicated by the fact, for example, that the dative clitic in (50) cannot move into the higher clause.

So perhaps instead we should parse these examples as in (53) below, and seek a way of relaxing the Specified Subject Constraint so that accusative Case-marked complements are allowed to move past the subject, but other complements are not.



Causatives are very similar to the cases we are concerned about. In our cases too, we find that an object DP is moving past a “subject” in order to get into Specifier of μP . Moreover, in both of these situations, the object is not moving past the verb whose subject they have skipped; this is what distinguishes these contexts from the ones where movement past a subject is not permitted, as in (41b).

This is where Chomsky’s proposal comes into play. His idea relies first on the assumption that Head Movement is a kind of adjunction, just as we have been assuming all along. This derives:

(54) X^0 Chains have at most two members.

To see this, consider the derivation sketched in (55)-(57), in which a head, X^0 , moves first to Y^0 and then onto Z^0 .

- (55)
- ```

 ZP
 |
 Z̄
 / \
 Z YP
 |
 Ȳ
 / \
 Y XP
 |
 X̄
 |
 X

```
- (56)
- ```

      ZP
      |
      Z̄
     / \
    Z   YP
        |
        Ȳ
       / \
      Y  XP
     / \  |
    X2 Y  X̄
          |
          X
          |
          e2
  
```
- (57)
- ```

 ZP
 |
 Z̄
 / \
 Z YP
 / \ |
 Y1 Z Ȳ
 / \ / \
X2 Y Y XP
 | |
 e1 X̄
 |
 X
 |
 e2

```

Remember that we are thinking of “chains” as the discontinuous terms that are spread across the positions by movement operations. Thus, when  $X^0$  moves and adjoins to  $Y^0$ , as in (56),  $X^0$ ’s chain is made up of the position that  $X^0$  moves into and the position it left. Let’s represent this as  $(X^0, e)$  — where “ $e$ ” is meant to represent the position



from which X moved. The step from (56) to (57) involves moving  $Y^0$  and adjoining it to  $Z^0$ . Although this movement also relocates the position of  $X^0$ , it is, technically, movement of  $Y^0$ , and thus a different chain is created:  $(Y^0, e)$ . In general, then, if Head Movement involves adjoining one head to another, then the derivation (55)-(57) represents the general case of iterative Head Movement. And, as can be seen, no chain formed will have more than two members.

Now the locality condition Chomsky formulates makes use of this idea about chains, and references a variety of terms which we define first.

(58)  $MAX(\alpha) =_{def.}$  the smallest maximal projection that contains  $\alpha$ .

(59) "Let us define the domain of CH [ $CH=(\alpha, e)$ ] to be the set of nodes contained in  $MAX(\alpha)$  and not containing any  $\bar{\alpha}$ ."

(basically Chomsky 1992, p. 19)

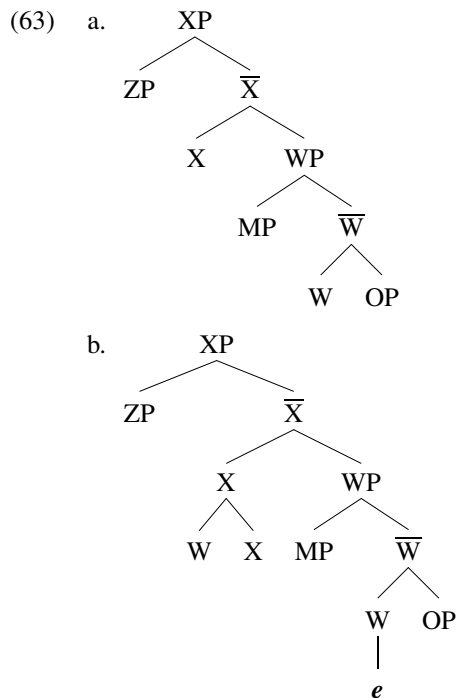
(60) "For any set S of categories, let us take  $MIN(S)$  (minimal S) to be the smallest subset K of S such that for any  $\alpha \in S$ , some  $\beta \in K$  reflexively dominates  $\alpha$ ."

(*ibid*, p. 16)

(61)  $\alpha$  can A-move to  $\xi$  in a structure  $\phi$  only if there is no  $\beta$  in  $\phi$  such that  $\xi$  c-commands  $\beta$  c-commands  $\alpha$  and  $\xi, \beta$  are not equidistant from  $\alpha$ .<sup>9</sup>

(62) (Only?) if  $\xi, \beta$  are in the same minimal domain, they are equidistant from  $\alpha$ .

To see what these definitions do, consider the trees in (63).

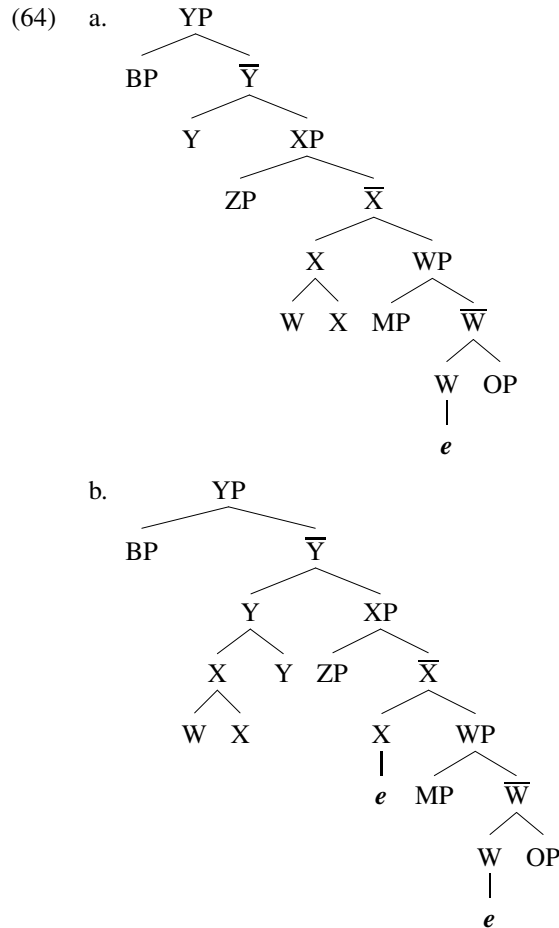


<sup>9</sup>Understand  $\beta$  to be an "argument position" for the cases that we are concerned with.

In (63a), the  $\text{MAX}(X)$  and  $\text{MAX}(W)$  are  $XP$  and  $WP$  respectively. The domain of  $X^0$  is  $\{ZP, WP, , MP, W, OP\}$  as well as everything that  $ZP$ ,  $MP$  and  $OP$  dominate. The domain of  $W^0$  is  $\{MP, OP\}$  and everything they dominate. The minimal domain of  $X^0$  is  $\{ZP, WP\}$ . And, finally, the minimal domain of  $W^0$  is just  $\{MP, OP\}$ . The condition in (61) will now prevent  $OP$  from A-moving to  $ZP$ , since  $MP$  and  $ZP$  are not equidistant to  $OP$ , and  $MP$  is “between” them, in the manner described. When  $MP$  is a “subject,” this will capture the effects of Specified Subject Constraint.

Consider now (63b). As before, the  $\text{MAX}(X)$  and  $\text{MIN}(\text{DOM}(X))$  are  $XP$  and  $\{ZP, WP\}$  respectively. But the  $\text{MAX}((W, e))$  is now  $XP$ . And as a consequence its Minimal Domain is going to be  $\{ZP, MP, OP\}$ . In particular, then, moving  $W$  to  $X$  has caused  $ZP$  to be included in  $W$ 's minimal domain. As a consequence,  $MP$  and  $ZP$  will be in the same minimal domain, and this will make them equidistant to  $OP$ . Thus, A movement from  $OP$  to  $ZP$  will be allowed.

Finally, consider a scenario in which there has been multiple head movement, as in (64).

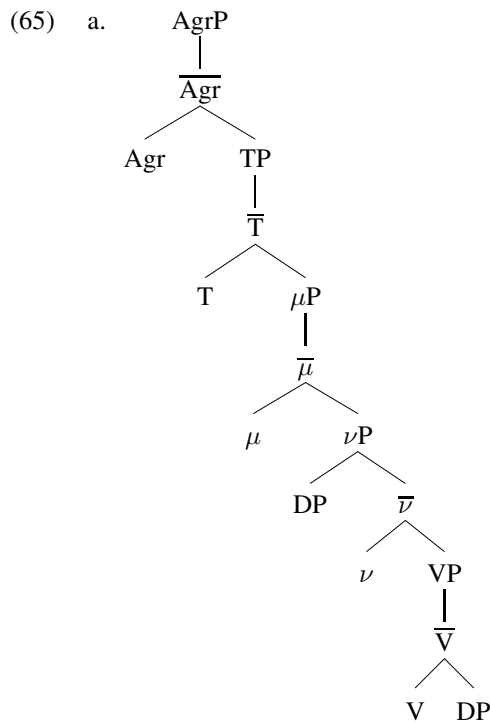


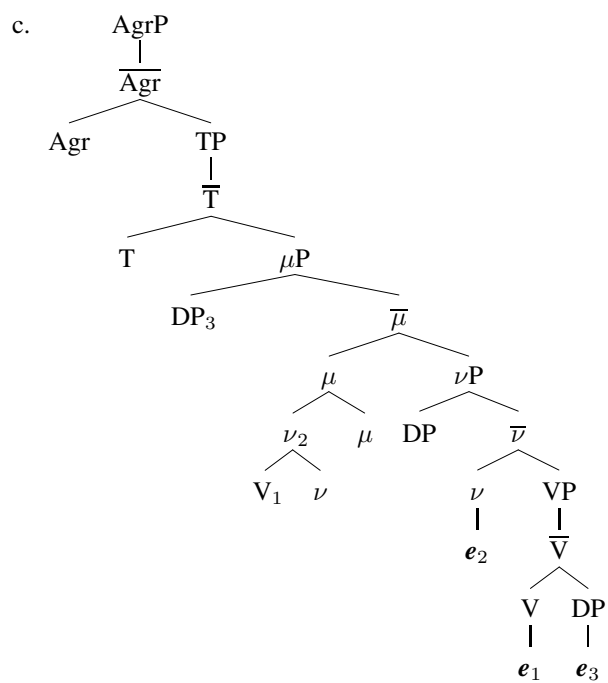
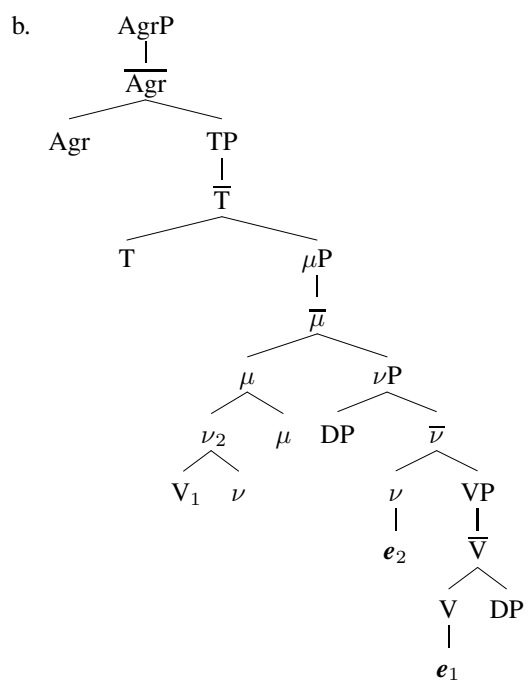
This derivation triggers in a crucial way the idea that chains which are formed by Head

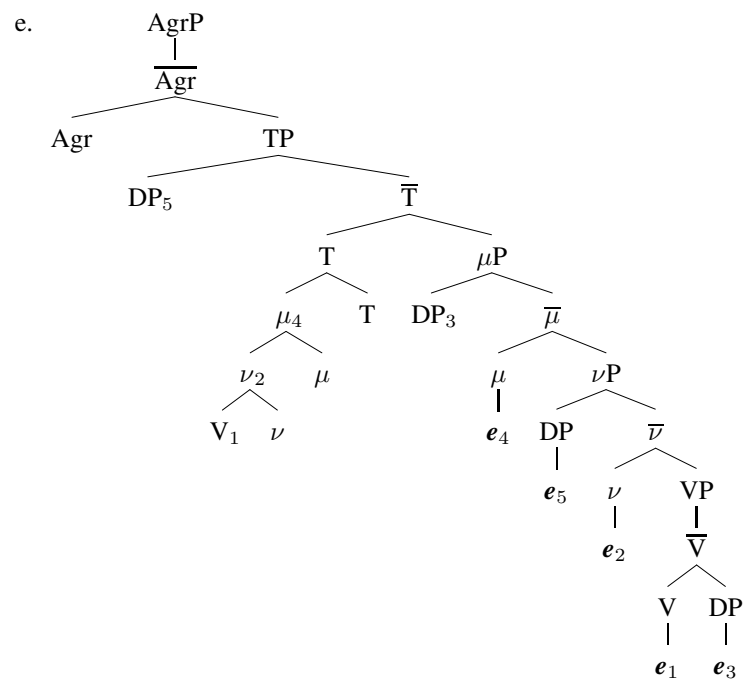
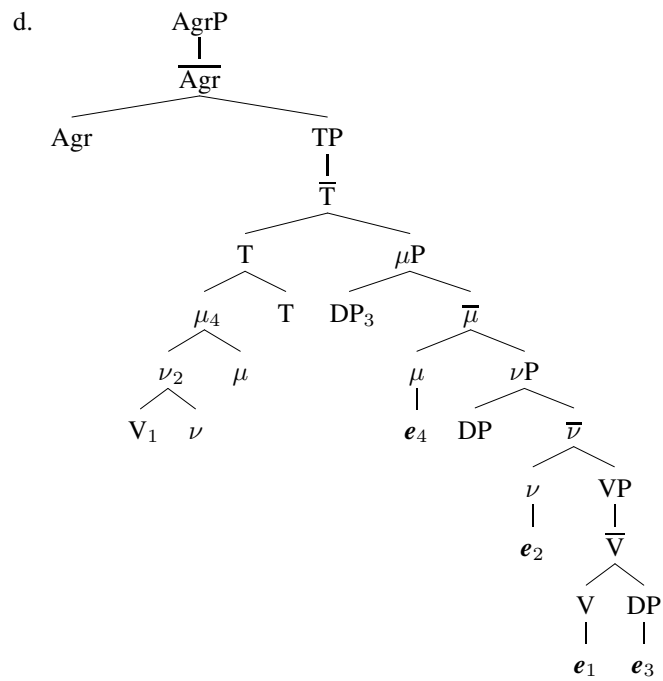
Movement have at most two members. It is this constraint that prevents (W, e) from having a larger minimal domain than that given to it in (63). Because the minimal domain of W doesn't grow as a function of X's movement, MP and BP will not be in the same minimal domain. As a consequence, (61) will prevent OP from moving past both MP and ZP into BP.

This is a general effect. And, as can be seen, it has the interesting consequence that it will derive the difference between the causative and non-causative clitic movement cases. Further, this proposal extends to our problem, and, interestingly, even might contribute to an explanation for why subjects and objects get the Cases they do. Let's see how.

Consider first how the derivation that brings the subject and object into the correct positions goes. This derivation move the object past the subject into Specifier of  $\mu$ P, and then take the subject past the object in Specifier of  $\mu$ P into Specifier of AgrP. This derivation is sketched in (65).



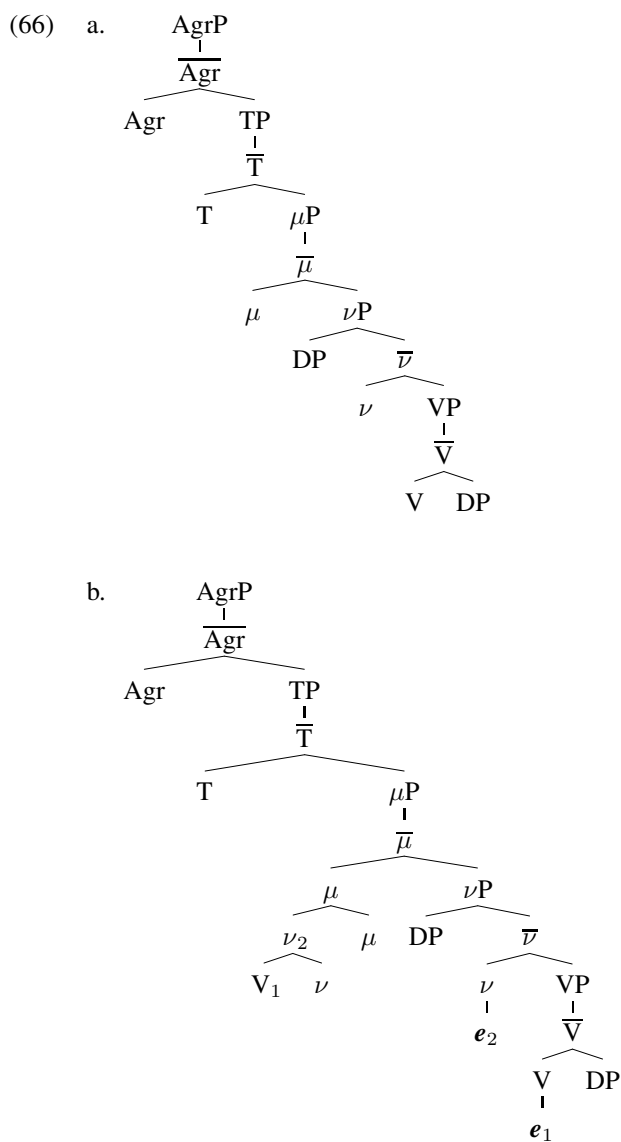


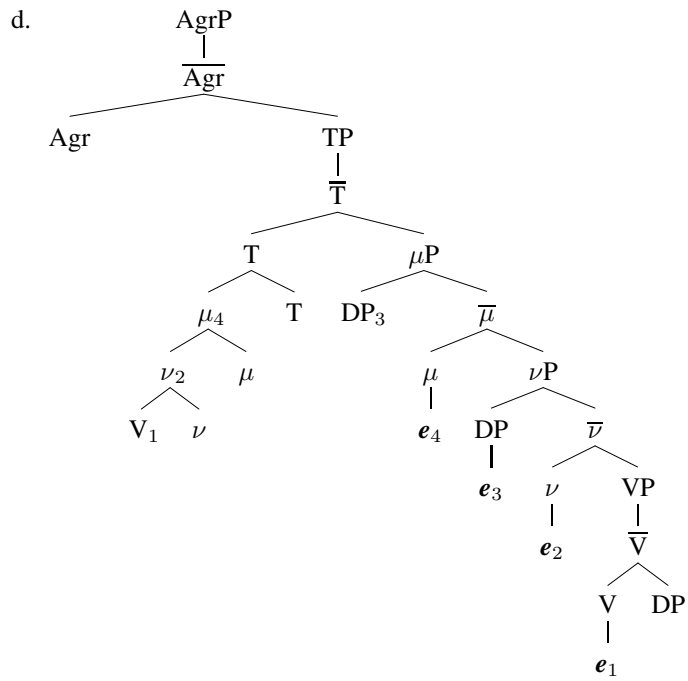
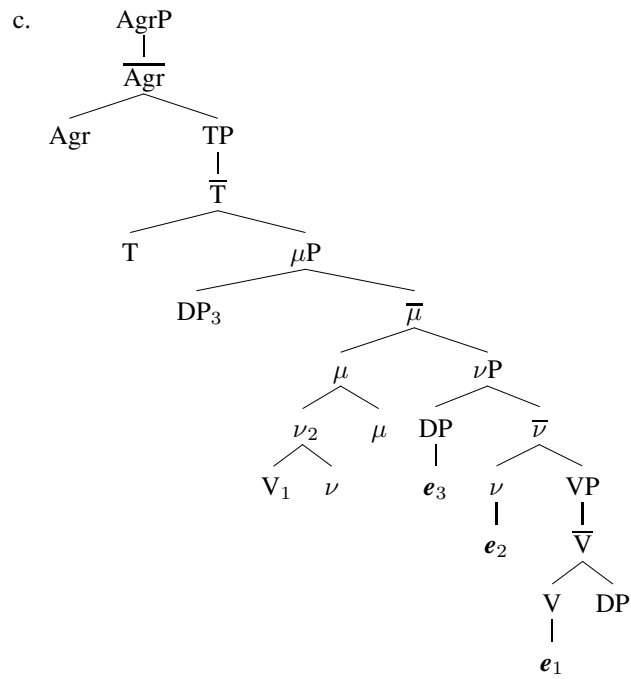


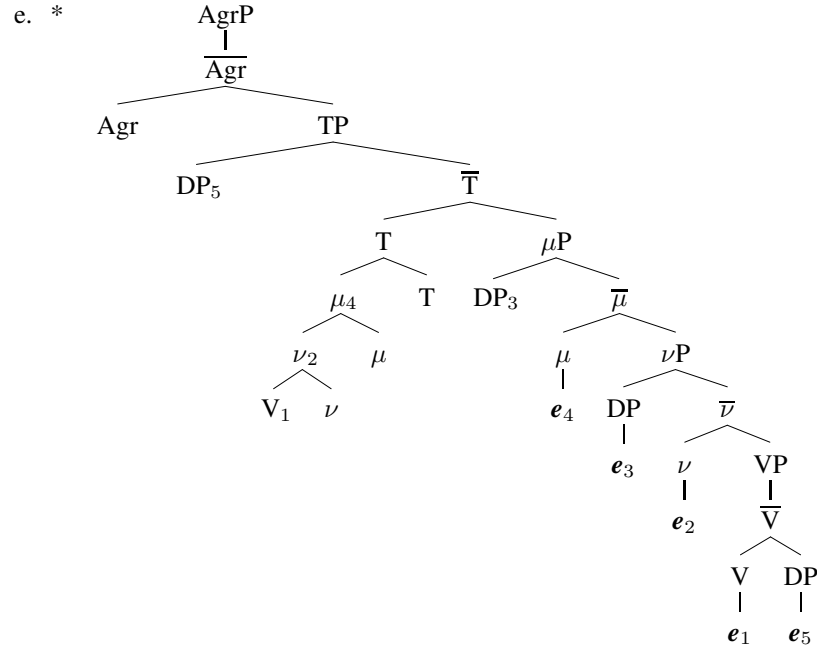
Movement of  $\nu$  to  $\mu$  causes the Specifier of  $\mu$ P and the Specifier of  $\nu$ P to be in the same minimal domain. Thus, movement of the object to Specifier of  $\mu$ P — the step

that takes (65b) to (65c) — will satisfy (61). (The c-commanding term that the object passes on this step is in the same minimal domain as the position it lands in.) Similarly,  $\mu$ 's movement to T, causes the Specifier of  $\mu$ P and the Specifier of TP to be in the same minimal domain. As a consequence, the subject is able to move past the Specifier of  $\mu$ P into the Specifier of TP without violating (25). From the Specifier of TP, the subject now has a clear path to Specifier of AgrP, where it gets Nominative Case.

Consider by contrast a derivation that delivers the subject and object into the wrong positions. Such a derivation might go something like (66).







In this derivation, the fatal step is the last one. Here the object has moved past both Specifier of  $\mu P$  and Specifier of  $\nu P$  into Specifier of TP. By moving into Specifier of TP, the object has moved past a c-commanding position (Specifier of  $\nu P$ ) which is not in the same minimal domain as the Specifier it has moved into.

As is clear from this derivation, then, this account must be augmented so that something is forced to be present in Specifier of  $\nu P$ , even when the subject that originates there has moved. This is, again, where the notion of a Chain comes into play. The rough idea is that even though the subject is spoken in the higher of the two positions it occupies through the derivation, it is actually occupying both of those positions throughout the derivation. This way of conceiving of the movement operation is very common and we shall see, soon, a couple of the facts that are credited to it.

One way this idea has been played out is to think of the movement operation as having two parts. It gives the moved term a new position, and it inserts a silent term that functions semantically like a variable bound to that moved term. Bound variables are referring expressions whose value is completely dependent on the interpretation given to the term it is bound to. These particular variables are commonly known as “traces.” So we define movement, now, as in (67).

(67) MOVE  $\alpha$

- a. Relocate  $\alpha$  from position  $\rho_1$  to position  $\rho_2$ .
- b. Place in  $\rho_1$  a trace bound to  $\alpha$ .

This has the nice auxiliary effect of offering a way to derive the Upwards Constraint. There is evidence that bound variable must be c-commanded by their binders, and many semantic theories of variable binding are designed to derive this. Thus, when a term



moves, it must relocate the term into a position from which it can bind the variable left behind, and this means that it will have to move to c-commanding positions.

There is a slight variant of this idea that is well represented in the literature. On this variant, movement continues to do nothing more than move terms from one position to another, leaving no traces. Instead there is a separate trace insertion operation which optionally applies. This operation is then forced by yet a third constraint. In *Lectures on Government and Binding*, the constraint that forces trace insertion is argued to be a strengthening of the Projection Principle. In particular, Chomsky proposes that the Projection Principle apply to every phrase marker in a derivation. We might formulate it as (68).

(68) THE PROJECTION PRINCIPLE

- a. Let  $\alpha$  have c-selection requirements, and  $\alpha'$  be an  $\bar{X}$  projection of  $\alpha$ . Then there must be an  $\alpha'$  in which only phrases satisfying  $\alpha$ 's c-selection requirements are found.
- b. If  $\alpha$  c-selects  $\beta$  in position  $\rho$  in one phrase marker in a derivation, then  $\alpha$  c-selects  $\beta$  in  $\rho$  in every phrase marker in that derivation.

When movement relocates a term from a position in which it satisfied some c-selection requirement, a violation of the Projection Principle will ensue. Inserting a trace, and letting this trace get the semantic value of the moved item, is intended to rectify this violation.

This variant differs from (67) in two ways. (67) says that traces should be found in every position from which a term moved, and no where else. But this alternative let's traces be inserted independently of the movement operation, making it possible that traces could exist in positions from which nothing has moved. And since the Projection Principle only imposes a requirement on c-selected positions, these are the only spots in which traces will be forced to occur. Unlike (67), the second variant allows movement to not leave a trace in its wake.

Either one of these versions of expressing the idea of Chain will serve the necessary role in (66). They will move require that the Specifier of  $\nu$ P be occupied at in every representation in the derivation. And this triggers a violation of Chomsky's locality condition.

Notice how this account prevents the subject from getting Accusative Case by blocking an object's movement to the Nominative Case marked position when this happens. That is, it blocks a derivation in which the subject erroneously moves into Specifier of  $\mu$ P by making it then impossible for the object to move into the higher Specifier of AgrP to get Case. Essentially, it rigs things in such a way that, because the subject and object DP arguments start out in hierarchically different positions, they have access to the Case marked positions that have parallel hierarchically different positions. But, crucially, it does this by making the object DP incapable of making it as far as the nominative Specifier of AgrP when a subject DP is also present. Nothing directly prevents the subject argument from moving into the accusative Case marked position. This leads to the prediction that subject arguments should be able to get accusative Case when no object DP is present. Thus we should expect to find cases where intransitive verbs of various sorts surface with accusative subjects. This never happens in English:

- (69) a. \* Him tried to speak.  
 b. \* Her talked to Jimbo.  
 c. \* Me ran.

Interestingly, this pattern does emerge in other languages. Unlike English, and all of IndoEuropean, which uses one Case for the subjects of transitive and intransitive verbs alike and has a different Case for objects, there are languages which use the same Case for objects and subjects of intransitive predicates, and a different Case for the subjects of transitive predicates. These are known as “Ergative” Case marking languages. The system proposed here, then, derives Ergative Case marking systems. What is left to be understood is how to control the difference between the two language groups so that the same constraint will allow for languages which don’t have Ergative Case marking systems.

Chomsky proposes that the difference be expressed in terms of a preference given to one or the other of these two Case assigners. We might express this as follows:

- (70) If  $\alpha^0$  assigns Case, then there must be something to bear it.  
 a. Delete  $\mu P$ .  
     (Nominative-Accusative systems.)  
 b. Delete AgrP.  
     (Ergative systems.)

In the cases we are concerned with, there are more Case assigners than there are phrases to bear those Cases. Thus, one or the other of AgrP or  $\mu P$  will violate (70). The two groups of languages make a different choice about how to alleviate that problem, removing one or the other of the Case assigners.

Let us adopt this as our method of guaranteeing that subject and object move into the right Case marked positions, though we may have a chance later to explore other possibilities. This provides a solution, then, to the most glaring problem with separating the accusative Case marker from the verb. But this isn’t the only problem we must find a solution for.

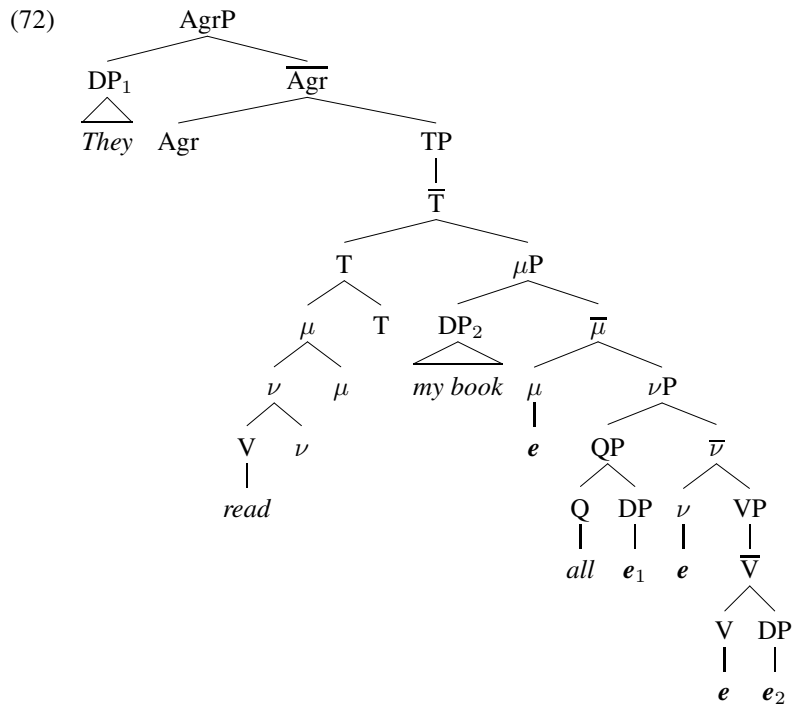
Another comes from letting the subject originate below the surface position of the object, and adopting Sportiche’s account of quantifier float. We should expect that a quantifier can be floated off of a subject and surface to the right of the object, as in (71), with the representation in (72) on the next page.

- (71) \* They read my book all.

(Note: in this, and all subsequent representations, I will use **e** to designate “traces.” That is, these are no longer mere typological devices to represent more than one phrase marker in a derivation. They are instead actually present in the phrase marker.) We also expect objects, now that they move overtly, to be able to strand a quantifier in their underlying position. This means that we should expect, contrary to fact, that sentences like (73) should be grammatical.

- (73) \* I read the books all.

But these sorts of problems exist for Sportiche’s account even before the innovations we’ve reviewed here. As Sportiche noted, we should expect passive sentences to allow



quantifiers to follow the verb, since, at least on the best established accounts of passive, this operation merely relocates an object from its underlying VP internal position to the Specifier of AgrP. But quantifiers cannot be stranded in this position in English:

- (74) \* The books were read all.

To square Sportiche's account with these data, we need to find something that derives the following generalization.

- (75) A quantifier cannot be stranded in a c-selected position.

Let's assume that we can find something with this consequence. (I will make a suggestion in a couple chapters.)

And, finally, as foreshadowed at the outset, there is a difficulty that accompanies the decision to define structural Case as Case that is assigned to Specifier position. This, recall, was one of the ideas that originally motivated taking Accusative Case to be assigned by  $\mu$ . But there is one place where structural Accusative Case does not seem to be assigned to Specifier position, and this you will remember is in infinitives such as (76).

- (76) I'd prefer [<sub>CP</sub> for [<sub>AgrP</sub> this to work out better ]].

Unless we can find an alternative treatment of these cases, it does not look like we want structural accusative Case to always be assigned by some head to its Specifier.

Let me sketch a solution to this problem that has been popular.<sup>10</sup> The essential idea in this approach is that the Agr<sup>0</sup> heading infinitival clauses is also capable of assigning a structural Case. In (76), that Case happens to be Accusative.

We initially adopted the view that it was *for* which assigns Case in (76) because of the correlation between the occurrence of this complementizer and the availability of Accusative Case on the following term. We shall have to find another way of expressing this correlation if it is Agr<sup>0</sup> that is responsible for assigning this Case. In fact, that *for* governs and assigns Accusative Case to the following Specifier position figured in a larger correlation involving the distribution of PRO. Recall that our typology of “transparent” and “opaque” infinitives involves, partly, the observation that only opaque infinitives are capable of housing PRO. We captured this by letting the distinction between “transparent” and “opaque” infinitives devolve to whether they are CPs or AgrPs. Then we devised the condition on PRO in (77) that is sensitive to this distinction.

(77) PRO RESTRICTION

PRO cannot be governed by a lexical item.

This accounts for why those Specifiers of infinitives we initially characterized as being targets for Case assignment by something outside the infinitive cannot host PRO. This is what the paradigm in (78) indicates.

- (78) a. I'd prefer [<sub>CP</sub>for [<sub>AgrP</sub> this to work out better ]].  
 b. I consider [<sub>AgrP</sub> this to be a bust ].  
 c. \*I'd prefer [<sub>CP</sub>for [<sub>AgrP</sub> PRO to work out better ]].  
 d. I consider [<sub>AgrP</sub> PRO to be a bust ].

This correlation is going to have to be found in some other way if Accusative Case is not assigned by *for* in (76)/(78a).

In fact, there is something slightly odd about accounting for the distribution of PRO with (77). Not only is (77) designed to account for the fact that only certain infinitives can have PRO in their Specifier position — blocking PRO from appearing in complement of V or P position, for instance — it is also designed to guarantee that PRO moves in those contexts where it is c-selected in one of its disallowed positions. Thus, for instance, in (79), PRO is driven from the object position of *promoted* into Specifier of infinitival AgrP.

- (79) Gary wanted [ PRO<sub>1</sub> to be promoted e<sub>1</sub> ].

The PRO Restriction prohibits PRO from remaining in the position governed by *promoted*, and so Argument Movement kicks in to generate an additional phrase marker in which PRO is not governed.<sup>11</sup> Thus, the engine that is driving Argument Movement when it applies to PRO is the PRO Restriction. But a precisely parallel derivation is fueled by the Case filter in those situations where an overt DP is involved, as in (80).

<sup>10</sup>See Bošković (1996, 1997).

<sup>11</sup>The EPP also requires an additional phrase marker to satisfy its requirements.

(80) Gary wanted [ Sally<sub>1</sub> to be promoted e<sub>1</sub> ].

The theory we have, then, says that the underlying engine for these two cases is independent. It's the Case filter that drives Argument Movement of overt DPs, and it's the PRO Restriction that drives Argument Movement for PRO.

The way Argument Movement applies to PRO and overt DPs is really exactly the same. It is the same, in particular, in a way that suggests that we do not want a different engine driving them. Recall that we formulated Argument Movement as follows:

- (81) ARGUMENT MOVEMENT  
Move a term from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory,  
only if:
- i. not doing so would violate the Case filter, and
  - ii.  $\beta$  c-commands  $\alpha$ , and
  - iii. there is no CP that contains  $\alpha$  but not  $\beta$ .

(93ii) is the Upwards Constraint, which we have now derived by way of the relationship between the moved item and the trace it must bind. We could probably simplify this to (82), therefore.

- (82) ARGUMENT MOVEMENT  
Move a term from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory,  
only if:
- i. not doing so would violate the Case filter, and
  - ii. there is no CP that contains  $\alpha$  but not  $\beta$ .

(82ii) captures the fact, recall, that Argument movement cannot take something out of a clause unless it's transparent. Thus, it captures the ungrammaticality of the cases in (83), where either a PRO or an overt DP have moved.

- (83) a. \*I wanted [ PRO<sub>1</sub> to be tried [CP to be promoted e<sub>1</sub> ]].  
b. \*I wanted [ Sally<sub>1</sub> to be tried [CP to be promoted e<sub>1</sub> ]].

This is one of the reasons, then, that we took Argument Movement to be responsible for resolving both the PRO Restriction and the Case filter in situations where one phrase marker could not do that.

But now consider (82i). This makes reference to the fact that it is the Case filter that is driving Argument Movement. It describes the ungrammaticality of cases like (84).

- (84) \*Jerzy<sub>1</sub> seemed to e<sub>1</sub> that Sal slept.  
*compare:*  
It seemed to Jerzy that Sal slept.

(84) is ungrammatical because *Jerzy* moves from the position Case marked by *to*. But we need the same constraint on Argument Movement when it applies to PRO, as in (85) shows.

- (85) \*I wanted [<sub>CP</sub> PRO<sub>1</sub> to seem to e<sub>1</sub> that Sal slept.

This isn't predicted by the current system. The Case filter can't be violated or satisfied by virtue of moving PRO, so its effects should not steer Argument Movement when Argument Movement is relocating PRO.

There are other problems with using the PRO Restriction to describe the distribution of PRO. For instance, recall that it allows PRO to reside in the Specifier of a root clause. In such a position, it will be the highest term in the sentence, and therefore there will be nothing that could govern it. And yet, PRO cannot reside in this position:

- (86) \*PRO was happy.

We closed this gap earlier by adding a prohibition against PRO residing in Specifier of finite AgrP. Clearly, it is not going to be straightforward to unite these two descriptions into a unified explanation for PRO's distribution.

And then there are problems that arise when the distribution of PRO is considered in closely related languages. We often find in these cases that sentences parallel to (76) are grammatical. In Icelandic, for instance, infinitival clauses with PRO subjects quite easily have an overt complementizer associated with them:

- (87) *María lofaði að PRO ekki lesa bókina.*  
 Mary promised that PRO not to read book-the  
 'Mary promised to not read the book.'

In this example *að* is the Icelandic complementizer — it is used in both infinitival and finite clauses.

The approach taken to the Accusative Case marking in examples like (88) and its complementary distribution with PRO, involves a theory of PRO that expresses its distribution in terms of Cases — rather than terms of being “ungoverned.” It will therefore directly address the problem posed by (85).

- (88) a. I'd prefer [<sub>CP</sub> for [<sub>AgrP</sub> this to work out better ]].  
 b. \*I'd prefer [<sub>CP</sub> for [<sub>AgrP</sub> PRO to work out better ]].  
 c. \*I tried [<sub>CP</sub> for [<sub>AgrP</sub> this to work out better ]].  
 d. I tried [<sub>CP</sub> for [<sub>AgrP</sub> PRO to work out better ]].

The leading idea is that PRO, like any other argument DP, requires a derivation in which it occupies a Case marked position. Unlike other DPs, it specifies which Case it must bear. Moreover, it requires a Case that is uniquely its own. Let's call it “invisible Case,” and the other Cases we've encountered (i.e., Accusative, Nominative, Genitive, etc.) “visible Case.” Thus, we should adopt a Case Theory like (89).

- (89) CASE FILTER
- a. Every overt DP must occupy a “visible” Case marked position and cannot occupy an “invisible” Case marked position.
  - b. Every PRO must occupy an “invisible” Case marked position and cannot occupy a “visible” Case marked position.

This will have the desired effect, note, of preventing PRO from occupying the Specifier position of finite AgrP, since this is a position to which a visible Case is assigned. In this way, (86) can be blocked.

Thus, we have now at least four kinds of Agr<sup>0</sup>:

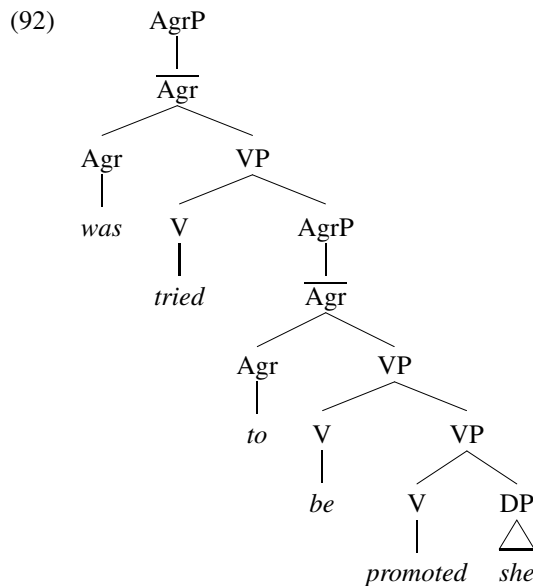
- (90)
- a. Agr<sub>fin</sub> assigns Nominative.
  - b. Agr<sub>inf<sup>a</sup></sub> assigns invisible.
  - c. Agr<sub>inf<sup>b</sup></sub> assigns Accusative.
  - d. Agr<sub>inf<sup>c</sup></sub> doesn't assign Case.

Verbs like *try* and *promise* select infinitival clauses headed by Agr<sup>0</sup><sub>inf<sup>a</sup></sub>. Thus in these infinitivals we find PRO subjects, and never overt ones.

Moreover, we no longer need to claim that these infinitival clauses have empty complementizers — they could simply be AgrPs. The only work that classing them as CP does now is to prevent Argument Movement from relocating things out of them. But this can now be derived from (89) and the Extended Projection Principle. To see this, consider how (91), an example of the putative constraint against A Movement out of CPs, will be derived.

- (91) \*She<sub>1</sub> was tried [ to be promoted e<sub>1</sub> ].

If we let the infinitival complement be just an AgrP, and not a CP, then we can take the underlying representation to be as in (92). (I've foreshortened lots of steps here — in particular, much of the structure belonging to the root clause has been suppressed.)



marked, now: the lower one gets invisible Case from the infinitival  $\text{Agr}^0$ , and the higher one gets Nominative Case from the finite  $\text{Agr}^0$ . As a consequence, as soon as *she* has moved into one of these positions, (82i) will prevent it from moving into the other.

Everything we've just described here will hold for finite AgrPs too. That is, (82i) will combine with the Extension to the Projection Principle to prevent Argument Movement out of a finite clause this is because finite clauses, like the infinitival clause in (92), assigns Nominative Case to its Specifier.

Indeed, we can now define opaque infinitives as ones that assign Case to their Specifiers. As we've just seen, this Case marked position is going to capture any DP that is attempting to Argument Move out of the infinitive. And, because one of the Cases that infinitival  $\text{Agr}^0$ s can assign is the invisible Case, we capture the fact that PRO is capable of living in only opaque clauses. Thus we capture one of the correlations that the opaque/transparent distinction was defined over: opaque infinitives can host PRO but are islands for Argument Movement, whereas transparent infinitives cannot host PRO and are not islands for Movement.

So, we can simplify the definition of Argument Movement once more, to:

(93) ARGUMENT MOVEMENT

Move a term from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory, only if not doing so would violate the Case filter.

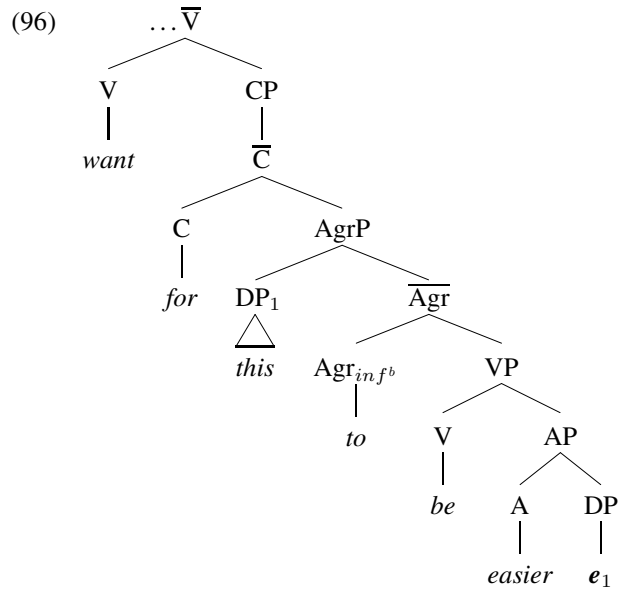
What about transparent infinitives? These come in two sorts, neither of which allow PRO inside them. These, then, are infinitives headed by  $\text{Agr}^0_{infc}$ : the  $\text{Agr}^0$  that does not assign a Case. As a consequence, not only can PRO not appear in these infinitives, but it is possible to Argument move something into Specifier of AgrP, satisfying the Extension to the Projection Principle, and then move back out. These infinitives, in other words, will not be islands for Argument movement. When these infinitives are c-selected by a verb that assigns and external  $\theta$ -role, then there will be an Accusative Case assigning  $\mu\text{P}$  in the higher clause available for an argument that might begin within the infinitive. This is what happens with *consider* or *believe*, for instance. The first sort, then, invokes a syntax like that in (94) on the facing page (in which the  $\nu\text{P}$  belonging to the higher clause has been suppressed, along with some material in the embedded infinitive).

The other sort of verb that c-selects a transparent infinitive is *seem*, a verb that has neither an external  $\theta$ -role nor  $\mu\text{P}$ . As a consequence, in these cases, if there is an argument within the infinitive Case it will not be able to move into the  $\mu\text{P}$  immediately above the infinitival clause, as happens in (94), because this will be missing. Instead, it will have to move farther seeking out a higher Case marked position. If the higher clause is a finite AgrP, then this position will be the Nominative Case Specifier of this AgrP.

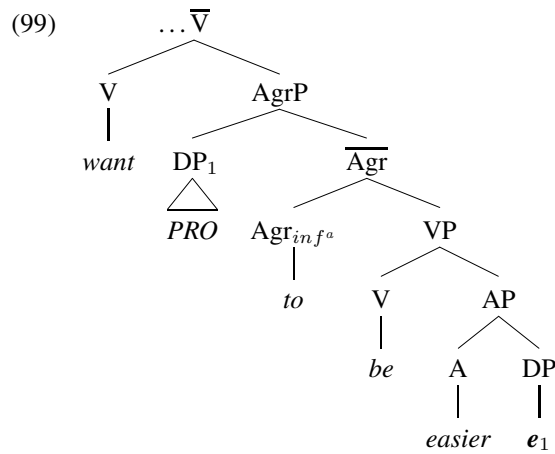
Transparent infinitives, then, are just those that don't assign Case to their Specifier position. We've got a new way, then, for capturing these generalizations. They all devolve to the Case-based condition on Argument Movement. The difference between the two infinitives being just whether they assign Case or not:







I suggest that we treat (98a) as a normal opaque infinitive, and (98b) as a normal transparent one. This amounts to the claim, then, that *want*, and the verbs like it, select any of the kinds of infinitival clauses that English has. The representation (98a) will get is parallel to that given to the complements of *try* or *promise*; it's something like (99) (where  $\text{Agr}_{inf^a}^0$  represents the  $\text{Agr}^0$  that assigns invisible Case).



(98b) will get a representation parallel to that we've given to *believe* in (94). This is different than the original account we entertained for this case. We adopted the view that the infinitive in (98b) is an opaque infinitive because of the inability of passivizing *want* to rob the Accusative DP of its Case. There is a difference between *want* and *believe* in this respect that will be lost on this proposal:

- (100) a. Mittie was believed to like chocolate.  
 b. \* This was wanted to be easier.

If (98b) really is a transparent infinitive, then this difference must be accounted for in some other way.

The advantage to seeing (98b) as a transparent infinitive is that there is some distributional evidence that the Accusative Case which the subject bears comes from the higher clause. The distributional evidence is this: the only terms that can select these infinitives with all three guises are verbs. As we've just reviewed, *want* is one of these, and so is *prefer*:

- (101) a. I prefer [for chocolate to be discussed].  
 b. I prefer [PRO to be discussed].  
 c. I prefer [chocolate to be discussed].

When adjectives select this type of infinitive, this last guise drops out:

- (102) a. It is possible [for chocolate to be discussed].  
 b. It is possible [PRO to be discussed].  
 c. \* It is possible [chocolate to be discussed].

Adjectives in English don't support Accusative Case — that is, they don't come with  $\mu$ Ps — and in this respect they differ from verbs. We can explain the contrast between (101c) and (102c) by way of this difference between adjectives and verbs if the Accusative Case that is assigned in (101c) comes from the higher clause. This is just what classing these as transparent infinitives does.

We've seen, then, that there are three types of infinitival AgrPs in English, and they differ with respect to the Case assigning properties of their heads. Moreover, we've seen that there are three types of predicates that c-select infinitives. There are those that uniquely select Agr<sub>inf<sup>a</sup></sub>Ps, whose heads assign the invisible Case. These predicates, *promise*, *try*, *glad*,<sup>12</sup> etc. select the opaque infinitives — those that have PRO in them and do not allow overt DPs to find their Case outside the infinitive. Then there are those predicates, *believe*, *seem*, *likely*, etc., that uniquely select AgrPs headed by Agr<sub>inf<sup>c</sup></sub>: the Agr<sup>0</sup> that assigns no Case. These are transparent infinitives; they cannot host PRO, but they do allow DPs within them to find Case from without. And then there are predicates that select any sort of infinitive: the two already mentioned and the CP infinitival, whose head, *for*, selects the AgrP which supports Accusative Case. These are *want*, *prefer*, *possible* and a few others.

<sup>12</sup>As, for instance, in "I am glad to meet you."



## Chapter 9

# Rethinking Things

We have a lot of balls in the air, at the moment. It will be useful, perhaps, to take a moment to see what we have decided on, and what is still left open. At a minimum this will give us a chance to see if there is a coherent image of our system that might help guide us as we go forward.

### 9.1 Cleaning up

There remain some unsolved problems from the decision to let Accusative DPs surface in a position higher than the underlying subject position. One of these is that it brings the concomitant conclusion that main verbs in English move rather far, and this is at odds with Pollock's conclusion. Recall that Pollock's conclusion is that main verbs in English never move overtly, and that this is what is responsible for the contrast between French and English illustrated by (1).

- (1) a. Jean embrasse souvent Marie.
- b. \* Jerry kisses often Mary.

But we've also seen that this is an incomplete comparison set; English behaves like French when the complement is not a DP:

- (2) Jerry spoke fervently to Mary.

The contrast, therefore, targets something much narrower than the relationship between a main verb and its complement — it has to do with the relationship between a main verb and its accusative Case marked DP. What we look for, then, is something that forces the verb and Accusative DP to be adjacent in English, and model French differently.

Why are the verb and Accusative DP adjacent in English? If  $\mu$  is the source of Accusative Case assignment, then it cannot be that the verb and DP must be adjacent because of a condition on Case assignment, as in Stowell's account. Instead it emerges for a conspiracy of reasons on our present account. The first contributing reason is that the verb surfaces in English in the head directly above  $\mu$ P. The second contributor is

that Accusative DPs surface in the Specifier of  $\mu P$ . Adjacency between these two is now guaranteed if:

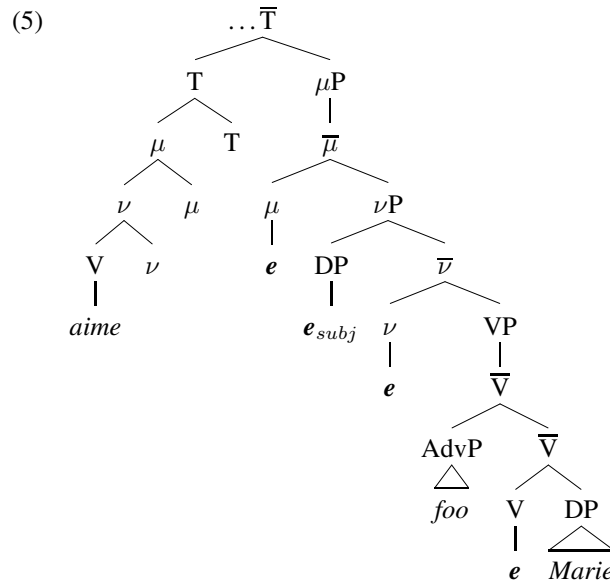
- (3) a. Nothing can adjoin (to the left) of  $\mu P$ , and
- b. Specifier of  $\mu P$  is linearized to be first.

If we prevent movement operations from adjoining things to the left of  $\mu P$ , then (3a) will follow from the conditions which position terms in D-structures. Since  $\mu$  c-selects only  $\nu P$ , the only other terms that can be in  $\mu P$  are modifiers, and they will be prevented from adjoining to  $\mu P$  by the modification rule:

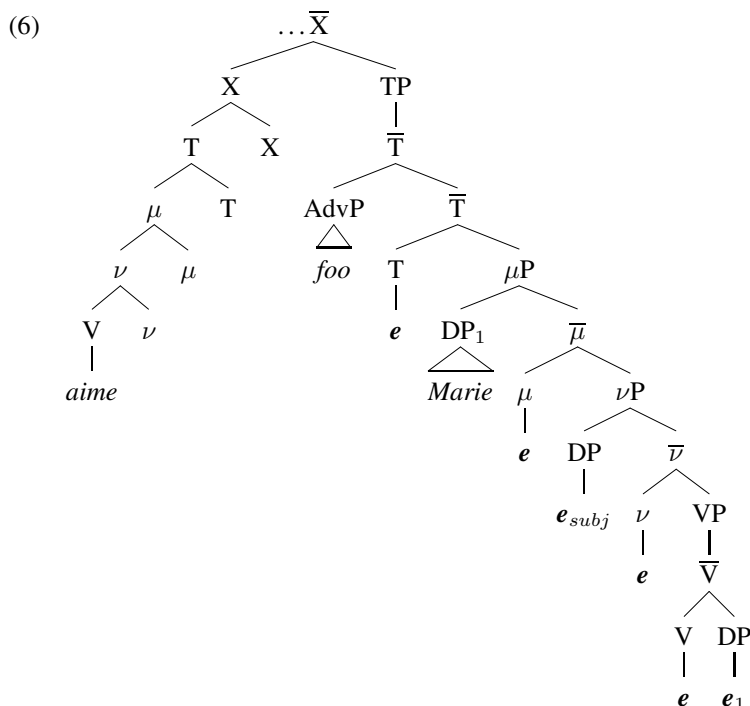
- (4)  $\{\bar{X}, \alpha\}$  iff  $\alpha$  modifies  $\bar{X}$ , where “ $\{\gamma \beta, \alpha\}$ ” indicates that  $\alpha$  and  $\beta$  are daughters of  $\gamma$ .

(17) requires that modifiers be adjoined to  $\bar{X}$ s.

If this is the reason that Accusative DPs and verbs are adjacent, then there are many potential differences between English and French that could produce Pollock’s difference. It could be, for instance, that French DPs are not forced to move overtly into Specifier of  $\mu P$  to get Case. This would allow them to have a lower position in English, and consequently a greater distance from the surface position of the verb. This would give to (1a) a parse like (5) below, for instance.



Or, alternatively, we could posit that verbs move in French beyond  $T^0$ , to some as of yet unknown head position. In this respect, the solution would be pointwise identical to Pollock’s proposal, though it would employ a different grammar of the ordering of complements than Pollock used. On this view, the sentence in (2a) might have a representation like (6), in which  $X^0$  represents the new mystery head to which French verbs move.



I don't know how to decide between these alternatives, or even whether we should. But the point is that Pollock's data do not present a problem for the decision to let main verbs in English move overtly: we have an embarrassingly rich assortment of possible accounts.

Another difficulty created by the decision to let objects surface higher than the underlying position of the subject that has gone untreated is its faulty interaction with Sportiche's account of quantifier float. We should expect quantifiers floated off of subjects to be able to surface to the right of objects, as in (7), whose parse is (8) on the next page.

(7) \* They read my book all.

Indeed, letting objects move brings the more direct problem that it predicts that quantifiers should be able to float off of objects, as in (9), and this too is incorrect.

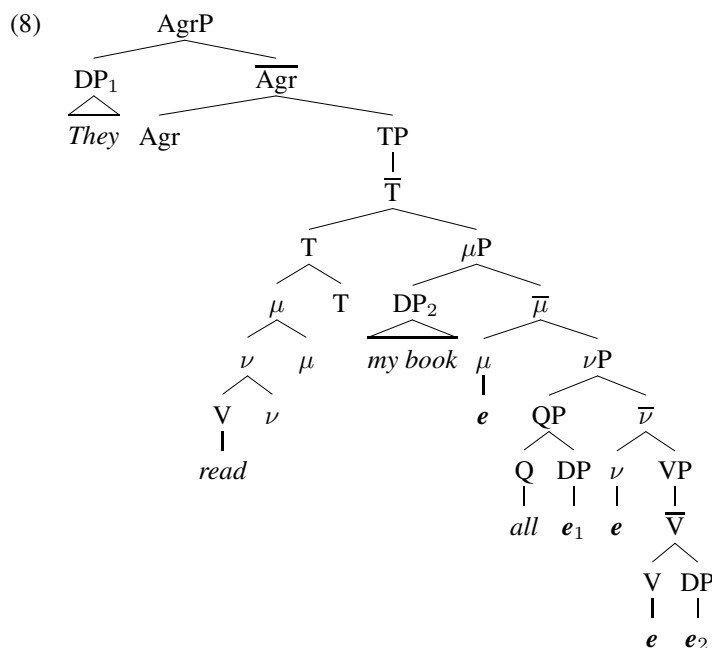
(9) \*I read the books all.

If we are to preserve both Sportiche's account of quantifier float and the present system of Case assignment, we must find some way of blocking these outcomes.

In fact, Sportiche's account of quantifier float already harbored this need. As he noted, his account predicts that objects which have moved in the Passive construction should be able to leave a floated quantifier in postverbal position, as in (10).

(10) \*The girls were visited all.

Sportiche's account requires something to block (10), and I suggest that it is also responsible for the ungrammaticality of (7) and (9). Something like (11) will get us a



long way in this direction.

- (11) A quantifier cannot be floated in its c-selected position.

While this will be sufficient to prevent (7), it still leaves one derivation available to produce the ungrammatical (9) and (10). The unwanted derivation that remains is one in which the object has moved into Specifier of VP, and from there, moves into its surface position. This derivation would allow the quantifier to be stranded in the Specifier of VP position, as in (12) on the facing page. Here, then, are some problems we must leave hanging: what is the source of (11), and how are derivations like (12) to be blocked.

With the problems put on hold, let's briefly bring together the central elements of the grammar. That portion of the grammar that was formerly expressed with phrase structure rules is expressed by way of the following set of well-formedness conditions:

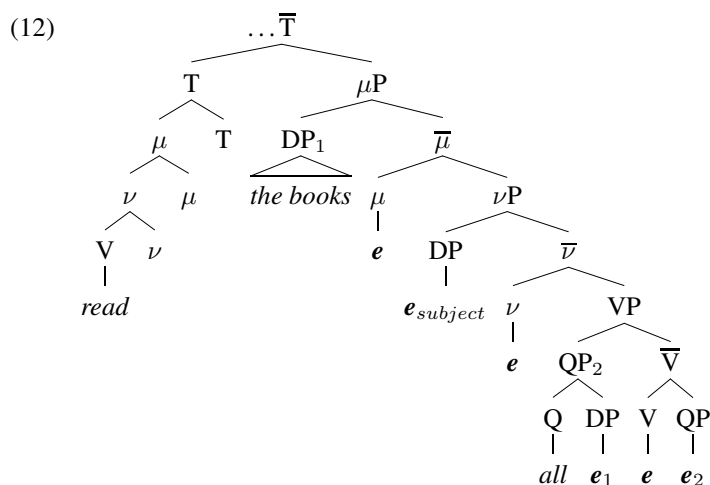
- (13)  $\bar{X}$  Skeleton:

- a.  $XP \rightarrow \{\alpha, \bar{X}\}$
- b.  $\bar{X} \rightarrow \{\bar{X}, \beta\}$
- c.  $\bar{X} \rightarrow \{X^0, \gamma\}$

- (14) Linearization:

- a. Specifier: [first, last]
- b. Projection of  $X^0$ : [first, last]





(15) PROJECTION PRINCIPLE I

Let  $\alpha$  have c-selection requirements, and  $\alpha'$  be some projection of  $\alpha$ . Then there must be an  $\alpha'$  such that all the phrases within  $\alpha'$  are c-selected by  $\alpha$ .

(16) THE THETA CRITERION

- i. For every argument, there is exactly one c-selected position.
- ii. For every c-selected position, there is exactly one argument.

(17) THE MODIFICATION RULE

$\{\overline{X}, \alpha\}$  iff  $\alpha$  modifies  $\overline{X}$ , where “ $\{\gamma \beta, \alpha\}$ ” indicates that  $\alpha$  and  $\beta$  are daughters of  $\gamma$ .

These conditions combine to determine how one phrase marker — the “D-structure” — looks. This is just one of a set of phrase markers that together make up a sentence’s derivation. A derivation is made necessary, in the situations we have examined, when the effects of the Projection Principle I and Theta Criterion conflict with the Case filter:

(18) THE CASE FILTER

Every DP must be assigned Case.

or when an “inflectional”  $X^0$  is generated with strong features that are matched by features carried on some other head. In these two scenarios, no single phrase marker can simultaneously satisfy all of the conditions, and so transformational rules will generate a series of them that, simultaneously, do. This series of phrase markers is defined below.

(19) A Derivation:

- a. Let  $\mathcal{R}$  be a transformational rule: that is, a function  $\mathcal{R}(P_i) = P_j$ , where  $P_i$  and  $P_j$  are phrase markers.

- b. Let *D-structure* be a phrase marker with lexical items that satisfies the  $\bar{X}$  Skeleton, Projection Principle I, the Theta Criterion and the Modification rule.
- c. Let *S-structure* be the phrase marker that satisfies the Case Filter and has no strong features in it, and is phonologically interpreted.
- d. A *Derivation* =<sub>def</sub> an ordered n-tuple of phrase markers,  $\mathcal{P} = (P_1, P_2, \dots, P_n)$ , such that:
  - i. Each  $P_i = \mathcal{R}(P_{i-1})$ , and
  - ii.  $P_1$  is a D-structure, and
  - iii.  $P_n$  is semantically interpreted (i.e., an LF), and
  - iv. There is exactly one  $P_i$  that is an S-structure, and
  - v. For every AgrP,  $\alpha$ , in the D-structure, there is some  $P_i$  which has a phrase in Specifier of  $\alpha$ .<sup>1</sup>
  - vi. If  $\alpha$  c-selects  $\beta$  in position  $\rho$  in the D-structure, then  $\alpha$  c-selects  $\beta$  in  $\rho$  in every  $P_i$ .

## (20) ECONOMIZE

If two derivations equally satisfy the same constraints, choose the one that minimizes the D-structure to S-structure portion of the derivation.

(21) Instances of  $\mathcal{R}$ 

- i. A Movement: Move an XP.  
Subject to the “minimal distance” version of the SSC & a condition that blocks movement from a Case marked position.
- ii. Head Movement: Move an  $X^0$ .  
Subject to the Head Movement Constraint and the Word Criterion.
- iii. Insert Trace: “Insert  $e$ ,” a silent bound variable.

Every  $\mathcal{R}$  is subject to:

## (22) THE LIKES ATTRACTS LIKES CONSTRAINT

An  $X^n$  can only adjoin to or substitute into an  $X^n$  position.

We can simplify the definition of S-structure if we let Case be expressed in terms of features, a suggestion that Chomsky (1995) makes. This will require us to change slightly the conditions under which features are checked off, allowing the Specifier-Head relation to be one way of checking them off. And we’ll have to also record the fact that “inherent” Case satisfies the Case filter in a way different than structural Cases. Inherent Cases, recall, are assigned under the same conditions that  $\theta$ -roles are assigned. If we are to express all instances of Case assignment in terms of features, then, it will be necessary to adopt something along the lines of (25).

- (23) Let every DP have a strong Case feature, and let  $\mu$  and Agr have strong Case features as well.

<sup>1</sup>This is the Extension to the Projection Principle.

- (24) A feature is deleted iff the term that bears it gets into a checking relation with a term that has a matching feature.  $\alpha$  is in a checking relation with  $\beta$  if  $\alpha$  is adjoined to  $\beta$ , or if  $\alpha$  is in  $\beta$ 's Specifier.
- (25) An “inherent” Case feature is deleted when c-selected by a term lexically specified as “assigning” that Case.

In addition to capturing the Case filter, (23) will also force Case assignment to be obligatory. And this derives the constraint prohibiting Argument movement from a Case marked position. To see this, there are two cases to be considered. The first arises when a DP with a Case feature  $\alpha$  moves into a Specifier of a head also bearing Case feature  $\alpha$  and then moves again into a Specifier of a head bearing Case feature  $\alpha$ . In these situations, once the DP has moved into the first Case marked position, (24) will cause the feature on the DP to be checked off as it checks off the feature on the head “assigning” Case. As a consequence, this DP will no longer have a Case feature on it, and consequently, cannot check off the Case feature on some higher head. The second scenario to consider is one in which a DP bearing Case feature  $\alpha$  moves first to a Specifier of a head bearing a different Case feature,  $\beta$ , and then from there moves onto a Specifier of a head bearing Case feature  $\alpha$ . In this scenario, the  $\beta$  Case feature will not be checked off, and since it is strong, will destroy the S-structure.<sup>2</sup>

Expressing the Case filter in terms of features now allows for a simplification of the definition of *S-structure*, as in (26).

- (26) Let *S-structure* be a phrase marker that has no strong features and is phonologically interpreted.

Chomsky (1995) has suggested tying the phonologically interpreted part together with the strong feature part more directly with:

- (27) A phrase marker with strong features cannot be phonologically interpreted.

And we change (19div) so that it reads:

- (28) Some  $P_i$  is phonologically interpreted.

We might similarly think of expressing the EPP, i.e. (19dv), in terms of features. Let's define a special feature, the “EPP feature,” which is checked off when it is in a checking relation with any category. Then we could express (19dv) with:

- (29) Agr has an EPP feature.

The Word Criterion, we might see as part of a definition of the mapping between positions in a phrase marker and the material that feeds them. That is, we might understand a fuller definition of the  $\bar{X}$  Skeleton to derive this.

And finally, note that the SSC and the Head Movement Constraint are very similar. We've spent some time seeing how the SSC might have a “weakening” which allows it to behave differently in contexts where head movement has applied, and we've

<sup>2</sup>This scenario could survive if some other  $\beta$ -bearing DP could occupy the Specifier of the head with the  $\beta$  Case feature. I have not been able to conceive of a situation where the ingredients necessary for such an outcome are present.

seen nothing similar in the case of the Head Movement Constraint. So, as it stands, there is at least this significant difference in them. But for the sake of discussion, let's factor these constraints out from the rules and relativize them appropriately:

- (30) ALMOST RELATIVIZED MINIMALITY<sup>3</sup>  
 $X^n$  cannot move past a c-commanding  $Y^n$ , unless  $X^n$  lands in a position that is in the same minimal domain that  $Y^n$  is in.

Okay, so putting all these changes together, we keep the  $\bar{X}$  Skeleton, the linearization principle, the modification rule, the Projection Principle I, the Theta Criterion and Economize. We dispense with the Case Filter, putting in place of this strong features. We dispense with the EPP as part of the definition of derivation, and let Agr have an EPP feature. And finally, we dispense with the rule specific constraints, and adopt Almost Relativized Minimality, and a definition of feature checking that derives the “can't move from a Case marked position” constraint. Our definition of a derivation and transformational rules now looks like this:

- (31) Derivations:
- a. Let  $\mathcal{R}$  be a transformational rule: that is, a function  $\mathcal{R}(P_i) = P_j$ , where  $P_i$  and  $P_j$  are phrase markers.
  - b. Let  $D\text{-structure}$  be a phrase marker with lexical items that satisfies the  $\bar{X}$  Skeleton, Projection Principle I, the Theta Criterion and the Modification rule.
  - c. A *Derivation* =<sub>def</sub> an ordered n-tuple of phrase markers,  $\mathcal{P} = (P_1, P_2, \dots, P_n)$ , such that:
    - i. Each  $P_i = \mathcal{R}(P_{i-1})$ , and
    - ii.  $P_1$  is a D-structure, and
    - iii.  $P_n$  is semantically interpreted (i.e., an LF), and
    - iv. Some  $P_i$  is phonologically interpreted, and
    - v. If  $\alpha$  c-selects  $\beta$  in position  $\rho$  in the D-structure, then  $\alpha$  c-selects  $\beta$  in  $\rho$  in every  $P_i$ .
- (32) Instances of  $\mathcal{R}$ :
- i. A Movement: Move XP.
  - ii. Head Movement: Move  $X^0$ .
  - iii. Insert trace.

## 9.2 Towards Deriving $\bar{X}$ Theory

### 9.2.1 Kayne's “Antisymmetry” hypothesis

There are some properties of this system that generate certain generalizations about derivations as they express themselves in English. One of these is that all movement

<sup>3</sup>The “real” Relativized Minimality inspires this, and can be found in Rizzi (1990).

operations, except for NP Shift, have relocated items leftwards. This effect is produced by setting the linearization values so that Specifiers come first, and so do projections of  $X^0$ , and then letting the fact that traces be in the scope of the phrase that binds them force movement to always be upwards. In fact, scope also has the property of overlaying onto linear order in English — as we saw in our examination of Larsonian Shells, if  $\beta$  is in the scope of  $\alpha$  then  $\alpha$  precedes  $\beta$ . Remember that I declared that scope is the semantic expression of c-command, so another way of putting this is:

(33) If  $\alpha$  c-commands  $\beta$ , then  $\alpha$  precedes  $\beta$ .

The fact that movement, and the other scope sensitive operations, express themselves linearly, then, derives from (33). And (33) is a function of the way Linearization has applied.

Interestingly, with respect to phrases, the English situation is hugely well attested, and the contrary is virtually nonexistent. Specifiers — at least in their guise as subjects — almost always canonically precede  $\bar{X}$ s. And movement of phrases is overwhelmingly to the left — NP Shift being the salient counterexample. In the case of heads, English is less representative. There are many languages that are head final — in fact in Greenberg's survey they are the most numerous. And we have seen at least one case that involved a plausible analysis of Head Movement moving something rightwards: this was the case of verbs in German moving to the right of *zu*. Still, as it stands, the examples of head movement going to the right are very rare relative to leftward moving head movement. German and Dutch have phenomena that are indicative of rightward movement,<sup>4</sup> But there are very few other examples. By contrast, there are very many analyses that posit leftward head movement: it's found in all the Germanic, Romance and Celtic languages, many Semitic languages, all of the Bantu languages that have been studied in depth (this is a huge language family), virtually all of the Baltic languages, probably the various Chinese dialects, a smattering of western African languages (e.g., Eʒe, Abe), and no doubt many others.

Kayne is impressed with this trend, and in Kayne (1994) he seeks to derive it. What he proposes is a linearization algorithm that will replace the linearization parameters (14). This algorithm is, in fact, something our grammar as it presently stands is in need of anyway. Like the phrase structure rules that they replace, the  $\bar{X}$  skeleton, linearization parameters and other principles that define a D-structure do not presently impose a linear order on the lexical items that fit into those phrase markers. The linearization parameters, note, impose a linearization on phrases, for the most part, and therefore don't express how the strings of words within those phrases will be linearized relative to the other strings in the sentence. In fact, all this grammar presently does is express groupings of lexical items in terms of sets. Thus, for instance, the string in (34) is given the structure in (35) (leaving out the detail of NumP, and the other functional projections within a DP).

(34) the coin under the chair

(35) {the, {{coin, {under, {{the, {chair}}}}}}}}

<sup>4</sup>We saw this phenomenon in the ordering of German infinitival marker *zu* and the infinitival verb in one of our homework assignments.

What is required is a procedure that projects the structure of the set in (35) onto a linear ordering that is expressed in (34). In particular, what is needed is something that specifies that by virtue of *the* preceding coin, under, the, chair — this what the English setting of the linearization parameters will specify — all of the lexical items in that set will follow *the*, and so on for each of the other word-set pairings in (35).

In what follows, let lower case letters (i.e., “x,” “y,” “z,” etc.) range over terminals (i.e. lexical items) and upper case letters (i.e., “X,” “Y,” “Z,” etc.) range over non-terminals. Then, if we understand “[ $\alpha$ ,  $\beta$ ]” to mean that  $\alpha$  precedes  $\beta$ , any linearization algorithm should have the consequences in (36).

- (36) a. If [ $x, y$ ] and [ $y, z$ ], then [ $x, z$ ]. (it’s Transitive).  
 b. For all distinct  $x$  and  $y$  in a phrase marker, then either [ $x, y$ ] or [ $y, x$ ]. (it’s total).  
 c. not ([ $x, y$ ] and [ $y, x$ ]) (it’s antisymmetric).

At least given the cases we have examined, this seems to be true of all sentences. We might imagine a linearization algorithm that meets these requirements, and also works with our Linearization parameters, to be something along the lines of:

- (37) Let  $X$  and  $Y$  be points on a phrase marker. If  $\{X, Y\}$  and [ $X, Y$ ], then [ $x, y$ ] for all  $x$  dominated by  $X$ , and all  $y$  dominated by  $Y$ .

This will achieve these constraints in a binary branching tree (i.e., a tree that has no more than two daughters per node), because the Linearization parameters will linearize every sister, and (37) will linearize all the terminals for each sister.

This would make the linearization of terminals in a phrase marker a virtual slave to the linearization parameters. The only thing (37) adds, beyond meeting the requirements in (36) is a prohibition against “crossing branches.” That is, it gives phrases an image in continuous strings of terminals. Otherwise, the way that terminals will be linearized follows entirely from the way that the linearization parameters linearize sisters. This makes the fact that c-command maps onto precede (or follows) a function of the linearization parameters.

Kayne’s idea is to reverse this. He proposes to make the way that sisters, and consequently terminals, are linearized a function of asymmetric c-command. First, he interprets the statements in (36) as constraints on a linearization,<sup>5</sup> and then defines a linearization as follows.

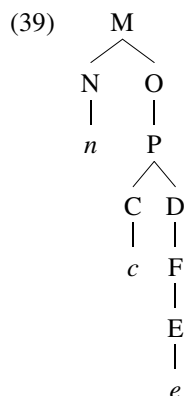
- (38) a.  $\alpha$  c-commands  $\beta$  iff every phrase dominating  $\alpha$  dominates  $\beta$  and  $\alpha$  does not contain  $\beta$ .  
 b.  $\alpha$  asymmetrically c-commands  $\beta$  iff  $\alpha$  c-commands  $\beta$ , and  $\beta$  does not c-command  $\alpha$ .  
 c.  $d(X) =_{def}$  the set of all terminals dominated by  $X$ .  
 d.  $d(\langle X, Y \rangle) =_{def}$  the set of all ordered pairs  $\langle x, y \rangle$  such that  $x$  is dominated by  $X$  and  $y$  is dominated by  $Y$ .  
 e.  $d(\langle X_1, Y_1 \rangle, \langle X_2, Y_2 \rangle, \dots, \langle X_n, Y_n \rangle) =_{def} d(\langle X_1, Y_1 \rangle) \cup d(\langle X_2, Y_2 \rangle) \dots \cup d(\langle X_n, Y_n \rangle)$ .

<sup>5</sup>Rather than, say, a consequence of the linearization procedure.

- f. Let  $A$  be the set of all  $\langle X, Y \rangle$  in some phrase marker such that  $X$  asymmetrically c-commands  $Y$ .
- g.  $d(A)$  is a linearization.

He calls (38g) the *Linear Correspondence Axiom* (LCA).

Let's see this work in an example:



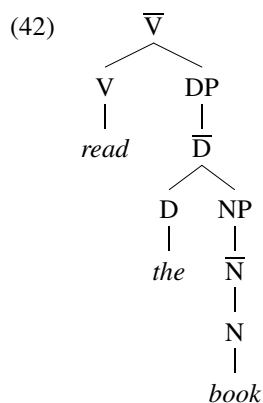
For this phrase marker we find the following values for  $A$  and  $d(A)$ :

- (40)
- a.  $A = \{ \langle N, P \rangle, \langle N, C \rangle, \langle N, D \rangle, \langle N, E \rangle, \langle N, F \rangle, \langle C, E \rangle, \langle C, F \rangle \}$
  - b.  $d(A) = \{ \langle n, c \rangle, \langle n, e \rangle, \langle c, e \rangle \}$

(40b) is total, transitive and antisymmetric — so it meets the constraints on a linearization. By “ $d(A)$  is a linearization,” we should understand this to mean that the ordered pairs can be understood as arranging their elements in a “precedes” or “follows” relationship. So, using our “[ ]” notation, (40) produces either (41a) or (41b).

- (41)
- a.  $\{ [n, c], [n, e], [c, e] \}$
  - b.  $\{ [c, n], [e, n], [e, c] \}$

A structure like that in (39) which might be found in nature is:

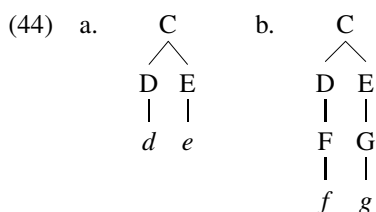


Which the LCA will linearize in one of the ways in (43).

- (43) a. read the book  
b. book the read

As can be seen in this example, then, terminals are linearized in a way that reflects the asymmetric c-command relations of the non-terminals that contain them. If  $\alpha$  is in a non-terminal that asymmetrically c-commands another non-terminal,  $\beta$ , then  $\alpha$  will either precede or follow every terminal in  $\beta$ .

The LCA has a partially welcome auxiliary result. It prevents any phrase from having daughters of the same projection level. To see this, consider the phrase markers in (44).



In both of these cases, the linearizations fail. As (45) shows, (44a) is (radically) not total, and (44b) is not antisymmetric.

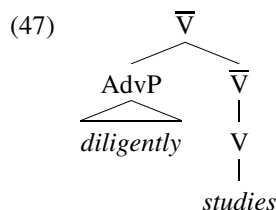
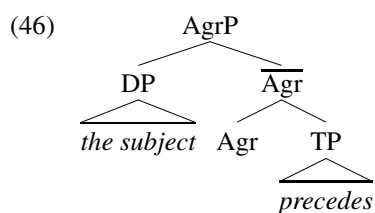
- (45) a. (44a):  
A = { }  
d(A) = { }  
  
b. (44b):  
A = {  $\langle D, G \rangle$ ,  $\langle E, F \rangle$  }  
d(A) = {  $\langle f, g \rangle$ ,  $\langle g, f \rangle$  }

Banning (44a) is welcome when C is not an  $X^0$ . This would prevent two heads from being sisters — that is, it will prevent a phrase from having two heads. Where we might want to allow (44a) is when one  $X^0$  has adjoined to another. Actually, we haven't seen many cases of this in which two terminals are involved — but we might want to consider instances of cliticization to involve such a structure. Let's come back to this.

Banning (44b) is also welcome, when C is the smallest  $\bar{X}$  in a projection. This will derive the fact that every phrase must have a head. When C is the smallest  $\bar{X}$ , then, the LCA has the welcome consequence of deriving that it have exactly one head, raising the possibility that this requirement could be removed from the  $\bar{X}$  Skeleton.

Structures like (44b) shouldn't always be banned, however. When C is a maximal projection, for instance, it is perfectly acceptable to have two phrasal daughters. Every AgrP, for instance, is fashioned in this way; and very many non-minimal  $\bar{X}$ s have this shape as well.





What unites these two environments is that one of the sisters is an  $\bar{X}$ . One way we could shore up the LCA for these cases, then, would be to make c-command sensitive to the projection level:

- (48)  $\alpha$  c-commands  $\beta$  iff  $\alpha$  is not an  $\bar{X}$ , and . . .

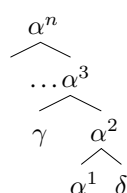
We'd want, I think, to search for a reason  $\bar{X}$  should not be able to c-command, however.

Kayne suggests something very close to (48), but his suggestion targets the case in (44a), when C is a head, as well. His idea takes as its starting position that there is no distinction between XPs and  $\bar{X}$ s. This decision will reek havoc with some of our principles (the modification rule, for instance), but let's be brave and forge ahead nonetheless. So, in place of (13), we have:

- (49) a.  $XP \rightarrow \{XP, YP\}$   
 b.  $XP \rightarrow \{X^0, (YP)\}$

And, over the structures formed by these, we define these terms:<sup>6</sup>

- (50) A CATEGORY is the set of  $\alpha$ 's in:



such that each  $\alpha^i$  is a

projection of  $\alpha^{i-1}$ , and  $\alpha^1$  is  $\alpha^0$ . Each  $\alpha^i$  is a SEGMENT of that category.

Kayne's method of allowing the phrase markers in (44) uses this segment/category distinction by way of making c-command sensitive to it, as in (51).<sup>7</sup>

<sup>6</sup>The notions "category" and "segment" come from May (1985), and were used extensively in Chomsky (1986a).

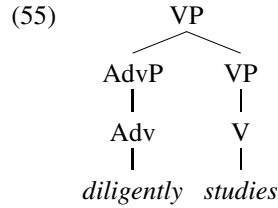
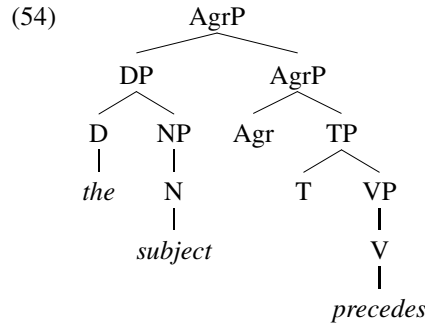
<sup>7</sup>Kayne has:

$\alpha$  c-commands  $\beta$  iff  $\alpha$  and  $\beta$  are categories and every category that *dominates*  $\alpha$  *dominates*  $\beta$ , and  $\alpha$  excludes  $\beta$ .

but this makes it difficult to define "dominates" for categories.

- (51)  $\alpha$  c-commands  $\beta$  iff  $\alpha$  and  $\beta$  are categories and every category that contains  $\alpha$  contains  $\beta$ , and  $\alpha$  excludes  $\beta$ .  
(Kayne 1994, (3), p. 16)
- (52) A category  $\alpha$  EXCLUDES  $\beta$  iff no segment of  $\alpha$  dominates  $\beta$ .
- (53)  $\alpha$  CONTAINS  $\beta$  if every segment of  $\alpha$  dominates  $\beta$ .

Let's consider how this will work in the problematic situations in (46) and (47), which would now have the representations in (54) and (55).



When we compute out the A and d(A) of these phrase markers, we find:

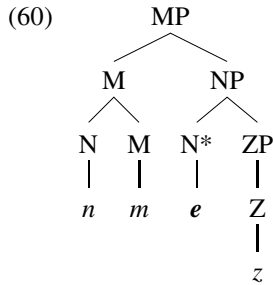
- (56) For (54):  
 $A = \{ \langle D, N \rangle, \langle DP, AgrP \rangle, \langle DP, Agr \rangle, \langle DP, TP \rangle, \langle DP, T \rangle, \langle DP, VP \rangle, \langle DP, V \rangle, \langle Agr, T \rangle, \langle Agr, VP \rangle, \langle Agr, V \rangle, \langle T, V \rangle \}$   
 $d(A) = \{ \langle the, subject \rangle, \langle the, precedes \rangle, \langle subject, precedes \rangle \}$
- (57) For (55):  
 $A = \{ \langle AdvP, V \rangle, \langle AdvP, VP \rangle \}$   
 $d(A) = \{ \langle diligently, studies \rangle \}$

These are total, transitive and antisymmetric linearizations, and will yield one of the outcomes given in (58) on the next page.

Notice that in (54), AgrP doesn't c-command the subject, and consequently doesn't asymmetrically c-command any of the material in the subject. This is because AgrP does not exclude the subject. As a result, the contents of the subject will be asymmetrically linearized with respect to the contents of TP. The same situation holds in (55) with respect to VP and the AdvP. Because VP does not exclude AdvP, it will not c-command it, and this means that the material in AdvP will get linearized relative to the rest of the VP only once.

- (58) For (56):  
 $\{[the, subject], [the, precedes], [subject, precedes]\}$   
 $\{[subject, the], [precedes, the], [precedes, subject]\}$
- (59) For (57):  
 $\{[diligently, studies]\}$   
 $\{[studies, diligently]\}$

The same effects described here for phrases carry over to heads. If one head adjoins to another, as in (60), then they will be linearized in an antisymmetric way.



The A and d(A) for (60) are as in (61).

- (61)  $A = \{ \langle N, M \rangle, \langle N, N^* \rangle, \langle N, ZP \rangle, \langle N, Z \rangle, \langle M, N^* \rangle, \langle M, ZP \rangle, \langle M, Z \rangle, \langle N^*, Z \rangle \}$   
 $d(A) = \{ \langle n, m \rangle, \langle n, e \rangle, \langle n, z \rangle, \langle m, e \rangle, \langle m, z \rangle \}$

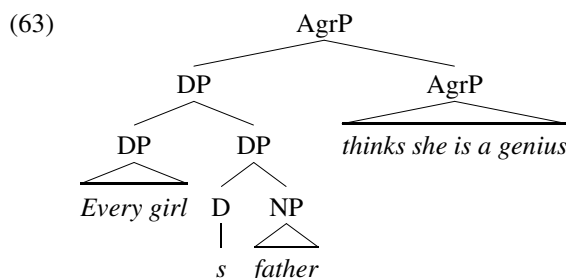
This linearization is also complete, transitive and antisymmetric. Notice in particular that because NP and N c-command each other, they are not in A, thus avoiding a symmetric linearization of the terminals they dominate.

Modifying c-command as in (51) to solve this problem is not innocent. It has consequences in a variety of areas. For example, it predicts that a DP in the a Specifier position of some phrase should c-command outside that phrase. Kayne argues that this is consequence is supported by the availability of binding in (62).

- (62) Every girl<sub>1</sub>'s father thinks she<sub>1</sub> is a genius.

Recall that pronouns can get bound to quantifiers only if they are in the scope of those quantifiers, and c-command is the syntactic expression of scope. Thus, the fact that *she* can be bound to *every girl*, suggests that *every girl* c-commands *she*. On (51), and the  $\bar{X}$ -free phrase markers we are examining, this indeed obtains, as (63) on the following page indicates.

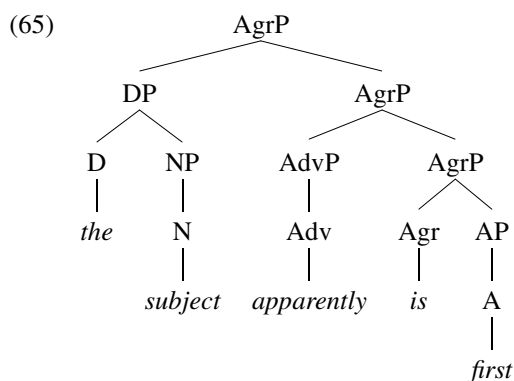
But there are other data that suggest that this is not a consequence we desire for c-command. Neither of the other two diagnostics of scope we have relied on suggest that the genitive DP is capable of c-commanding out of the DP it is embedded within. There is no disjoint reference effect in (64a), nor can the anaphor be bound in (64b).



- (64) a. Her<sub>1</sub> father likes Jill<sub>1</sub>.  
 b. \* Her<sub>1</sub> father likes herself<sub>1</sub>.

Without knowing how to resolve these conflicting facts, it's hard to know whether (51) is supported or undermined by these examples.

Another, perhaps unwelcome, consequence of this solution is that no phrase can have more than one Specifier or Adjunct. To see this, consider the scenario in (65).<sup>8</sup>



The A and d(A) of (65) are given in (66) below. This violates antisymmetry, by virtue

- (66)  $A = \{ \langle D, N \rangle, \langle DP, AgrP \rangle, \langle DP, Adv \rangle, \langle DP, Agr \rangle, \langle DP, AP \rangle, \langle DP, A \rangle, \langle AdvP, D \rangle, \langle AdvP, NP \rangle, \langle AdvP, N \rangle, \langle AdvP, AgrP \rangle, \langle AdvP, Agr \rangle, \langle AdvP, AP \rangle, \langle AdvP, A \rangle, \langle Agr, A \rangle \}$

$d(A) = \{ \langle the, subject \rangle, \langle the, apparently \rangle, \langle subject, apparently \rangle, \langle apparently, the \rangle, \langle apparently, subject \rangle, \langle apparently, is \rangle, \langle apparently, first \rangle, \langle is, first \rangle \}$

of  $\langle the, apparently \rangle$  and  $\langle apparently, the \rangle$ , among others. In essence, then, this way

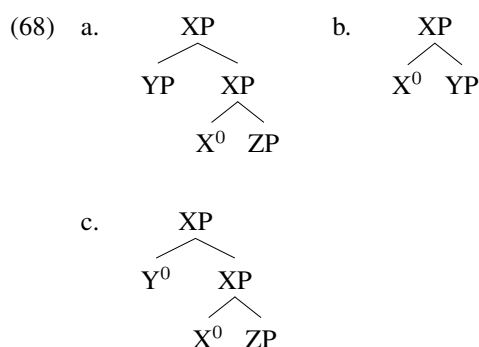
<sup>8</sup>As in many of the phrase markers in this chapter, I've ruthlessly truncated structure needless to the point at hand.

of implementing the LCA robs the recursivity of our phrase building rules from having much utility. It is no longer possible to introduce an indefinite number of terms into a phrase by way of recursivity. Some other way of capturing these facts will have to be found.

The LCA, then, has far reaching consequences for  $\bar{X}$  Theory. In fact, it derives much of the information that is coded into the  $\bar{X}$  Skeleton — so much so, that it becomes possible to entertain reducing the  $\bar{X}$  Skeleton to (67).

$$(67) \quad \alpha P \rightarrow \{\alpha', (\beta)\}, \text{ where } \alpha' \text{ is } X^0 \text{ or } XP.$$

This, together with the LCA and the definition of category and segment, will allow just the shapes in (68).



The first two of these are well attested, of course; but (68c) needs to be blocked. Kayne suggests a principle which insists that Specifiers be phrasal. In any case, if some way of blocking (68c) can be found, then the LCA allows us to dispense with most of  $\bar{X}$  Theory. The cost is seriously attenuated phrase markers, ones that are too small to accommodate the parses we've given up to now.

But what of the goal set out at the outset? How does the LCA derive the fact that “up” translates as “precedes” in so many cases? Kayne suggests deriving this from the LCA and the stipulation that  $d(A)$  always translates into a “precedes” linearization:

$$(69) \quad \langle \alpha, \beta \rangle =_{def} [\alpha, \beta]$$

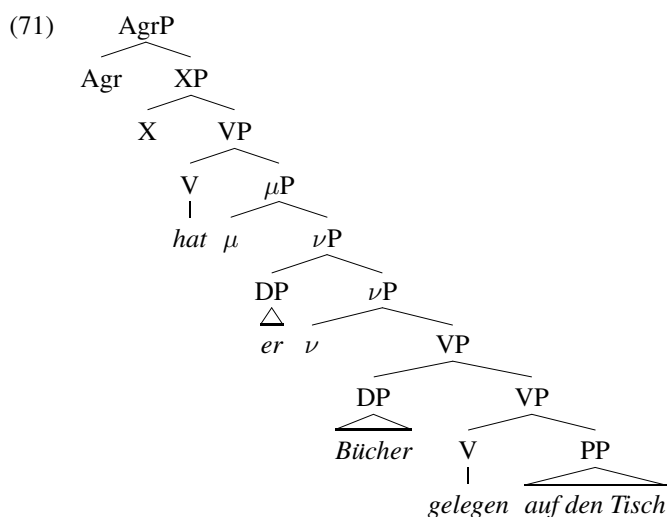
Once the force that blocks (68c) is identified, only (68a,b) are permitted and if (69) holds, then these will be linearized with the head or Specifier preceding the rest. Since movement is always to a head or Specifier, this will derive the consequence that phrasal and head movement tends to be leftwards. Similarly, if the other scope sensitive phenomena involve relations between phrases — which means that these will always be relations between a phrase in a Specifier position and a phrase within that Specifier's sister — then these phenomena too will end up putting the c-commanding item to the left of the c-commanded term.

As Kayne notes, this method of deriving the left-to-right relation that scope seems to map onto has the rather radical consequence of allowing only underlying representations in which Specifiers and heads both precede complements. So far as I know, there are not many situations which threaten the Specifier/complement part of this prediction; languages do seem to frequently line up subjects, and other canonical

Specifiers, so that they precede the rest of the material in the phrase they are Specifiers for. But the prediction for verb and complement runs into scores of apparent counterexamples. We've already encountered one language which is apparently incompatible with this prediction: German.

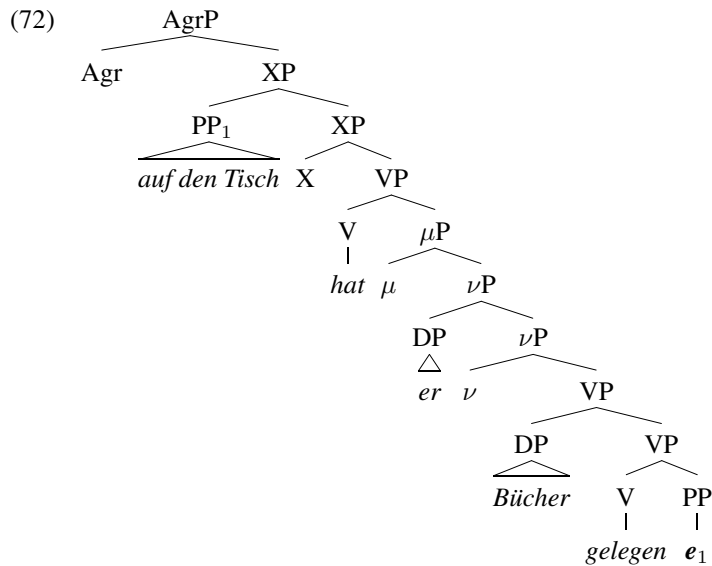
Of course, Kayne is aware of German, and the many other languages which appear to be head final. He suggests that these languages are, indeed, head initial underlyingly, but that a suite of movement operations systematically obscures this at S-structure. To get an image of how this might work, let's reconsider German word-order from this perspective. We might give the sentence whose surface form is (70) would have an underlying representation like that in (71).<sup>9</sup>

- (70) ... weil Er auf den Tisch Bücher gelegen hat  
 ...since he on the table books put has  
 ... 'since he has put books on the table'

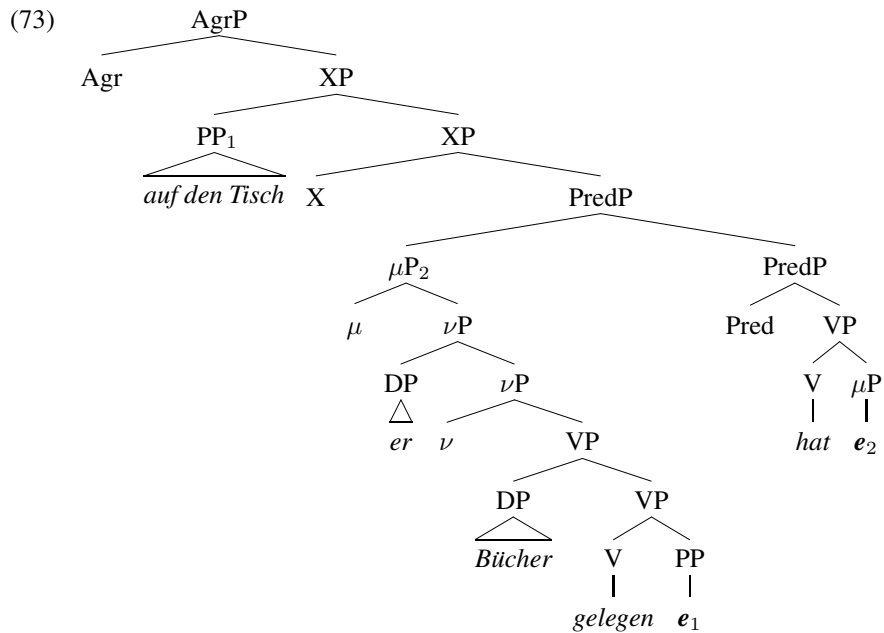


Imagine that Scrambling is obligatory, and relocates *auf den Tisch* to XP, as indicated in (72).

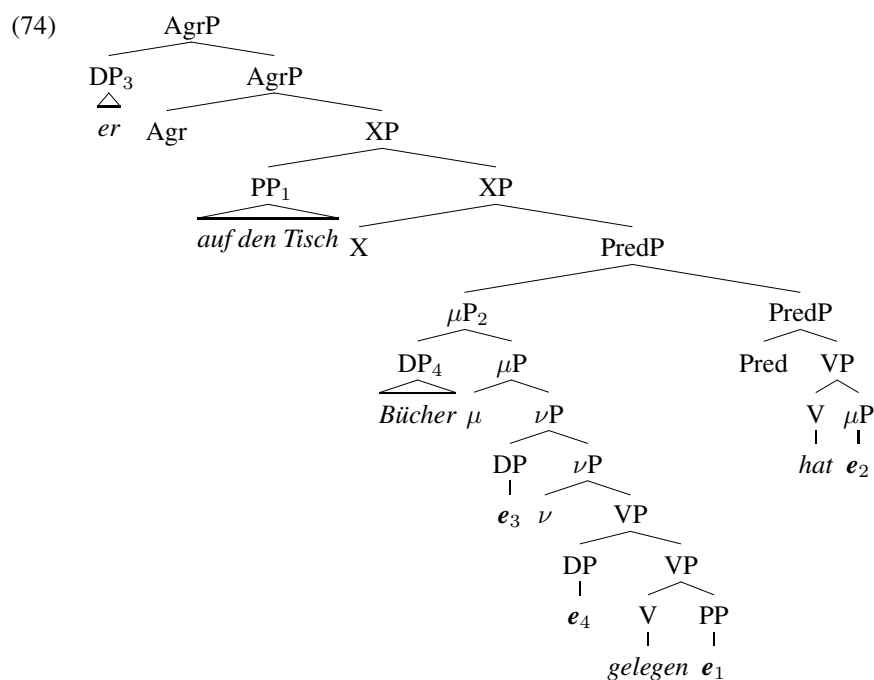
<sup>9</sup>I'll ignore in these parses TP — it's possible that the XP in these phrase markers could, in fact, be TP.



Making Scrambling obligatory, then, has the potential of bringing all the complements to the left of the verb that c-selects them. Now what's necessary is to get XP to precede *hat*. This can be achieved by moving XP leftwards. Zwart (1997) argues for something along these lines, and suggests that there is a functional phrase — “Predicate Phrase” — which attracts  $\mu$ P to its Specifier. If Predicate Phrase is immediately above all the VPs, then movement into it will yield something like (73).

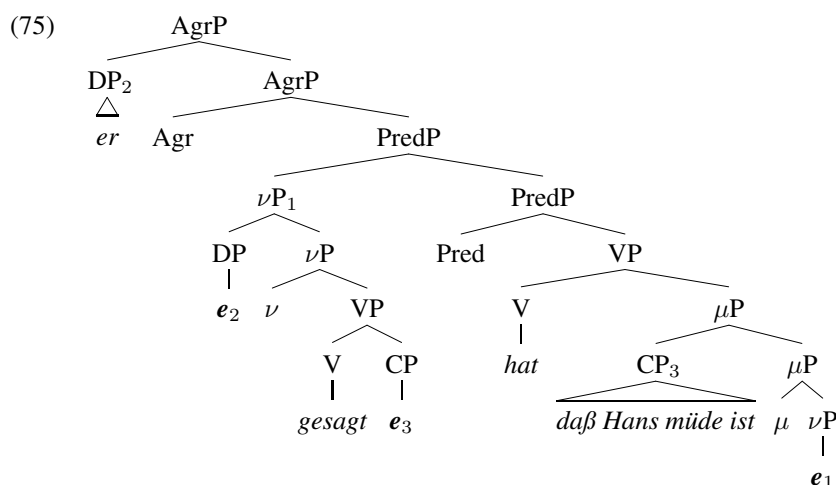


Once the two DPs have moved into their Case marked positions, we'll get the correct word order, shown in (74) below.

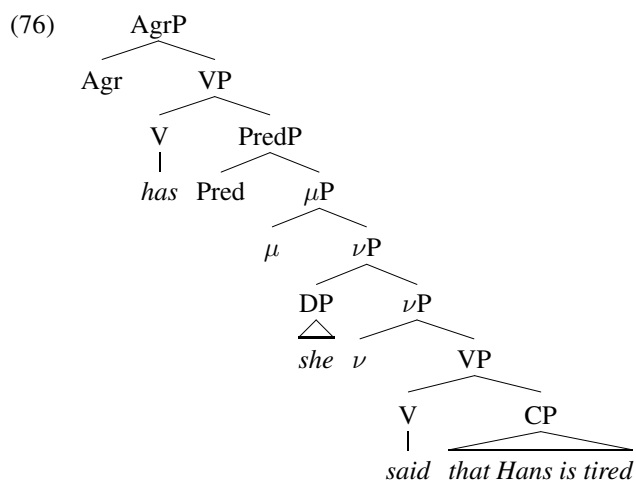


Cases where a German verb has a clausal complement which, as you'll recall, surfaces at the right edge of the sentence, might have a derivation in which the phrase that moves to Specifier of PredP is smaller than  $\mu$ P. This would allow the clausal complement to move out of the phrase that moves into Specifier of PredP, and consequently be stranded behind. The surface parse for such a case might look as in (75).

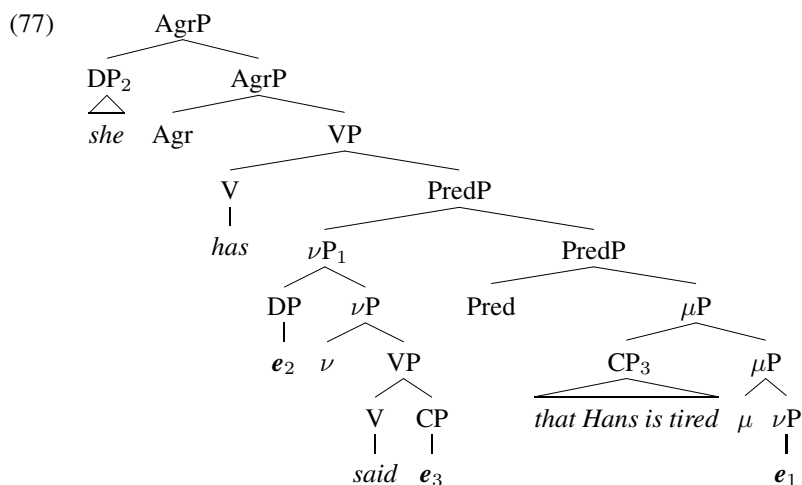




These derivations may look odd, but they would have to be mirrored in English as well in contexts where NP Shift arises. Because the Kayne rigs the LCA so that it prevents rightward movement, the phenomena that “NP Shift” describes cannot be characterized in the terms that Ross did. On the other hand, Larson’s treatment of the phenomena is fully consistent with Kayne’s program. Recall that on Larson’s view, NP Shift arises when a  $\bar{V}$  moves leftward past the phrase that is “NP Shifted.” There are a variety of possible ways to implement that idea here, but one which would make English NP Shift constructions look rather like what we have just reviewed for German would give English sentences an underlying representation like that in (76).



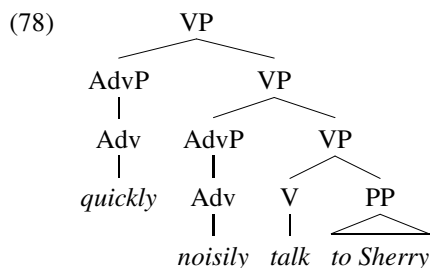
This is very similar to the underlying representation given to German for the parallel sort of case; the differences are that there is no XP — so there is no position above the auxiliary verb to which things move — and PredP is below the auxiliary verb. As a consequence, the steps that produced the German form in (75) will create in English a form like (77).



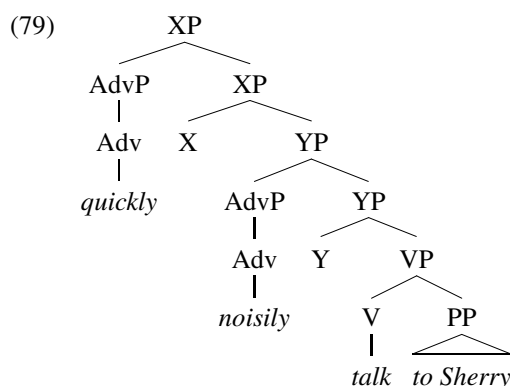
It is still possible, in other words, to give a parallel syntax to the German and English instances of NP Shift.

In general, then, Kayne's claim is that all verb final configurations come about as a consequence of movements that distort the underlying head initial relations. On this schema, then, the head initial/head final linearization parameter emerges as differences in the conditions that are forcing movement of internal arguments and VPs. To show that this is the correct way of characterizing the head-initial/head-final parameter is a huge task, and it remains quite unfinished. There have been, however, rather surprisingly good progress made in the Germanic languages — there are some puzzles which find an explanation on this view. To date, much less success has been garnered in rigidly head final languages, such as Korean and Japanese, however. This remains an open project.

Another difficulty for Kayne's implementation of the LCA is, as we've seen, that it destroys the ability to recursively introduce modifiers into phrases. It would prevent VPs of the following sort, for instance:

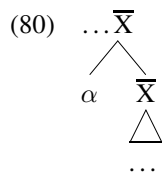


Kayne's suggestion about these sorts of cases is that the recursive devices of the rule building phrase markers is simply used differently. Rather than introducing phrases without end, to which modifiers are adjoined, it introduces heads without end, as in (79) on the next page. On this view, then, it is possible for the phrase marker building procedure in (67) to insert contentless heads, the "X" and "Y" in (79), in order to make room for modifiers. A close alternative is to imagine that the "X" and "Y" in (79)



dominate something that has semantic content, but which has no phonological reflex (in English). Cinque (1999) has proposed such a position, based on a cross-linguistic examination of the syntax and morphology of adverbial modification. He suggests, in particular, that there are phrases whose heads select, in a way, adverbs of certain semantic types. Sometimes these heads have a phonological reflex — sometimes they manifest themselves as inflections on nearby verbs, for instance — and sometimes they don't.

Cinque's proposal provides a way for capturing some of the information that is lost by abandoning the  $XP/\bar{X}$  distinction. In particular, recall that we've restricted to position of modifiers to positions that have  $\bar{X}$ s as sisters and mothers. The modification rule forces the modification relation to obtain only in the position  $\alpha$  stands in in (80).



Part of the role of this rule was to prevent expletives from being in these positions. When this is combined with the Projection Principle, recall, the result is that expletives are possible only in Specifier positions, and this accords with what we discover empirically. But this rule also prevented modifiers from themselves being in Specifier positions. Without the  $XP/\bar{X}$  distinction, it is no longer possible to capture these facts with such a rule. But if modifiers are, like adverbs, only possible in positions which are adjoined to Cinque's hidden phrases, then we can express these restrictions in these terms. Imagine that the heads of these phrases require that the phrases that are adjoined to them have the semantic content that modifiers do, and that they bring together this semantic content in the appropriate way with the VP, NP, or other phrase that is being modified. This will block expletives from these positions, since they have no semantic content. Similarly, if we imagine that the semantic content of modifiers cannot combine with the phrases they are modifying without the help of Cinque's phrases, then modifiers will be blocked from appearing anywhere but in these phrases; this will prevent

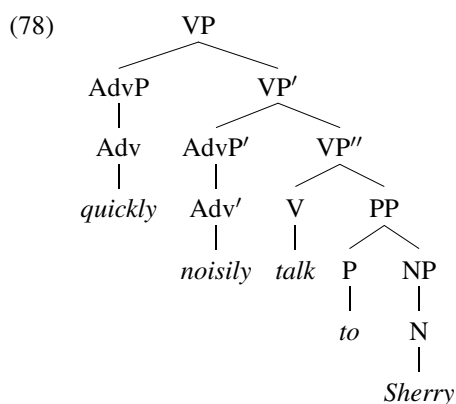
them from appearing in Specifier positions, then.<sup>10</sup>

Kayne's proposal for capturing the correlation between precedence and c-command, then, require quite far-reaching changes to the grammar we've been developing. There is presently considerable controversy about the likely success of the changes that would seem to be required. It's useful, perhaps, to observe that some of the consequences of Kayne's proposal are a function of his particular implementation. In fact, it is possible to keep the LCA, and its consequences for deriving the "precedes=higher" equation, but not run into all of the controversial results for the shape of phrase markers that Kayne discusses. Many of these consequences derive from the particular way in which Kayne suggested for overcoming the problem posed by the configurations in (44). (That is, for the fact that on a standard definition of c-command, the LCA prevents a term from having two daughters that are both phrases or heads.) His solution was to tinker with the definition of c-command in such a way that the LCA does not prevent Specifiers from having a phrasal sister, for instance.

But the definition of c-command he chooses is not the only conceivable one. Even slight differences in how we define c-command to overcome this problem will have very different consequences for the shape of phrase markers. Suppose, for instance, that we design c-command as follows:

- (81)  $\alpha$  c-commands  $\beta$  iff every phrase that dominates  $\alpha$  dominates  $\beta$ , and no segment of the category  $\alpha$  belongs to dominates  $\beta$ .

Unlike Kayne's definition of c-command, this will allow multiple adjunction. To see this, reconsider (78), repeated below (with some helpful diacritics added).



Under Kayne's definition,  $\text{AdvP}'$  asymmetrically c-commands the  $\text{Adv}$  dominating *quickly*, and  $\text{AdvP}$  asymmetrically c-commands the  $\text{Adv}$  dominating *noisily*. As a result, the linearization that results is not antisymmetric, and fails. But under (81) this is no longer true; in particular,  $\text{AdvP}'$  does not c-command the  $\text{Adv}$  dominating *quickly*. As a result, the A and d(A) of (78) are as in (82), and satisfy the complete, transitive and antisymmetric conditions on a linearization.

<sup>10</sup>For a semantic treatment of Cinque's phrases that has this outcome, see Morzycki (2002).

- (82) a.  $A = \{ \langle \text{AdvP}, \text{VP}' \rangle, \langle \text{AdvP}, \text{AdvP}' \rangle, \langle \text{AdvP}, \text{Adv}' \rangle, \langle \text{AdvP}, \text{VP}'' \rangle, \langle \text{AdvP}, \text{V} \rangle, \langle \text{AdvP}, \text{PP} \rangle, \langle \text{AdvP}, \text{P} \rangle, \langle \text{AdvP}, \text{NP} \rangle, \langle \text{AdvP}, \text{N} \rangle, \langle \text{AdvP}', \text{VP}'' \rangle, \langle \text{AdvP}', \text{V} \rangle, \langle \text{AdvP}', \text{PP} \rangle, \langle \text{AdvP}', \text{P} \rangle, \langle \text{AdvP}', \text{NP} \rangle, \langle \text{AdvP}', \text{N} \rangle, \langle \text{V}, \text{P} \rangle, \langle \text{V}, \text{NP} \rangle, \langle \text{V}, \text{N} \rangle \}$
- b.  $d(A) = \{ \langle \text{quickly}, \text{noisily} \rangle, \langle \text{quickly}, \text{talk} \rangle, \langle \text{quickly}, \text{to} \rangle, \langle \text{quickly}, \text{Sherry} \rangle, \langle \text{noisily}, \text{talk} \rangle, \langle \text{noisily}, \text{to} \rangle, \langle \text{noisily}, \text{Sherry} \rangle, \langle \text{talk}, \text{to} \rangle, \langle \text{talk}, \text{Sherry} \rangle, \langle \text{to}, \text{Sherry} \rangle \}$

Another detail of Kayne's implementation concerns interpreting the ordered pairs that  $d(A)$  produces as translating into a uniform "precedes" linearization. This has the desirable effect of predicting that Specifiers will always precede their sisters — this does seem to be a strong cross-linguistic trend — but it also has the more controversial effect of forcing all languages to be head initial. It's possible, however, that how  $d(A)$  is interpreted could vary as a function of the projection level of the terms it is ordering. For instance, we might imagine a scheme like that in (83) which would permit right headed languages that have initial Specifiers.

- (83)  $\langle \alpha^0, \beta \rangle =_{def.} [\beta, \alpha]$  (Korean/Japanese)  
 $\langle \alpha^0, \beta \rangle =_{def.} [\alpha, \beta]$  (English)  
 $\langle \alpha P, \beta \rangle =_{def.} [\alpha, \beta]$  (Universal?)

The theory Kayne provides us, then, has some flexibility. It is possible to preserve his account for the coincidence of precedes=c-commands, and even derive the simplification of the  $\bar{X}$  Skeleton, without embracing all of the consequences for phrase markers that are discussed in Kayne (1994).

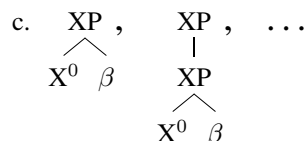
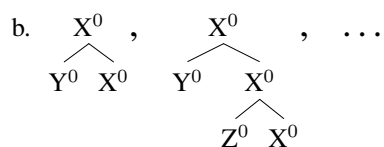
### 9.2.2 The "Bare Phrase Structure" proposals

Kayne's project gives a way of explaining the coincidence of precedence and c-command; but it also, as we have seen, reduces the  $\bar{X}$  Skeleton to something like (84).

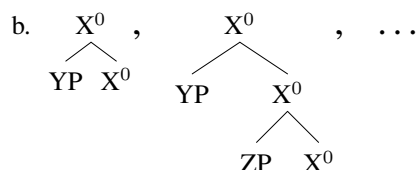
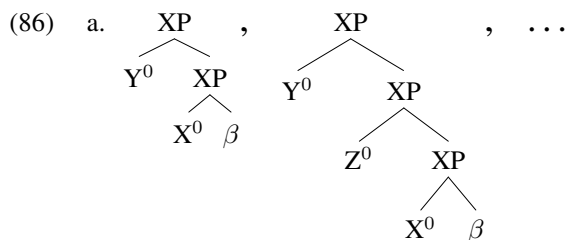
- (84)  $\alpha P \rightarrow \{ \alpha', (\beta) \}$ , where  $\alpha'$  is  $\alpha^0$  or  $\alpha P$ .

The LCA, along with some decisions about how to define ccommand, derives much of the information that was originally in the  $\bar{X}$  Skeleton, and, to a certain extent, goes beyond this. If we adopt the definition of c-command in (81), rather than Kayne's, the LCA and (84) permit only the shapes in (85), where " $\beta$ " is, in each case, optional.

- (85) a.  $\begin{array}{c} \text{XP} \\ \swarrow \searrow \\ \text{YP} \quad \text{XP} \\ \quad \swarrow \searrow \\ \quad \text{X}^0 \quad \beta \end{array}, \begin{array}{c} \text{XP} \\ \swarrow \searrow \\ \text{YP} \quad \text{XP} \\ \quad \swarrow \searrow \\ \quad \text{ZP} \quad \text{XP} \\ \quad \quad \swarrow \searrow \\ \quad \quad \text{X}^0 \quad \beta \end{array}, \dots$



If we let (84) have the role that the  $\bar{X}$  Skeleton had in defining D-structures, then D-structures will be confined to the forms in (85). Although Kayne let the LCA hold of every phrase marker in a derivation, but if (84) is restricted to just producing D-structures, then its influence on constraining phrase markers will be lost. As a consequence, there are possibilities available to S-structures that aren't available to D-structures. These are in (86), and, at least so far as we've seen up to now, should be prevented.



Since these structures would be formed by a movement operation adjoining a head to a phrase (in (86a)) or a phrase to a head (in (86b)), the Likes-Attract-Likes constraint will prevent them.

It might be desirable, however, to think of the Likes-Attract-Likes constraint itself as deriving from constraints on phrase markers. One might see it, for instance, as a kind of “faithfulness” constraint: a constraint that prevents movement from creating phrase markers that could not be formed underlyingly. Indeed, the earliest version of the Likes-Attract-Likes constraint said exactly this. Emonds (1972) proposed a “Structure Preservation” constraint, a constraint which prohibits structures from being created in the derivation that are not capable of being produced by the phrase structure rules for the language. If this way of thinking about the Likes-Attract-Likes constraint is on the right track, we should find a way of letting the influence that (84) has on D-structures be present throughout the derivation.

This is one of the consequences of the proposals in Chomsky's "Bare Phrase Structure" paper. Chomsky's aim in this paper is to explore reducing the  $\bar{X}$  Skeleton beyond even what is achieved in Kayne's work. In particular, he seeks to reduce, or at least reorganize, some of the information in  $\bar{X}$  theory that Kayne presupposes.

If we examine the aspects of  $\bar{X}$  theory that Kayne's LCA requires, we'll see that they are based on the following:

- (87) a. The notion of *projection*, e.g., an  $X^0$  projects an XP.
- b. The *segment/category* distinction (see (50)).
- c. A distinction between heads ( $X^0$ ) and phrases (XP).
- d. The notion of *sets*, or *dominates*, that (84) expresses, i.e., that one, or two, terms can form another.

The heart of Chomsky's idea is to express the distinction in (87c) in terms of (87d). This, in turn, will require a new formulation of *projection*.

He begins by following ? and calling into question the distinction between  $X^0$  and the terminal item that  $X^0$  dominates. Indeed, this is something of a holdover from phrase structure rules. In grammars with phrase structure rules — like the one we began with — these rules express, in templatic form, the possible arrangements of categories permitted in the language. Particular instantiations of these arrangements are then expressed by way of inserting words into the templatic form that the phrase structure rules provide. When we dispensed with phrase structure rules, however, we removed from  $\bar{X}$  theory all category specific information; this got factored into c-selection requirements, constraints on modification and principles such as the Case filter. There is no longer any need, therefore, to think of  $\bar{X}$  theory as providing a template into which lexical insertion sticks terminals. Chomsky suggests that we simply equate  $X^0$  with the terminal it would dominate.  $\bar{X}$ s, then, are simply terminals. Phrases, he suggests, are terminals that have undergone the set forming operation that (87d) makes reference to. Moreover, following an idea in Muysken (1982), he suggests that the set forming operation be prevented from applying to a single terminal. As a consequence, "heads" will always be single terminals, and "phrases" will always be pairs of terminals grouped into a set. In place of (84), then, Chomsky proposes a set forming operation that is defined for two terminals. He calls this operation MERGE.

- (88)  $\text{MERGE}(x)(y) =_{\text{def}} \{x, y\}$ , where  $x$  and  $y$  are terminals or  $\{x, y\}$ .

This would give to a string like (89) a representation like that in (90) on the following page.

- (89) the fiery wreckage

One thing left out of (90) that seems necessary is that  $\{\text{fiery}, \text{wreckage}\}$  is a projection of *wreckage*, and that  $\{\text{the}, \{\text{fiery}, \text{wreckage}\}\}$  is a projection of *the*. Chomsky suggests that we think of "projection" as a labeling of the sets formed by Merge. To represent that  $\{\text{fiery}, \text{wreckage}\}$  is a projection of *wreckage*, then, can be done by letting  $\{\text{fiery}, \text{wreckage}\}$  have *wreckage* as its label. Chomsky proposes the unfortunate means of representing this in (91).

- (91)  $\{\text{wreckage}, \{\text{fiery}, \text{wreckage}\}\}$

$$(90) \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{the} \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{fiery} \quad \text{wreckage} \end{array} \end{array} \equiv \{the \{fiery, wreckage\}\}$$

This doesn't allow us to distinguish the result of MERGE from the result of labeling, however, so I suggest we adopt instead the notational convention in (92).

$$(92) \quad \{wreckage \text{ fiery, wreckage} \}$$

The notion of *projection*, then, can be expressed in terms of a labeling function. Modifying somewhat what Chomsky proposes, we can define this with (93).

$$(93) \quad \text{PROJECT}(\{x,y\}) =_{def} \{\alpha x,y\}, \alpha \text{ the LABEL}(\{x,y\}),$$

$$\text{where } \begin{cases} \alpha = x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha = \text{LABEL}(x) \text{ or LABEL}(y) & \text{for } x, y \text{ phrases} \end{cases}$$

Now, the representation that (89) receives is (94).

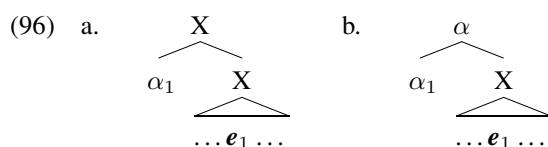
$$(94) \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{the} \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{the} \quad \text{wreckage} \\ \diagup \quad \diagdown \\ \text{fiery} \quad \text{wreckage} \end{array} \end{array} \equiv \{the \text{ the}, \{wreckage \text{ fiery, wreckage} \} \}$$

Of course, as far as PROJECT is concerned, (89) could also get the representation in (95).

$$(95) \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{fiery} \quad \begin{array}{c} \diagup \quad \diagdown \\ \text{the} \quad \text{fiery} \\ \diagup \quad \diagdown \\ \text{fiery} \quad \text{wreckage} \end{array} \end{array} \equiv \{fiery \text{ the}, \{fiery \text{ fiery, wreckage} \} \}$$

This representation, and others like it, will have to be blocked in the same way that such representations were blocked under the old system. That is, the category-specific information that is present in modification and c-selection blocks such cases. In (95), for instance, because *fiery* is an adjective, it will not meet the c-selection requirements of *the*. With regard to D-structures, then, PROJECT can be left free to choose the label from the elements Merge has combined, the correct choice arising as a function of c-selection and modification. Something different is needed when PROJECT is determining the label for a phrase that has been formed by movement. In this scenario, we don't want the phrase that has moved to project, as in (96b), but instead the other phrase should, as in (96a).

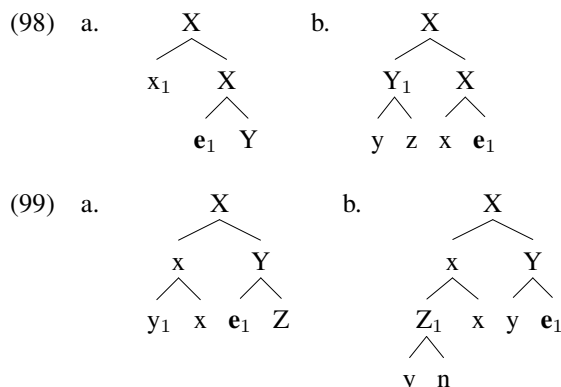




What we look for is something that derives (97).

(97) A moved term cannot PROJECT.

These definitions allow for an approach to the Likes-Attract-Likes constraint — or the Structure Preservation constraint — that wasn't available before. Chomsky notes that defining the head/phrase contrast in terms of being a terminal or a set of terminals, when combined with these definitions of PROJECT, sometimes causes the status of a moved term to change through the derivation. To see this, consider the scenarios in (98) and (99).



In (98) are scenarios in which something has moved and Merged with a term that started out as a phrase; in (99) are the same situations but where the moved item Merges with something that started out as a head. In the a-cases, the moved term starts out as a head; and in the b-cases, the moved item starts out as a phrase. What Chomsky observes is that in the a-cases, the item that is moving has changed its status with respect to PROJECT. The moved item starts out in a position in which it is the smallest item sharing a label, but it ends up in a position in which it is the largest item sharing a label. Perhaps this “Projection status” matters; he proposes (100).

(100) An item cannot move from one position to another if it has a different project status in those two positions.

With respect to movement that targets a phrase as landing site, this has the right consequence; it permits (98b) and blocks (98a). But with respect to movement targeting a head as landing site, this badly misfires, blocking the expected outcome in (99a) and allowing the unexpected one in (99b).

In fact, the scenarios in (99) bring into focus a difficulty for the more general thesis that the head/phrase distinction can be expressed solely in terms of Merge. Recall that one of the constraints we have relied on to guide derivations is the Word Criterion,

which prevents an  $X^0$  from dominating more than (or less than) one word. With the loss of  $X^0$ , this principle is lost as well. What the Word Criterion expresses is that the scenarios in (99) have a special status with respect to morphology. If we are to find a way of reexpressing the Word Criterion in the terms made available by “Bare Phrase Structure,” we might say something like (101).

- (101) THE NEW WORD CRITERION  
A terminal can project at most one non-word.

This allows Merge to put a terminal together with another term, and for a non-word phrase just once. Thus, for instance, a terminal can combine with its complement to form a phrase; but from then on out, anytime that terminal projects, the resulting phrase will have to be a word. Chomsky suggests that something along these lines — he doesn’t propose (101) — “overrides” (100) in scenarios like (99).

How do these proposals interact with Kayne’s LCA? There are incompatibilities, and Chomsky sketches some ways to smooth them. Chomsky’s scheme requires that the first terminals that are put into a phrase always take the form in (102).

- (102) 
$$\begin{array}{c} x \\ \diagup \quad \diagdown \\ x \quad y \end{array}$$

The terminals in this structure c-command each other, and as a consequence the  $d(A)$  of this tree will include neither of them. If  $d(A)$  is a linearization, then it will be dramatically non-total in this case, and this phrase marker should fail. To allow for phrases of this sort, Chomsky suggests changing the criterion of Totality so that it holds only of phonetically interpreted terminals:

- (103) For all distinct, phonetically overt terminals,  $x$  and  $y$ , in a phrase marker, either  $[x,y]$  or  $[y,x]$ .

This will allow (102) just in case one of  $x$  or  $y$  is phonetically null. One way this can be achieved, at least if we abandon Kayne’s view that the LCA must hold of every phrase marker in a derivation, is by moving one of  $x$  or  $y$ . Because the trace that will occupy the position vacated is phonetically null, the LCA will be able to interpret the result and still be Total. Another way this might be achieved is by letting  $x$  or  $y$  be a phonetically null terminal to begin with as; for instance, the functional heads holding agreement, tense or number features have this status.

Consider next the scenario in (104).

- (104) 
$$\begin{array}{c} x \\ \diagup \quad \diagdown \\ y \quad x \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ y \quad z \quad x \quad n \end{array}$$

This is the scenario that Kayne labored so mightily with. He, recall, wanted to prevent  $\{_x x,n\}$  from c-commanding  $\{_y y,z\}$ , as this will, wrongly, force  $d(A)$  to put the terminals  $x,n$  into a symmetric linearization with  $y,z$ . Kayne’s solution, recall, is to invoke the segment/category distinction, and to restrict c-command to categories. Because  $\{_x$

$x, n\}$  is not a category, then, it will not c-command anything, and because the category made up of  $\{x, x, n\}$  and  $\{x, \{y, y, z\}, \{x, x, n\}\}$  does not exclude  $\{y, y, z\}$ , it will not c-command  $\{y, y, z\}$  either.

Chomsky proposes a similar strategy for (104), though he wishes to preserve the distinction between “adjuncts” and “specifiers” that Kayne dispenses with. The representations we have considered up to now are those in which Chomsky proposes we understand Specifiers to be defined in. Thus, for instance, in (104),  $\{y, y, z\}$  is a Specifier. For the LCA to apply with the outcome Kayne desires for these cases, Chomsky suggests restricting c-command with (105).<sup>11</sup>

(105) Only the minimal and maximal projections can c-command.

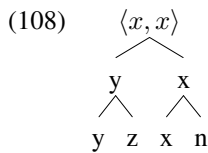
Adjunction structures, he suggests, differ from (104) by virtue of the label on the phrase that Merge forms. There is, in other words, other outcomes for PROJECT, as indicated in (106).

(106)  $\text{PROJECT}(\{x, y\}) =_{\text{def}} \{\alpha x, y\}, \alpha \text{ the LABEL}(\{x, y\}),$

$$\text{where } \begin{cases} \alpha = x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha = \text{LABEL}(x) \text{ or } \text{LABEL}(y) & \text{for } x, y \text{ phrases, or} \\ \alpha = \langle x, x \rangle \text{ or } \langle y, y \rangle & \text{for } x, y \text{ terminals, or} \\ \alpha = \begin{matrix} \langle \text{LABEL}(x), \text{LABEL}(x) \rangle \\ \text{or} \\ \langle \text{LABEL}(y), \text{LABEL}(y) \rangle \end{matrix} & \text{for } x, y \text{ phrases} \end{cases}$$

He then suggests that phrases with the  $\langle x, y \rangle$  label dominate only the phrase that they are projected from. If we take c-command to be defined as in (107), consider how these assumptions will operate in an adjunction environment like (108).

(107)  $\alpha$  c-commands  $\beta$  iff every phrase dominating  $\alpha$  dominates  $\beta$ , and neither  $\alpha$  nor  $\beta$  dominates the other.



Because the phrase labeled  $\langle x, x \rangle$  dominates the phrase labeled  $x$ , but not the phrase labeled  $y$ , not every phrase that dominates  $\{x, x, n\}$  dominates  $\{y, y, z\}$ . As a consequence,  $\{x, x, n\}$  does not c-command  $\{y, y, z\}$ , as desired.

As with Kayne’s treatment of the scenarios in (104) and (108), there is no particular reason we should favor this implementation of what’s needed over other possible ones. Whether a distinction between Specifiers and Adjuncts is needed will depend on how the differences between modifiers and arguments end up being captured. Similarly, whether a phrase may have more than one adjunct or Specifier per phrase will

<sup>11</sup>So far as I can see, this simply smuggles in the  $XP/X^0$  distinction that he otherwise tries to redefine in terms of Merge.

again turn on how it turns out we need to treat these phrases. There is nothing, in other words, built into the LCA which decides these matters. By contrast, the ban against structures like (102) is embedded within the LCA; and, interestingly, the availability of these structures is just as embedded within the Bare Phrase Structure reduction of  $\bar{X}$  theory. This is where the two ideas find their strongest conflict, one that is perhaps resolved by Chomsky's suggestion.

In what follows we shall adopt both the LCA and the Bare Phrase Structure hypotheses in the hope that their differences can be resolved. Because of the relative plasticity of the outcomes these hypotheses have for phrase structure, however, we shall not feel constrained to build representations that, for instance, conform strictly to the S-V-O pattern that Kayne advocates. We will, however, fully embrace the reduction of to  $\bar{X}$  theory these proposals afford. In particular, let us abandon the  $\bar{X}$  Skeleton, and adopt in its stead MERGE, with the consequences that the LCA imposes on the resulting parses. Concretely, let's adopt the following conventions:

- (109)
- a. No distinction between XP and  $X^0$ .
  - b. Multiple adjuncts/Specifiers are permitted.
  - c. Terminals will be represented under  $X^0$ s.
  - d. Phrases will be given the traditional labels, e.g., "DP," "VP," etc.

These representations are those that are consistent with the system we entertained as a variation of Kayne's at the end of section 9.2. Nonetheless, I shall view the notational conventions in (109c) and (109d) as merely a more familiar packaging for the Bare Phrase Structure proposal. That is, though the graphs I will draw from here on out will put terminals within an  $X^0$ , let us understand the  $X^0$  to merely be the label for the terminal — it has no real status in the structure. Similarly, though I will label phrases with "VP," "DP" and so on, let this simply be an alternative rendering of the labels that PROJECT would produce. Thus, (109) is consistent with the Bare Phrase Structure proposals, with the exception of ignoring the Specifier/Adjunct distinction.

## Chapter 10

# Quantifier Raising

In the revisions to the theory we have just concluded, a large role was played by the scope asymmetries that emerge in the material which follows the verb. We began this revision, recall, in an effort to understand why when one complement precedes another, that complement tends to have scope over the second, but not the other way round. The reason we thought that this scope asymmetry existed is because of a variety of phenomena that I declared are sensitive to scope relations — the bound variable reading of pronouns, the relationship between a reciprocal and its “antecedent,” and the disjoint reference effect that arises between pronouns and name-like terms. Assuming that these phenomena do, in fact, reveal something about scopal relationships, I then declared that c-command was the syntactic relation that determined when one thing has another in its scope, and this led the way to the revisions we have now completed.

It’s time now to examine the reasons for the declarations I relied on earlier. We would like to substantiate that the phenomena used in determining scope are indeed good indicators of scopal relationships, and, further, we would like to know to what extent c-command is the correct syntactic reflex of scopal relationships. We will begin this task by examining the bound variable reading that pronouns get.

Recall that pronouns are able to receive an interpretation in which their value varies with the value given to some nearby quantifier. In (1), for example, the pronouns *his* and *her* can be understood as variables for the quantifiers *every boy* and *which girl*.

- (1) a. Every boy<sub>1</sub> knows his<sub>1</sub> mother.
- b. Which girl<sub>1</sub> visited her<sub>1</sub> father?

(I will use numerical subscripts to indicate referential dependencies. In general, when two DPs bear the same subscript, or index as they are sometimes known, this should be understood to indicate that those DPs will refer to the same things.) On this reading, (1a) can be seen as a shorter way of saying the rather cumbersome set of conjunctions in (2).

- (2) Max<sub>1</sub> knows his<sub>1</sub> mother, and
- Jerry<sub>2</sub> knows his<sub>2</sub> mother, and
- ⋮

This series continues until all of the boys under discussion are named. Similarly, the question in (1b), when it gives the bound variable interpretation to the pronoun, can be thought of as a long set of conjunctions like the following:

- (3) Sally<sub>1</sub> visited her<sub>1</sub> father, and  
 Jill<sub>2</sub> visited her<sub>2</sub> father, and  
 Sandy<sub>3</sub> visited her<sub>3</sub> father, and  
 Sherry<sub>4</sub> visited her<sub>4</sub> father, and  
 ⋮

As in (2), this series continues until all of the girls under discussion are named. Unlike (1a), however, this series of sentences shouldn't be thought of as capturing an assertion that the speaker of (1b) is making, but rather should be thought of as posing a set of possible answers to the question that the speaker of (1b) is asking. Informally, then, we can think of questions as being sentences that pose a selection of answers — or a form for an answer — that the hearer is invited to choose from.

Let's call the sentences from which these series of conjunctions are formed, the quantifier's scope. In (1), then, the phrase of which the quantifiers are subjects are their scopes. So, if we decide that  $\nu P$  is the phrase that contains both the subject of a clause and the VP in which all the complements of a clause are found, then we can see the  $\nu P$ s of (1) as the scopes of the quantifiers. These  $\nu P$ s are within the material that the quantifiers in (1) c-command, and so these cases are consistent with the declaration that c-command determines what some term has scope over.

We learn from (4) that for a pronoun to get a reading that allows it to function as a variable bound to a quantifier, it must fall within the scope of that quantifier.

- (4) a. Her mother asked who left.  
 ( $\neq x$ 's mother asked for which  $x$ ;  $x$  a person [ $x$  left])  
 b. Who asked her mother to leave?  
 (= for which  $x$ :  $x$  a person [ $x$  asked  $x$ 's mother to leave])?

In (4a), the quantifier *who* has scope just over the embedded clause, and as a consequence cannot bind the pronoun that lies in the higher sentence. In (4b), by contrast, *who* has the entire sentence within its scope, and as a consequence can bind *her*. So, we see here evidence for:

- (5) For something to be a variable bound to a quantifier, it must be within that quantifier's scope.

This is not sufficient, however, as the paradigm in (6) indicates.

- (6) a. Who<sub>1</sub> did [ [PRO<sub>1</sub> seeing his<sub>1</sub> father]] bother t<sub>1</sub>?  
 b. \* Who<sub>1</sub> did [ [Mary's seeing his<sub>1</sub> father]] bother t<sub>1</sub>?  
 c. \* Who<sub>1</sub> did [ [PRO<sub>arb</sub> seeing his<sub>1</sub> father]] bother t<sub>1</sub>?  
 d. \* Who<sub>1</sub> did [his<sub>1</sub> father] bother t<sub>1</sub>?

In all of these examples, *his* falls within the scope of *who*, and so we should expect it to be able to function as a variable bound to *who*. In fact, however, this is only possible

in (6a), where, interestingly, there is a PRO also bound to *who*. The other examples in (6) are said to instantiate instances of “Weak Crossover,” a term coined for these situations by Paul Postal. If we put together what distinguishes the good (6a) from the remaining cases in (6) with the good example in (4b), we find that what they share is the existence of some other variable between the quantifier and the pronoun. That is, it seems that there is an additional condition on the bound variable reading for pronouns whose effects might be described with (7).

(7) WEAK CROSSOVER

If an overt pronoun is to be the bound variable to  $\alpha$ , then it must be c-commanded by something else that functions as a variable bound to  $\alpha$ .

In (6a), the other variable is PRO. In (1b), the other variable could be the “trace” left by moving *who*; and in (1a), it could be the trace that every boy left in its  $\theta$ -marked position (Specifier of  $\nu$ P). So, for this description of these examples to go through, we must imagine that the position a moved wh-phrase vacates is occupied by a variable. Some have construed this as evidence for the existence of “traces.”

Now the paradigm in (8) suggests that quantificational DPs can have a scope wider than the c-command domain that their surface position indicates, and so here is our first problem for the thesis that c-command and scope are related.

- (8) a. PRO<sub>1</sub> seeing his<sub>1</sub> mother pleased every boy<sub>1</sub>.  
 (= for every  $x$ ,  $x$  a boy, [ $x$  seeing  $x$ ’s mother pleased  $x$ ].)  
 b. PRO<sub>arb</sub> seeing his<sub>1</sub> mother pleased every boy<sub>1</sub>.  
 ( $\neq$  for every  $x$ ,  $x$  a boy, PRO<sub>arb</sub> seeing  $x$ ’s mother pleased  $x$ )

The pronoun, *his* can be understood as a variable bound to *every boy* in (8a), but not (8b). (8a) corrects for the effects of Weak Crossover, and we can see, therefore that *his* must be within the scope of *every boy*, since otherwise we would not expect it to be able to function as a variable bound to *every boy*.

We see evidence for a similar conclusion in the reading that (9) gets.

- (9) A different girl visited every boy.

The word *different* requires a context in which a set of individuals are established from which the NP that it modifies is compared. So, *different girl*, for instance, is felicitous only in situations where there are a set of girls established from which the girl that *different* distinguishes can be drawn. In (9), the set of girls is brought about through the action of *every*. As noted above, we want sentences with the quantifier *every* in their subject position to be understood as a series of conjoined sentences, which differ only by virtue of the value given to their subject (and to any variable which that subject might bind). It appears, now, that we want sentences of the form (10) to be semantically equivalent to the series in (11), where “ $x$ ” refers to a member of the set denoted by  $X$ , and  $DP_n$  refers to a member of the set denoted by  $DP$ .

- (10)  $X$  verbed every  $DP$ , where “ $X$ ” is an expression like *a girl*

- (11)  $x$  verbed  $DP_1$ , and  $x$  verbed  $DP_2$ , and  $x$  verbed  $DP_3$ , and ...

This series of sentences must exhaust the (relevant) set of individuals picked out by  $DP$  in (10). Now as the series in (11) makes plain, there are a family of “ $x$ ”s possible, and

it is this set that *different* can access. Thus in (9), the set of girls over which *different* ranges is established by way of *every boy* having *A different girl visited x* in its scope.

That *different* is given the reading it has in (9) by virtue of being within the scope of *every boy* can be confirmed by comparing (12).

- (12) The girl who everyone likes introduced a different boy.

In this example, *everyone* does not have a *different boy* within its scope — this statement gets matched with the series of statements in (13) — and as a consequence the set of boys over which *different* ranges has to be established through some other means. (This sentence is felicitous, then, only in a discourse where some set of boy(s) has been established.)

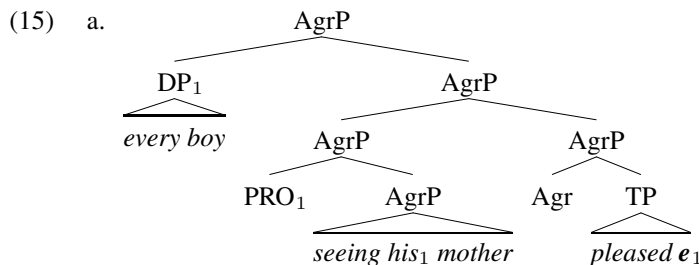
- (13) the girl who DP<sub>1</sub> likes introduced a different boy.  
 the girl who DP<sub>2</sub> likes introduced a different boy.  
 the girl who DP<sub>3</sub> likes introduced a different boy.

These data, then, pose something of a dilemma with regard to the question whether c-command is the relevant syntactic correlate of scope. Looking at the difference between (1) and (4), it looks like the thesis that the scope of some quantifier is the material it c-commands is confirmed. Similarly, the inability of *everyone* in (12) to have a *different boy* in its scope is consistent with this thesis: *a different boy* is not c-commanded by *everyone*. But in (8) and (9) we have counterexamples to the thesis.

Robert May<sup>1</sup> suggests a method for rescuing the c-command thesis that employs the covert portion of the derivation that we have already invoked. He proposes that there is a rule that moves quantificational DPs, assigning them a new c-command position. He calls this the *Quantifier Rule*, and formulates it as follows.

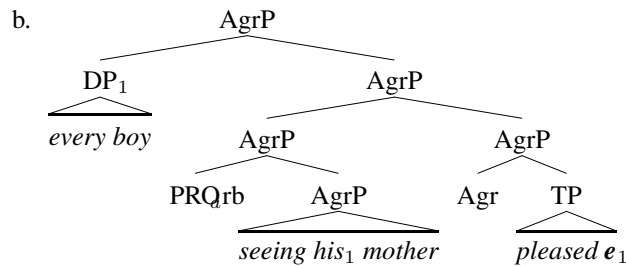
- (14) QUANTIFIER RULE (QR)  
 Adjoin a quantificational DP to AgrP in the post S-structure part of the derivation.

It forms the representations in (15) from (8a,b).



<sup>1</sup>See May (1977) and May (1985).





If we model the covertness of QR in terms of derivations, like we have previously, then we can think of QR as forming a level of representation, let's call it Logical Form (or, LF) from S-structures. Because this level of representation isn't S-structure, its effects are not "visible." This is all identical to the treatment we've adopted for covert verb movement. Now we can reframe (5) so that it holds of the output of QR, and thereby capture the range of cases of bound variable readings for pronouns.

- (16) For  $\alpha$  to be a variable bound to  $\beta$ , it must be within  $\beta$ 's scope at Logical Form.

If QR is to be a solution to the dilemma we've just encountered, we can't let it mess up too badly the c-command relations that hold of S-structures. For instance, we don't want *everyone* in (12) to be moved into a position in which it c-commands *a different boy*, for this would wrongly allow this sentence to get the interpretation for *different* that it doesn't have. We need to place a bounding constraint on QR, preventing it from moving quantifiers too far. The need for this bounding effect can be seen in examples like (17) as well.

- (17) a. Seeing his mother means that every boy can go on the field trip.  
 ( $\neq$  [PRO<sub>1</sub> seeing his<sub>1</sub> mother] means [that every boy<sub>1</sub> can go on the field trip].)
- b. A different girl knows that every boy went on the field trip.  
 (*different* not licensed by *every boy*.)

In neither of these cases is *every boy* able to take the entire sentence as its scope. As a consequence, *different* does not get the meaning it can otherwise have when in the scope of an *every* quantifier, and *his* cannot be understood as a variable bound to *every boy*. In view of these facts, let's adopt the preliminary (18) as our bounding constraint on QR.

- (18) QR cannot move something out of a CP.

This will also correctly prevent everyone from gaining too large a scope in (12).

May's solution to the problem these examples pose for the c-command thesis, then, is to posit a covert representation, formed by QR, at which the scopes of the quantifiers involved is in fact their c-command domain. A conceivable alternative might be that there is a more sophisticated semantic treatment of these cases which will allow us to abandon the c-command thesis. This more sophisticated semantics would have to give quantifiers meanings that allow object quantifiers to compose with the surrounding material in such a way that these quantifiers gain scope over the sentence they are objects of. Is there a way of telling which of these directions is correct? Because

this is a syntax-semantics question — that is it falls under the second of the two goals identified at the outset for syntax: How is it that syntactic form encodes meaning? — we can approach this question from two angles. Is there a way of formulating a semantics that has the right consequences, and if so, is there support for that semantics? And, from the syntax side: Is there evidence for the representations that QR forms?

We shall examine here only the syntactic evidence.

There are, I believe, two clear pieces of evidence that QR is correct. One supports that part of the hypothesis that says that a movement operation is responsible for giving quantifiers their surprising scope. And the other supports the existence of the representations formed by QR.

The first piece of evidence is found in examples like (19).

- (19) a. A unicorn is likely to be discovered.  
b. A unicorn is anxious to be discovered.
- (20) a. It is likely that a unicorn is discovered.  
b. There is a unicorn which is likely to be discovered.
- (21) a. \*It is anxious that a unicorn is to be discovered.  
b. There is a unicorn which is to be discovered.

There is an ambiguity in (19a) which is absent in (19b). (19a) can have a reading in which it is not a statement about some unicorn or other, but is instead a statement about the likelihood of discovering a unicorn — this is the reading that (20a) aims to paraphrase. On this interpretation of (19a), the speaker of this sentence is not asserting that unicorns exist. As a consequence, it is perfectly consistent for a speaker of such a sentence to also deny the existence of unicorns. For instance, there is nothing odd about this speech:

- (22) I've never been convinced that unicorns exist, despite the frequent reference to them in myth and folklore. After all, in several centuries there's never been a confirmed sighting of a unicorn; and let's face it, given their nature, a unicorn is likely to be discovered.

This reading reflects the fact that *a unicorn* in (19a) can be interpreted within the embedded infinitival clause. That is, it can be interpreted in roughly the same position that we see *a unicorn* in (20a). These cases are known as instances of "Quantifier Lowering."

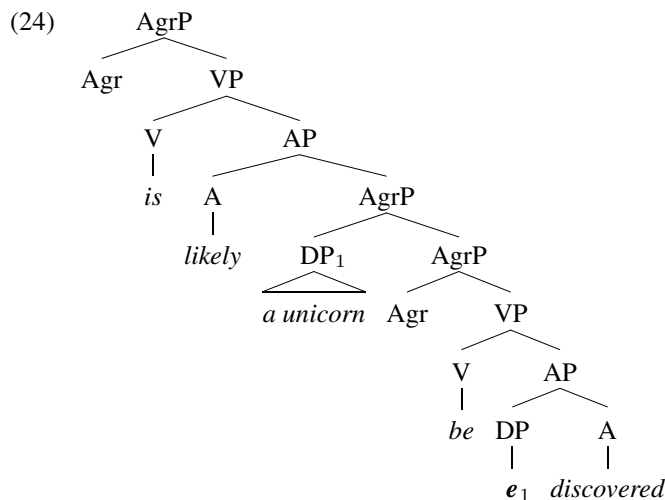
A parallel reading is not available in (19b). The speaker of (19b) is asserting the existence unicorns; this can be seen by noting that (19b) is odd if it is uttered by the same speaker who simultaneously denies their existence. The following speech, for example, differs in felicity from (22).

- (23) I've never been convinced that unicorns exist, despite the frequent reference to them in myth and folklore. After all, in several centuries there's never been a confirmed sighting of a unicorn; and let's face it, given their reported loneliness, a unicorn is anxious to be discovered.

This reflects the fact that *a unicorn* in (19b) cannot be interpreted within the infinitival clause.

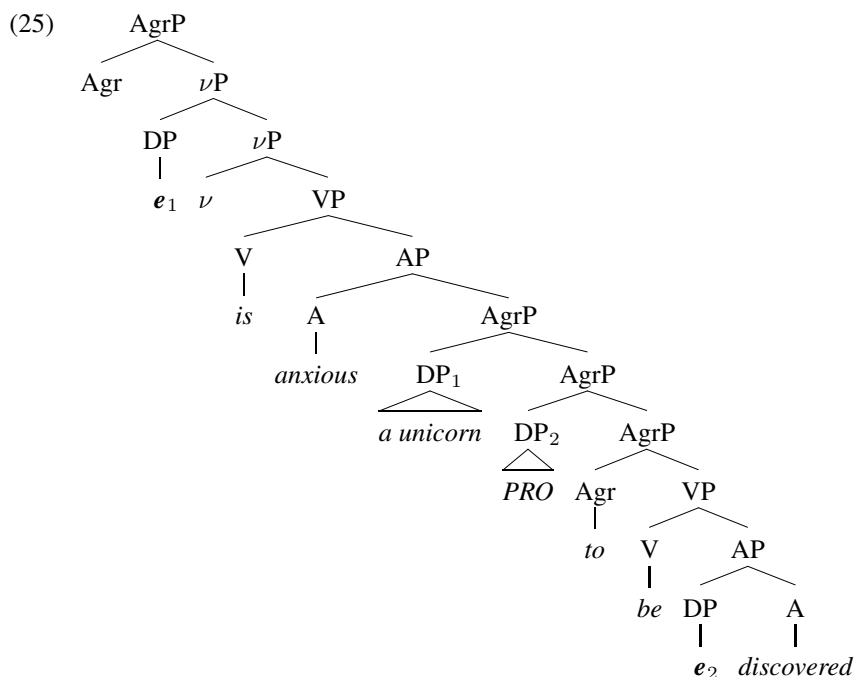
The relevant difference between (19a) and (19b) is that the infinitive following likely is transparent, whereas the infinitive following (19b) is opaque. In our model of this difference, the subject DP in (19a), but not (19b), has moved out of the embedded infinitival. In (19b), the understood subject of the embedded infinitival is PRO — construed as coreferent with the higher subject — not the DP *a unicorn*. Instead, *a unicorn* in (19b) gets its  $\theta$ -role from the  $\nu$ P of the higher clause.

Given this model of the difference between transparent and opaque infinitives, the QR solution to our scope problem predicts this contrast. At least it does so when combined with the Projection Principle. Recall that the Projection Principle requires that phrases moved from  $\theta$ -marked positions leave a trace in that position which is construed as a variable bound to the moved item. Consider, then, the representations that QR would give to the two sentences in (19) if it were to place *a unicorn* within the embedded infinitival clause. (I'll simplify these representations, removing for the time being, the material that expresses accusative Case assignment and tense.) In (24) is the case involving a transparent infinitive; and in (25) on the next page is the opaque clause case.



Note that in (24b) there is a trace left in the root clause, in the Specifier of  $\nu$ P position that this DP gets its  $\theta$ -role from. The problem with (24b), then, is that this trace is not c-commanded by *a unicorn* and, consequently, is not within the scope of *a unicorn*. But traces are construed as variables bound to the term that bears the  $\theta$ -role assigned to the position they occupy, so this runs counter to the requirements imposed on traces. By contrast, there is no such trace in the root clause of (24a). This is because *a unicorn* in this example comes from a  $\theta$ -marked position that is in the embedded infinitive. But this trace will be c-commanded by, hence in the scope of, *a unicorn* even if QR moves it down into the embedded clause.

What these examples indicate is that we want the scopes of quantifiers to be very sensitive to the syntactic representations of the clauses holding those quantifiers. In particular, we want it to see the difference between transparent and opaque infinitives. The QR method of assigning scope does just that, exploiting principles of the



analysis of the transparent/opaque contrast that are independently motivated. That is why these phenomena can be construed as evidence for the QR model.

The second sort of evidence for QR comes in the arena of VP Ellipsis. You'll recall that this is a process which forms sentences such as (26).

- (26) Sally read the books you recommended, though Sammy didn't [<sub>VP</sub> ].

In the second clause there is a missing VP, as indicated, that is understood to be identical to the VP in the first clause; that is, the second clause means something like: "Sammy didn't read the books you recommended."

What should we understand the missing VP in these examples to be? Maybe VP Ellipsis should be seen as a kind of (silent) pro-VP, akin, perhaps, to:

- (27) Sally read the books you recommended, though Sammy didn't do so.  
Sally read the books you recommended, though Sammy didn't do it/that.

Maybe, that is, there is some principle of English that allows, let's say, *do so* to go unpronounced. Thus, VP Ellipsis is just the *do so* construction, but with the *do* and *so* unspoken.

But there are some differences between these constructions which makes this look, initially, unlikely. For example, it is possible to understand there to be a variable bound by an interrogative quantifier inside an elided VP, as in (28).

- (28) a. I know which book Max read, and which book Oscar didn't  $\Delta$ .  
 b. This is the book of which Bill approves, and this is the one of which he doesn't  $\Delta$ .

(Fiengo and May 1994, (99a,c):229)

But this isn't possible in cases of *do so* or *do it*:

- (29) a. \*I know which book José didn't read for class, and which book Lulumae did it for him.  
 (compare: I know that José didn't read this book for class, but that Lulumae did it for him.)  
 b. \*This is the book which O.J. Berman reviewed, and this is the one which Fred won't do so.  
 (compare: O.J. Berman reviewed this book but Fred won't do so.)

Let's reject this interpretation of VP Ellipsis, therefore.

Another possibility is that the VP that is missing in VP Ellipsis is just a plain old VP, with all the material VPs normally have, and all that is exceptional is that this material is not spoken. Because these elided VPs are understood to be the same as some other VP in the area, we might imagine that the process that allows one to fail to pronounce VPs requires that the VP which goes silent is identical to some other VP. All that is required on this conception, is to spell out what "identical" means in this context. Unfortunately, this is not a trivial task, as there are cases of slight mismatch between the elided VP and its antecedent. In (30), for example, the antecedent VP has a trace (=variable) in it that is not understood to be part of the elided VP.

- (30) a. This is just the kind of thing that Harris could have suggested *e*. And in fact, he did  $\Delta$ .  
 b. Harry is someone they would like to send *e* to the Olympics. And they will  $\Delta$  too, if they can finance it.

(Hardt 1993, (21-2):15-6)

In (30a), for example, the antecedent VP is *suggested e*, but the VP which is elided is *suggested it*. Similarly, in (31) the elided and antecedent VPs need not be identical.

- (31) We like our friends and they do  $\Delta$  too.

In this case, we can understand the elided VP to be *like their friends*, which is different from the apparent antecedent VP: *like our friends*.

These cases suggest, then, that if we are to understand a missing VP to be just a normal VP that has gone unpronounced because it is the "same" as some other VP, we will have to allow "sameness" to be something looser than strict identity. Nonetheless, let's assume that VP Ellipsis is just a silent, but otherwise ordinary, VP which is the "same," as some other VP, and hope that we can find a characterization of "same" that allows for the slight differences that we have witnessed. The argument I'm about to report for QR, unfortunately, really requires that we know more about what it means for an elided VP to be the same as its antecedent. The two cases we've seen above in which

antecedent and elided VP don't match both involve cases in which there is a mismatch in the value of a pronoun or trace. So, as a rough guide, let's understand "sameness" in this context to mean identity, modulo the value that these sorts of terms get. Pronouns and traces are the two things we've seen can get a bound variable interpretation, and in fact, in our cases of mismatch above, that is the interpretation they could be getting. So, let's adopt a definition of "sameness" that allows only variables to have a different expression.<sup>2</sup>

- (32) Two VPs must be the same if one is to go unpronounced. Two VPs are the same if they are identical up to the variables they contain.

The argument I'm about to report on behalf of QR from VP ellipsis requires that this guess about what "sameness" is be not too far off the mark. Thus, the argument should be understood as having this important qualification built into it.

The argument on behalf of QR from VP Ellipsis arises in contexts like:

- (33) Jerry likes everything that you do  $\Delta$ .

These are cases of "Antecedent Contained Deletions" (or ACD). They are problematic because it is difficult to find an antecedent VP for them that has anything like the right form. What we want in (33), for instance, is a VP of the form in (34).

- (34) everything that<sub>1</sub> you do [<sub>VP</sub>like  $e_1$ ]

But there appears to be no VP like this in (33).

There are two popular strategies for solving this problem. One involves QR; let's save that for last. The other makes use of another process, known as Pseudogapping, following the term Levin (1986) introduces.

- (35) While I don't like asparagus, I do  $\Delta$  chocolate.

Suppose this process licenses an ellipsis of all the material in a VP except the object, or indirect object, which remains as remnant. If Pseudogapping is what underlies ACDs, then we might see (33) has having a source like:

- (36) Jerry likes [<sub>DP</sub>[<sub>CP</sub>that you do everything]].

While this is a viable analysis of the ACD in (33), it is not a possible analysis for all ACDs. This is because there are some situations in which Pseudogapping is not licensed, but ACD is. For example, as (37a) demonstrates, Pseudogapping cannot delete a double-object verb and the first object that follows it, and strand the second object. And yet, as (37b) this is a context in which ACD is licensed.

- (37) a. \* While I didn't give Jerry crayons, I did  $\Delta$  toys.  
b. Liz gave me every toy that you did  $\Delta$ .

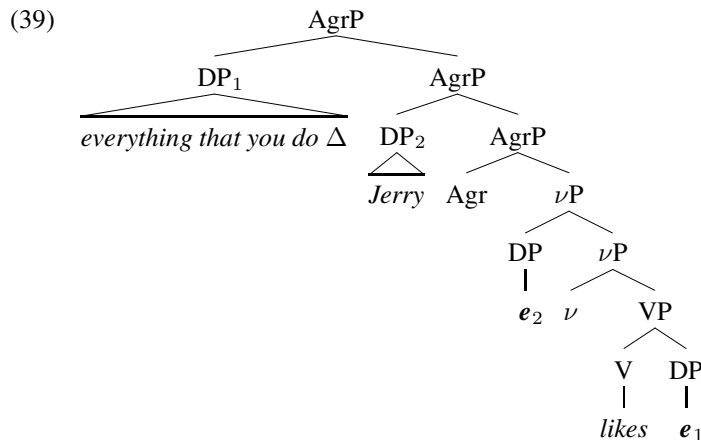
The elided VP in (37b) is understood to be: "you (did) [VP give me  $e$ ]." So, if Pseudogapping is to produce this sentence, it would have to delete the struck-out material in (38).

<sup>2</sup>This is roughly what Sag (1980) argues for. For a more recent exploration of the relevant condition, with many relevant considerations, see Fiengo and May (1994), Tomioka (1997), Fox (2000) and Merchant (2001).

(38) Liz gave me [ [that you did [~~VP give me~~ every toy]].

And this is precisely what (37a) shows us isn't possible.

So, let's abandon the Pseudogapping approach to these cases, and turn to the other popular account, one that makes use of QR. If we let the antecedent to an elided VP be formed by QR, then we will create VPs with a form needed to be antecedents for the ellipsis. For example, once QR has applied to (33), the LF representation will be as in (39).



By moving the object out of the VP, we have created something, namely: *likes e* that is precisely the form we need to license the ellipsis. So, if we let the elided VPs be licensed when they are the “same” as VPs that exist at LF, QR will rescue us from the problem that ACDs pose. Here then is the evidence for the LF structures that QR produces.

That the rule which fixes the scopes of quantifiers is also the process that forms the VPs we need as antecedents in cases of ACD is indicated by various similarities in these two processes. For example, we have seen that quantifiers cannot have a scope that is larger than the CPs they are in at S-structure. We captured this fact by placing a bounding constraint on QR that prevents it from moving quantifiers out of CPs. We can see the same locality condition at play in instances of ACD too. It is not possible for an elided VP to take as its antecedent some VP that contains it, if that VP embeds a CP which contains the ellipsis. The ungrammaticality of (40) is illustrative.

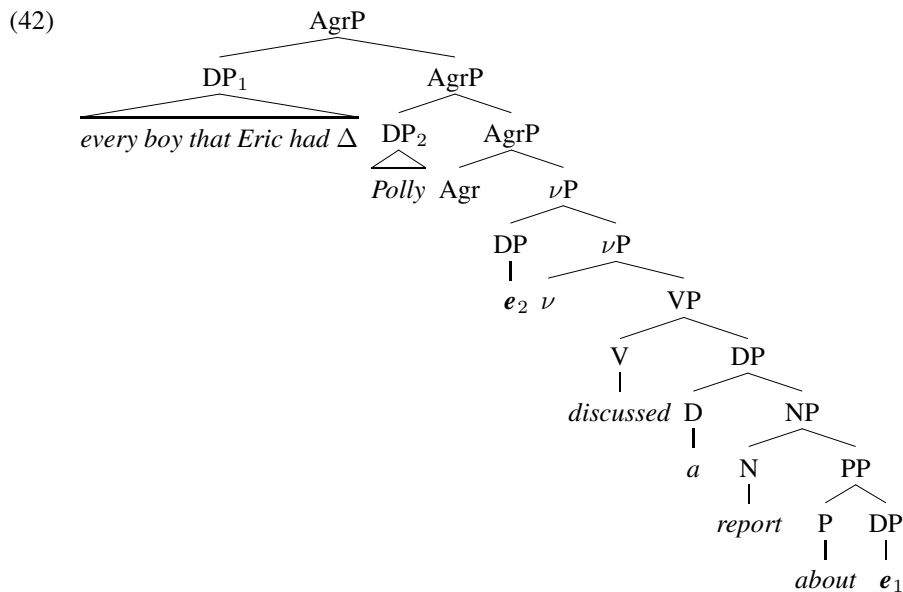
(40) \*Dulles believed that [everyone Hoover did Δ] is a spy.

If QR cannot move a quantifier out of a CP, then this will follow. There is no LF for (40) in which the following VP is manufactured: “believed that *e* is a spy.” To get that LF, it would have to be possible for *everyone that Hoover did Δ* to move beyond the CP that contains it. Since *believe that e is a spy* is the VP that is elided in (40), there is no VP, even at LF, that is the “same,” and the sentence is ungrammatical.

Further, there is some reason to think that the process which forms antecedent VPs for ACDs correlates with the scope of the DPs involved. In (41a), for instance, *every boy* must have wider scope than *a report*, if it is to get an interpretation in which the elided VP is: *discussed a report about e*.

- (41) a. Polly discussed a report about every boy that Eric had  $\Delta$ .  
 b. Polly discussed a report about every boy that Eric had discussed.

QR would give (41a) the LF representation in (42).



And this both creates an antecedent VP of the right form, but also places a *report* within the scope of *every boy*.

Note, incidentally that this is also an example where a Pseudogapping analysis isn't possible. As (43) illustrates, Pseudogapping does not seem able to leave behind a portion of a DP it has elided.

- (43) \*Polly didn't discuss a report about James, but she did  $\Delta$  Bill.

This is precisely what would be needed, however, if Pseudogapping were to be the source of (41a). So again, we see that Pseudogapping can't be the means for resolving all instances of ACD.

One final case in which the scopes of the DPs holding an ACD seem to correlate with the size of the VPs that have elided is found in cases where the subjects of infinitives that get Case from the higher clause. In these cases, (44) is an example, the subject of an infinitival clause moves into the higher clause to get Case. Now, the first thing to notice about these examples is that there are two readings attendant to the scope of the accusative Case-marked subject. In (44), for example, there are two meanings that are associated with whether the scope of *nobody* includes the root clause or includes just the embedded clause.

- (44) Dulles believed nobody to be a spy.

May (1985) devises a clever way to bring out these readings which, unfortunately for us, involves a separate occurrence of VP ellipsis. The way involves introducing a before clause with an ellipsis, as in (45).



- (45) Dulles believed nobody to be a spy before Angleton did  $\Delta$ .  
(from May (1985), but see also Fiengo and May 1994, sec. 6.2.)

This has two readings dependent on the scope of *nobody*; they are:

- (46) a. Before Angleton came to believe that no one is a spy, Dulles came to that belief.  
b. There is no one whom Dulles came to believe is a spy before Angleton reached that belief.

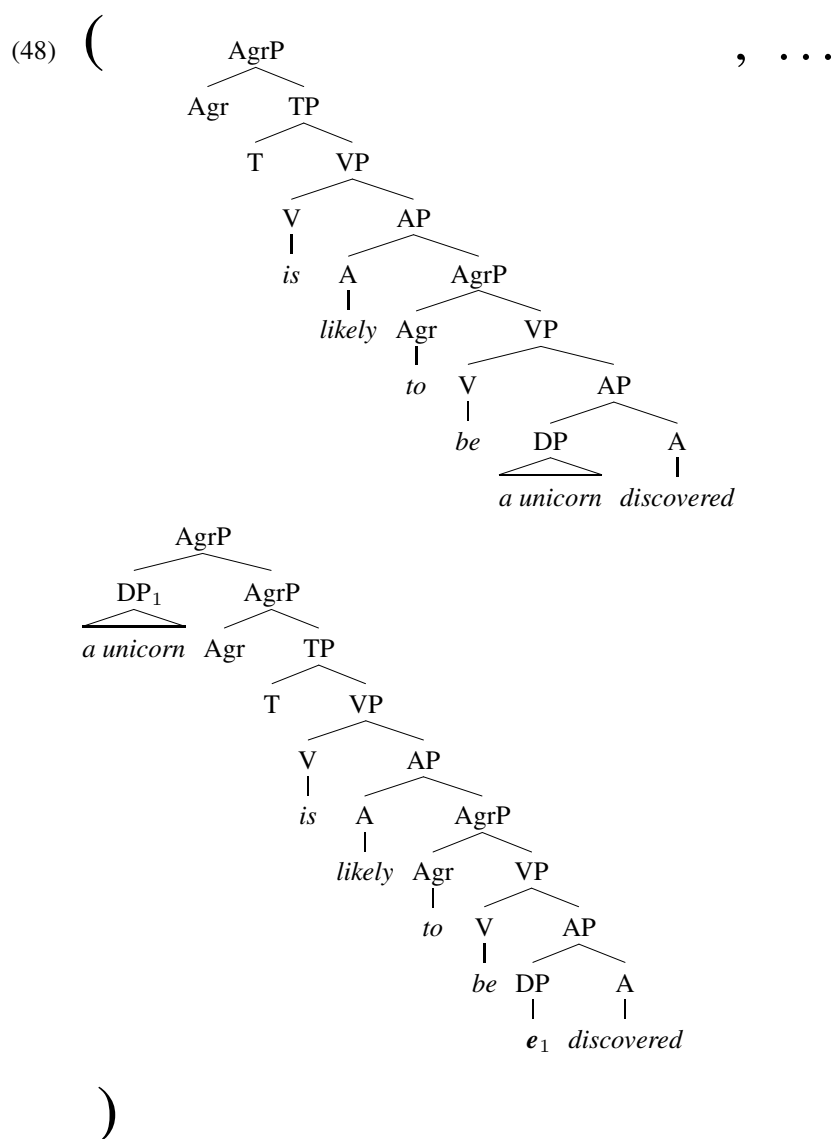
Now, interestingly, when the *nobody* argument contains an ellipsis, only the wide-scope reading in (46) is possible.

- (47) Dulles believed [nobody that Hoover did  $\Delta$ ] to be a spy before Angleton did  $\Delta$ .  
  
 $\neq$  Before Angleton came to believe that no one Hoover had some belief about is a spy, Dulles came to that belief.  
  
 $=$  There is no one Hoover believed to be a spy whom Dulles came to believe is a spy before Angleton reached that belief.

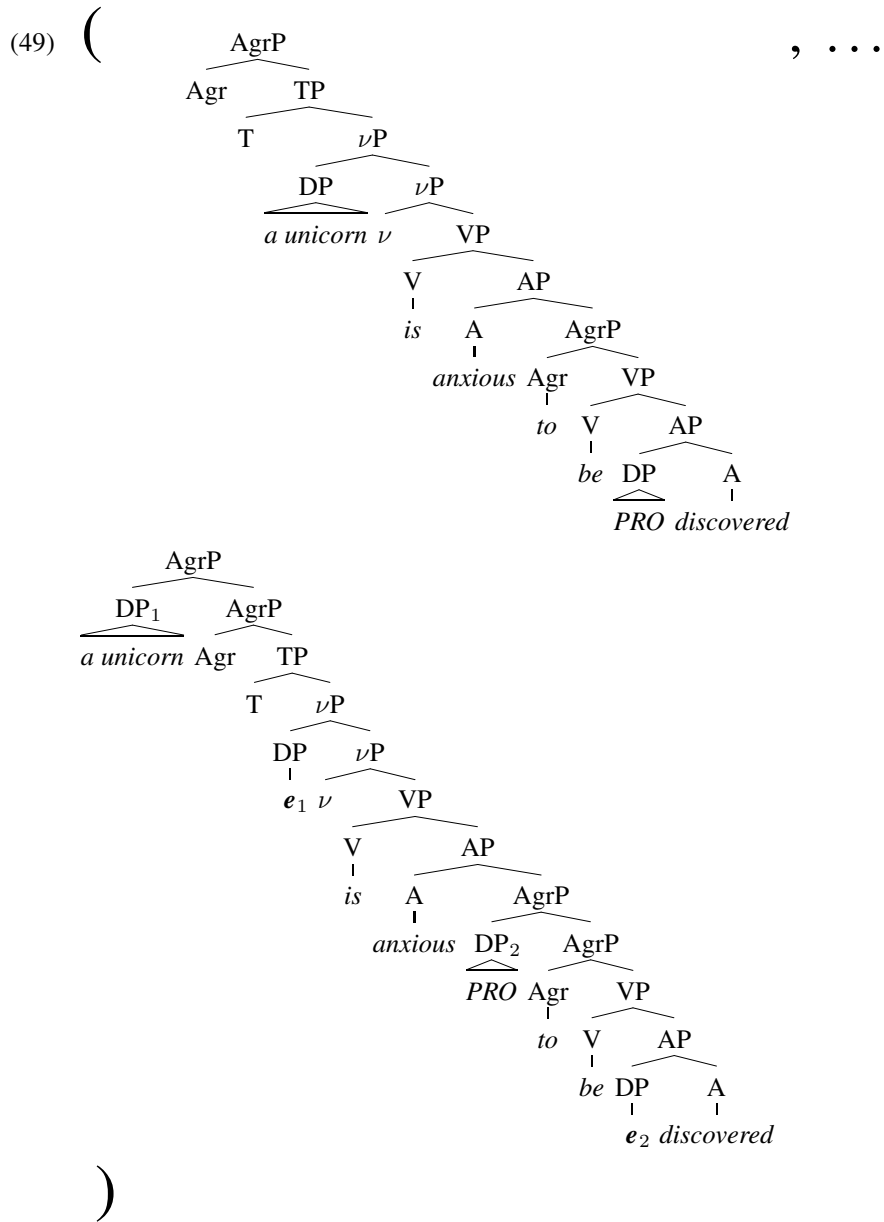
This is just what we expect if QR produces both the antecedent VPs for ACD and fixes the scopes of quantifiers.

We have, then, two kinds of phenomena which speak on behalf of the QR model of fixing quantifier scope. The distribution of quantifier lowering phenomena is understood once QR and the Projection Principle are combined. And these instances of ACD, and the match they illustrate between the scope of the quantifier containing the ellipsis and the size of the VP that can serve as antecedent, also is correctly captured under a QR interpretation of scope. To the extent that QR is correct, so too is the thesis that scope maps off of c-command relations. From here on out, then, let's accept the thesis that the scope of something is just its c-command domain at the LF phrase marker, and not necessarily at the S-structure phrase marker.

There is another way of interpreting the quantifier lowering and ACD cases that dispenses with QR, but nonetheless preserves the c-command = scope thesis. This alternative makes heavy use of the more articulated representations we've developed for complements, and expresses the idea that it is not just one member of a derivation which can be semantically interpreted. To see how this might address the issue, consider again the issue of quantifier lowering. Compare the derivations that lead to the S-structures of the opaque and transparent clauses we used to bring out this phenomenon.

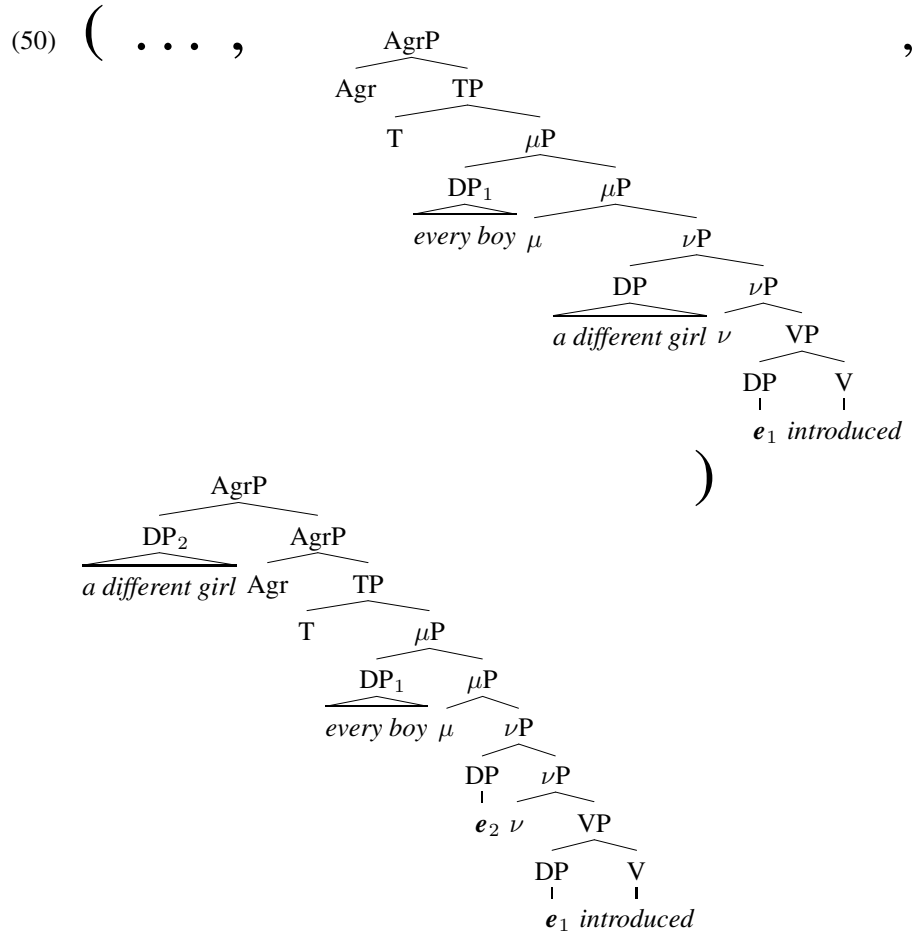


If we can let either of the two phrase markers indicated in this derivation be interpreted, then the ambiguity we've seen for this example will be correctly modeled. The case involving an opaque clause will have a derivation whose parallel phrase markers are shown in (49).



Letting either of these two phrase markers be interpreted will not create the ambiguity that we see in the transparent example. In particular, *a unicorn* will never be interpreted within the embedded infinitival clause.

Consider next a case in which the subject falls within the scope of the object. A derivation of (9) has the following two phrase markers in it:



Here too it would be sufficient to let the semantic component have access to more than just the LF phrase maker.

We don't want the semantic component to take the entire derivation, whatever that would mean, as its input. Instead, it seems to be necessary for it to have just a single element of the derivation as its input. This can be seen by considering cases such as (51).

- (51) A unicorn seems to his friends to be unhappy.  
 = there is a unicorn<sub>1</sub> which seems to his<sub>1</sub> friends to be unhappy.  
 ≠ It seems to his<sub>1</sub> friends that a unicorn is unhappy.

What needs to be captured here is that if the pronoun is to be understood as a variable bound to *a unicorn*, then *a unicorn* cannot be interpreted within the infinitival complement. We do not, in other words, want to allow the semantic component to interpret *a unicorn* in both of the positions it occupies.

What these cases reveal, then, is that the information needed to capture the scope ambiguities, at least in these cases, is present already in the derivation. There

is no need to produce more elements in the derivation, and so the use of QR to do so is suspect. A number of people have argued that in fact there is evidence against the QR method of expressing these ambiguities, and that instead a method like the one just described — a procedure of interpreting derivations that allows access to any of the phrase markers — is to be preferred. Norbert Hornstein might have produced the most systematically complete portfolio for this position, so we shall examine some of his arguments.<sup>3</sup> As we'll see, the evidence is somewhat mixed.

A suite of Hornstein's arguments target the fact that on the QR approach, for a subject to fall within the scope of an object, that subject can be in its surface position, whereas on his approach, the subject must be in  $\nu P$ . If we can find diagnostics for being within  $\nu P$ , then it should be possible to test for these different outcomes. Hornstein proposes a variety of such diagnostics. He suggests, for instance, that certain adverbials must be higher than  $\nu P$ , and therefore fall beyond the scope of the subject in its underlying position. It's a consequence of this, he suggests, that it is difficult to get a reading for (52) that simultaneously gives the subject lower scope than the indirect object and allows it to bind the pronoun in the adverbial clause.

- (52) Someone<sub>1</sub> danced with every woman before he<sub>1</sub> left the party.  
(Hornstein 1995b, (24a,c): 160)

This follows, in other words, if the subject must lower in order to fall within the object's scope.

Another similar argument turns on the observation that a VP which has a subject lowered into it is not an appropriate antecedent for an elided VP with a different subject. So a sentence like (54) should have only the wide-scope reading for someone, unlike (53).

- (53) I got someone to interview every candidate that I interviewed.  
(54) I got someone to interview every candidate that I did  $\Delta$ .  
( $\Delta$  = [interviewed *t*]) (Hornstein 1995b, (34): 163)

This is predicted, on Hornstein's story, because the LFs for the narrow and wide scope readings should be as in (55a), which allow for a proper antecedent VP only in the case where the subject hasn't lowered.

- (55) a. someone<sub>1</sub> [ [every candidate that I did  $\Delta$ ]<sub>2</sub> [<sub>VP</sub> e<sub>1</sub> interview e<sub>2</sub> ]]  
Scopes: *someone* wider than *every candidate*  
 $\Delta$  = [ e<sub>1</sub> interview e<sub>2</sub> ]  
b. [ [every candidate that I did  $\Delta$ ]<sub>2</sub> [<sub>VP</sub> someone interview e<sub>2</sub> ]]  
Scopes: *every candidate* wider than *someone*  
\* $\Delta$  = [someone interview e<sub>2</sub>]

I am somewhat skeptical of these judgements, to be honest. Consider by contrast the examples in (56).

- (56) a. I got a different detective to interview every suspect that I had  $\Delta$ .  
b. She made at least one candidate talk to every guest that I had  $\Delta$ .

<sup>3</sup>See also Kayne (1998). For Hornstein's arguments, see Hornstein (1994, 1995a).

So, I'm not sure there is an argument from these data for Hornstein's interpretation of QR.

Hornstein offers a variant of this argument, one which uses examples where the ellipsis is within a subject. If lowering brings subjects into the VP, and Lowering is the only way to give an object wide scope, then we should not expect a subject containing an ellipsis to have narrow scope relative to an object. A test case is (57).

- (57) a. [A man that should  $\Delta$ ] has eaten every chocolate bar.  
 b. \* [ [every chocolate bar]<sub>2</sub> [<sub>VP</sub> [a man that should  $\Delta$ ] eat  $e_2$  ]]

And, in fact, Hornstein reports that in examples like (57), the subject cannot be understood as falling within the scope of the object.

Again, however, I am somewhat skeptical of these judgements; consider the examples in (58).

- (58) a. At least one man that SHOULD  $\Delta$  HAS eaten every chocolate bar.  
 b. At least one woman that tried to  $\Delta$  has reached every mountain top.

So, again, I am not sure there is an argument from these examples.

In any case, that something is wrong with the assumption that Lowered Subjects necessarily come back into VP (or that the identity condition on VP Ellipsis is constrained in the way we assumed above) is further indicated by examples such as (59), from Pustejovsky and Sells (1982).

- (59) A Canadian flag was hanging in front of most windows, and an American one was  $\Delta$  too.

In this case, *most windows* can be understood as having wider-scope than the indefinite subjects in both clauses. Maybe we should adopt the view that Subjects can Lower to a position below the object, but still outside of the VP. Indeed, this is what the parse we adopted earlier would permit.

Because on Hornstein's thesis, the scopal positions of an argument are fixed by Argument movement, we expect to see the constraints on Argument movement also showing up on scope taking. One of his arguments is built on this expectation. He points to paradigms like (60).

- (60) a. A different girl seemed to be reviewing every report.  
 b. Someone<sub>1</sub> seemed to his<sub>1</sub> boss to be reviewing every report.  
     ?Clearer in: *A different girl<sub>1</sub> seemed to her<sub>1</sub> boss to be reviewing every report.*  
 c. Someone<sub>1</sub> seemed to himself<sub>1</sub> to be reviewing every report.  
     (Hornstein 1995b, (22b,c): 160)

As (60a) illustrates, the subject of the higher clause can, in this raising context, have scope lower than the object of the embedded infinitival. In this example, *different* can be licensed by the quantification that *every report* invokes. On Hornstein's account, this can only happen if *a different girl* is lowered by into the embedded clause. This is because *every report* Argument Moves just as far as the accusative Case marked

position which licenses it, and this position is within the embedded clause. That this is a correct prediction is seen by considering (60b) and (60c). In these cases, there is a term in the embedded clause that must be within the scope of the higher subject — the anaphor *himself* must fall within the scope of its antecedent, recall, and if we understand *his* as a variable bound to *someone*, so must *his* — thus, lowering of the subject will be blocked in these examples. And, as expected, *someone* cannot fall within the scope of the embedded object. In general, Hornstein’s approach requires that an object will never include in its scope anything higher than the  $\mu$ P that it gets Case from. This requires for a subject to fall within an object’s scope that it be construed within this  $\mu$ P. The cases (60) fall within this expected pattern.

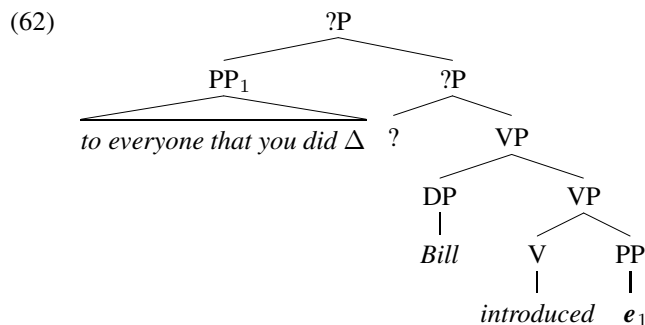
My own assessment, then, is that the arguments on behalf of Hornstein’s revision are that the locality condition on QR matches the locality condition we’ve seen on Argument movement, and the fact involving adverbials in (52).

Interestingly there are a variety of phenomena that argue against Hornstein’s revision. One is simply that complements other than accusative Case marked DPs can have scope wider than the VP they are construed in. There are example like (61), for example.

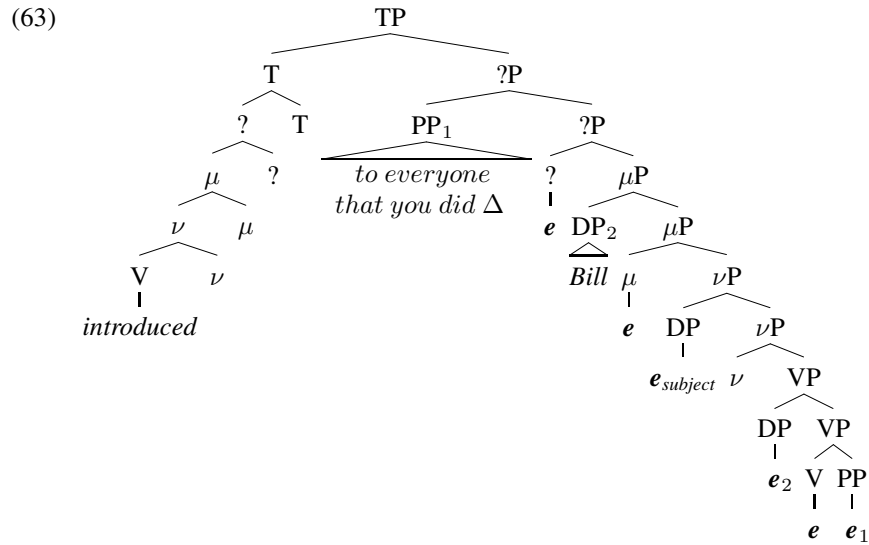
- (61) a. John introduced Bill to everyone that you did  $\Delta$ .  
 b. A different girl introduced Bill to every boy.  
 c. Betsy told me about everything that you did  $\Delta$ .  
 d. A different girl told me about every cat.

As before, the apparent instances of ACD in these cases require that we find a parse that produces VPs of the form “introduced Bill to  $e$ ” and “told me about  $e$ ” to serve as antecedents for the ellipses. In these cases, however, the ACD is contained not in a direct object, but in an indirect object. So, we cannot exploit in these situations the movement to the accusative Case marked spot to resolve these ellipses. Similarly, we need to let the *every*-phrases have a *different girl* in their scope if *different* is to get the reading licensed by the universal quantifier. But, again, this cannot be because these terms left adjoined to  $\mu$ P.

In fact, Hornstein proposes that indirect objects *do* move in the same way that accusative Case-marked objects do, but into a position in which “indirect object Case” is licensed. To illustrate: this would give (61a) a representation for like that in (62). (Understand ?P to mark the position in which indirect object Case is marked.)



If Hornstein's view is to be made workable in view of the locality condition on Argument Movement that we adopted from Chomsky, we are going to have to be careful where ?P is placed. Hornstein supposes that ?P is above the surface position of the accusative Cased-marked position, and so if the underlying position for subjects is beneath  $\mu$ P, we might get a representation for (61a) something like (63).



But the movement of the indirect object into Specifier of ?P indicated in (63) would violate the locality condition on A movement that we have adopted. Recall that that condition prohibits A movement of a term past more than one filled Specifier position, and that is precisely what has happened in (63). The only way to avoid this is to deny that the underlying position of subjects is beneath ?P — then movement to Specifier of ?P would pass only Specifier of  $\mu$ P — but this would give us no way of accounting for how *a different girl* in (61b) manages to fall within the scope of *to every boy*. Hornstein's account, then, would require some rethinking of the locality condition on A Movement.

Further, if indirect objects move leftwards in this fashion to form S-structures, we will have to rethink our method of ensuring that only accusative DPs can surface to the left of particles. Recall that we accounted for this fact by restricting the range of things that can move leftwards to just accusative DPs.

Here, then, are two difficulties with Hornstein's revision that arise when we embed them into our present theory of word-order.

But there are harder problems for Hornstein's proposal from examples such as (64) and (65).

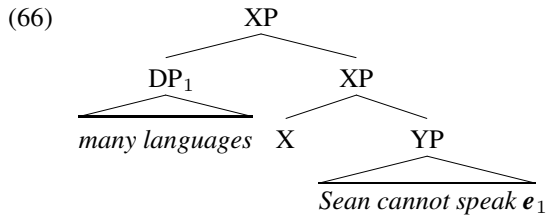
(64) Sean can't speak many languages.

(65) She's bought a book about every subject.

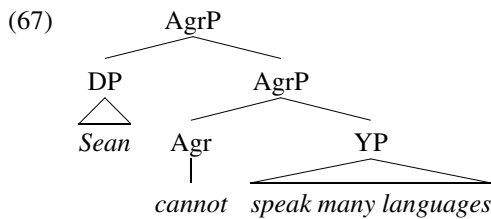
The sentence in (64) is ambiguous. On one reading, *many languages* is outside the scope of *can't*. On this interpretation, (64) says that the number of languages which



Sean can't speak is large. To get this interpretation, we need to find an LF that has the shape roughly in (66).

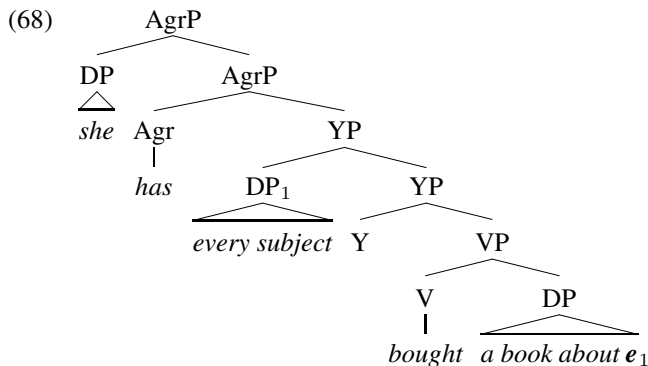


On the other reading, *many languages* is inside the scope of *can't*; on this interpretation, (64) says that the number of languages that Sean can speak is small. For this interpretation, we need an LF like that in (67).



On Hornstein's view, the object's scope is fixed in the position it gets Case from. But this example suggests that an object can have two scope positions, and, moreover, that one of those is higher than the position of negation and modals. This isn't what's expected, then.

The problem for Hornstein in (65) is that this sentence has as its most natural interpretation one in which *every subject* has wider scope than the object it is within. That is, the values given to *a book* can vary as a function of the values given to *every subject*. This suggests we need to find for (65) an LF like that in (68).



Cases such as these — one's in which the contents of a DP scope out of that DP — are called instances of “inverse linking.” That they call for an LF with the form indicated by (68), we've already seen in the guise of ACD in (41).

LFs such as these do not look like they could be the products of A Movement. So far as I know, there are no instances of A Movement that relocate a term out of a DP,

as has happened in these examples; and, moreover, in (68), the DP that has moved has relocated from a Case marked position, something we have seen earlier is not possible for A Movement.

It's reasonably clear, then, that A Movement cannot express all of the scopal ambiguities that arise. We will need an independent process of QR. Nonetheless, we can learn an important lessons from Hornstein's proposal: some of the scopal ambiguities we have examined — namely those that involve accusative DPs — can be resolved in terms of the position they A Move into.

Let me sketch a suggestion here that will have the desired consequences, and that makes use of some ideas in the literature. The first part of the suggestion is to adopt a proposal of Chomsky's concerning the interpretation of derivations. We have seen that, for a certain range of cases, the principles we have adopted up to this point produce elements of a derivation that would be suitable for deriving the semantic interpretation. As we've seen, for this range of cases, we seem to want to let the semantic component use these representations to form the interpretation. Chomsky proposes a method of implementing this idea that still preserves the hypothesis that the final member of the derivation is the only phrase marker that is interpreted. His idea changes how movement is defined, so that it does not actually **change** the position of a term, instead it merely **adds** positions to that term. On this view, then, Another way of meeting these goals is to understand movement as a two part operation: one which copies the term being moved into a new position, and another operation which determines which copy is pronounced:

(69) Move  $\alpha =_{def}$  copy  $\alpha$  into a new position.

Pronounce  $\alpha$ : in a chain of copies, pronounce the one that c-commands all the others.

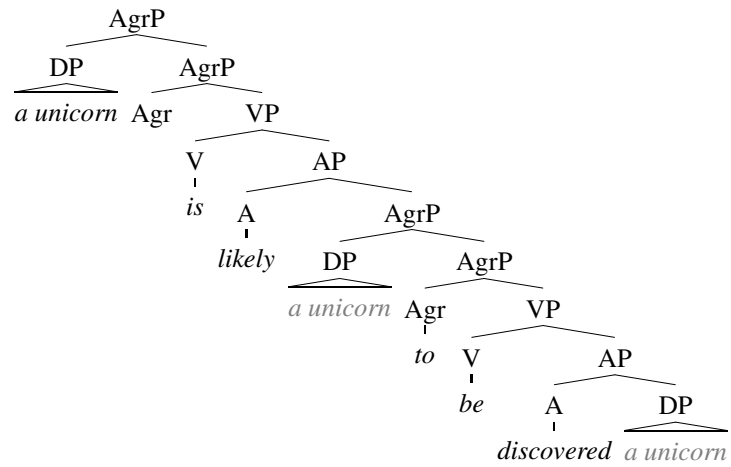
With this interpretation of "movement," namely one in which it generates a menagerie of copies of some phrase, we must ensure that we do not simultaneously semantically produce a series of copies. That is, we don't literally want to produce copies of the affected term; what we want is merely to generate images of a single term in a variety of positions. We need a way of letting all these copies be interpreted semantically only once. This can be done with (70).

(70) Exactly one copy may be semantically interpreted.

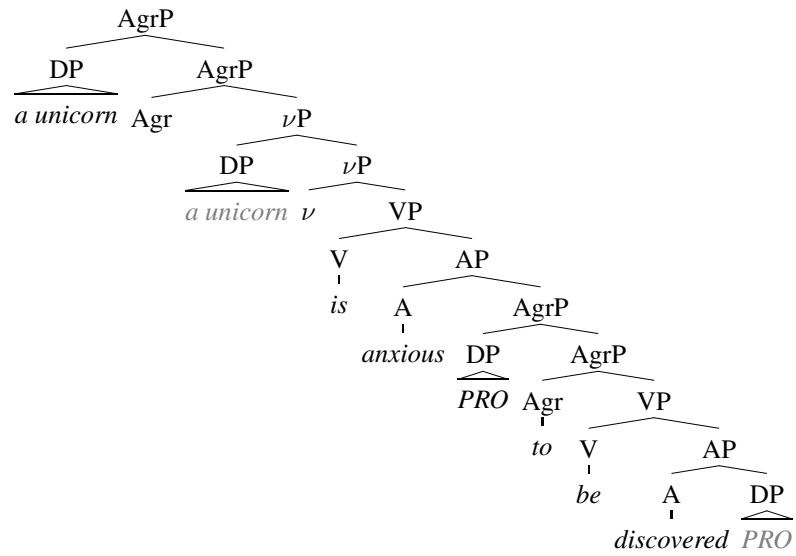
This, in fact, is the implementation of the idea that any representation in a derivation can be semantically interpreted that Hornstein adopts.

Now consider what this interpretation of movement will do in those cases we examined earlier involving quantifier lowering. Holding constant the Projection Principle, the examples in (19) will yield S-structure representations like those in (71) on the facing page. From here on out, let's adopt the convention, employed in these representations, of shading the copies that are not pronounced. Because *a unicorn* has moved out of the infinitive in the first case, but not the second, this new interpretation of movement will allow a copy of *a unicorn* to be inside the first infinitive but not the second. When this copy is semantically interpreted, we get the quantifier lowered reading. In this way, then, this construal of movement mimics the effects of QR in those situations in which the quantifier is interpreted in a position lower than its surface one.

(71) a.

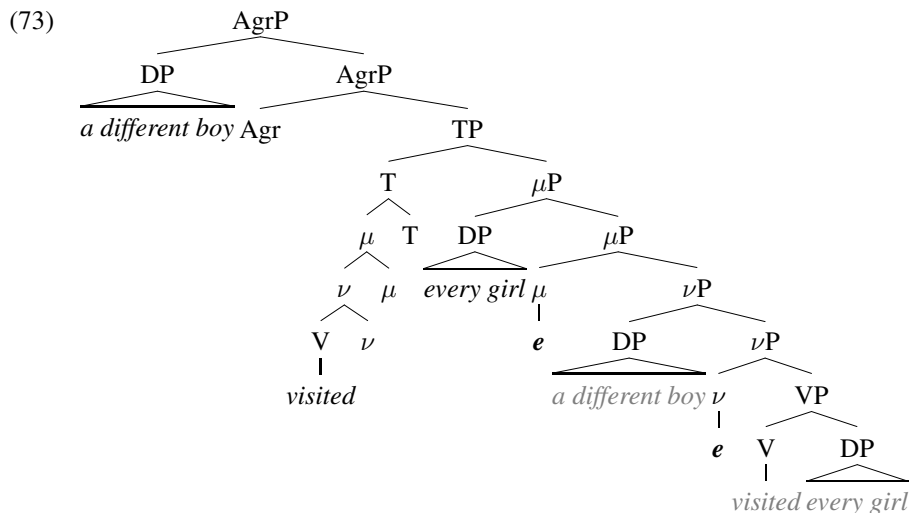


b.



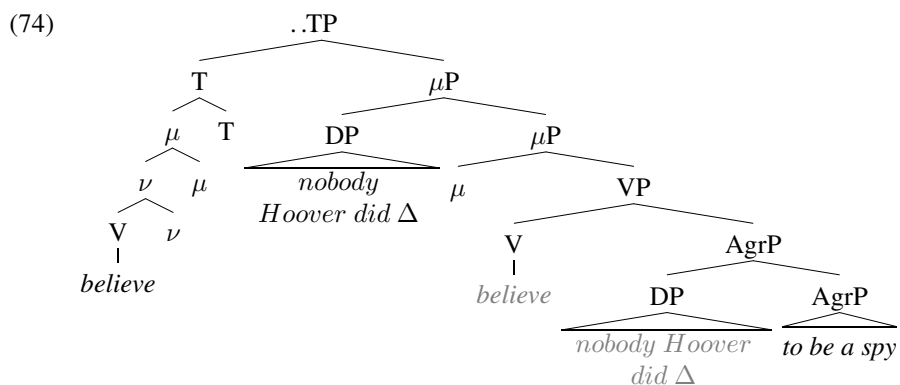
To account for simple cases where the subject appears to fall within the scope of the object, this method can be combined with our model of accusative Case assignment to produce the relevant LFs. Thus, for example, combining these ideas would give to the sentence in (72) the surface representation in (73).

(72) A different boy visited every girl.



If the higher copy of *every girl* and the lower copy of *a different boy* are semantically interpreted we will get the reading in which *a different boy* falls within the scope of *a every girl*.

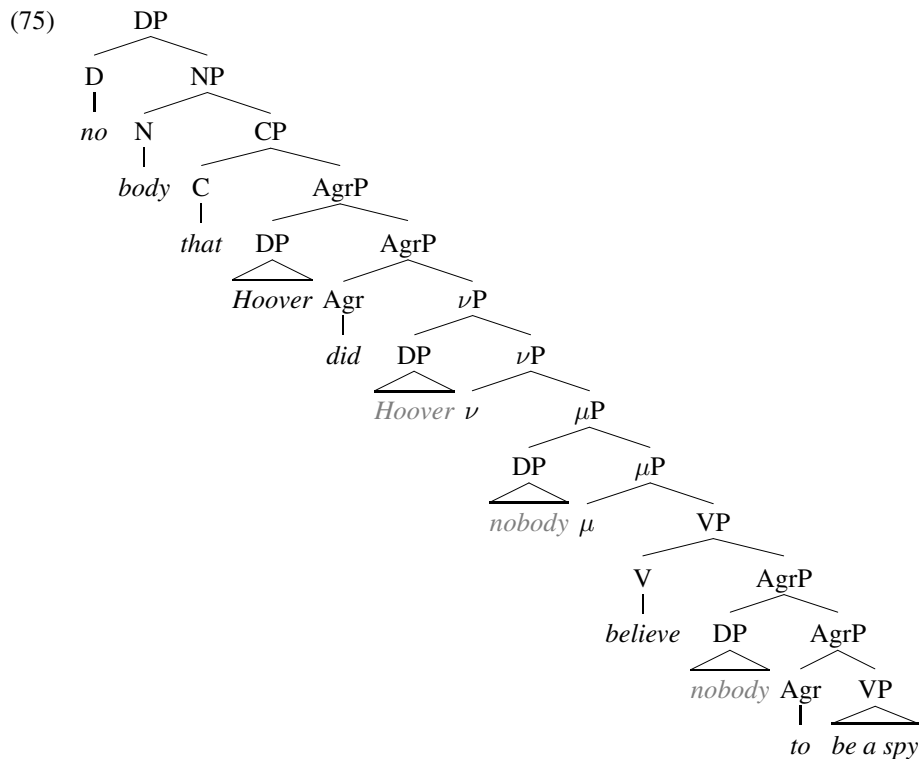
Consider finally the instances of ACD we've been investigating. These cases will require an additional step. Let's concentrate on the case in (47), which illustrates the dependence on ACD resolution to scope. Under the present set of assumptions, this example will get the representation in (74).



In forming this parse, I have suppressed considerable detail, including the presence of another copy of the *nobody Hoover did* phrase within the embedded VP. I've also adopted the convention of using our old typological device of "e" in those positions

where a copy will not be interpreted. In this case, those are the positions where  $\mu$  and  $\nu$  originate. We'll return immediately to why I decide that these are not positions in which the moved verb can be interpreted.

The elided VP in this example includes a structure in which movement has applied, and its syntax will therefore be affected by the decision to render movement operations in terms of copies. If we assume that relative clauses are formed by moving the noun that they modify out of the clause, then the *nobody Hoover did* DP will have the structure in (75).



So, we need to find a way of letting the two VPs in (76) be equivalent.

- (76) a. [<sub>VP</sub> believe *nobody* to be a spy ]  
 b. [<sub>VP</sub> believe *nobody Hoover does* to be a spy ]

Part of the solution to this dilemma must involve the fact that the *nobody* and *nobody Hoover does* copies in each of these VPs are not semantically interpreted as introducing the quantification. In the relative clause it is clear that the *nobody* is interpreted as the “head” of the DP that the relative modifies. And the *nobody Hoover does* DP is semantically interpreted in a rather high position; recall that the point of this example was to illustrate that *nobody Hoover does* is interpreted in a position that places it higher than the *before* clause in (45), repeated below.

- (45) Dulles believed nobody to be a spy before Angleton did Δ.

If we let these copies be interpreted as variables, or some other semantic term that is anaphoric on a higher copy, then we might be able to fit the two VPs in (76) to the identity condition on VP Ellipsis in (32).

The proper way of achieving this remains somewhat controversial. It appears that we do not want to render these copies as variables because of a kind of example discussed in an unpublished paper by Chris Kennedy.<sup>4</sup> What's central to Kennedy's examples is that they put copies that are very distinct semantically in the positions that the *nobody* DPs occupy in (76); (77) is one such example.

(77) \*Polly visited every town near a lake that Eric did  $\Delta$ .

This example is ungrammatical in part because it does not have the interpretation in (78).

(78) Polly visited every town near a lake that Eric visited.

This means that moving the *every town* DP out of the VP headed by *visited* does not seem capable of creating an antecedent of the form necessary to elide *visit x*. Keeping the copy-theory of movement in mind, the representation the two VPs would get is (79).

- (79) a. [<sub>VP</sub> visit every town near a lake that Eric did ]  
       b. [<sub>VP</sub> visit a lake ]

If we were permitted to interpret copies as simply variables, then these two VPs should be able to both be interpreted as *visit x*, and the ellipsis should be licensed. We want the fact that *lake* and *town* do not match to have a cost in this case.

To see that it really is the mismatch between *town* and *lake* that is responsible for the ill-formedness of (77), compare it to the minimally different (80), and example discussed by Uli Sauerland in a SALT paper.

(80) Polly visited every town near the one that Eric did  $\Delta$ .

This is grammatical with the interpretation in (81), and therefore it must be possible for the VPs in (82) to count as equivalent.

(81) Polly visited every town near the one that Eric visited.

- (82) a. [<sub>VP</sub> visited every town near the one that Eric visited ]  
       b. [<sub>VP</sub> visited the one ]

Sauerland even reports that examples such as (83) do not merit the badness that (78) has, which suggests that the VPs in (84) might come close to being equivalent.

(83) ??Polly visited every city near the town that Eric did.

- (84) a. [<sub>VP</sub> visited every city near the town that Eric did ]  
       b. [<sub>VP</sub> visited the town ]

<sup>4</sup>See Heim (1997) for discussion of the paper, and for an examination of the paradox.

There is much left to be done, then, in understanding how to describe the conditions that match an elided VP and its antecedent in situations where they contain different copies. This remains an open problem, so far as I know. Let us cheat, then, and adopt the too simple solution that these copies get interpreted as simple variables. There are a variety of ways to achieve this end.

Let's use a modification of the trace theory we've already sketched. If we define the LF parse as the phrase marker in a derivation that is semantically interpreted, we can do this with the following system.

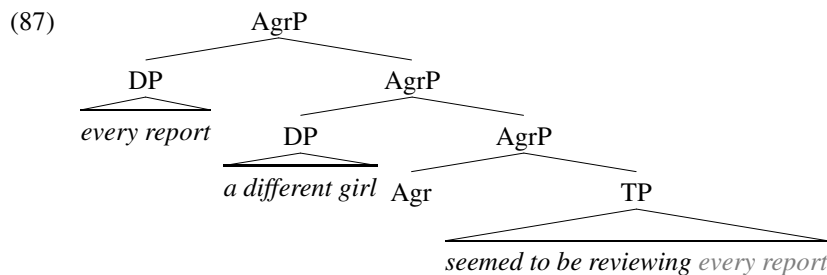
- (85)
- a. Move  $\alpha =_{def}$  Copy  $\alpha$  into a new position.
  - b. Delete every copy but one by LF.
  - c. Insert a "e," and interpret it as a variable bound to the term whose copy it replaces.

For the antecedent VP in (74) to be the same as what has elided, then, it will be necessary to delete the lower two copies of *nobody Hoover did* and replace these with traces. Once these copies are deleted, only the higher one will survive to LF, and this is the one that is interpreted. This will preserve the correlation between scope and ACD resolution that we discovered in connection with this example.

With this copy theory of movement in place, we can now return to the problem we started with: how are we to satisfy Hornstein's arguments against QR and at the same time capture the cases which seem to require QR. The situations in which QR seems to be prevented from applying are just those examples in which understanding the subject in a lower position would be sufficient. For example, recall that in (86), it appears that for the subject to fall within the scope of the object it must be interpreted within the embedded infinitival.

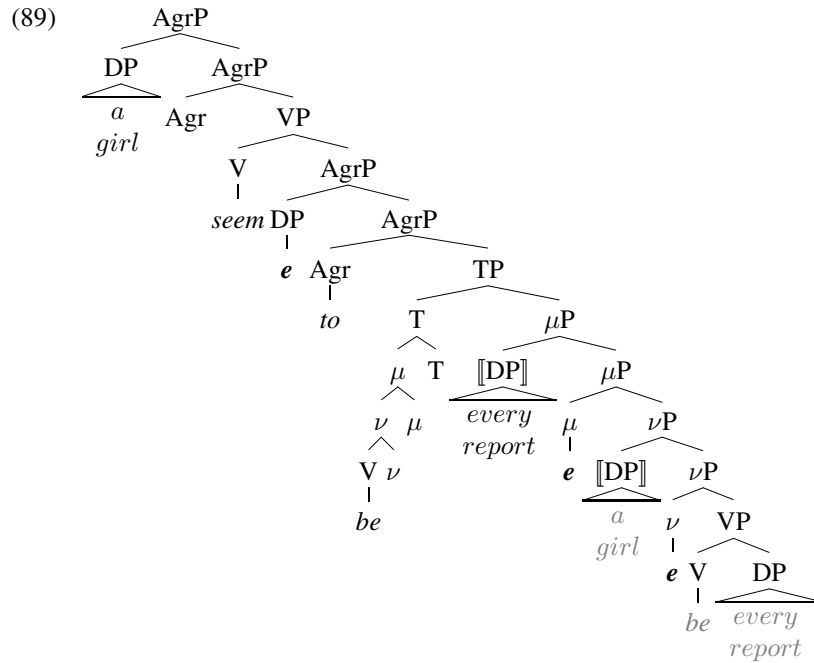
- (86) A different girl seemed to be reviewing every report.

Here's an example where QR must be prevented from producing an LF like that in (87), for otherwise it should be possible for *a different girl* to bind the pronoun in (88) when it falls within the scope of the object.



- (88) \*A girl<sub>1</sub> seemed to her<sub>1</sub> boss to be reviewing every report.

Instead, the only way the subject seems capable of being interpreted within the scope of the object is for it to be interpreted in one of the positions it has gathered as a function of the copy operation. That is, this interpretation arises just when the lowest copy of *a girl* is the one that gets semantically interpreted (indicated here with the "interpretation function" brackets: "[ ]"), in the LF:



What we see in this case, and all of the other situations in which QR seems to be blocked, is that the interpretation QR would have provided is available already by virtue of the copies.

By contrast, the situations in which QR seemed to be necessary were just those in which the derivations we've already devised did not provide an LF with the relevant interpretation. These were all cases, in other words, in which only QR could have put the relevant DP into its interpretable position.

Thus, a way of characterizing when QR is possible and when it is blocked goes as follows: only QR to get some interpretation when it isn't already made available. This can be expressed as a generalization of Economize:<sup>5</sup>

- (90) With respect to a semantic interpretation,  $\psi$ , form the smallest derivation that produces a phrase marker with  $\psi$ .

We used Economize previously to minimize the size of the derivation as it produces the phonologically interpretation. In particular, in examining instances of covert verb movement, we invoked Economize to favor derivations that move things after the sentence is phonologically interpreted rather than moving them before they are interpreted. We could state this in a way that parallels (90) with no loss of coverage:

- (91) With respect to a phonological interpretation,  $\phi$ , form the smallest derivation that produces a phrase marker with  $\phi$ .

Because the derivation of the phrase marker with  $\phi$  is always a subset of the derivation of the phrase marker with  $\psi$ , these two conditions do not compete with each other.

<sup>5</sup>This is essentially what Fox (1995) argues for, though on different grounds.



In particular, once the derivation to  $\psi$  is minimized, there is no minimization of the derivation to  $\phi$  that will add to  $\psi$ 's derivation. We can put these two statements together, now, with (92).

(92) ECONOMY

Let the selection of lexical items from which a sentence is to be constructed be  $\mathcal{N}$  (the “Numeration”). Minimize the derivation from  $\mathcal{N}$  to the phrase markers that yield  $\psi$  and  $\phi$ .

Although this will correctly prevent QR except where it is the only way to give an interpretation, and therefore navigates both the evidence for and against QR that we have seen, it isn't entirely accurate yet. Because it measures Economy relative to a semantic or phonological interpretation, it does not compare derivations that lead to different interpretations. As a consequence, it doesn't yet prevent the interpretation that is represented in (88). The smallest derivation that leads to an interpretation of this sentence in which the pronoun is bound by the subject and the subject is within the object is in fact one that invokes QR. Because there is no way of getting to an LF with this interpretation without QR, Economy should allow QR to apply and we should therefore expect this interpretation to be available.

What's needed is a way of individuating the meanings that Economy compares — the  $\psi$  in (91) — so that the bound variable interpretation of pronouns is not involved. That is, we want both the interpretation in (93a) and the interpretation in (93b) to be the same  $\psi$  for Economy.

- (93) a. A girl<sub>1</sub> seemed to  $x_1$ 's boss to be reviewing every report<sub>2</sub>.  
b. A girl<sub>1</sub> seemed to her<sub>2</sub> boss to be reviewing every report<sub>3</sub>.

How to do this, or even whether it can be done, involves caching out what exactly the “semantic interpretations,” i.e.,  $\psi$ s, are. There is no consensus on this that I know of — indeed, it is a question that has only just barely been formulated presently. We must leave this problem unsolved, then.

There is another problem that we will have to set aside as well, and this one too involves how the system we've just developed controls the interpretations that pronouns can get. This problem arises in the special cases of pronouns that require antecedents within their sentences — so-called anaphors. The behavior of anaphors was also one of our diagnostics for scope in our examination of complement structure. As we shall see in greater detail in the following chapter, anaphors must be found within the scope of their antecedent. This is demonstrated for reciprocals, one kind of anaphor found in English, by contrasts like those in (94).

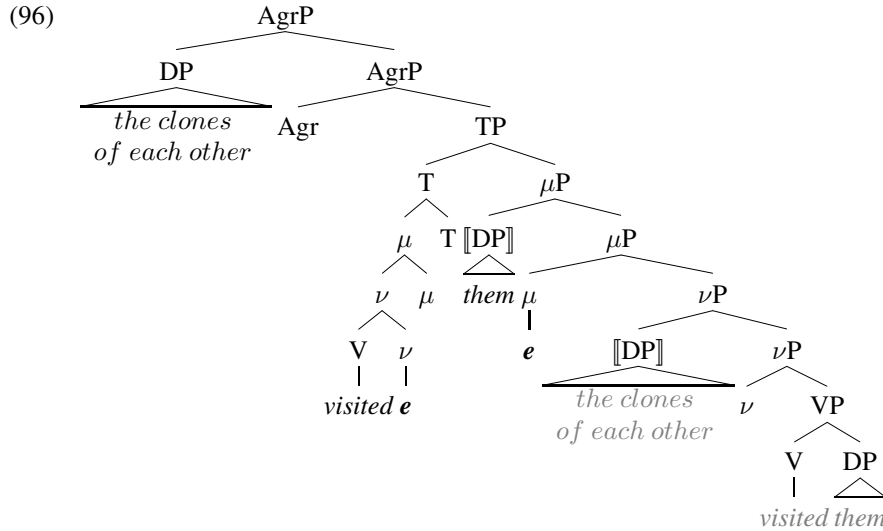
- (94) a. The children<sub>1</sub> like each other<sub>1</sub>.  
b. \* The children<sub>1</sub>'s mother likes each other<sub>1</sub>.

The ungrammaticality of (94b) stems from the fact that *the children* is buried within the subject and therefore does not c-command *each other*. As a consequence, *each other* is not within the scope of *the children* in (94b), and the example is ungrammatical. By contrast, *each other* is c-commanded by *the children* in (94a), and therefore satisfies its requirement that it fall within the scope of its antecedent.

To see the problem that arises with anaphors, and our present treatment of scope, consider a sentence like (95).

(95) The clones of each other visited them.

Under our present assumptions, this sentence should have an LF representation like (96).



If the object is interpreted in its higher position, and the subject in its lower position, i.e. if the DPs within “ $\llbracket \ ]$ ” are interpreted, then *each other* should fall within the scope of *them*. This leads to the expectation that *them* should be a potential antecedent for *each other* in (95); but, in fact, this sentence is ungrammatical precisely because *them* can't be an antecedent to *each other*.

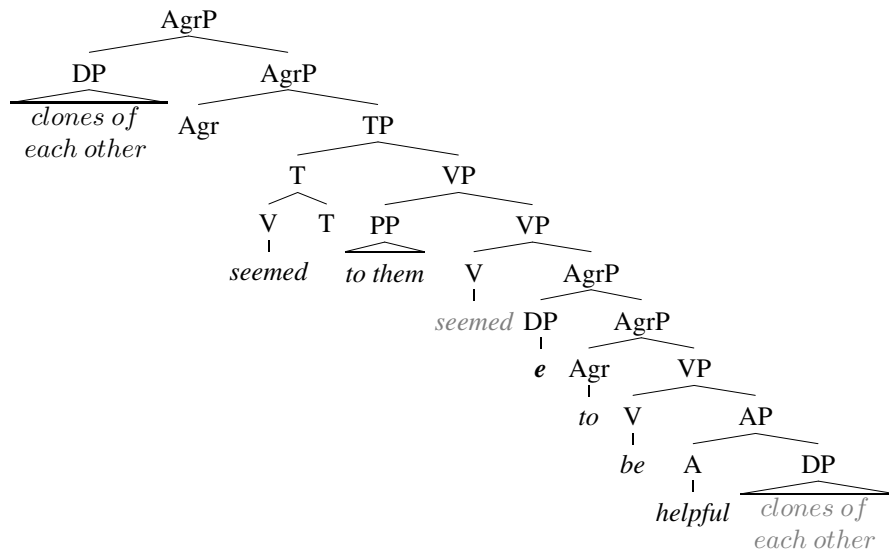
What makes this problem especially difficult to solve is that there are other cases in which the present system seems to make the correct prediction about the antecedents for anaphors. In (97), for instance, it does seem that *them* can be the antecedent to *each other*.

(97) Clones of each other seemed to them to be helpful.

Under the present system, this sentence will have an LF like that in (98) on the next page. It would seem from this example that the system we have devised is necessary, for only if the subject can be interpreted within the lower clause can the reciprocal fall within the scope of *them*. There is something of a paradox, then, in these two examples; (95) seems to lead to the conclusion that we do not want anaphors to be able to find antecedents under “quantifier lowering,” or “reconstruction,” as it is sometimes called, whereas (98) seems to suggest that we do.

This is a problem that is restricted to anaphors, and so we might imagine that there is some property peculiar to them that is responsible for their duplicitous behavior. In the next chapter we'll take a closer look at the behavior of anaphors.

(98)





## Chapter 11

# Introduction to the Binding Theory

In our discussion of the principles that determine where arguments are placed within syntactic representations, one of the cases we spent some time examining is the “silent” argument found in infinitives. These invisible DPs have a very narrow distribution, as we’ve now seen in some detail. They can only occur in the Specifiers of some non-finite clauses, as the paradigm below indicates.

- (1) a. Samantha tried [PRO to like natto].
- b. \*Samantha tried [PRO to like PRO].
- c. \*Samantha tried [PRO to leave with PRO].
- d. \*PRO tried [PRO to like natto].

In (1b-c), there is no hidden term, like there is in (1a).

This hidden term has a particular range of interpretations, by which it can be recognized. It is either understood to be obligatorily the same in reference as some other term in the sentence, as in the case in (1a), where PRO refers to the same individual that *Samantha* does. It can also get a “generic” interpretation, as in the following example.

- (2) Samantha wondered how PRO to eat natto.

PRO under its first interpretation is known as “Controlled”; and in its second use it is called “Arbitrary.”

Of course, one should be skeptical about the existence of things which do not impinge on the senses. And so, we might entertain the thesis that there is in fact, just as we see, nothing in the Specifier position of these non-finite clauses. Of course, we must still find an explanation for the differences that (1) illustrates. But now the task could fall under a different description: why are only certain “subjects” able to go unclothed by a syntactic realization.

Note that there is a related phenomena that we should view this matter in the light of. Recall that we have discovered that the Specifiers of AgrP are positions that

always surface with something in them. But if there are no subjects present in the syntactic representations of (1) and (2), then we will have to find a way of accounting for why this is possible for finite clauses but not non-finite ones. And this method should connect with our explanation for why only external  $\theta$ -roles can be unexpressed in the syntax. On the view that there is something syntactically present in the Specifier position of the non-finite clauses, we should discover why these “silent” subjects are possible only in non-finite rather than finite clauses. In the next couple days I want to present the case for the view that there is a subject in these situations. And so our task will come down to figuring out both why these subjects can appear only in Specifier positions, and why they may appear only in non-finite clauses. For the other direction, one that seeks an explanation for this paradigm under the view that there is no syntactic subject in these cases, see Chierchia (1984).

The evidence that I will provide comes by way of the influence these “silent” subjects have on the referential dependencies carried among the terms in the sentences in which they are found. So we begin by examining some of the factors that seem to be involved in the syntax of these referential dependencies. This aspect of syntax goes by the name “Binding Theory.”

The kinds of referential dependencies that we are interested in can be illustrated with cases like the following.

- (3) a. The woman likes herself.
- b. The women like each other.
- c. The woman likes the woman.

The cases in (3a) and (3b) differ from those in (3c) and (3d) with respect to the referential relationship that is asserted between subject and object. In (3a) and (3b), the message that subject and object have the same referent is conveyed, (3c) and (3d) convey the message that subject and object have different referents. There are two popular methods for representing this difference. In one, advocated by Chomsky (1981) which follows an idea flowing from Lasnik (1976), every argument is adorned with a “referential index.” We can summarize this method in the following way.

- (4) a. If  $\alpha$  and  $\beta$  bear the same referential index, understand the sentence they stand in to assert that they have the same referent.
- b. If  $\alpha$  and  $\beta$  bear different referential indices, understand the sentence they stand in to presuppose that they have different referents.

Another method, one championed by Reinhart (1983) and Higginbotham (1983) places (the same) indices on arguments that have a bound variable interpretation, but otherwise does not represent referential dependencies by way of a syntactic device. On this view, the non-coreference that is conveyed in (3c) and (3d) comes by way of a pragmatic inference. Something like the following, perhaps, will work.<sup>1</sup>

- (5) a. If  $\alpha$  and  $\beta$  bear the same referential index, understand the sentence they stand in to assert that one’s referent derives from the other.

<sup>1</sup>There are a number of ways of expressing this view; the one (5) is based on comes closest to a proposal in Fiengo and May (1994).

- b. Understand two arguments to have different referents, unless their coreference is asserted by the sentence they stand in.

In what follows, let us adopt the first method for representing the differences in (3), though keep in the back of your mind the other method. If both the “coreferent” and “non-coreferent” readings map on to indexation relations, then we seek a theory that explains which distribution of indices are grammatical and which are not.

Our first observation is that the relation of c-command appears to be critical in determining these indexation patterns. Compare (3) with (6).

- (6) a. \* [The woman<sub>1</sub>’s father]<sub>2</sub> likes herself<sub>1</sub>.  
 b. \* [The women<sub>1</sub>’s father]<sub>2</sub> likes each other<sub>1</sub>.  
 c. [The woman<sub>1</sub>’s father]<sub>2</sub> likes her<sub>1</sub>.  
 d. [The woman<sub>1</sub>’s father]<sub>2</sub> likes the woman<sub>1</sub>.

Whereas in (3a/b) coreference could be asserted, i.e., *the woman* could be coindexed with *herself* or *each other*, this is not the case in (6a,b). And, where neither (3c) nor (3d) could assert coreference between *the woman* and *her* or *the woman*, i.e., these terms could not be coindexed, this is possible in (6c,d). Note further than when *herself* or *each other* is unable to be understood as coreferent, the sentence falls into ungrammaticality. A similar contrast emerges in (7).

- (7) a. \* Clones of herself<sub>1</sub> liked the woman<sub>1</sub>.  
 b. \* Each other<sub>1</sub>’s fathers liked the women<sub>1</sub>.  
 c. Her<sub>1</sub> father liked the woman<sub>1</sub>.  
 d. The woman<sub>1</sub>’s father liked the woman<sub>1</sub>.  
 (8) a. \* I believe herself<sub>1</sub> to like the woman<sub>1</sub>.  
 b. \* I believe each other<sub>1</sub> to like the women<sub>1</sub>.

In these cases too, coindexation is permitted (7c) and (7d), but blocked in (7a) and (7b).

In general, what appears to be true is that reflexive pronouns and the reciprocal, which together are sometimes referred to as “anaphors,” must have “antecedents,” that is: they must be coindexed with terms that are prominent to them.<sup>2</sup> It appears that prominence is defined in terms of the geometrical relations that, in particular, c-command, defined again below, seems to be the relevant relation.

- (9)  $\alpha$  c-commands  $\beta$  iff every phrase containing  $\alpha$  contains  $\beta$  and  $\alpha$  doesn’t dominate  $\beta$ .

Now we can define the condition governing anaphors as:

<sup>2</sup>Some have attempted to define “prominence” in terms of the argument structures of the predicates involved: so, for example, we might imagine that  $\alpha$  is prominent to  $\beta$  iff  $\alpha$  and  $\beta$  are arguments of the same predicate and  $\alpha$  bears a  $\theta$ -role higher on the list in (1) than  $\beta$ ’s  $\theta$ -role.

- (1) Agent  $\zeta$  Cause  $\zeta$  Theme  $\zeta$  Goal  $\zeta$  Location

This idea has an early champion in Jackendoff (1972), and has been pursued most recently in a series of papers by Edwin Williams.

- (10) An anaphor must be BOUND.  $\alpha$  is BOUND by  $\beta$  iff  $\beta$  c-commands  $\alpha$  and they are coindexed.

So, BINDS collapses the prominence requirement with the antecedence requirement, and (10) requires both of these for anaphors. It looks like the indices borne by personal pronouns and “full” DPs have the complementary requirement, so we have:

- (11) a. Pronouns must not be bound.  
b. “Referential DPs” must not be bound.

These conditions then allow pronouns and referential DPs, sometimes abbreviated to “r-expressions,” to be coindexed (hence coreferential) only with terms that don’t c-command them. These three conditions are called Principles A, B and C in *Lectures on Government and Binding*.

Consider next (12).

- (12) a. The woman<sub>1</sub> said that Martha<sub>2</sub> likes her<sub>1</sub>.  
b. \* The woman<sub>1</sub> said that Martha<sub>2</sub> likes the woman<sub>1</sub>.  
c. \* The women<sub>1</sub> said that Martha<sub>2</sub> likes each other<sub>1</sub>.  
(13) a. The woman<sub>1</sub> considers Martha<sub>2</sub> angry at her<sub>1</sub>.  
b. \* The woman<sub>1</sub> considers Martha<sub>2</sub> angry at the woman<sub>1</sub>.  
c. \* The woman<sub>1</sub> considers Martha<sub>2</sub> angry at herself<sub>1</sub>.  
d. \* The women<sub>1</sub> consider Martha<sub>2</sub> angry at each other<sub>1</sub>.  
(14) a. The woman<sub>1</sub> bought Martha<sub>2</sub>’s picture of her<sub>1</sub>.  
b. \* The woman<sub>1</sub> bought Martha<sub>2</sub>’s picture of the woman<sub>1</sub>.  
c. \* The woman<sub>1</sub> bought Martha<sub>2</sub>’s picture of herself<sub>1</sub>.  
d. \* The women<sub>1</sub> bought Martha<sub>2</sub>’s picture of each other<sub>1</sub>.

These cases indicate that there is a locality constraint on Principles A and B that is not found for Principle C. Further, these three sets of data make it appear that the same locality constraint governs both Principles A and B. Also notice that *Martha* is close enough to the pronoun or anaphor for Principles A and B to be satisfied. So the locality condition involved seems to carve out the smallest maximal projection containing the anaphor or pronoun as the region in which Principles A and B must be satisfied. Further, if (14) is compared to (15), it looks like the presence of the intervening argument determines whether this smallest XP counts in the locality condition.

- (15) a. ?? The woman<sub>1</sub> bought a picture of her<sub>1</sub>.  
b. The woman<sub>1</sub> bought a picture of herself<sub>1</sub>.  
c. The women<sub>1</sub> bought pictures of each other<sub>1</sub>.

Let’s call the region in which Principles A and B apply, a BINDING DOMAIN. Then what we search for is a definition of Binding Domain which captures the facts just presented. One idea would be to define Binding Domain as the first phrase that contains the anaphor or pronoun and a potential antecedent for that term. This would capture the dependence on the genitive that Binding Domains appear to have, but it would have to be fixed so as to correctly apply to cases like (16).



- (16) a. The woman<sub>1</sub> told Martha<sub>2</sub> about herself<sub>1</sub>.  
 b. \* The woman<sub>1</sub> told Martha<sub>2</sub> about her<sub>1</sub>.

One response to these cases has been to exploit the fact that the potential antecedent in (16) is an object, whereas the terms in (12)-(14) which have this effect are, loosely speaking, subjects. Chomsky's *Conditions on Transformations* introduces a method for defining Binding Domains that rests on the hypothesis that this is the relevant difference in these cases. The constraint he devised was meant to govern a whole host of processes — for example, it was used to govern the distance that A Movement can relocate things, as we've seen — it said roughly that the processes it applies to cannot span a distance that includes what he called a "Specified Subject," and the constraint he named "The Specified Subject Constraint." I'll present the more recent version of this approach, fashioning my discussion on the presentation in *Lectures on Government and Binding*.

Our first step, then, might be as follows.

- (17) THE PRINCIPLES  
 A. An anaphor must be bound in its Binding Domain.  
 B. A pronoun must be free in its Binding Domain.  
 C. An r-expression must be free.

$\alpha$  is BOUND in  $\Phi$  iff there is a  $\beta$  contained in  $\Phi$  that c-commands and is coindexed with  $\alpha$ .

$\alpha$  is FREE in  $\Phi$  iff  $\alpha$  is not bound in  $\Phi$ .

$\Phi$  is a BINDING DOMAIN for  $\alpha$ , iff  $\Phi$  is the smallest phrase that contains  $\alpha$  and a SUBJECT.

A SUBJECT =<sub>def</sub> an argument in a position to which an external  $\theta$ -role can be assigned.

It's not just any subject, however, that creates a Binding Domain; rather, as (18) indicates, it's only subjects which c-command the term whose Binding Domain being computed that have that effect.

- (18) The woman<sub>1</sub> told [Martha<sub>2</sub>'s parents]<sub>3</sub> about herself<sub>1</sub>.

According to our definitions above, VP should be the Binding Domain for *herself* in this example because it contains the subject *Martha*. But that would wrongly mark this sentence ungrammatical, since the antecedent for *herself* is not within this VP. So let's make this change.

- (19)  $\Phi$  is a BINDING DOMAIN for  $\alpha$ , iff  $\Phi$  is the smallest phrase that contains  $\alpha$  and a SUBJECT with c-commands  $\alpha$ .

The decision to let Principles A and B be governed by the same locality constraint screws up in a couple places in English. One of those is (20).

- (20) a. The woman<sub>1</sub> likes [her<sub>1</sub> father]<sub>2</sub>.  
 b. The women<sub>1</sub> like [each other<sub>1</sub>'s fathers]<sub>2</sub>.

In *Lectures on Government and Binding*, it is decided that (20b) should be seen as “normal,” and (20a) as somehow the exceptional case. There is some plausibility to this, as in many languages the correlate to (20a) would be ungrammatical. In fact, (20b) follows if we understand dominance to be reflexive.

- (21) Everything dominates itself.

Because dominance precludes c-command, this will prevent the subjects in (20) from making the DPs which contain the pronoun or anaphor Binding Domains. As for (20a) Chomsky suggests that the English genitive reflexive is homophonous with the genitive personal pronoun; that is, he relates this to the unavailability in English of forms like *herself's*. Indeed, in languages where (20a) is ungrammatical, there is generally a genitive form of the reflexive pronoun (as in the mainland Scandinavian languages).

This decision creates a problem, however, in cases like the (22).

- (22) \*Pictures of each other were on sale.

Because Principle A puts the requirement that anaphors must have antecedents within the scope of the Binding Domain constraint, it will not require anaphors to have antecedents in cases where no Binding Domain is defined. This is one such case: the only subject in this example is the entire subject *pictures of each other*, and this DP doesn't c-command *each other* because it dominates it. The response to this problem in *Lectures on Government and Binding* is to set up a default clause which guarantees that every term has a Binding Domain.

- (23)  $\Phi$  is a BINDING DOMAIN for  $\alpha$ , iff  $\Phi$  is the smallest phrase that contains  $\alpha$  and:  
 a. a SUBJECT which c-commands  $\alpha$ , or  
 b.  $\Phi$  is the root AgrP.

Another possibility, of course is to reframe Principle A so that the requirement that anaphors have antecedents is not within the scope of the locality condition. But, as we shall see, Chomsky has a plan for this solution.

Consider now the contrasts in (24) and (25).

- (24) a. \* The women<sub>1</sub> believe that each other<sub>1</sub> are intelligent.  
 b. The women<sub>1</sub> believe each other<sub>1</sub> to be intelligent.  
 c. The women<sub>1</sub> consider each other<sub>1</sub> intelligent.  
 (25) a. The woman<sub>1</sub> believes that she<sub>1</sub> is intelligent.  
 b. \* The woman<sub>1</sub> believes her<sub>1</sub> to be intelligent.  
 c. \* The woman<sub>1</sub> considers her<sub>1</sub> intelligent.

Our present system does not correctly distinguish the a-cases from the others. There are four surface differences we might exploit in searching for a characterization of these

contrasts. In the a-cases there is agreement morphology within the “clause” containing the pronoun or anaphor, but not so in the b or c cases. In the a-cases, the pronoun or anaphor is in a nominative Case marked position, but not so in the b or c cases. In the a-cases, the “clause” containing the pronoun or anaphor contains tense. And, finally, in the b and c cases, the surface position of the embedded subject is the accusative Case-marked position in the higher clause. The first three of these differences have been exploited at one point or another by Chomsky. In *Constraints* he uses the difference in Tense, and proposes what he called the Tensed S Condition, which we might frame for the special case of Binding Theory in the following way.

(26) THE TENSED S CONDITION

A Binding Domain for  $\alpha$  is the smallest phrase that contains  $\alpha$  and Tense.

But in *On Binding*, he notes that this screws up in cases like the (27).

(27) The women<sub>1</sub> believe that [each other<sub>1</sub>’s pictures]<sub>2</sub> are on sale.

And opts for exploiting the difference in Case. The condition he formulates he calls the Nominative Island Constraint, which we can formulate as follows.

(28) THE NOMINATIVE ISLAND CONSTRAINT

An anaphor cannot be assigned Nominative Case.

In *Lectures on Government and Binding*, he uses the difference in agreement morphology. If we are to exploit the fourth difference — that in the b and c cases, the accusative subject is in the higher clause, we will want to refit the definition of Binding Domain so that it is no longer sensitive to the c-command restriction on Subjects, and this will require a rethinking of cases like (20).

Before we do that, however, let’s revisit the question we began with: what is the nature of the “silence” of the subjects understood to stand in (certain) infinitival clauses? Is it merely phonetically absent, and there is otherwise something present syntactically? Or is it absent both phonetically and syntactically and present only in the interpretation? With the eyes the Binding Theory has provided us, consider (29) - (31).

- (29) a. The woman<sub>1</sub> promised Martha<sub>2</sub> [PRO<sub>1</sub> to examine herself<sub>1,\*2</sub>].
- b. The woman<sub>1</sub> promised Martha<sub>1</sub> [PRO<sub>1</sub> to examine her<sub>\*1,2</sub>].
- (30) a. The woman<sub>1</sub> told Martha<sub>2</sub> [PRO<sub>2</sub> to examine herself<sub>\*1,2</sub>].
- b. The woman<sub>1</sub> told Martha<sub>2</sub> [PRO<sub>2</sub> to examine her<sub>1,\*2</sub>].
- (31) a. The woman<sub>1</sub> wondered [how PRO<sub>2</sub> to examine \*herself<sub>1</sub>/oneself<sub>2</sub>].
- b. The woman<sub>1</sub> wondered [how PRO<sub>2</sub> to examine her<sub>1</sub>/\*one<sub>2</sub>].

What these examples show is that the understood subject of these infinitivals both has the ability to produce a Binding Domain and to provide a term against which pronouns and anaphors have their reference checked by Principles A and B. If we were to adopt the position that these subjects are not present in syntactic representations, but only present semantically, we would have to recast our principles here so that they made reference not to syntactic entities, but instead to semantic ones as well. Thus, for example, we might define a Binding Domain in terms not of subjects, but rather “external

$\theta$ -roles.” And we might similarly recast Principles A and B so that the referential dependencies they govern can be on terms that appear in the interpretation alone, rather than only on the syntactic representations which are later given an interpretation.

But both of these possible reformulations face difficult problems. The example in (32) might suggest that subjects need not have any semantic content, thus making a redefinition of Binding Domain along those lines difficult.

- (32) a. ?\* The women<sub>1</sub> decided that it seems to each other<sub>1</sub> that Martha has left.  
 b. The women<sub>1</sub> consider it obvious to them<sub>1</sub> that all primes are odd.

Here, the expletive *it* causes the clause embedded under *decide* and *consider* to be a Binding Domain for anaphors and pronouns therein. Because expletives have no semantic content, it therefore doesn’t look like we would want to frame the definition of Binding Domain so as to make it sensitive to whether there is a meaningful term which functions as Subject. Rather, what these examples suggest is that Binding Domains are those phrases which have something that functions as a syntactic subject in them.

And the phenomenon of Null Complements, as illustrated in (35), provide cases which suggest that antecedents for Principles A and B must be syntactically clothed.

- (33) a. Jerry<sub>1</sub> talks to one<sub>2</sub> about oneself<sub>2</sub>.  
 b. \* Jerry<sub>1</sub> talks – about oneself<sub>2</sub>.

In both of the examples in (33), the Goal of *talk* is part of the meaning of the sentence. In (33a), this Goal is expressed with the phrase *to one*; but in (33b) it goes unexpressed. We might imagine in these circumstances that the  $\theta$ -role Goal that *talks* has is simply optional. When that option is expressed, the Projection Principle (and Theta Criterion) will force it to be borne by an argument in the syntactic representation. But when that option is not expressed, the meaning of the verb will continue to involve a Goal, but it will go unexpressed in the syntax. In such cases we can say that the Goal argument is implicit. What the contrast in (33) suggests, then, is that Principle A can only be satisfied by arguments that get syntactic expression. The implicit argument of *talk* in (33b) cannot serve as an antecedent to a reflexive pronoun. So, if this is true of (33), then it suggests that there is something playing the role of subject in infinitival clauses, even when they are not visible.

Note, as an aside, that these examples pose a problem for our decision to express prominence in terms of c-command — a problem we have noted before. The problem with a c-command interpretation of “prominence” that arises in cases like (33) results from the presence of a PP dominating *Martha*. Some have suggested that cases such as these demand a definition of prominence that is defined in terms of the  $\theta$ -roles borne by the DPs whose referential relation is being judged. So, for example, Jackendoff (1972), and many subsequent approaches of this sort, place constraints on the antecedents of anaphors, requiring that the antecedent bear a  $\theta$ -role higher on some hierarchy than the anaphor. The problem with theories like these, however, is their generality. The relation of “prominence” that we are trying to define does not hold of coarguments, as seen for example from cases like “Mary believes herself to be intelligent,” and these theories can only get off the ground for cases where the two DPs are coarguments.

Another response to the problem of prominence posed by (33) is to not abandon a geometric definition of this relation, but to modify it so that this particular case is handled correctly. This basically results in a definition like our c-command, but with projections of P exempted.<sup>3</sup> Still another response has been to hold constant the definition of c-command and revise the parse that we give to cases like (33). Typically this maneuver treats the preposition as a Case particle, much like those found in Japanese, which is adjoined to the relevant DP. For concreteness, let's adopt this second response. A variant of this approach is the dramatically different complement structure that David Pesetsky advocates, and which we briefly reviewed in Chapter 8.

So we've seen that the syntactic presence of PRO can be discovered by the effect it has on the distribution of referential indices our theory uses to fix the referential dependencies among DPs. Hence, in addition to the other reasons we have posited PRO, we now have this one. As can be seen, there are a cluster of properties, then, whose correlation must be explained: apparent exception to the Theta Criterion, apparent exception to the EPP, and, finally, the apparent exception to these two properties of the binding theory. PRO is one explanation for the fact that the subject position of certain infinitives are exceptions to all of these generalizations.

The formulation of "Binding Domain" that is used in this theory is built on the premiss that coreferent pronouns and bound anaphors are in complementary distribution. We should never find that a coreferent pronoun and a bound anaphor can stand in the same structural position and have the same antecedent. The thesis expressed in Principle A and B is that the Binding Domain that determines the locality condition found with pronouns is the same as that found for anaphors; and this derives their complementary distribution. However, as we've seen, there are a narrow range of cases where this complementary distribution does not hold. Most of these situations involve places where we used the c-command restriction on Subjects to disqualify otherwise candidate Binding Domains. Some examples follow.

- (34) a. They<sub>1</sub> like their<sub>1</sub> pictures.  
 b. They<sub>1</sub> like each other<sub>1</sub>'s pictures.  
 c. They<sub>1</sub> said that pictures of each other<sub>1</sub> had arrived.  
 d. The said that pictures of them<sub>1</sub> had arrived.

We don't wish to abandon a theory that accounts for the near complementary distribution of bound pronouns and anaphors; that is, we would like to have an explanation for why it's this particular class of examples where they fail to have the same Binding Domain. One way of approaching this goal is to maintain the thesis that the Binding Domains are defined in the same way for Principles A and B, but introduce context sensitivity to the definition so that it is enabled to distinguish this case and apply Principles A and B differently. James Huang in a squib in *Linguistic Inquiry*, expanded in the fifth chapter of his dissertation, expresses an idea of this sort; I will discuss Chomsky's formulation of the idea, found in the *Knowledge of Language* reading.

This theory is quite a shift from the Specified Subject Constraint-based accounts we've been discussing. Indeed, it abandons the reference to subject altogether, and sets out a different criterion for being a potential Binding Domain. The central idea

<sup>3</sup>See Lasnik (1976) for one such definition.

it expresses – and this is Huang’s contribution – is that the Binding Domain increases up to the smallest phrase in which the pronoun or anaphor’s requirements could be satisfied. It explains why the environments in which the Binding Domains differ for pronouns and anaphors are just those like (34) since these are the very environments in which the differing requirements of pronouns and anaphors become salient. These are environments in which — to use the language of the Binding Theory that is to be replaced — the Subject that would create the Binding Domain fails to c-command the term whose Binding Domain is being computed. This matters, it seems, for anaphors, but not for pronouns. That is, the object DP in (34a,b) does not have a Subject that c-commands the pronoun or anaphor because, as it happens, the Subject IS the pronoun or anaphor. For this reason, the Binding Domain for the anaphor is a larger phrase. But for the pronoun, this is immaterial; the Binding Domain is this object DP no matter what the status of its Subject. Similar remarks hold for the case in (34c,d). Here too, the embedded clause has a Subject that does not c-command the pronoun or anaphor, in this case because it contains the anaphor or pronoun. This results in the Binding Domain being a larger phrase for the anaphor, but again it has no effect on the eligibility of the embedded clause as a Binding Domain for the pronoun. Why should anaphors care whether the Subjects making their Binding Domains c-command them, when pronouns don’t? Perhaps it is related to the fact that anaphors are looking for c-commanding antecedents, and pronouns aren’t. This is the central idea behind Huang’s proposal.

We can express Chomsky’s formulation of this idea in the following preliminary form.

- (35)  $\Psi$  is a BINDING DOMAIN for  $\alpha$  iff  $\Psi$  is the smallest phrase for which there is a Binding Theoretic Compatible (BT-Compatible) Indexing for  $(\Psi, \alpha)$ .

An indexing  $\mathcal{I}$  is BT-Compatible for  $(\Psi, \alpha)$  if

- i.  $\alpha$  is an anaphor and bound in  $\Psi$  under  $\mathcal{I}$ , or
- ii.  $\alpha$  is a pronoun and free in  $\Psi$  under  $\mathcal{I}$ .

Consider how this will capture the standard case in (36), and the previously problematic case in (37).

- (36) a. They<sub>1</sub> said that Mary likes each other<sub>1</sub>.  
       b. They<sub>1</sub> said that Mary likes them<sub>1</sub>.  
 (37) a. They<sub>1</sub> like their<sub>1</sub> pictures.  
       b. They<sub>1</sub> like each other<sub>1</sub>’s pictures.

The convention that terms don’t c-command themselves, the case in (37b) will require the Binding Domain to expand, since the DP which contains the genitive is able to satisfy the Binding Theoretic needs of the pronoun only.

A similar process, driven in this case by the anti-dominance clause in the definition of c-command, will account for the differing Binding Domains that pronouns and anaphors get in cases such as (38).

- (38) a. They<sub>1</sub> said that [<sub>IP</sub>[pictures of each other<sub>1</sub>]<sub>2</sub> were on sale].  
 b. They<sub>1</sub> said that [<sub>IP</sub>[pictures of them<sub>1</sub>]<sub>2</sub> were on sale].

In (38a), the embedded clause does not constitute a Binding Domain for the anaphor because indexing the anaphor with the subject in this clause (giving it the index “2”) would not cause it to be bound. Thus, only the root clause constitutes a Binding Domain for the anaphor.

The theory has a number of difficulties. The first is exemplified by cases like (39).

- (39) They<sub>1</sub> introduced Mary to each other<sub>1</sub>.

By (35), we should expect the VP to constitute a Binding Domain for the *each other* in this example. Chomsky proposes that there is a separate and additional constraint on which phrases may be Binding Domains. He defines it in terms of “Grammatical functions”; I’ll render it here in terms of  $\theta$ -roles.

- (40)  $\Psi$  is a BINDING DOMAIN for  $\alpha$  iff  $\Psi$  is the smallest Complete Functional Complex (CFC) in which there is a BT-compatible Indexing for  $(\Psi, \alpha)$ .  $\Psi$  is a CFC if:
- i. For all  $\theta$ -positions,  $\delta$ , in  $\Psi$ ,  $\delta$ ’s  $\theta$ -role assigner is in  $\Psi$ , and
  - ii. For all  $\theta$ -role assigners,  $\gamma$ , in  $\Psi$ , all positions  $\gamma$  assigns a  $\theta$ -role to are in  $\Psi$ .

The idea here is to restrict the phrases that have the potential to be Binding Domains to just those phrases that are in some sense, thematically complete. In this way, the VP in (39) is to be disqualified as a potential Binding Domain, since it does not include the subject  $\theta$ -marked position. Chomsky’s formulation of CFC predates our present syntax for external  $\theta$ -roles; in particular, it is not designed for the  $\nu$ P+VP series that we have put transitive clauses into. We should find another way of expressing this idea.

There are some other facts that should be considered in this connection. Certain PPs are not Binding Domains for either pronouns or anaphors; some examples are in (41) and (42).

- (41) a. \*They<sub>1</sub> introduced Mary to them<sub>1</sub>.  
 b. \*They<sub>1</sub> danced with them<sub>1</sub>.  
 c. \*She<sub>1</sub> talked about her<sub>1</sub>.  
 d. \*He<sub>1</sub> thought of him<sub>1</sub>.
- (42) a. They<sub>1</sub> introduced Mary to themselves<sub>1</sub>.  
 b. They<sub>1</sub> danced with themselves<sub>1</sub>.  
 c. She<sub>1</sub> talked about herself<sub>1</sub>.  
 d. He<sub>1</sub> thought of himself<sub>1</sub>.

These are all argument PPs, notice. When we look at canonical modifying type PPs, by contrast, we find that often they constitute Binding Domains for pronouns, though still not for anaphors. Compare the examples in (43) and (44) with those above.

- (43) a. Sam<sub>1</sub> pulled the covers over him<sub>1</sub>.  
 b. Jane<sub>1</sub> smeared the lotion all over her<sub>1</sub>.  
 c. Chris<sub>1</sub> scheduled the speaker just before him<sub>1</sub>.
- (44) a. Sam<sub>1</sub> pulled the covers over himself<sub>1</sub>.  
 b. Jane<sub>1</sub> smeared the lotion all over herself<sub>1</sub>.  
 c. Chris<sub>1</sub> scheduled the speaker just before himself<sub>1</sub>.

This is not a perfect correlation, however. There are locative argument PPs which behave like the modifiers in (43) and (44):

- (45) a. Sally<sub>1</sub> put the light next to her<sub>1</sub>.  
 b. Sally<sub>1</sub> placed the box on her<sub>1</sub>.
- (46) a. Jerry<sub>1</sub> put the light next to himself<sub>1</sub>.  
 b. Jane<sub>1</sub> placed the box on herself<sub>1</sub>.

The PPs that constitute Binding Domains for pronouns are candidates for being CFCs, whereas those which do not are apparently disqualified from being so.

The generalization, it seems to me, is that those PPs which are narrowly c-selected by the verb are not CFCs. For instance, the PPs in (45) and (46) can be replaced by any PP which refers to a location, and the sentences will behave the same.

- (47) a. Sally<sub>1</sub> put the light behind her<sub>1</sub>.  
 b. Sally<sub>1</sub> put the light above her<sub>1</sub>.  
 c. Sally<sub>1</sub> put the light beneath her<sub>1</sub>.  
 d. Sally<sub>1</sub> put the light in front of her<sub>1</sub>.  
 etc.

By contrast, there is no preposition that can be substituted for *to* in (41a).

- (48) a. I introduced Mary to myself.  
 b. \*I introduced Mary at myself.  
 c. \*I introduced Mary in myself.  
 etc.

There is some more intimate connection between the verb and the preposition in this case than there is in (47). I suggest that we equate this intimate connection with the equally mysterious intimate connection that holds between  $\nu$  and the VP it selects. It's over this mysterious connection that I suggest we define CFC.

Because I don't know how to make the mysterious intimate connection that is involved in these cases explicit – I don't know what it is – the definition of CFC will remain incomplete. Let's call the intimate connection that holds of verbs and certain PPs, and is also found between  $\nu$  and VP, " $\theta$ -linking." Then we can define CFC as in (49).



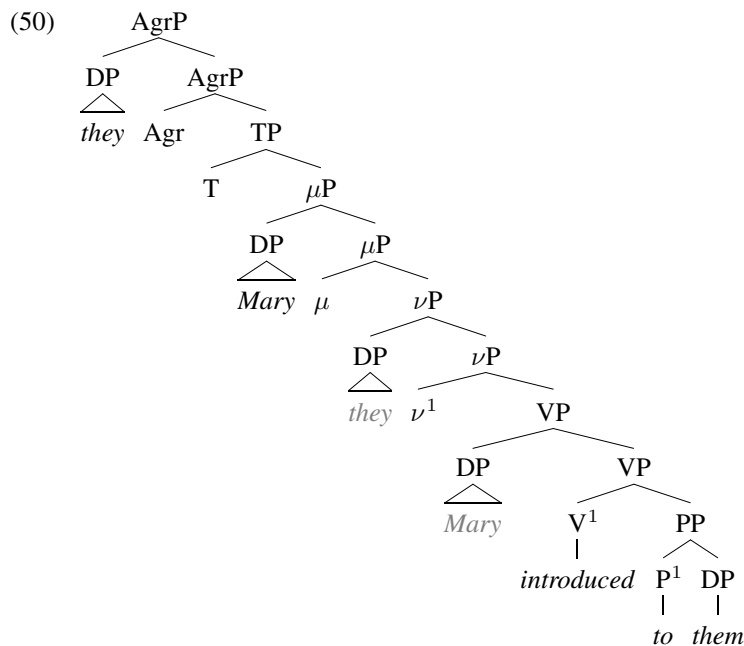
## (49) COMPLETE FUNCTIONAL COMPLEX

$\Phi$  is a Complete Functional Complex iff:

- i. For every  $\theta$ -marked term,  $\delta$ , in  $\Phi$ , all of the heads that are  $\theta$ -linked to the term that  $\theta$ -mark  $\delta$  are in  $\Phi$ , and
- ii. For every  $\theta$ -linked term,  $\gamma$ , in  $\Phi$ , all of the terms that are  $\theta$ -marked by something  $\gamma$  is  $\theta$ -linked to are in  $\Phi$ .

Let's consider, now, how this will apply to the cases we have looked at.

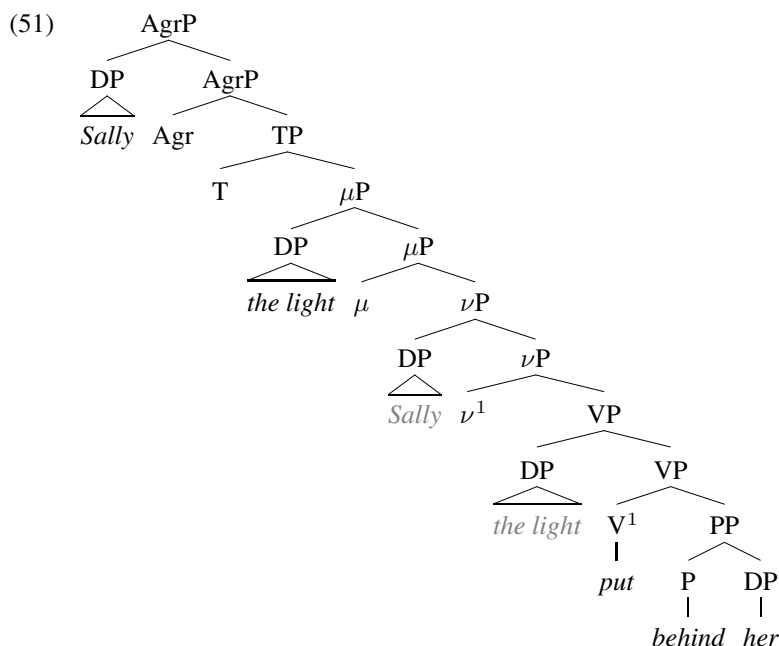
An example like (41a), which indicates that VPs are not Binding Domains, will have a representation like that in (50).<sup>4</sup>



Understand the co-superscripted heads to be  $\theta$ -linked. The smallest CFC in this graph is  $\nu P$ . There is no smaller phrase that contains all the arguments of the  $\theta$ -linked  $\nu$ +introduced+to cluster. In these situations, then, the CFC requirement prevents the definition of Binding Domain from picking out a phrase that would be too small.

Consider next how it will apply in situations in which a non- $\theta$ -linked PP is involved, such as (47a), whose surface parse is as in (51) on the following page. Because *behind* is not  $\theta$ -linked to *put* or  $\nu$ , the smallest CFC that contains *her* is the PP, *behind her*. And because a BT-compatible indexing is possible within this PP, this will also constitute the Binding Domain for *her*. As a consequence, *her* can corefer with *Sally*. If *her* were to be replaced with an anaphor – such as *herself* – then no BT-compatible Indexing will arise within the PP, and as a consequence, this PP will not be a Binding Domain for an anaphor within it, and the next smallest CFC will

<sup>4</sup>This parse does not reflect Verb movement.



be considered. The next smallest CFC is  $\nu P$ , and because there is a BT-compatible Indexing for an anaphor in this phrase, it will be that anaphor's Binding Domain.

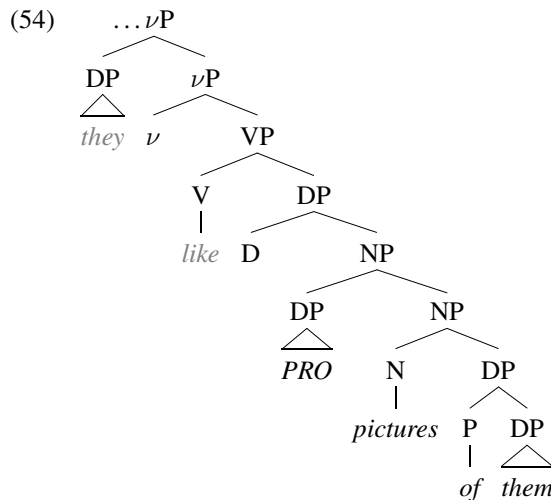
The largest problem for this approach is that it cannot distinguish (52) from (53).

- (52) a. They<sub>1</sub> like each other<sub>1</sub>'s pictures.  
 b. They<sub>1</sub> like their<sub>1</sub> pictures.
- (53) a. They<sub>1</sub> like pictures of each other<sub>1</sub>.  
 b. ?? They<sub>1</sub> like pictures of them<sub>1</sub>.

Because Binding Domains are no longer sensitive to “subjects,” this subject/non-subject asymmetry is no longer captured. How is this contrast to be accounted for?

Chomsky makes use here of the idea that there is a hidden PRO in the Specifier of these DPs. If we suppose that *of* is a Case prefix, then (53b) gets a parse like that in (54) on the next page.<sup>5</sup> Recall that we are treating *of* as a kind of object Case mark on DPs within NPs. As a consequence, *of them* is not a PP, and not a potential Binding Domain as a consequence. In fact the smallest CFC that contains *of them* is the NP headed by *pictures*, and this will be its Binding Domain as well. As a consequence, Principle B will prevent *of them* from bearing the same index as PRO. Chomsky suggests that this is in fact the correct outcome, and that the disjoint reference effect we

<sup>5</sup>The parse of the object DP in (54) ignores most of the detail we learned about the internal structure of DPs. In this case, the details might, in fact, matter — it will depend on how a variety of assumptions about the underlying position of “subject” arguments are positioned within DPs. Nonetheless, I think the points I will make here remain valid.



perceive is not between the pronoun and the higher subject, but rather with PRO, and PRO is taken to be coreferent with the subject of the higher verb. As support for this interpretation of the facts, he cites the contrast in (55).

- (55) a. ?? They<sub>1</sub> told rumors about them<sub>1</sub>.  
 b. They<sub>1</sub> heard rumors about them<sub>1</sub>.

Presumably it is easier to understand the rumors that someone tells to “belong” to that teller; whereas rumors one hears are more likely to belong to some other person. If we understand the hidden PRO in these DPs to refer to the possessor of the NP, then these contrasts are accounted for. Note that (55a) improves if it is understood in such a way that rumors being told do not “belong” to the tellers. That is, if it has a meaning closer to “They recounted rumors about them.”

But is there any reason to believe that DPs can host PRO, as we have claimed infinitival clauses to? We might not expect this possibility, given the “invisible” Case account of where PRO can appear. In fact, however, there is some evidence that PRO can reside within DPs.

Recall that one of the reasons we were led to posit PRO in infinitival clauses is because it offered a simple way of maintaining the generalization that the Theta Criterion/Projection Principle expresses when confronted with examples like (56).

- (56) They<sub>1</sub> tried [PRO<sub>1</sub> to leave].

This reasoning, however, does not seem extendable to DPs; indeed, the Theta Criterion/Projection Principle doesn’t appear to extend to DPs. There are many cases which make it appear that nouns do not have to assign their  $\theta$ -roles.

- (57) a. She examined the patients  
 \*— examined the patients  
 \*She examined —

- b. Her examination of the patients was swift.  
the examination was swift.
- c. I assigned unsolvable homework problems.  
\*— assigned unsolvable homework problems.  
\*I assigned —.
- d. My assignment of unsolvable homework problems has been noted upon.  
the assignment has been noted upon.
- e. Jerry expressed his feelings.  
\*— expressed his feelings.  
\*Jerry expressed —.
- f. Jerry's expression of his feelings surprised us.  
The expression surprised us.

But recall that the noun shifts slightly in meaning depending on whether or not it combines with arguments; this is clearest, perhaps, in (57d) and (57f). Lees (1963) called the argument taking versions of these nouns “process” nominals, and those that don't express arguments, “result” nominals. There are ways of disambiguating these nouns in the direction of the “process” reading. When they are so disambiguated, Grimshaw (1990) argues that they are required to express their arguments. One way of doing this is to give these nominals an Agent, as in (58).

- (58) a. Her intentional examination of the patient was swift.  
\*Her intentional examination was swift.
- b. My intentional assignment of unsolvable homework problems has been noted upon.  
My intentional assignment has been noted upon.
- c. Jerry's deliberate expression of his feelings surprised us.  
Jerry's deliberate expression surprised us.

Another is to modify these nouns with adjectives which select the process meaning, as in (59).

- (59) a. her frequent examination of the patient
- b. \* The frequent examination
- c. my constant assignment of unsolvable homework
- d. \* the constant assignment
- e. Jerry's sudden expression of his feelings
- f. \* the sudden expression

Now, interestingly, when the process reading is selected, it appears that only the object argument becomes obligatorily expressed:

- (60) a. the frequent examination of the patient
- b. the constant assignment of unsolvable homework
- c. the sudden expression of one's feelings

So, with respect to process nominals — the ones where the Theta Criterion/Projection Principle does seem to hold — we do have an argument for PRO in DPs like the one we had for infinitival clauses.

Roeper makes another argument for the existence of PRO within DPs, that does not rely on the process reading.<sup>6</sup> This one is based on the need rationale clauses have for a “controller,” as we saw in the contrast below.

- (61) a. The ship was sunk [PRO to collect the insurance].  
b. \* The ship sank [PRO to collect the insurance].

Rationale clauses seem to find a controller in apparently subjectless DPs:

- (62) a. the destruction of the city [PRO to prove a point]  
b. the review of a book [PRO to prove a point]

That the controller in these examples is, as expected, in the Specifier of NP is suggested by the contrast this example has with:

- (63) a. \* the city’s destruction [PRO to prove a point]  
b. \* a book’s review [PRO to prove a point]

There can’t be a PRO in the Specifier of NP in these examples because it’ll compete for this position with the overt possessive.

Finally, we might note that Principle A effects also suggest that DPs can have a PRO in them.<sup>7</sup>

- (64) a. \* Oneself’s purchase of these drugs is not recommended.  
b. The purchase of these drugs for oneself is not recommended.  
c. \* Their purchase for oneself is not recommended.

And Principle C effects also point in the same direction, witness the examples in (65).<sup>8</sup>

- (65) a. \* To realize that John<sub>1</sub> is unpopular upsets him<sub>1</sub>.  
b. \* The realization that John<sub>1</sub> is unpopular upsets him<sub>1</sub>.

We learn from (65a) that the PRO subject of an infinitive in this position favors being controlled by the object, here *him*. As a consequence, when *John* is coindexed with *him*, so will PRO be. but if PRO and *John* are coindexed, we get a Principle C violation. Because (65b) behaves in precisely the same way, it would appear that there is a PRO subject within the subject of this example as well.

So, perhaps it is safe to assume that DPs can have a PRO subject in them. On the Case based approach to the distribution of PRO, this entails that there must be an “invisible” Case assigner within DPs.

But even if DPs can indeed hold PRO, I’m not sure this can be successfully exploited to explain the weakish disjoint reference effect in (53b). The problem, first in noted Williams (1985), is that we should expect PRO to be optional in these cases, as it must be for (66) to work as it does.

<sup>6</sup>See Roeper (1985, 1987).

<sup>7</sup>These examples, and their relevance to locating PRO within DPs, are from Giorgi and Longobardi (1991) — their examples are Italian, but English works the same way.

<sup>8</sup>These examples are from Ross (1967).

(66) They<sub>1</sub> told me<sub>2</sub> rumors about myself<sub>2</sub>.

And if it is optional, there is no way to make it responsible for the disjoint reference effects it has been blamed for. So let's abandon this solution to the problem, and hope that another emerges.

The definition of Binding Domain I've given in (40) isn't the one that Chomsky gives, and doesn't actually express Huang's idea. The definition expresses the notion that the Binding Theoretic needs of the relevant term are what's governing how that term's Binding Domain is figured by simply copying Principles A and B in. Thus, Principles A and B are stated independently as conditions on pronouns and anaphors, and as ingredients in computing Binding domains for pronouns and anaphors. Their statement is not necessarily linked, and this formulation therefore permits the conditions on anaphors and pronouns to differ, say, from that we find for English while the conditions that govern the determination of their Binding Domains remains the same. This isn't the idea behind Huang's proposal. Chomsky's formulation of the Binding Theory links Principles A and B to their Binding Domains. He does this by causing the Binding Principles to hold of indexings rather than on pronouns and anaphors. That is, he converts Principles A and B into conditions on indexings rather than conditions on pronouns and anaphors. Then the conditions on anaphors and pronouns (what we formerly called Principles A and B) point to these indexings. Still paraphrasing Chomsky somewhat, what this conversion comes to is (67).

- (67) An indexing  $\mathcal{I}$  is BT-compatible for  $(\Phi, \alpha)$  if
- i.  $\alpha$  is an anaphor and bound in  $\Phi$  under  $\mathcal{I}$ , or
  - ii.  $\alpha$  is a pronoun and free in  $\Phi$  under  $\mathcal{I}$ .

For some  $\Phi$  such that (i) or (ii),  $\mathcal{I}$  must be BT-compatible with  $(\Phi, \alpha)$ :

- i.  $\alpha$  is an anaphor or pronoun and  $\Phi$  is the smallest CFC for which there is an indexing BT-compatible with  $(\Phi, \alpha)$ ,
- ii. If not (i), then  $\Phi$  is the root clause containing  $\alpha$ .

(Understand CFC's to be defined as before.) As in the LGB Binding Theory, the clause in (ii) is required to guarantee that the anaphor has a Binding Domain in (68).

(68) Each other's pictures arrived.

This, recall, flows from the decision to express the requirement that anaphors require antecedents in the same statement that voices the various requirements about which terms can be their antecedent. An alternative would be to see the condition that anaphors have antecedents as being independent of the definition of Binding Domain — maybe flowing, for instance, from their semantic interpretation.

# Bibliography

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral Dissertation, Massachusetts Institute of Technology.
- Aissen, Judith. 1974. Verb raising. *Linguistic Inquiry* 5:325–66.
- Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–416.
- Baker, Mark. 1988. *Incorporation*. Chicago, Illinois: University of Chicago Press.
- Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13:1–38.
- Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17:347–54.
- Beck, Sigrid, and William Snyder. 2001. The resultative parameter and restitutive ‘again’. In *Auditor vox Sapientiae: a festschrift for Arnim von Stechow*, ed. Caroline Fery and Wolfgang Sternefeld, 48–69. Berlin: Akademie Verlag.
- Belletti, Adriana. 1982. On the anaphoric status of the reciprocal construction in Italian. *The Linguistic Review* 2:101–38.
- Belletti, Adriana. 1990. *Generalized verb movement*. Rosenberg & Sellier.
- Bernstein, Judy. 1991. DPs in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus* 3:101–26.
- Bernstein, Judy B. 1993. Topics in the syntax of nominal structure across romance. Doctoral Dissertation, City University of New York.
- den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In *On the Formal Syntax of the Westgermania*, ed. W. Abraham. Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Bošković, Željko. 1996. Selection and the categorial status of infinitival complements. *Natural Language and Linguistic Theory* 14:269–304.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Massachusetts: MIT Press.

- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.
- Burzio, Luigi. 1986. *Italian syntax*. Reidel Publishers.
- Chierchia, Gennaro. 1984. Topics in the syntax and semantics of infinitives and gerunds. Doctoral Dissertation, University of Massachusetts, Amherst.
- Chomsky, Noam. 1957. *Syntactic structures*. Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Chomsky, Noam. 1977a. Conditions on transformations. In *Essays on form and interpretation*, 81–162. New York, New York: Elsevier North-Holland, Inc.
- Chomsky, Noam. 1977b. *Essays on form and interpretation*. Amsterdam, The Netherlands: North-Holland Publishing Company.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.
- Chomsky, Noam. 1986a. *Barriers*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1986b. *Knowledge of language*. New York, New York: Praeger Publishers.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417–54. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1992. *A minimalist program for linguistic theory*, volume 1. MIT, Cambridge Massachusetts: MIT Working Papers in Linguistics.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row, Publishers.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.
- Contreras, Heles. 1984. A note on parasitic gaps. *Linguistic Inquiry* 15:698–701.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press.
- Diesing, Molly, and Eloise Jelinek. 1995. Distributing arguments. *Natural Language Semantics* 3:123–76.
- Dryer, Matthew S. 1989. Plural words. *Linguistics* 27:865–95.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68:81–138.



- Emonds, Joseph. 1972. Evidence that indirect object movement is a structure preserving rule. *Foundations of Language* 8:546–61.
- Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York, New York: Academic Press.
- Fiengo, Robert, and Robert May. 1994. *Indices and identity*. Cambridge, Massachusetts: MIT Press.
- Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3:283–341.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.
- Gerdts, Donna. 1981. Object and absolutive in Halkomelem Salish. Doctoral Dissertation, University of California, San Diego.
- Gibson, Jean. 1980. Clause union in Chamorro and in universal grammar. Doctoral Dissertation, University of California, San Diego.
- Giorgi, Alessandra, and Giuseppe Longobardi. 1991. *The syntax of noun phrases*. Cambridge: Cambridge University Press.
- Giusti, Giuliana. 1990. Floating quantifiers, scrambling and configurationality. *Linguistic Inquiry* 21:633–41.
- Green, Georgia M. 1974. *Semantics and syntactic regularity*. Bloomington, Indiana: Indiana University Press.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.
- Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Massachusetts: MIT Press.
- Gruber, Jeffrey S. 1965. Studies in lexical relations. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Hale, Ken, and Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, Massachusetts: MIT Press.
- Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The view from building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 53–110. Cambridge, Massachusetts: MIT Press.
- Hardt, Dan. 1993. Verb phrase ellipsis: Form, meaning and processing. Doctoral Dissertation, University of Pennsylvania.
- Harris, Zellig. 1946. From morpheme to utterance. *Language* 22:161–83.
- Heim, Irene. 1997. Predicates or formulas? evidence from ellipsis. Cornell University: CLC Publications.

- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.
- Higginbotham, James. 1983. Logical form, binding, and nominals. *Linguistic Inquiry* 14:395–420.
- Hornstein, Norbert. 1994. An argument for minimalism: The case of antecedent-contained deletion. *Linguistic Inquiry* 25:455–80.
- Hornstein, Norbert. 1995a. *Logical Form: from GB to Minimalism*. Cambridge, Massachusetts: Basil Blackwell.
- Hornstein, Norbert. 1995b. Putting truth into Universal Grammar. *Linguistics and Philosophy* 18:381–400.
- von Humboldt, Wilhelm. 1836. *Über die verschiedenheit des menschlichen sprachbaues und ihren einfluss auf die geistige entwicklung des menschengeschlechts*. Berlin: Royal Academy of Sciences of Berlin.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, Ray. 1977. *X' syntax*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, Ray. 1990. On Larson's treatment of the double object construction. *Linguistic Inquiry* 21:427–55.
- Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht Holland: Foris Publications.
- Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9:577–636.
- Kayne, Richard S. 1975. *French syntax: the transformational cycle*. Cambridge, Massachusetts: MIT Press.
- Kayne, Richard S. 1984a. *Connectedness and binary branching*. Dordrecht, Holland: Foris Publications.
- Kayne, Richard S. 1984b. Unambiguous paths. In *Connectedness and Binary Branching*, 129–64. Dordrecht: Foris Publications.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Kayne, Richard S. 1998. Overt vs. covert movement. *Syntax* 1.
- Kayne, Richard S. 2000. *Parameters and universal grammar*. Oxford: Oxford University Press.
- Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1:111–36.

- Koster, Jan. 1978. Why subject sentences don't exist. In *Recent Transformational Studies in European Languages*, ed. Samuel Jay Keyser. Cambridge, Massachusetts: MIT Press.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–37. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–92.
- Larson, Richard. 1990. Double objects revisited: reply to Jackendoff. *Linguistic Inquiry* 21:589–632.
- Larson, Richard, and Robert May. 1990. Antecedent containment or vacuous movement: reply to Baltin. *Linguistic Inquiry* 21:103–22.
- Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2:1–22.
- Lasnik, Howard. 1981. Restricting the theory of transformations. In *Explanation in Linguistics*, ed. Norbert Hornstein and David Lightfoot. London: Longmans.
- Lees, Robert. 1963. *The grammar of english nominalizations*. The Hague: Mouton.
- Levin, Nancy. 1986. *Main-verb ellipsis in spoken English*. New York: Garland Publishing, Inc.
- May, Robert. 1977. The grammar of quantification. Doctoral Dissertation, Massachusetts Institute of Technology.
- May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, Massachusetts: MIT Press.
- Merchant, Jason. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Morzycki, Marcin. 2002. Mediated modification: Functional structure and the interpretation of modifier position. Doctoral Dissertation, Graduate Linguistics Students Association, University of Massachusetts at Amherst.
- Muysken, Pieter. 1979. Quechua causatives and logical form: A case study in markedness. In *Theory of Markedness in Generative Grammar*, ed. Adrianna Belletti, Luigi Brandi, and Luigi Rizzi. Scuola Normale Superiore, Pisa.
- Muysken, Pieter. 1981. Quechua word structure. In *Binding and Filtering*, ed. Frank Heny. Cambridge, Massachusetts: MIT Press.
- Muysken, Pieter. 1982. Parametrizing the notion 'head'. *Journal of Linguistic Research* 2:57–75.
- Pesetsky, David. 1982. Paths and categories. Doctoral Dissertation, Massachusetts Institute of Technology.

- Pesetsky, David. 1995. *Zero syntax*. Cambridge, Massachusetts: MIT Press.
- Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20:365–424.
- Postal, Paul M. 1969. On so-called ‘pronouns’ in English. In *Modern Studies in English*, ed. D. Reibel and Sanford Schane. Englewood Cliffs, New Jersey: Prentice-Hall.
- Pustejovsky, James, and Peter Sells, ed. 1982. *VP deletion and across-the-board quantifier scope*. GLSA.
- Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago, Illinois: University of Chicago Press.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: MIT Press.
- Roberts, Ian. 1987. *The representation of implicit and dethematized subjects*. Dordrecht, Holland: Foris Publications.
- Rochemont, Michael, and Peter W. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Roeper, Thomas. 1985. Copying implicit arguments. 273–83. Stanford: Stanford Linguistics Association.
- Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18:267–310.
- Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.
- Ross, John. 1974. Nouniness. In *Three Dimensions of Linguistic Theory*, ed. O. Fujimura, 137–257. Tokyo: TEC Corporation.
- Runner, Jeffrey. 1995. Noun phrase licensing and interpretation. Doctoral Dissertation, University of Massachusetts at Amherst, Amherst.
- Safir, Kenneth. 1987. The syntactic projection of lexical thematic structure. *Natural Language and Linguistic Theory* 5:561–601.
- Sag, Ivan. 1980. *Deletion and logical form*. New York: Garland Publishing.
- Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. Doctoral Dissertation, Massachusetts Institute of Technology.
- Saito, Mamoru, and Naoki Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29:439–74.
- Sapir, Edward. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace and Co.

- von Stechow, Arnim. 1996. The different readings of *wieder* "again": A structural account. *Journal of Semantics* 13:87–138.
- Stowell, Timothy. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.
- Tomioka, Satoshi. 1997. Focussing effects in VP ellipsis and NP interpretation. Doctoral Dissertation, University of Massachusetts, Amherst.
- Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral Dissertation, Massachusetts Institute of Technology.
- Valois, Daniel. 1991a. The internal syntax of DP. Doctoral Dissertation, University of California, Los Angeles.
- Valois, Daniel. 1991b. The internal syntax of DP and Adjective placement in French and English. 367–82. Université du Québec à Montréal: The Graduate Students Linguistic Association, The University of Massachusetts-Amherst.
- Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11:203–38.
- Williams, Edwin. 1985. PRO and subject of NP. *Natural Language and Linguistic Theory* 3:297–316.
- Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer Academic Publishers.